

SPECT imaging
and
Automatic Classification Methods
in
Movement Disorders

David John Towey

A thesis submitted for the degree of Doctor of Philosophy

Institute of Clinical Science

Imperial College London

I hereby declare that this thesis and the work described within, is entirely my own work except where specifically stated or referenced.

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

This work investigates neuroimaging as applied to movement disorders by the use of radionuclide imaging techniques. There are two focuses in this work:

- The optimisation of the SPECT imaging process including acquisition and image reconstruction.
- The development and optimisation of automated analysis techniques

The first part has included practical measurements of camera performance using a range of phantoms. Filtered back projection and iterative methods of image reconstruction were compared and optimised. Compensation methods for attenuation and scatter are assessed.

Iterative methods are shown to improve image quality over filtered back projection for a range of image quality indexes. Quantitative improvements are shown when attenuation and scatter compensation techniques are applied, but at the expense of increased noise.

The clinical acquisition and processing procedures were adjusted accordingly.

A large database of clinical studies was used to compare commercially available DaTSCAN quantification software programs.

A novel automatic analysis technique was then developed by combining Principal Component Analysis (PCA) and machine learning techniques (including Support Vector Machines, and Naive Bayes).

The accuracy of the various classification methods under different conditions is investigated and discussed.

The thesis concludes that the described method can allow automatic classification of clinical images with equal or greater accuracy to that of commercially available systems.

Table of Contents

List of figures.....	5
List of Tables	9
Acknowledgements	10
Published work.....	11
Glossary of terms	12
1. Introduction	15
1.1 Parkinsonian syndromes, clinical presentations and neurophysiological basis..	16
1.2 Parkinson's – Plus syndromes – presentation, clinical features and neurophysiological basis	17
1.3 Other neurological disorders.....	19
1.4 Neurotransmitters.....	22
1.5 Imaging of movement disorders and associated conditions	24
1.6 Imaging with other modalities and non-imaging techniques.....	30
1.7 SPECT imaging	30
1.8 Image reconstruction	37
1.9 Image analysis and classification	38
1.10 Aims of this work.....	43
2 Acquisition Parameters	44
2.1 Gamma camera manufacturer/model.....	46
2.2 Energy windows.....	46
2.3 Comparison of LEHR, VXGP and MEGP Collimators	50
2.4 Pixel size / Image Matrix	58
2.5 SPECT gantry orbits	60
2.6 SPECT Gantry movement.....	62
2.7 The number of projections.....	63
2.8 Time per projection	64
2.9 Conclusions.....	64
3 Image reconstruction and correction techniques	66
3.1 Filtered Back Projection.....	68
3.2 Iterative reconstruction	70
3.3 Image corrections	72
3.4 Image quality metrics.....	81
3.5 The choice of images for clinical visual reporting.....	86
3.5.1 The Choice of images for clinical visual reporting- Methods.....	86
3.5.2 The choice of images for clinical visual reporting - results.....	87
3.6 Practical investigations of the properties of reconstruction models using phantom data	87
3.7 Results: Phantom experiments with FBP.....	96
3.8 Results: Phantom experiments with OSEM.....	111
3.9 Compensation methods for attenuation.....	132
3.10 Compensation methods for scatter - discussion.....	138
3.11 Overview of reconstruction results.....	140
4 Clinical Data	143
5 Development of a PCA based classification system	148
5.1 Feature Extraction and machine learning	148
5.2 Machine learning and classification methods	151

5.3	Image processing and preparation.....	158
5.4	Classification models.....	168
5.5	Confounding features.....	173
5.6	Developing and Testing the PCA based classification technique.....	174
5.7	Developing and optimising the SVD/PCA based system using clinical dataset B - methods.....	181
5.8	Overview of the SVD and machine learning systems.....	197
6	Commercially Available Methods for Analysing DAT SPECT imaging.	200
6.1	Manually placed ROI/VOI techniques.....	200
6.2	Automatically Placed ROI/VOI Techniques.....	203
6.3	Size and Shape Analysis.....	206
6.4	Fractal Analysis.....	206
6.5	Voxel-by-voxel Analysis – Statistical Parametric Mapping.....	207
6.6	The BRASS package.....	209
6.7	The QuantiSPECT packages.....	214
6.8	Comparison of datasets.....	222
6.9	Reconstruction setting comparisons.....	222
6.10	Method comparisons - sensitivity and specificity.....	223
6.11	Variation with Age.....	225
6.12	Variation with age – discussion.....	227
6.13	Commercially available methods – discussions.....	228
7	Comparison of Classification systems.....	230
7.1	Overall accuracy.....	230
7.2	Variation with image smoothness.....	230
7.3	Statistical and clinical significance of these results.....	231
7.4	Overview and Conclusions.....	232
7.5	Future work:.....	233
	Bibliography.....	235
	Appendix.....	246

List of figures

Figure 1-1. Schematic diagram of synapse between two neurons. Vesicles in pre-synaptic neuron release a neurotransmitter (in this case Dopamine) into the cleft.....	22
Figure 1-2. Synthesis of Dopamine from Tyrosine	23
Figure 1-3 Example transverse slice from ¹²³ I-FP-beta-CIT SPECT scan with corresponding slice from T2weighted MRI.....	26
Figure 1-4 Schematic diagram of gamma camera.....	31
Figure 1-5 Schematic diagram of collimator	32
Figure 1-6. Emission probabilities of 123-Iodine. Principal gamma energy 158.99keV omitted for clarity. Uncertainties in the published data too small to be displayed.....	36
Figure 1-7 A subject positioned ready for a brain SPECT acquisition.....	37
Figure 1-8 The patient positioning as seen through the gantry	37
Figure 1-9 A series of transverse slices from a normal 123I-FP-B-CIT SPECT scan	40
Figure 2-1 Results of a Monte Carlo simulation. Showing the number of detected photons as a function of energy.....	48
Figure 2-2 The energy acceptance windows compared to the energy of direct and scattered photons.....	49
Figure 2-3 Resolution as a function of distance from the collimator for three collimators	53
Figure 2-4 FWTM Resolution as a function of distance from the collimator for three collimators	54
Figure 2-5 Resolution as a function of distance from collimator in scattering material. The FWHM and FWTM are shown with and without TEW scatter correction.	54
Figure 2-6 Line profiles across line sources at various depths in scattering material.....	55
Figure 2-7 Relative sensitivity with distance for photopeak window and TEW corrected photopeak window	56
Figure 2-8. The radius of rotation as measured for a phantom experiment. The scaled phantom image is also included.....	61
Figure 3-1 Projection images and sinogram data.....	67
Figure 3-2 Example Butterworth filters.....	69
Figure 3-3. Example central transverse slice reconstructed using OSEM with depth dependent resolution recovery and a range of iterations and subsets. The numbers of equivalent iterations (iterations*subsets) are 10, 16, 20, 24, 40, 64, 96, 128 and 200 from left to right.	71
Figure 3-4 Four possible photon paths	72
Figure 3-5 A central slice from a number of reconstructions using FBP and a range of Butterworth filters of order 10. From left to right, f_{co} =0.5, 0.7, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8 and 2.0 cycles /cm. Data from Phantom run number 2, at 13cm ROR.....	91
Figure 3-6 Transverse slice through the ideal count distribution (left) and an example reconstruction (right)	94

Figure 3-7 Two line profiles drawn across a transverse slice with corresponding profile chart displaying counts per pixel as a function of pixel number.....	96
Figure 3-8 Contrast as a function of the cut-off frequency for FBP reconstructed phantom data	97
Figure 3-9 Recovery coefficient as a function of cut-off frequency. The variation between recovery coefficients is shown on the secondary axis.....	98
Figure 3-10 Noise in the background uniform section of the phantom, as a function of the filter cut-off frequency.....	98
Figure 3-11 Contrast as a function of noise for FBP reconstructions.....	99
Figure 3-12 Recovery coefficient as a function of noise	99
Figure 3-13 Signal to noise ratio as a function of cut-off frequency for FBP reconstructed data.....	100
Figure 3-14 A central transverse slice reconstruction using FBP and a range of frequency cut-offs (from left to right, 0.5, 0.8, 1.2, 1.6 and 2.0 cycles /cm).	101
Figure 3-15 NMSE, SNR and Noise as a function of filter cut-off frequency for FBP reconstructed data.....	103
Figure 3-16 Line profiles of FBP reconstructed data using a range of filter cut-off frequencies.....	105
Figure 3-17 NMSE calculated for the line profiles as a function of the cut-off frequency.....	106
Figure 3-18 Line profile metrics as a function of reconstruction filter	107
Figure 3-19 FBP contrast as a function of acquisition radius of rotation	109
Figure 3-20 Contrast as a function of image refreshes for different OSEM based reconstructions	112
Figure 3-21 Signal as a function of Noise for different Iterative reconstructions	113
Figure 3-22 Contrast as a function of noise for different OSEM based reconstructions with post-reconstruction filter	114
Figure 3-23 SNR as a function of image refreshes	115
Figure 3-24 contrast for different phantom runs using FBP and OSEM.....	118
Figure 3-25 recovery coefficient as a function of image refreshes or cut-off frequency.....	119
Figure 3-26 Contrast as a function of image refreshes for varying extent of resolution model.....	121
Figure 3-27 Central transverse slice from DDRR-OSEM reconstructed volume using wide Gaussian modelling (x3z2). The image on the right has been smoothed post-reconstruction Butterworth filter	123
Figure 3-28 Line Profile NMSE minimums for a range of iterative reconstructions and FBP with and without post-reconstruction filtering (F /NF)	124
Figure 3-29 Line profile NMSE scores for OSEM and FBP reconstructions	125
Figure 3-30 Line NMSE score for the Putamen region for iterative reconstructions using different sizes of resolution model. The lower chart includes a post-reconstruction smoothing filter.....	126
Figure 3-31 Line NMSE score for the background region for iterative reconstructions using different sizes of resolution model. The lower chart includes a post-reconstruction filter	127

Figure 3-32 Contrast measurements using different acquisition radii, as a function of cut-off filter used for FBP (top) and image refreshes for DRR-OSEM (bottom).....	129
Figure 3-33 Contrast as a function of image refreshes when different ROR distances are used.....	131
Figure 3-34 The position of points of interest within a brain outline used for attenuation effect calculations a string of points are shown in the left striatum along with two separate regions used for non-specific binding estimates.....	135
Figure 3-35 Elliptical outlines based on head size measurements	136
Figure 3-36 Calculation of the attenuation effect at three points.....	137
Figure 5-1 Example data for normally distributed population of 100 subjects, with mean of 0 and standard deviation of 1 for both axes. The circles represent 1, 2 and 3 standard deviations from the mean.	152
Figure 5-2. Two example data sets are shown along with a test patient in green. The group-prototype method is used here to classify the test patient in the normal (blue) group. The decision line for the Group prototype method is shown and is defined as the being equidistant from the two group centres.	153
Figure 5-3. Two example data sets are shown with dissimilar amount of variation in the two components. The concentric ellipses represent the extent of the 1 st , 2 nd and 3 rd standard deviations for the two groups.	155
Figure 5-4. The decision lines when using different prior probabilities in a Bayesian Classifier	156
Figure 5-5 Two groups being separated by three straight decision lines	157
Figure 5-6 The optimum decision line as picked by SVM the parallel lines either side show the so-called 'fat margin' the width of which is maximised in the SVM process.....	158
Figure 5-7 Six example slices from the first four components.....	177
Figure 5-8 The component weights for 115 patients for the first three components. A test patient projected into this feature space is also shown. PS is Parkinsonian syndrome and NPS is non-Parkinsonian	178
Figure 5-9 The Area Under Curve from ROC analysis SVD based classification	179
Figure 5-10 ROC analysis for Naive Bayes and Group prototype classifiers, using 3 components.....	180
Figure 5-11 Six Transverse slices through averaged image volumes of non-PS patients (top row) and PS patients (bottom row).....	180
Figure 5-12 The central slices of the locally derived ¹²³ I-FP-β-CIT template.....	184
Figure 5-13 central slice from local template and T1 MR Image	184
Figure 5-14 fraction of total variance described in each component for three types of image pre-processing	187
Figure 5-15 NB classifier accuracy for different levels of pre-processing.....	189
Figure 5-16 Naive Bayes classifier accuracy as a function of image filtering levels	190
Figure 5-17 Accuracy as a function of number of components	193
Figure 5-18 Naive Bayes analysis using Non-Symmetrical Classification	197
Figure 5-19 Maximum-Hemisphere aligned data classified using SVM, NB and GP systems.....	197
Figure 6-1 The three VOIs definition used in the QuantiSPECT software	200

Figure 6-2 Positioning and size of the VOIs used in the BRASS analysis package	204
Figure 6-3 Example misregistration of scan using the BRASS software	211
Figure 6-4 BRASS-Caudate uptake as a function of post reconstruction filter ...	212
Figure 6-5 BRASS-Putamen uptake ratio as a function of post-reconstruction filter	212
Figure 6-6 BRASS ROC analysis results as function of post reconstruction filter	213
Figure 6-7 QuantiSPECT positioning uncertainty	215
Figure 6-8 2-Box Striatal Binding Index as a function of the post reconstruction filter	216
Figure 6-9 2-Box ROC analysis as a function of post-reconstruction filter	216
Figure 6-10 3-Box TBPI as a function of post reconstruction filter setting	217
Figure 6-11 3-Box Asymmetry index as a function of post reconstruction filter	218
Figure 6-12 3-Box ROC analysis as a function of filter cut-off.....	218
Figure 6-13 Crescent method striatal uptake as a function of reconstruction filter	220
Figure 6-14 Crescent method putamen uptake as function of reconstruction filter	220
Figure 6-15 Putamen to Caudate ratio as a function of reconstruction filter	221
Figure 6-16 ROC results for the crescent method.....	221
Figure 6-17 AUC as a function of filter setting for a range of analysis techniques	223
Figure 6-18 Maximum achievable accuracy for each classification model as a function of filter frequency	225
Figure 6-19 Age dependence of 2Box striatal binding index.....	227
Figure 6-20 Age Dependence of Crescent-Putamen uptake ratio.....	227

List of Tables

Table 1-1 Summary of radiopharmaceuticals that can be used for imaging the dopaminergic pathway.....	25
Table 2-1 Published recommendations for acquisition settings.....	45
Table 2-2 Results of scatter correction using Monte Carlo simulations.....	49
Table 2-3 Pixel Sizes for SPECT mode acquisition.....	59
Table 2-4 Comparisons of local settings with published guidelines.....	65
Table 3-1 Resolution characteristics for reconstruction software.....	77
Table 3-2 Summary of phantom filling parameters and acquisition times for anthropomorphic phantom investigations	88
Table 3-3 Calculated blurring function at different distances from collimator..	132
Table 3-4 Head sizes for a test population.....	135
Table 3-5 Calculated attenuation affects in different positions within different sized heads.....	137
Table 4-1 Demographics of patients scanned using original parameters	146
Table 4-2 Demographics of patients scanned using optimised parameters.....	146
Table 4-3 Breakdown of the diagnosis subtypes for the two datasets	147
Table 6-1 Summary of alignment errors for BRASS program	211
Table 6-2 Comparison of optimised VOI methods using dataset B.....	224
Table 6-3 Comparison of optimised VOI methods using both datasets.....	225
Table 6-4 Age dependence of uptake figures	226

Acknowledgements

There are numerous people to whom I am deeply indebted that have helped me in so many ways during my studies.

I begin with my supervisors Dr Kuldip Nijran, Dr Peter Bain and Prof Glen Blake. All three have provided their broad knowledge and expertise in my project and without their guidance the work would not have happened. Dr Nijran has helped in every aspect of the project over these long six years advising on techniques, and understanding of the results. Dr Bain has been a wonderful help with all of the clinical details that this lowly physicist is still striving to understand. And Prof Blake has been a guiding force throughout the project and in particular during the troublesome writing up period.

My thanks to Dr John Frank, who supported my project both financially and politically when I was trying to start this work.

My friends and colleagues at ICHNT and elsewhere who have helped with the work in so many ways. There are too many to mention, but all been there to discuss the problems, add insight with awkward questions, and share a coffee or two. In particular I must thank Sean Davies, for his ability to translate my ugly code into a more elegant equation.

Last, but in no way least, to my wife. Her constant support, understanding and love continue to amaze me. I only hope that I can be as good to her as she is to me.

Published work

The thesis includes data and extracts from the following article, included with permission of the publisher.

Automatic classification of 123I-FP-CIT images ¹

Towey DJ, Bain PG, Nijran KS, Nucl Med Commun. 2011 Aug;32(8):699-707

Some of the data and findings have been presented within the following conference proceedings:

Data-lead Classification of ioflupane images²

Towey DJ, Creative Counting in Nuclear Medicine, IPEM meeting, 9th Feb 2010

Data-lead Classification of ioflupane images ³

Towey DJ, Bain PG, Al-Nahhas A, Nijran KS, Eur J Nucl Med Mol Imaging 2009; OP189

Optimizing iterative reconstruction technique for 123-I Ioflupane (DaTSCAN) imaging⁴

Towey DJ, Bain PG, Al-Nahhas A, Nijran KS, Eur J Nucl Med Mol Imaging 2008;35 supp2:S336

Incidence of bilateral disease on DaTSCAN in assessment of patients with unilateral Parkinsonian symptoms ⁵

Szyszkowski T, Noronha M, Muthu S, Towey DJ, Khan S, Al-Nahhas A. Eur J Nucl Med Mol Imaging 2008;35 supp2:S308

Potential of SIMIND Monte Carlo software for evaluation of scatter correction techniques in 99mTc-MDP and 123I-ioflupane imaging⁶

Towey DJ, Moysiadou F., Nijran KS. Nucl Med Commun 2008, 29(5) 471-472

Essential Tremor and features of Parkinsonism: How best to analyse Dopamine Transporter images?

Towey DJ, Bain PG, Hensman DJ, Frank JW, Meades R, Kalogianni E, Nijran KS
BNMS autumn meeting 2007

Inter-operator variation in quantiSPECT assessment of DaTSCAN patient images
Gannon KE, Clarke EA, Towey DJ, Deeb J, Nucl Med Communications 2007, 28:A19-P3

The effect of reconstruction techniques on quantification in Datscan imaging

Kalogianni E, Towey DJ, Bain PG, Frank JW, Nijran KS
BNMS Autumn Meeting 2006

Glossary of terms

	Definition
¹²³ I-FP-β-CIT	Radiopharmaceutical used for imaging Dopamine transporters (DAT), replaced ¹²³ I-β-CIT
¹²³ I-β-CIT	Radiopharmaceutical used for imaging Dopamine transporters (DAT), replaced by ¹²³ I-FP-β-CIT due to improved imaging characteristics
AD	Alzheimer's Disease, a common degenerative dementia
Caudate	Small curved structure near centre of brain, part of Striatum
CT	Computed Tomography, clinical imaging technique which utilizes a rotating x-ray source and detectors to acquire transmission projection data which can be combined to give transverse slices (see also FBP)
D2	Dopamine receptor type 2, expressed on post-synaptic neuron
DA	Dopamine, a neurotransmitter
DAT	Dopamine Transporter, expressed on pre-synaptic neuron
DaTSCAN	Trade name for kits for the preparation of ioflupane/FP-β-CIT
DDRR	Depth Dependent Resolution Recovery
DLB	Dementia with Lewy Bodies, a degenerative dementia
EANM	European Association of Nuclear Medicine
ET	Essential Tremor, benign movement disorder
FA	Factor Analysis, A range of processing techniques applied to principal components
FBP	Filtered Back Projection, a reconstruction method that creates transverse slices from a series of projection images
FDG	FluoroDeoxyGlucose, glucose analogue used in PET imaging, labelled with ¹⁸ F
FDOPA	Fluorodopa, a PET radiopharmaceutical used to image the Dopamine synthesis and storage, labelled with ¹⁸ F
FP-β-CIT	Ligand used for imaging of Dopamine Transporters (DAT), normally labelled with ¹²³ I
FRR	Fixed Resolution Recovery

FT / IFT	Fourier Transform / Inverse Fourier Transform
FWHM / FWTM	Full Width at Half Maximum, Full Width at Tenth Maximum, ways of defining the width of a peak, often used in resolution measurement
IBZM	IodoBenZaMide, radiopharmaceutical used in SPECT imaging of the D2 Receptors, labelled with ¹²³ I
ioflupane	Chemical name for ¹²³ I-FP-β-CIT
IPD	Idiopathic Parkinson's disease, the most common form of Parkinsonism
LEHR / VXGP	Two collimators designed used for low energy high resolution gamma camera imaging
MLEM	Maximum Likelihood - Expectation Maximisation - an iterative reconstruction method that creates transverse slices from a series of projection images
MRI	Magnetic Resonance Imaging, clinical imaging technique which utilises the magnetic properties of tissues to form images
MSA	Multiple System Atrophy
MSA-P	Parkinsonian variant of MSA
NEMA	National Electrical Manufacturers Association
NMSE	Normalised Mean Square Error, method of comparing two sets of data
NPS	Non-Parkinsonian Syndrome, classification group used in this work to distinguish from PS
OSEM	Ordered Subset Expectation Maximisation, an accelerated form of ML-EM reconstruction
PCA	Principal Component Analysis
PD	Parkinson's Disease, a movement disorder
PET	Positron Emission Tomography, clinical imaging technique allowing imaging of positron emitters labelled to various pharmaceuticals/molecules.
PS	Parkinsonian syndromes, the group of Parkinson's like diseases including IPD, PSP and MSA and others
PSP	Progressive Supranuclear Palsy, a Parkinsonian syndrome
Putamen	Small curved structure near centre of brain, part of Striatum

PVE	Partial Volume Effect, image artefact seen in all digital imaging modalities
rCBF	regional Cerebral Blood Flow
ROI	Region of Interest, cluster of pixels analysed as a group
ROR	Radius of Rotation, the distance from the gamma camera to the centre of rotation during SPECT acquisition
SNR	Signal to Noise Ratio
SPECT	Single Photon Emission Computed Tomography, a mode of operation for gamma cameras which allows 3D volume data to be acquired
SPET	See SPECT
SPM	Statistical Parametric Mapping, set of computational algorithms developed for analysis of groups of brain images
SSM / CSSM	Step and Shoot Mode, Continuous Step and Shoot Mode, different gantry motions in SPECT imaging
Striatum	Small structure near centre of brain, formed of the Caudate and Putamen, part of Basal Ganglia
SVCD	Small Vessel Cerebrovascular Disease
SVD	Singular Value Decomposition, computational technique for calculating eigenvalues, eigenvectors and singular values
SVM	Support Vector Machine, an example of a supervised learning technique to classify data points in 2 or more dimensions
TEW	Triple Energy Window, a scatter compensation technique used in gamma camera imaging
VOI	Volume of Interest, cluster of voxels analysed as a group

1. Introduction

This work investigates neuroimaging as applied to movement disorders by the use of radionuclide imaging techniques. There are two focuses in this work:

- The optimisation of the SPECT imaging process
- The development and optimisation of automated analysis techniques

Before discussing these ideas, the nature of the clinical conditions being investigated will be described.

1.1 Parkinsonian syndromes, clinical presentations and neurophysiological basis

Parkinson's Disease was first described by James Parkinson in 1817 and originally named the shaking palsy ⁷. It is the second most common neurological disorder after Alzheimer's disease, and the most common movement disorder ⁸. The term Parkinsonian syndromes (PS) encompass a range of related diseases including amongst others Parkinson's Disease, Progressive Supranuclear Palsy and Multiple System Atrophy. Idiopathic Parkinson's disease (IPD) accounts for around 80% of cases.

1.1.1 IPD – presentation, clinical features, neurophysiological basis

Idiopathic Parkinson's Disease normally presents in the form of motor effects, most commonly a tremor, but also rigidity, involuntary movements, abnormal posture and Bradykinesia (slowness of movement) ⁹. Non-motor dysfunction is also widely seen in Parkinson's Disease, with a range of Neuropsychiatric (depression, anxiety, apathy and cognitive impairment), anosmia, sleep dysfunctions (insomnia, limb movement, restless leg syndrome) and autonomic symptoms (Urinary Bladder, Nocturia, Sexual dysfunction)¹⁰⁻¹³. In advanced disease around 80% of patients will display cognitive impairment, e.g. memory loss, and almost half will show signs of depression ¹².

Post-mortem investigations find damage and loss of pigmented cells in the substantia nigra, including cytoplasmic inclusions, known as Lewy bodies. The cells in the substantia nigra extend to the striatum and supply it with Dopamine

(DA). The overall effect is a depletion of DA in the striatum, globus pallidus and substantia nigra. Within the striatum, the posterior of the putamen is normally affected first followed by the caudate. The loss of DA from the basal ganglia causes them to become overactive leading to the clinical motor features ¹⁴.

Diagnosis from clinical examination has around 80% sensitivity, but only 30% specificity when compared to post-mortem examination ¹⁵.

1.2 Parkinson's - Plus syndromes - presentation, clinical features and neurophysiological basis

Within the Parkinsonian syndromes there exists a range of so-called Parkinson's-plus syndromes ¹⁶. These typically present with similar motor features to IPD, but can vary.

1.2.1 Progressive Supranuclear Palsy

This disease shares some of the problems with movement seen in IPD, including a shuffling gait and accounts for around 6-10% of movement disorder cases ^{16,17}. The main differences that are seen are problems with eye movement. Typically, this begins with impairment of the down gaze and progresses to general ophthalmoparesis. Patients with PSP will tend to have an increased number of falls, which can be due to a combination of gait and eye-movement problems. However, these effects may not be seen at first presentation.

As with IPD, post-mortem examination shows damage and loss to the substantia nigra ¹⁷ that gives rise to the Parkinsonian motor changes. Neurofibrillary tangles are seen in many areas of the brain including the brain stem. The brainstem changes give rise to the eye movement effects.

Neurofibrillary tangles are also seen in other diseases, most notably Alzheimer's Disease (AD). However, the characteristic amyloid plaques of AD are absent.

1.2.2 Multiple System Atrophy

Multiple System Atrophy (MSA) combines three previously separate disorders of Shy-Drager syndrome, olivopontocerebellar atrophy and striato-nigral degeneration^{16,18-20}. Clinically the Parkinsonian symptoms are accompanied by autonomic failures. Two subcategories exist depending on which symptoms predominate, MSA-P where the symptoms are predominantly Parkinsonian, and MSA-C in the case of cerebellar ataxia. The former is more common in western populations, but considerable overlap between the two groups exist²⁰.

The physiological basis is glial cytoplasmic inclusions. Neuronal loss is seen in the basal ganglia along with loss in the cerebellum, pons, inferior olivary nuclei, and spinal cord.

1.2.3 Dementia with Lewy Bodies

Dementia with Lewy Bodies (DLB) is a specific condition first defined in the late 1970s^{16,21,22}. The main clinical presentation is dementia.

Lewy bodies are cytoplasmic inclusions common in many neurological conditions (including PD and AD) and exist in two forms: Cortical Lewy Bodies, and Brain stem Lewy Bodies. DLB is characterised by the presence of diffuse cortical Lewy bodies. Cortical changes give rise to the neurological symptoms. As the disease progresses these bodies may be seen in the substantia nigra. Damage and loss of cells here leads to Parkinsonian movement features. More

recently, DLB and the dementia variant of PD (PDD) have been viewed as different presentations of the same disease ^{23,24}.

The three conditions, IPD, MSA and DLB belong to a group of diseases known as α -synucleinopathies as the characteristic Lewy Bodies are formed from aggregated α -synuclein.

1.3 Other neurological disorders

1.3.1 Alzheimer's disease

Alzheimer's Disease (AD) is the most common neurological disorder and accounts for around half of all dementias. It was originally described as "A unique illness involving the cerebral cortex" by Alois Alzheimer ²⁵ who detailed the characteristic neurofibrillary tangles and amyloid plaques.

Clinical symptoms at presentation are mainly related to short-term memory problems and other neurological symptoms. However, differentiating this from other dementias such as DLB can be problematic. Such differentiation is important as patients with DLB can be highly sensitive to the neuroleptic (or antipsychotic) treatments often given in AD ²².

1.3.2 Tremors

The term tremor is associated with many movements of the body, as well as diseases and disease groups. A consensus statement of the movement disorders society ²⁶ set out to define and standardise the descriptions of movements so that better comparisons could be made between diagnostic and research groups. Defining the tremors by frequency, and activation cause (i.e. rest, posture or intentional movement), can direct the definition of some of the tremor

syndromes. The following descriptions of tremor groups are based on the disease definitions given there.

1.3.2.1 Essential Tremor

Essential Tremor (ET) is an involuntary tremor or shaking of the limbs or head for which no specific cause can be identified. It is not linked to any neurodegenerative processes such as Lewy bodies. It is the main differential diagnosis with Parkinsonian Syndromes as the disorders of movement can be difficult to separate. ET can have a genetic link as it can occur in families, known as Familial Essential Tremor (FET), but will often occur with no obvious familial history – known as Sporadic Essential Tremor (SET).

1.3.2.2 Psychogenic Tremor

These tremors occur in patients suffering from psychological problems. The physical symptoms do not originate from physiological changes, but are a manifestation of the psychological problems. A common example being a stress induced tremor.

1.3.2.3 Holmes Tremor

This tremor is generally related to brainstem/cerebellum or thalamic lesions. It presents clinically with a slow frequency tremor in both resting and intention movement ²⁷.

1.3.3 Vascular diseases

Disturbances in the blood supply to the brain are collectively known as cerebrovascular diseases. The disturbances can be transient or permanent,

reductions or occlusions and can vary in size and number. A number of subtypes are recognised including Multi-infarct dementia (MID) and Small Vessel Cerebrovascular Disease (SVCD). The area of brain affected can be permanently damaged and treatment is targeted on preventing further problems by limiting vascular risk factors. The area of brain affected will obviously determine the clinical manifestation. Vascular Dementia is the second most common form of dementia, and is difficult to distinguish clinically from AD ²⁸. A related condition is Cerebro-Vascular Parkinson's Syndrome (CVPS) ²⁹ where the vascular disease affects the striata producing Parkinsonian features.

1.3.4 Drug Induced Parkinsonism

There are a number of prescription drugs that can affect the action of dopamine in the brain. In particular, some neuroleptic drugs, often prescribed to control dementia and schizophrenia, can cause severe parkinsonian symptoms. In general, withdrawing these drugs will lead, in time, to a reversal of the Parkinsonian symptoms ³⁰. However, patient specific risk-benefit analyses have to be performed to ensure the best clinical outcome.

1.3.5 Other Neurological conditions

Because of the nature of neurological symptoms, there exist many overlaps between normal variants and the clinical features of many diseases. Other diseases types that can present as movement disorders include, but are not limited to, Multiple sclerosis, Huntington's disease, Fragile-X syndrome and Wilson's disease.

1.4 Neurotransmitters

Neurotransmitters are a group of chemicals utilised in the path of signals between neurons in the central nervous system. Signals pass along neurons as waves of polarisation. The passage of these signals across synapses – the joins between individual neurons - is facilitated by neurotransmitters.

The neurotransmitters are stored within a great number of vesicles within the terminal of the pre-synaptic neuron. When the action potential reaches the synapse, these vesicles eject their transmitters into to the synaptic cleft. The transmitters will bind to and activate some of the receptors on the post-synaptic neuron. Activated receptors act as open ion channels, changing the potential on the post-synaptic membrane. Once sufficiently polarised, this produces an action potential in the post-synaptic neuron and the signal is passed on.

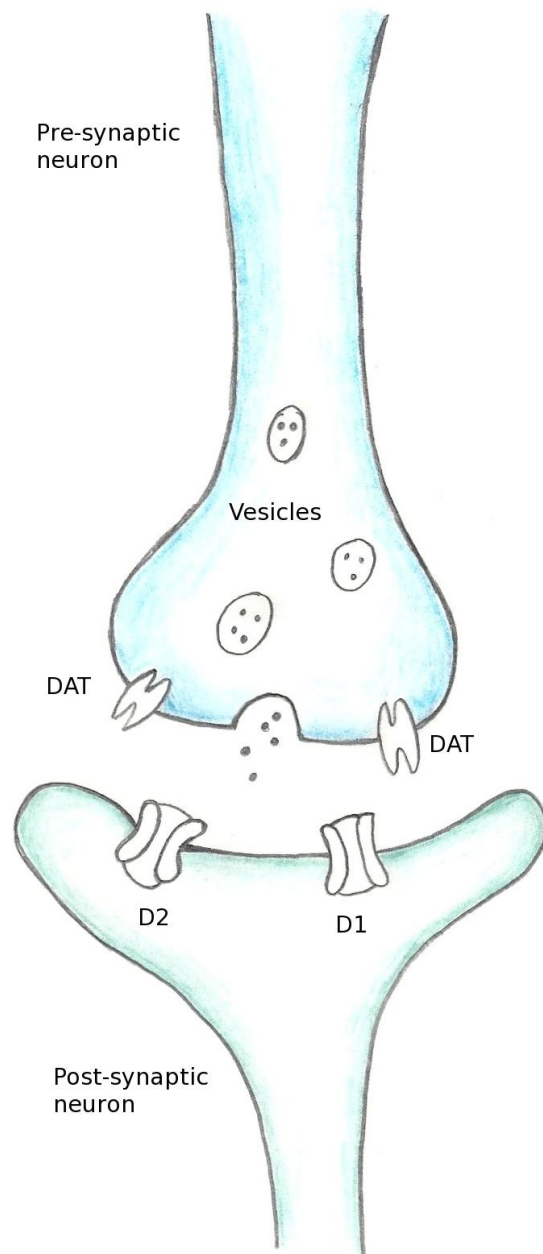


Figure 1-1. Schematic diagram of synapse between two neurons. Vesicles in pre-synaptic neuron release a neurotransmitter (in this case Dopamine) into the cleft.

The transmitters only act on the post-synaptic receptors for a brief time as they are quickly broken down by enzymes. Unbound transmitters in the cleft are reabsorbed by transporters in the pre-synaptic neuron.

1.4.1 Dopamine

Dopamine (DA) is a neurotransmitter that plays a role in many neurological processes including movement, cognition and mood. It is also important in reward and pleasure ¹⁴.

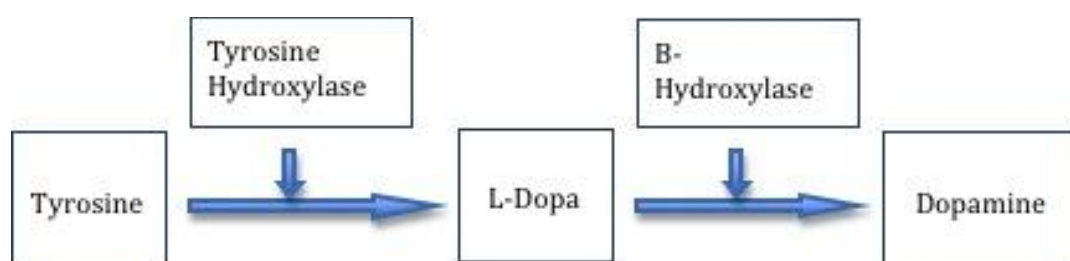


Figure 1-2. Synthesis of Dopamine from Tyrosine

Dopamine is synthesised as part of a series of enzyme mediated actions starting with Tyrosine. This is oxidised by tyrosine hydroxylase to give L-Dopa, which is then converted to DA by dopamine B-hydroxylase. In some cell types, DA is further processed to give Norepinephrine and Epinephrine. Tyrosine and L-Dopa can pass through the blood-brain-barrier, whereas dopamine cannot. Dopamine production occurs within the pre-synaptic neuron where it is stored within the vesicles ready for release into the synapse.

To date, at least five DA receptor subtypes have been identified ³¹. They can be split into two sub groups, D1-like receptors (D1 and D5) and D2-like receptors (D2, D3 and D4). D1 and D2 dopamine receptors are the more common forms and are found in particularly high concentrations in the striatum. The D2

subtype is seen predominantly in the post-synaptic membrane of the nigro-striatal synapses.

The DA Transporter (DAT) is the site on the pre-synaptic neurons where DA is re-absorbed for storage ready for the next excitation.

Both the D2 receptor and the DA Transporter are important targets for imaging studies of Parkinsonian syndromes.

1.5 Imaging of movement disorders and associated conditions

Functional imaging with radiopharmaceuticals for movement disorders can be divided into two groups depending on whether they target the pre-synaptic or post-synaptic neurons. Further divisions can be made on process or feature being imaged and with which modality it can be imaged, either single Photon Emission Computed Tomography (SPECT) or Positron Emission Tomography (PET). An overview of these tracers is given in Table 1-1, which also includes some of the main references on their use.

In general, PET imaging is superior to SPECT imaging in humans due to the improved resolution and accuracy of quantification. As PET involves imaging from 360 degrees simultaneously, kinetic measurements can be made on tracers with faster biokinetics than for SPECT. For these reasons PET is normally the preferred choice in the research setting. However, due to cost, availability and practical considerations, SPECT imaging is the more widely used clinical investigative imaging tool. An introduction to SPECT imaging is given in section 1.7.

Table 1-1 Summary of radiopharmaceuticals that can be used for imaging the dopaminergic pathway

		SPECT	PET
Pre-synaptic imaging	DA synthesis /storage		¹⁸ F-DOPA ²³
Pre-synaptic imaging	DA Transporter	¹²³ I-β-CIT ^{32-34 35}	¹¹ C / ¹⁸ F -β-CIT
		¹²³ I-FP-β-CIT ^{34 36,37}	¹¹ C / ¹⁸ F -FP-CIT
		^{99m} Tc-TRODAT ³⁸	
		¹²³ I-PE2I ³⁹	¹¹ C / ¹⁸ F -PE2I
Post-synaptic imaging	D2 receptor	¹²³ I-IBZM ^{40,41}	¹¹ C-raclopride ⁴¹
		¹²³ I-Epidepride ⁴²	¹¹ C / ¹⁸ F -Fallypride

1.5.1 Imaging Dopamine synthesis and storage

The design of radiopharmaceuticals for brain imaging is hampered by the blood-brain barrier (BBB). This membrane protects the brain by preventing larger molecules and objects such as bacteria from entering the brains extracellular fluid. DA cannot cross the BBB and so cannot be used as an imaging tracer. However, ¹⁸F-DOPA mimics the action of L-DOPA and can cross the blood-brain barrier. Its uptake will reflect the locations of DA production and storage. The uptake of ¹⁸F-DOPA will be dependent on a number of factors including the rate of transfer across the BBB, the rate of synthesis within the neuron and its retention within the neuron ^{31,43}.

1.5.2 Imaging the Dopamine Transporter

The radiopharmaceuticals used to image the dopamine transporter (DAT) are based on analogues of cocaine. Cocaine binds to the DAT blocking the re-uptake

of DA, subsequently increasing DA levels within the synapses producing the 'high'. The most common pharmaceuticals that have been used to study DAT using SPECT are ^{123}I - β -CIT^{32,33,44} and ^{123}I -FP- β -CIT⁴⁴⁻⁴⁸. An example transverse slice from a ^{123}I -FP- β -CIT scan is shown in Figure 1-3. The image shows high uptake in the caudate and putamen. The corresponding slice from a T2 weighted MRI scan is shown for anatomical localisation. This patient shows normal uptake of ^{123}I -FP- β -CIT.

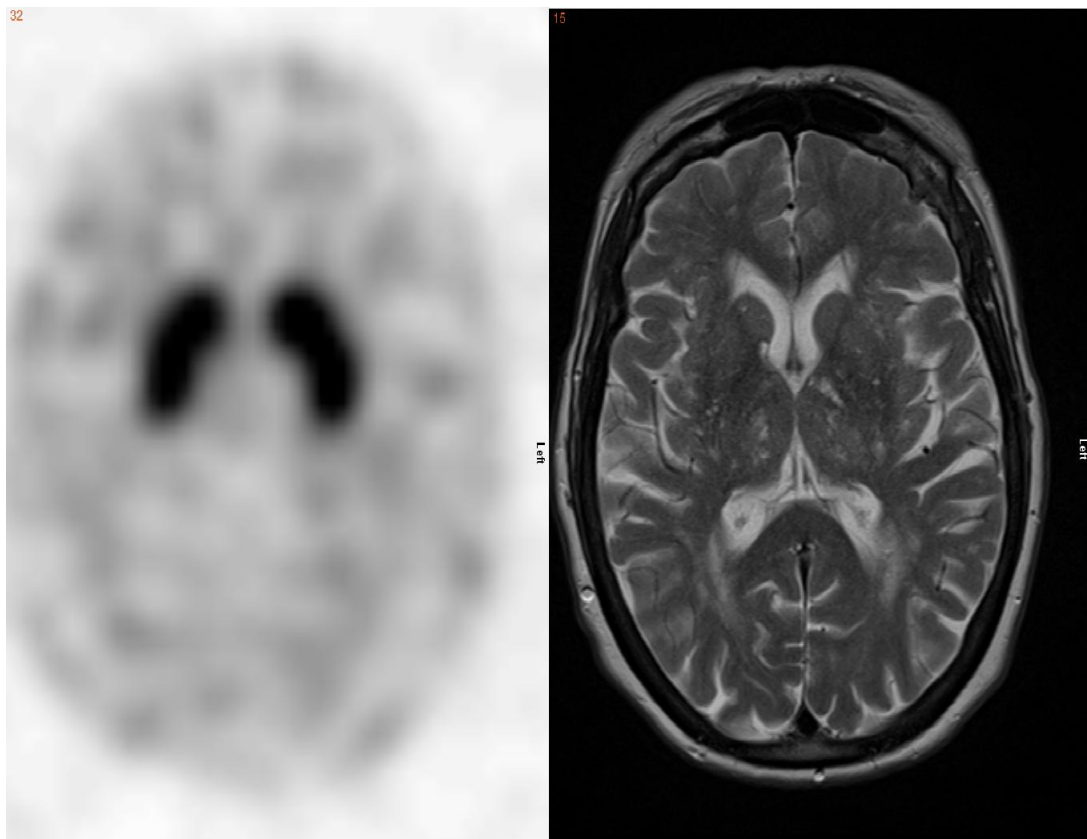


Figure 1-3 Example transverse slice from ^{123}I -FP-beta-CIT SPECT scan with corresponding slice from T2weighted MRI

The biokinetics of ^{123}I - β -CIT in the brain and its binding to the dopamine and serotonin transporters was described by Brucke et al.³². Further work by Innis et al.³³ showed that loss of striatal dopamine transporters could be detected

using ^{123}I - β -CIT imaging. They suggested that it could be used for both monitoring disease, and distinguishing between idiopathic Parkinson's and other disorders.

The binding of β -CIT based ligands (including ^{123}I - β -CIT, ^{123}I -FE- β -CIT and ^{123}I -FP- β -CIT) to dopamine, serotonin and norepinephrin transporters were compared by Okada ⁴⁴. This work and others lead to the use of ^{123}I -FP- β -CIT for DAT imaging.

Booij et al. ⁴⁵ described the bio-kinetics of ^{123}I -FP- β -CIT, showing a pronounced decrease in uptake in Parkinson's disease compared to age matched healthy volunteers. The bio-distribution and kinetics have also been described elsewhere ⁴⁷. The faster kinetics of ^{123}I -FP- β -CIT compared to ^{123}I - β -CIT, allow imaging at 3-6hrs as opposed to 24hrs, with higher specific to non-specific binding ratios. However, the relative changes in uptakes between normal and Parkinson's patients for the two tracers are equivalent ⁴⁶.

In retrospective analyses, the use of ^{123}I -FP- β -CIT has been shown to have a large impact on diagnosis and clinical management in clinically uncertain Parkinsonian syndromes ⁴⁹, dementia and DLB ⁵⁰. Another retrospective study showed a normal ^{123}I -FP- β -CIT image to give a high negative predictive value ⁵¹.

A study of ^{123}I -FP- β -CIT and ^{18}F -FDOPA ⁵² showed that ^{123}I -FP- β -CIT gave similar accuracy for discrimination between Parkinsonian and non-Parkinson's patients. The authors also described an age dependency of the uptake of ^{123}I -FP- β -CIT.

Comparisons between ^{18}F -FP- β -CIT and ^{11}C -FP- β -CIT suggest the former is superior due to the lower production of metabolites ⁵³. The increased half-life is

also beneficial, as it allows better imaging of the biokinetics. No direct comparisons of ^{18}F and ^{123}I labelled FP- β -CIT have been published.

More recently, ^{123}I labelled PE21 has been proposed as an alternative due to its lower affinity with Serotonin Transporters (SERT) ³⁹.

$^{99\text{m}}\text{Tc}$ TRODAT-1 is a relatively new agent that can be used to image DAT ³⁸. The use of $^{99\text{m}}\text{Tc}$ as the radioisotope as opposed to the cyclotron-produced ^{123}I may allow more widespread use.

1.5.3 Imaging the Dopamine receptors

Dopamine receptors are found in both pre- and post-synapse neurons, but the vast majority of D2 receptors are seen in the post-synaptic neurons. Hence, the D2 receptor has been used as post-synaptic imaging target. Most routine clinical imaging of D2 has been done with ^{123}I -IBZM ⁵⁴. ^{11}C -raclopride and ^{11}C / ^{18}F - fallypride are important PET research probes, but due to cost, availability and practicalities they have not found their way to clinical use.

Koch et al. ⁵⁵ demonstrated that combining pre- and post- synaptic imaging could improve diagnostic power. Financial and logistical issues have prevented this becoming the clinical routine.

1.5.4 Differentiation of parkinsonian syndromes using SPECT

1.5.4.1 MSA and PD

Nocker et al. ⁵⁶ investigated the progression of MSA-P and IPD. They found significant reductions of DAT in the brainstem of MSA-P patients compared to IPD at baseline. The reduction of uptake was greater in MSA-P patients compared to IPD. The brainstem uptake in the MSA-P patients did not decline

further after baseline. Their work showed distinct differences in the rate of progression of disease between MSA-P and IPD. Their results also suggest differences in the pattern of uptake between the two disease groups. In another publication by some of the same authors ⁵⁷, similar differences between MSA-P and IPD were highlighted. Changes in glucose metabolism in MSA have also been shown ⁵⁸.

1.5.4.2 PSP and PD

The clinical differences between PSP and IPD are hard to distinguish at early stages of disease. However, some work ⁵⁹ has suggested that the rapid progression of PSP, along with its more symmetrical progression, allows differentiation of these diseases using ¹²³I-FP- β -CIT imaging.

1.5.4.3 DLB and PD

Using both rCBF and DAT images, Rossi et al. ⁶⁰ showed no statistically significant differences between DLB and Parkinson's Disease -dementia. Others have reported on regional cerebral blood flow differences in PS ⁶¹ and DLB ⁶² when compared to healthy volunteers.

1.5.4.4 Holmes Tremor

In the case of Holmes tremor reported changes in ioflupane imaging have been inconsistent with some workers reporting marked dopaminergic loss ⁶³ and others minimal changes ⁶⁴.

1.5.4.5 Wilsons disease

In Wilson's disease (a copper deposition disorder) effects to both the pre- and post- synaptic neurons have been shown using ^{123}I - β -CIT and ^{123}I -IBZM respectively ⁶⁵.

1.6 Imaging with other modalities and non-imaging techniques

CT imaging is widely used to identify areas of vascular disease, trauma and malignancies, however its use in movement disorders is limited. Likewise, MR imaging has found little use so far. Ultrasound, though difficult in the brain due to the fact it is enveloped in bone, has been used with some success, however it could not be used to define severity of disease ⁶⁶.

Due to the early onset of anosmia in many patients, olfactory testing has been suggested as a useful adjunct to DAT imaging in early Parkinson's Disease ⁶⁷. However, the low positive predictive value would rule out its use in population screening on its own.

1.7 SPECT imaging

Single Photon Emission Computed Tomography (SPECT) imaging utilises radiopharmaceuticals to allow visualisation and quantification of physiological processes. After injection, the radiopharmaceutical circulates around the body and localises according to its biochemistry. A gamma camera is then used to detect the gamma emissions coming from the radioactive isotope. To perform tomographic imaging, the gamma camera is rotated around the patient acquiring images at equidistant projection angles. These projection images must then be

reconstructed to produce the transverse slices that represent the activity distribution in the patient.

In this section, the operation of a gamma camera is described, followed by the acquisition and reconstruction of SPECT images.

1.7.1 Gamma camera design and operation

The basic gamma camera design was first demonstrated by Anger⁶⁸ as a method of imaging single photon emissions and replaced the rectilinear scanner. A schematic diagram is given in Figure 1-4.

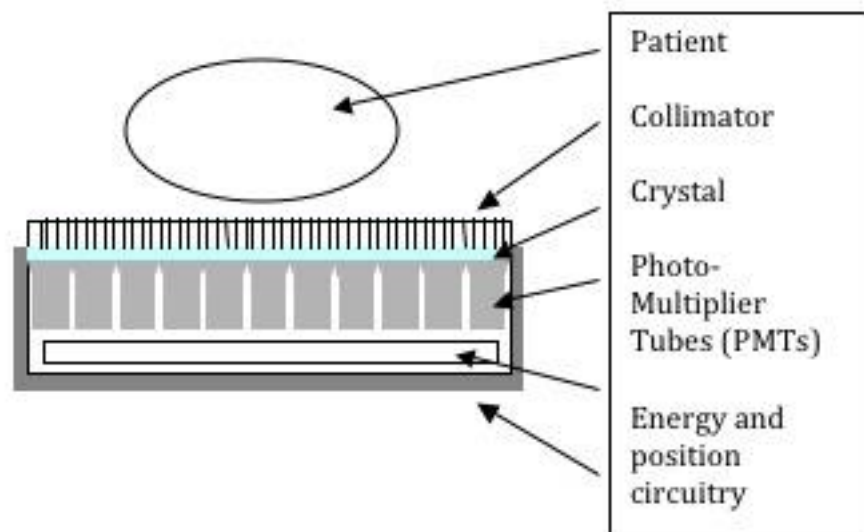


Figure 1-4 Schematic diagram of gamma camera

Gamma photons emitted from within the patient travelling in the direction of the gamma camera first encounter the collimator. These are normally constructed from folded sheets of lead, to form a series of parallel holes or tubes. They have also been made by drilling into blocks and are sometimes made of tungsten (Figure 1-5).

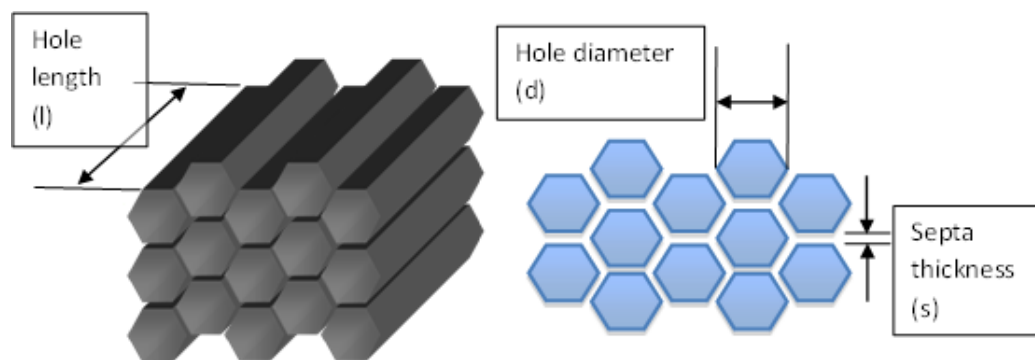


Figure 1-5 Schematic diagram of collimator

The walls between these holes, known as septa, absorb photons that are not travelling perpendicular to the face of the detector (or within a certain acceptance angle). The ratio of the width (d) and depth (l) of these holes define the acceptance angle, the range of angles of photon paths that can still reach the detector (Figure 1-5). A reduced acceptance angle will improve the geometric resolution of the collimator, but will also reduce the sensitivity. Conversely, a wider acceptance angle will improve the sensitivity, but degrade the resolution. The septa of the collimator will also affect the sensitivity as they cover an area of the detector proportional to their thickness. The required thickness of the septa will be related to the range of energy of the photons being imaged. As the number of gamma rays passing through a material falls exponentially with thickness, no amount of material would attenuate all the photons; hence, collimator designs are defined for energy ranges that allow less than 5% transmission. This situation can be complicated where isotopes have multiple higher energy emissions.

The detector is a scintillation crystal (NaI doped with Tl), coupled with an array of Photo-Multiplier Tubes (PMTs). Gamma rays incident on the crystal will be absorbed by photoelectric absorption or undergo Compton scattering. After a

short delay, the energy absorbed by the crystal is re-emitted as a group of visible light photons. The number of photons emitted is proportional to the energy deposited in the crystal.

Some of the emitted photons will hit the photocathodes of the PMTs. This interaction causes the emission of electrons from the photocathode. The electrical potential between the dynodes cause an acceleration of the electrons and results in a cascade of electrons along the dynodes, resulting in an electrical signal. The signals from all the PMTs can be combined to give the position of the original crystal interaction. The sum of the signals is relative to the energy deposited. An energy acceptance window (the photopeak window) can then be used to discriminate between scattered and non-scattered photons, since scattered photons will be of lower energy. The energy of the scattered photon ($h\nu'$) is related to the scattering angle (θ) and the initial energy ($h\nu$) by the constant $m_e c^2$ (Equation 1-1).

$$\frac{1}{h\nu'} - \frac{1}{h\nu} = \frac{1}{m_e c^2} (1 - \cos \theta)$$

Equation 1-1

The energy resolution of a modern gamma camera is around 10% (Full Width at Half Maximum - FWHM). Hence, non-scattered photons will be detected at a range of energies, and so compromise must be made between widening the photopeak window to accept sufficient non-scattered photons, but not to include too many scattered photons. Scatter compensation methods are discussed further in Chapter 2.

1.7.2 Resolution with distance

The resolution of a gamma camera is the error in the positioning of an emission. It is formed of two components; the first is the intrinsic resolution of the detector R_i . This is positioning error of the interaction within the detector. It is dependent on a number of factors including the crystal thickness, the light output of the crystal, and PMT size and arrangement. The second component, R_c is a geometric uncertainty of the origin of the emission relating to the acceptance angle of the collimator.

As the geometric resolution of the collimator is based on an acceptance angle, its width will be dependent on the collimator-source distance. For a circular holed collimator, R_c can be approximated as shown in Equation 1-2.

$$R_c \approx d \frac{(l_{eff} + b)}{l_{eff}}$$

Equation 1-2

Where d is the diameter or width of the hole, b is the distance from the source to the collimator and l_{eff} the effective length of the holes. The effective length of the holes l_{eff} is slightly shorter than the physical holes (l) due to septal penetration. The effective length is given by Equation 1-3⁶⁹ where μ is the linear attenuation coefficient for the collimator material at the energy being imaged.

$$l_{eff} = l - 2\mu^{-1}$$

Equation 1-3

For a hexagonal collimator the resolution will have an angular dependency that follows the width of the hole.

The intrinsic and collimator resolution effects combine in quadrature to give the total system resolution R_{Tot} (Equation 1-4).

$$R_{Tot} = \sqrt{(R_c^2 + R_i^2)}$$

Equation 1-4

1.7.3 Septal penetration

The collimator design is optimised for a particular energy ranges. Typically, this will involve 'Low' (<160-180keV), 'Medium' (160-280keV) and 'High' (>280keV) energy ranges. The isotope of interest here is Iodine-123, which has main photon emission at 158.99keV (emission probability 0.9922), which would imply low energy collimators. However, the emission spectrum from ^{123}I is complex ⁷⁰ and contains a number of higher energy emissions (see Figure 1-6). The probability of emissions above 159keV is 3%, with emission above 500keV at 2% probability. These higher energy photons will not be attenuated as much by the collimator septa. This septal penetration can lead to reduced contrast and increased noise ⁷¹. Some level of septal penetration is always present in gamma camera imaging, however, when higher energy emissions are present, as is the case for 123-Iodine, a greater level of septal penetration will be present. Thus, the choice of collimators is not straightforward ⁷¹.

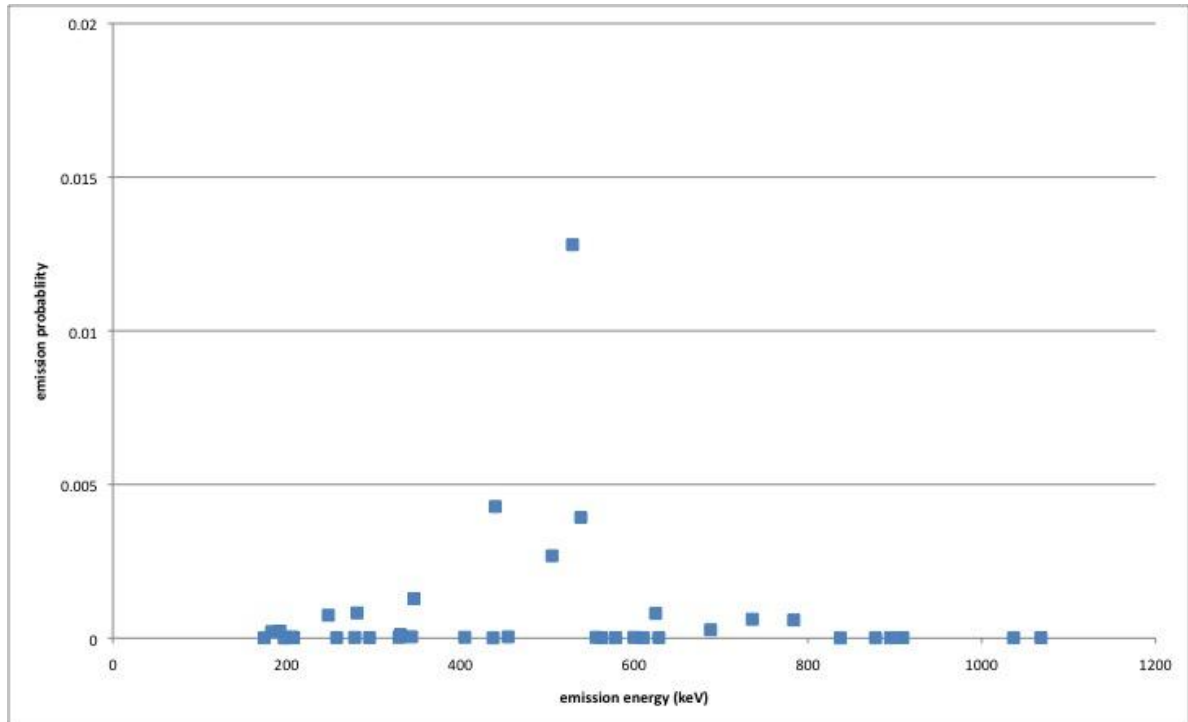


Figure 1-6. Emission probabilities of 123-Iodine. Principal gamma energy 158.99keV omitted for clarity. Uncertainties in the published data too small to be displayed.

1.7.4 SPECT mode acquisition

Single Photon Emission (Computed) Tomography (SPECT or SPET) is an acquisition mode in which the gamma camera is rotated around the object, acquiring a series of static images from a number of discrete, uniformly spaced angles. To allow quicker scans modern scanners will have multiple detectors acquiring at different angles at the same time. Often this is with two detectors but three and four detector systems have also been used. These static images are known as projections or projection images. The projection images can be combined using reconstruction algorithms to produce transverse slices or 3D volume data.

Figure 1-7 and Figure 1-8 show the positioning of a patient for a SPECT brain scan using a dual headed gamma camera. The patients' head is held in a special

cradle to allow the detectors to be positioned as close as possible, avoiding the shoulders and couch. Forehead and chin straps are used to minimise patient movement.



Figure 1-7 A subject positioned ready for a brain SPECT acquisition



Figure 1-8 The patient positioning as seen through the gantry

1.8 Image reconstruction

The two categories of algorithm are analytical and iterative reconstruction. A brief description of these models is given here, but a more complete description can be found in the literature ^{72,73}.

1.8.1 Analytical reconstruction

The most common form of analytical reconstruction program, and the one still widely used in nuclear medicine, is Filtered Back Projection (FBP). Here the projection data are convolved with a windowed ramp filter. The ramp filter is required to limit the star artefact. The windowing filter will normally be a low-pass filter (such as a Butterworth filter). Such windowing is required to prevent the amplification of high frequency data in the images since high frequency data is made purely of noise. Once convolved with a suitable filter, the projection data can be back projected across the image space.

1.8.2 Iterative reconstruction

Iterative reconstruction models generally work by estimating the activity distribution, calculating the projection data that would arise from this estimate and then comparing the estimated projection data with the measured projection data. This comparison then allows the distribution estimate to be updated.

Thus, there are three main parts of the iterative algorithm:

- The system model (which is used to calculate the projections from the distribution estimate)
- How the differences between the measured and calculated projections are defined
- How these differences are used to update the distribution estimate

A common implementation of iterative reconstruction is Maximum likelihood – Expectation Maximisation (ML-EM) ⁷⁴.

1.9 Image analysis and classification

When trying to classify a medical image the aim is to combine all the information in an image into some meaningful description. For example the images in this work, the starting point is a patients' scan of around 500,000 voxels, and the finish point we hope to reach is a meaningful diagnosis (Parkinson's disease or tremor) or some level of disease progression (early – late disease). The normal process is to first identify the salient features of the image. The second stage is to use these features to classify the image by comparing them to known features or patterns.

This combination of feature extraction followed by classification occurs in all the assessment techniques discussed here.

1.9.1 Visual assessment

The main assessment technique used for the majority of medical imaging tasks is visual assessment by an experienced reporter. They look for features in the images that are typical of the healthy or diseased states. Allowances are made for known image artefacts and limitations of the modality being used. The evaluation of SPECT images is performed using specialised nuclear medicine reporting software that allows some basic manipulation of the images.

The reporting clinician is presented with images that have been through a series of preparatory steps:

- Image reconstruction (which may include a range of image correction factors – see Chapter 3)
- Image alignment (often performed manually but can also be automatic registration to a template)
- Display colour table windowing – there are three general techniques to this, scaling to some predetermined absolute figure (e.g. an SUV of 7 in PET imaging), scaling to the highest pixel, or scaling based on the non-specific binding (e.g. with the maximum being some multiple of the background region). The first two options are the simplest to implement, however, the third may be easier to interpret for some image types.

Each clinician will have their own way of visually analysing the images - the consensus involves looking for two properties: firstly, the image quality (was there patient movement? Have the images been re-aligned correctly?). The second stage is to examine the images for signs of disease. Such steps are

obviously pharmaceutical specific. The normal variants with which they compare may be also be age and sex specific.

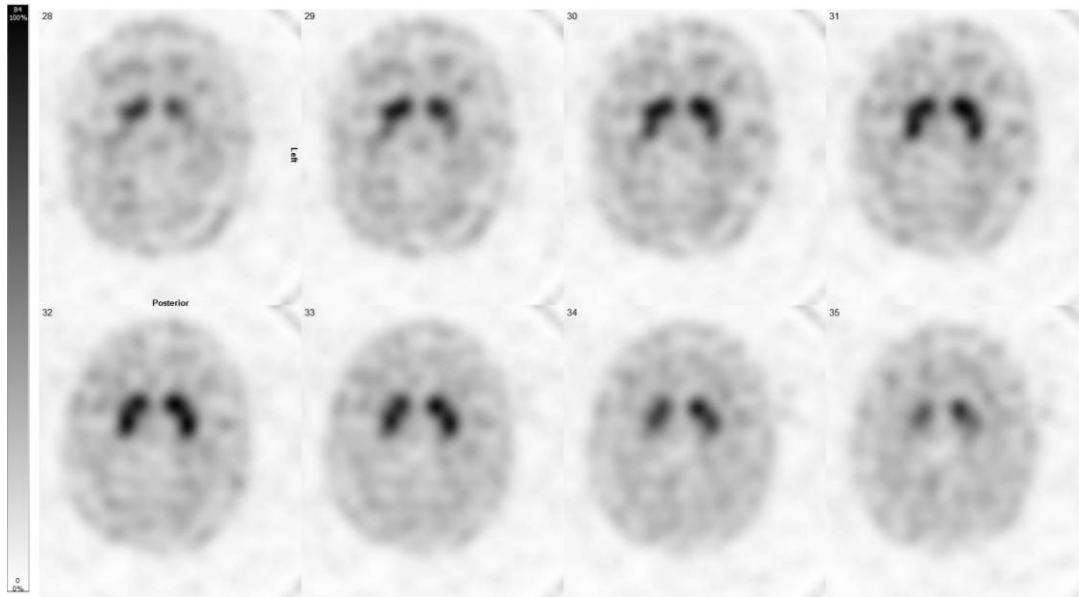


Figure 1-9 A series of transverse slices from a normal ^{123}I -FP-B-CIT SPECT scan

Using ^{123}I -FP- β -CIT SPECT imaging as an example, and with reference to the example transverse slices shown in Figure 1-9, the following characteristics would be looked for:

- Symmetry – assuming there are no technical problems with the acquisition and the images have been aligned properly, non-symmetrical images would suggest disease
- Shape of the uptake – the classic shape of striatal uptake is the ‘dot-and-comma’. Disease progression is typically seen as loss of signal from the putamen so a change in shape towards a dot rather than commas would be suggestive of disease
- Over-all uptake – a uniformly reduced uptake is suggestive of disease, this can be assessed visually by comparing to background/non-specific

binding areas, but this is not easy to see without some form of quantisation or a normalised way of applying the colour table

- Comparison with previous scans, including previous ^{123}I -FP- β -CIT imaging, CT and MR where available
 - Comparison with clinical History. If symptoms are unilateral, concordance with uptake patterns may add confidence to a particular report. However, image and symptomatic concordance is not guaranteed
- 5.

Good sensitivity and specificity have been found when using visual assessment to differentiate between groups of subjects with parkinsonism and either essential tremor or healthy volunteer ^{75,76}.

1.9.2 Regions and Volumes of Interest (ROI/VOI)

Regions and Volumes of Interest are widely used in most imaging modalities. The VOI is used to define a group of voxels (and/or partial voxels) and the average or total signal within the volume gives some measure of the tissue or organ contained within it. These figures need to be processed further, either by comparing to a normal range/cut-off or by combining it with data from another volume.

VOI sizes typically range from individual voxels up to sub-organ or organ sized volumes. The shapes can be defined manually or automatically using some form of edge detection/region growing algorithm. Often standard sized VOIs will be defined based on templates or structural images.

VOI can be placed manually by an operator or automatically. Automatic positioning can be based on image registration, edge detection or image maximums/minimum values.

One advantage of using VOI data is that (assuming appropriate corrections can be applied) the counts/signal in the VOI will often directly represent a clinically relevant physical value, or can be used to create such a figure. For example, the counts within a VOI on ^{123}I -FP- β -CIT images will be proportional to the number of transporter sites within that volume.

^{123}I -FP- β -CIT imaging lends itself nicely to quantification techniques as it is based on measuring uptake in known structural components rather than trying to identify increased or decreased uptake across the whole body, as is often the case in tumour imaging. Overviews and comparisons of various techniques have been published ^{77,78} along with calls for the widespread implementation of semi-quantitative analysis ⁷⁹. The European guidelines for dopamine transporter imaging ⁸⁰ state that semi-quantitative analysis may be useful, and where performed should include standardised ROIs and have a normal range preferably based on age-matched controls. They recommend the use of either occipital or cerebral regions for count normalisation.

A number of VOI based methods have been suggested for quantification in dopamine receptor and transporter imaging. These methods can be split into two subgroups depending on whether the placement of the VOIs is manual or automatic. Three of the techniques using manually based VOIs have been combined into a software package (QuantiPSECT, Mirada solutions / GE-medical Oxford UK).

1.10 Aims of this work

The main aim of this work is to investigate the use of classification techniques in movement disorders. However to ensure the best performance of the process each step must be optimised and hence there are a number of aims:

- Investigate the gamma camera performance characteristics when using ¹²³Iodine
- Optimise the acquisition settings taking into account scanner performance, patients' needs and pharmaceutical characteristics
- Investigate and optimise the image reconstruction and image correction techniques for use with ¹²³I-FP-β-CIT SPECT imaging
- Develop an automatic classification technique
- Compare the new automatic technique with other published techniques
- Using a database of clinical scans, test the following hypothesis:

“An automatic classification tool based on machine learning techniques will give better performance than currently available techniques.”

2 Acquisition Parameters

The performance of gamma cameras for imaging ^{123}I -Iodine SPECT varies greatly and the exact characteristics will affect the quality of the images produced, and potentially the likelihood of an accurate diagnosis. In this chapter, a range of scanner performance characteristics will be measured. These results will be used to define image acquisition parameters most suitable for clinical dopamine transporter imaging.

The following parameters will affect the imaging performance:

- Gamma camera manufacturer / model
- Collimator
- Pixel size and Field of view/zoom setting
- Number of projections
- Orbit shape (circular/non-circular)
- Orbit type (continuous / step and shoot)
- Time per projection
- Energy window selection

The original EANM procedure guidelines for imaging of dopamine transporters published in 2002 ⁸¹ were updated following developments and comments and were published in updated form in 2010 ⁸⁰. They were produced by the EANM Neuroimaging committee and summarize the recommendations for the imaging of Dopamine Transporters using both ^{123}I - β -CIT and ^{123}I -FP- β -CIT. As the

recommendations are designed to cover a wide range of equipment and institutional clinical practices they are necessarily broad.

The manufacturer of ^{123}I -FP- β -CIT (under the tradename 'DaTSCAN') has also produced basic guidelines covering recommended acquisition and processing procedures. Unlike the EANM guidelines, the manufacturer's recommendations give specific camera settings for individual scanners; the data quoted here refer to the gamma cameras used in this research. The three sets of guidelines are summarised in Table 2-1.

Table 2-1 Published recommendations for acquisition settings

Parameter	EANM 2002	EANM 2010	^{123}I -FP- β -CIT manufacturer (GE)
Delay to imaging	3-6hours	3-6hours	3-6hours
Imaging device	Minimum 2 detectors	Minimum 2 detectors	2 detectors
Collimators	LEHR or LEUHR Fan beam preferred	LEHR or LEUHR Fan beam preferred	VXGP / LEHR
Orbit style	Not stated	Not stated	Circular
Orbit mode	Step and shoot, (but continuous may reduce time and reduce mechanical wear)	Step and shoot, (but continuous may reduce time and reduce mechanical wear)	Not specified
Radius of rotation	Smallest possible	Smallest possible	Smallest possible, 11-13cm
Matrix	128*128	128*128	128*128
Angular sampling	3degrees	3degrees	3degrees
Pixel size / Zoom	Between $\frac{1}{2}$ and $\frac{1}{3}$ expected resolution	Between $\frac{1}{2}$ and $\frac{1}{3}$ expected resolution	No zoom, ~4.6mm
Total counts	>3million	>3million	>1.5million
Total time	40-50mins typical for triple headed	30mins typical for triple headed system	30sec/projection (30mins total)

2.1 Gamma camera manufacturer/model

There are currently three main manufactures producing gamma cameras for clinical use and a number of smaller companies producing specialist equipment. However all the patients used in this work were scanned at one hospital, using one of two cameras. The cameras are identical (Forte, ADAC Laboratories, California, USA) except for the collimators used on them. Due to time constraints, and the nature of this project, only these two cameras were investigated.

2.2 Energy windows

The equipment available in our institution limits the number of energy windows that can be acquired simultaneously. Up to three non-overlapping windows can be acquired per detector at the same time. The width of the photopeak window will affect the resolution and sensitivity of the scanner. Increasing the width of the window will mean more counts will be included in the image. However, the wider window will take in a greater proportion of scatted photons. Scattered photons are imaged in erroneous positions, degrading the resolution of the image. The energy resolution of the detector determines its ability to distinguish between scattered and non-scattered photons. The energy resolution is dependent on the crystal (material, doping levels, quality and thickness), the optical coupling of the crystal to PMTs, and the PMTs and associated electronics. The sensitivity-resolution trade off within the collimator will have to be taken into account when optimising the photopeak widths. Detector uniformity and energy maps will be affected by the choice of photopeak width and changes to it will require further corrections maps.

The main photopeak window was fixed at 20% symmetrical window, centred on the main photopeak, as per the manufacturers recommended procedure. This was done to ensure the system correction files were appropriate.

2.2.1 Energy window simulation

A series of Monte Carlo simulations were performed to investigate the contributions of scattering events and collimator interactions. The SIMIND package ⁸² was configured to model the scanners used in this work. A range of basic phantoms were used and tested against real-life measurements and found to match real scanner performance to within experimental error ⁶.

Figure 2-1 shows the energy spectra output from a simple head sized cylinder (15cm diameter, 20cm length, uniform activity distribution, uniform material) for direct photons, and photons that have been scattered when imaging with a LEHR collimator on the ADAC forte gamma camera. Only a single projection was simulated.

The results of the simulation show that a large number of the photons detected within the photopeak window will have undergone at least one scattering interaction. A number of scatter correction methods have been proposed, some of which use secondary energy windows as estimates of the scatter fraction within the photopeak. Smaller windows, positioned very close to the main photopeak are likely to give a better estimate, however, the photon flux away from the main peak is much lower, so the use of very thin scatter windows will result in noisy estimates.

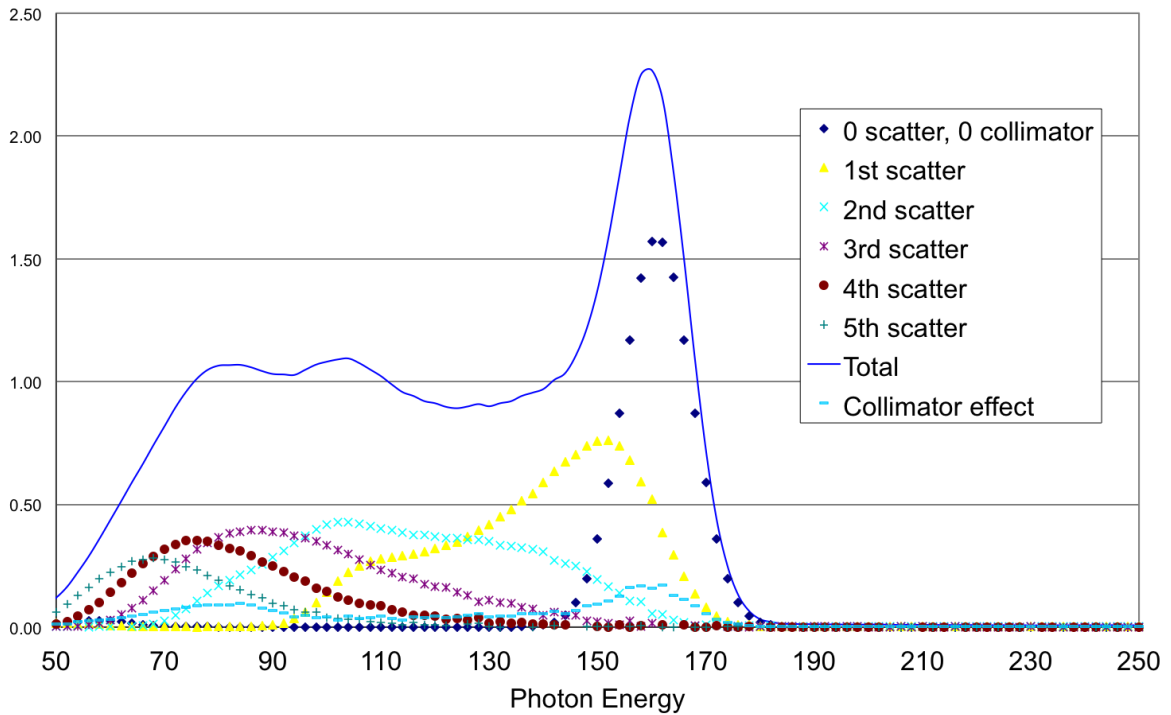


Figure 2-1 Results of a Monte Carlo simulation. Showing the number of detected photons as a function of energy

A number of window widths have been used in published data. In triple energy window corrections both 3%^{83,84} and 7%⁸⁵ have been used. For dual energy window corrections, with the scatter window below the photopeak these widths may be increased to 20% Dobbeleir et al. ^{71,86} or 40%⁸⁷. A slightly different approach was taken by Small et al. ⁸⁸, who focused on the down scatter from higher energy emissions and used a 20% window above the photopeak.

A very simple test was performed using the spectra output from this Monte Carlo model. Using the energy windows proposed by Iida et al. ⁸⁵ the number of direct (non-scattered) photons, and total photons (scattered, non-scattered, and collimator penetration and scatter) were calculated for the three energy windows. These are shown in Figure 2-2 as direct photons, scattered photons, and total photons detected. Three energy windows are also shown representing

a 20% photopeak, and two 7% scatter windows. These counts are shown in Table 2-2.

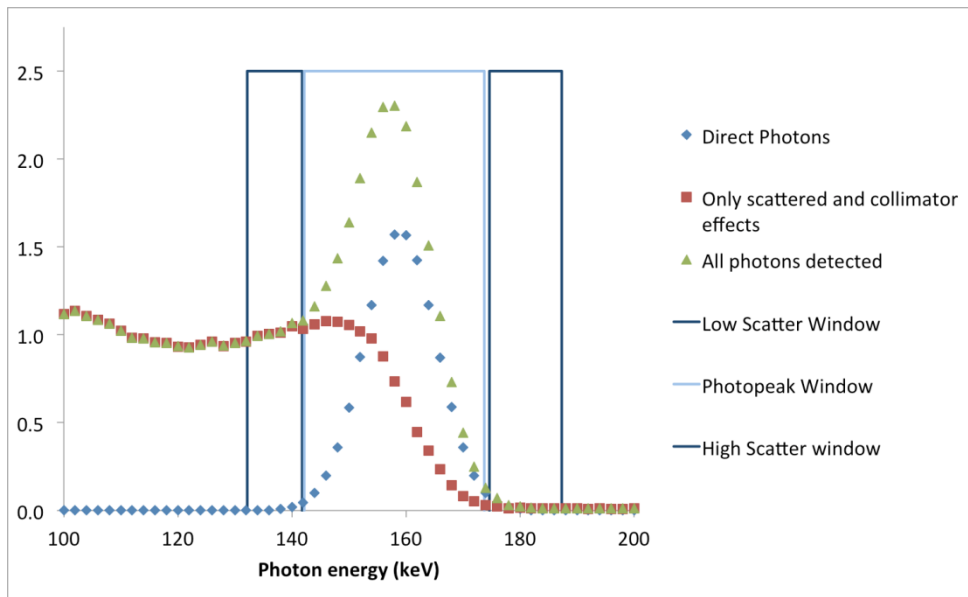


Figure 2-2 The energy acceptance windows compared to the energy of direct and scattered photons

Table 2-2 Results of scatter correction using Monte Carlo simulations

Window	Direct (Non-scattered)	Total (Direct and scattered)	TEW corrected	Error
Low scatter	7	517		
Photopeak	1250	2333	1272	2%
High scatter	36	168		

In this very simple test of the TEW scatter correction, the windows proposed by Iida gave very reasonable results, with an error of 2%. However, the application of this technique will need to be further investigated using more realistic data.

The optimum scatter window width will be dependent on the scatter correction

technique used, the scattering conditions of the clinical procedure being undertaken, and the scanner design including energy resolution of the detectors and scattering properties of the collimators. Only three energy windows could be applied simultaneously using the equipment available in this work, so a retrospective analysis of all techniques would not be possible. A full investigation into the optimum widths and positioning of these windows was considered to be beyond the scope of this work so a choice was made on based on the published data. The windows were chosen to match those in Iida et al.⁸⁵ with two 7% scatter windows positioned either side of the photopeak. These widths were using the following collimator comparison tests. Scatter correction is discussed further in chapter 3.

2.3 Comparison of LEHR, VXGP and MEGP Collimators

The design of collimators is a compromise between resolution and sensitivity, with sufficient thickness of septa to prevent excessive penetration. Their design is optimised based on the energy of photon being imaged. Until recently, ¹²³I was not a regularly used isotope, and hence it is rare for collimators to be optimized for it. The main emission energy of ¹²³I is 159keV, but it also emits a number of higher energy photons (Figure 1-6). These factors complicate the choice of collimator⁷¹.

The presence of higher energy photons will degrade the image in two ways. The higher energy photons may not be attenuated by the septa of the collimator. A second effect occurs when higher energy photons undergo Compton scattering. The energy loss means they fall within the photopeak acceptance window. For both these effects, the resulting scintillation event in the detector will give

erroneous positional information. They both add a background “fog” to the image, increasing the noise, reducing the contrast and affecting the quantisation. Macey et al.⁸⁹ investigated collimator choice for whole body imaging. Improved resolution was found with the low energy high-resolution collimators, but medium energy collimators gave better quantitative accuracy. They concluded that the MEGP collimators were superior for whole body imaging. However, no attempts to correct for the septal penetration or scatter were included in that work.

Gillard et al.⁹⁰ investigated measurement of size and activity in SPECT when using low and medium energy collimators. As well as the characteristic dependency of recovery coefficient on object size, they found that septal penetration in low energy collimators could lead to reduced contrast, and image artefacts. The most obvious artefact is activity being positioned outside of the emitting body, seen as a hot ring around the reconstructed volume. This “hot ring” artefact was also described by Macey et al.⁸⁹.

More recently, corrections for down scatter and septal penetration have been investigated when using LEHR collimators^{82,86,88,91}. They have been shown to significantly improve the image quality by compensating for the image degradation caused by scatter and septal penetration. Allowing imaging to be performed using LEHR collimators taking full advantage of the superior resolution.

An investigation was performed to describe the resolution performance of the various collimators available for clinical use. Three types of collimator are available at this institution, two low energy high-resolution collimators (LEHR, VXGP) and medium energy general-purpose collimators (MEGP). The resolution

and sensitivity for ^{123}I can be measured directly using techniques described by National Electrical Manufacturers Association (NEMA) ⁹² and can be further investigated by changing the source to collimator distances.

2.3.1 Collimator comparison - Method

Two capillary tubes were filled with ^{123}I and placed in parallel with 8cm separation. The tubes were suspended between the two camera heads parallel to the long axis of the camera. As the internal diameter of the capillary tubes is very small compared the resolution of the camera, the resulting images are good approximations to the line spread function of the camera.

A 256x256 matrix was used with a detector mask of 25.4cm (i.e. 2.19 zoom) to produce a pixel size of 1.1mm. A series images was acquired with a range of collimator to source distances ranging from 5cm to 15cm. The acquisitions were performed on each of the three collimator styles available and profiles drawn across the images. The timing was adjusted to ensure a minimum of 1000 counts in the highest pixel of the profile. Repeat measurements were taken over two years.

A second series of images were acquired using Perspex sheets to act as a scattering medium. A total thickness of Perspex of 18cm was placed gently on the collimator face. The capillary tubes were positioned between the Perspex sheets at different depths into the Perspex.

The relative sensitivity of the collimators was measured as a function of distance from the collimator. This was performed using the total counts in-air resolution images after correcting for radioactive decay and differences in acquisition

length.

2.3.2 Collimator comparison - Results

The resolution measured in air are shown in Figure 2-3 and Figure 2-4 as a function of collimator to source distance in terms of full width at half maximum (FWHM) and Full Width at Tenth Maximum (FWTM) respectively. The manufacturers' specifications for ^{99m}Tc imaging are also shown along with distance of interest for SPECT imaging of the striatum. No specification for ^{123}I imaging is quoted by the manufacturer.

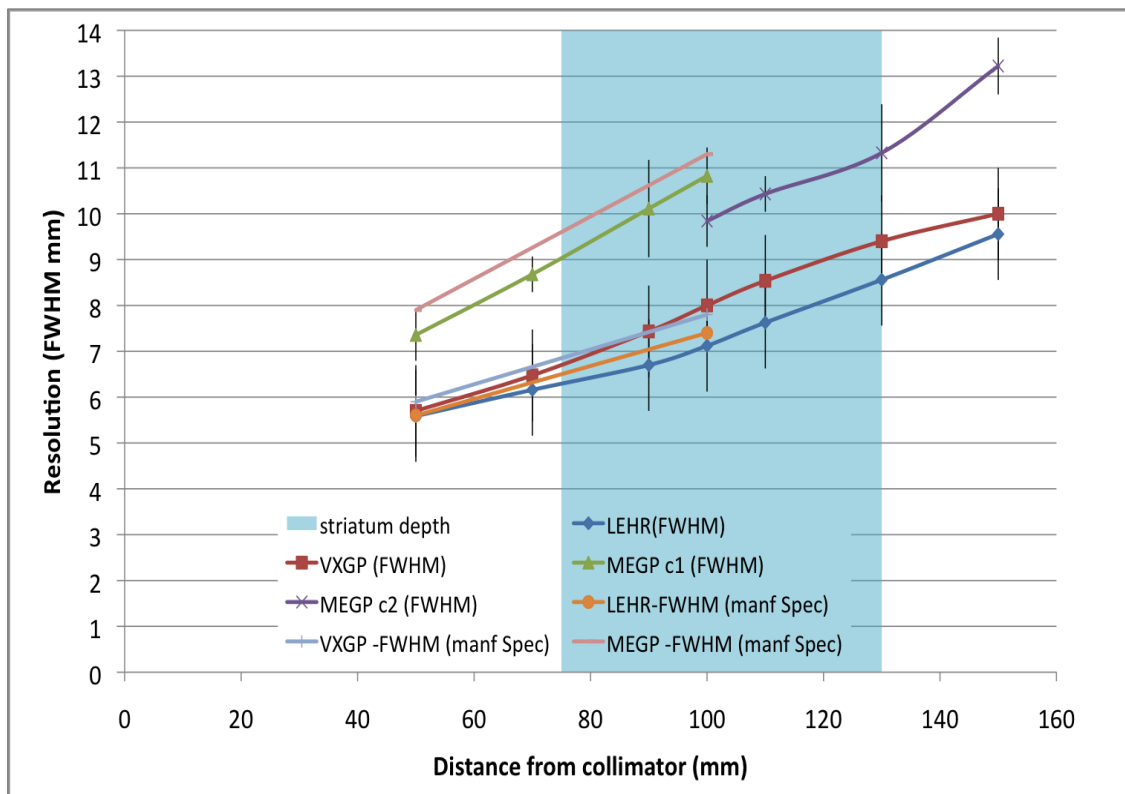


Figure 2-3 Resolution as a function of distance from the collimator for three collimators

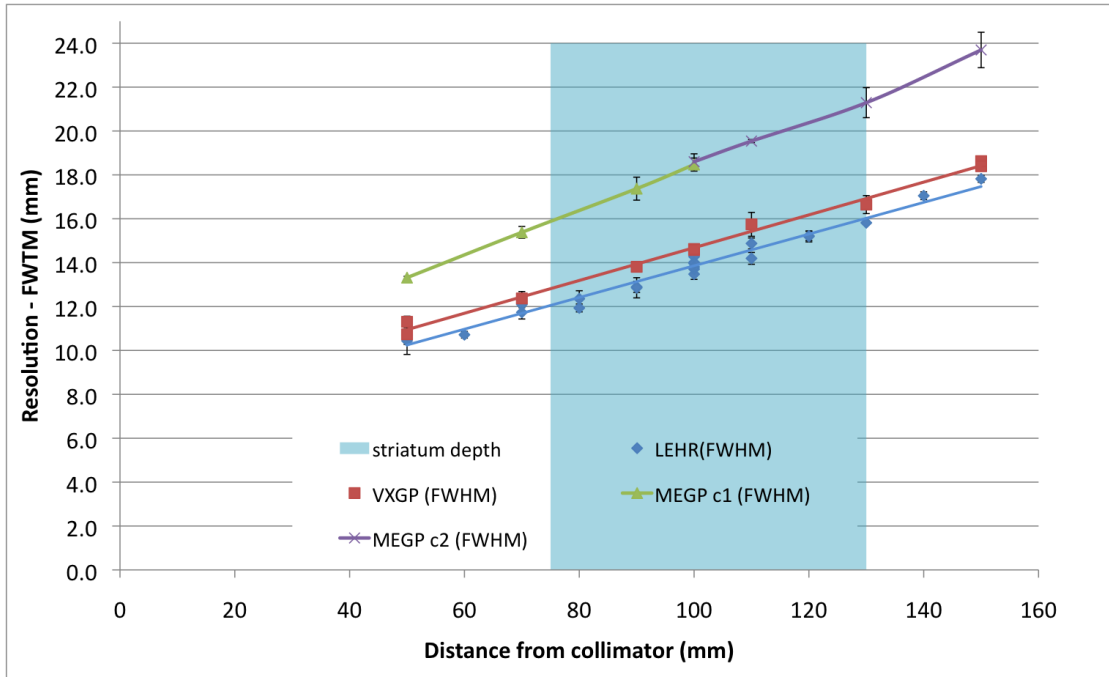


Figure 2-4 FWTM Resolution as a function of distance from the collimator for three collimators

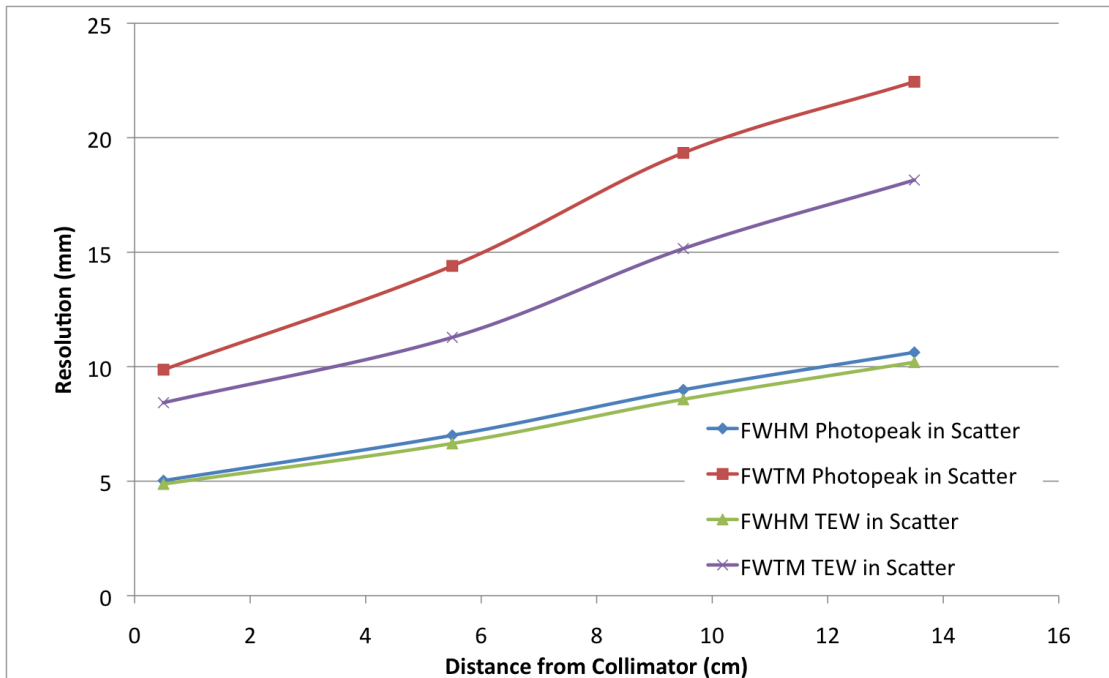


Figure 2-5 Resolution as a function of distance from collimator in scattering material. The FWHM and FWTM are shown with and without TEW scatter correction.

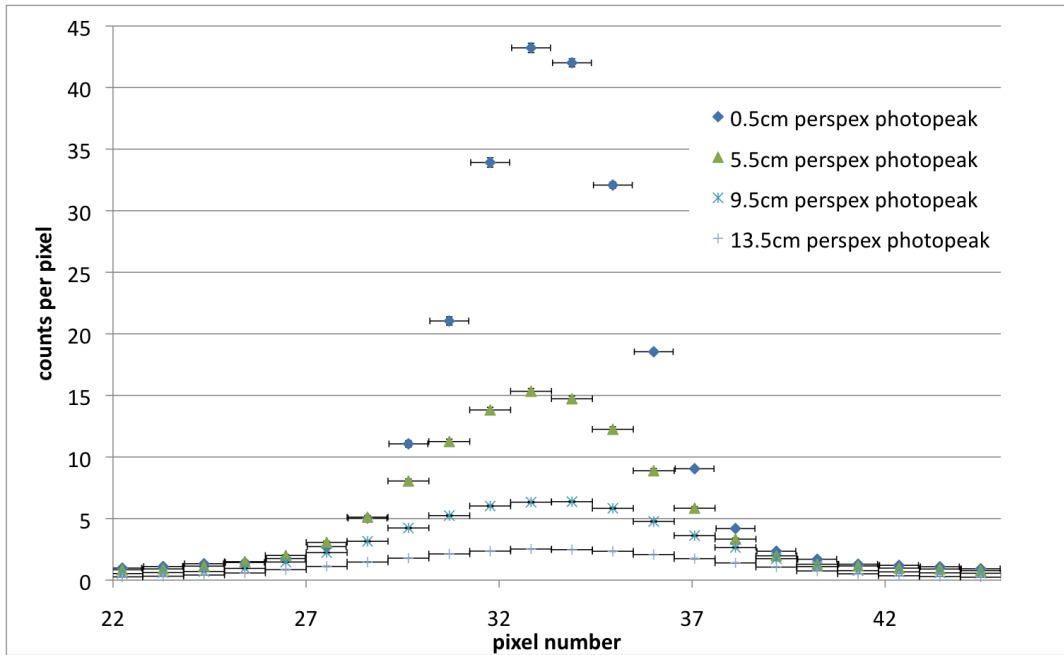


Figure 2-6 Line profiles across line sources at various depths in scattering material

Figure 2-6 shows the profiles of the line sources acquired in a scattering material. Counts per second per pixel vs x position (mm). Y-error given as +/- square root of counts/pixel. The x-error is +/- half pixel. The counts per pixel have been rescaled according to the length of acquisition. As the line sources are positioned deeper into the material, further away from the detector, the signal is spread out and attenuated.

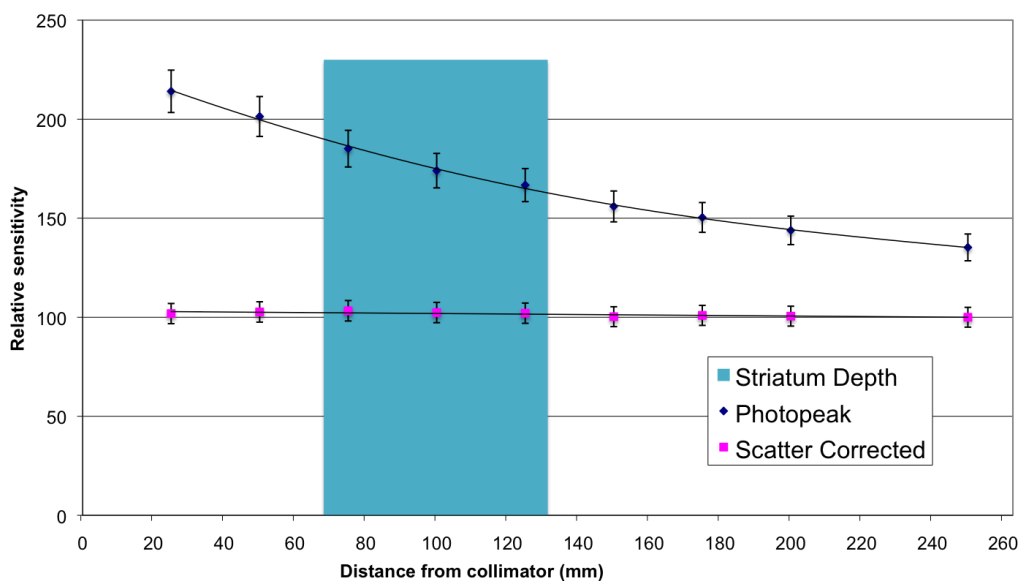


Figure 2-7 Relative sensitivity with distance for photopeak window and TEW corrected photopeak window

Figure 2-7 shows the relative sensitivity as a function of distance from the collimator. The data has been decay corrected and then normalised to the sensitivity at 25cm with scatter correction.

In SPECT imaging for average sized patients with no positioning difficulties, the striatum occupies a distance of up-to 4.5cm from the centre of rotation. This equates to between 6.5 – 10.5cm from the nearest collimator. This distance has been shown on Figure 2-3 and Figure 2-7.

All collimator performance measurements were within manufacturers' specifications.

2.3.3 Collimator comparisons - Discussion:

As can be clearly seen in these graphs, improved resolution is seen with the two low energy collimators compared with the medium energy. There is a small

improvement in resolution when imaging with LEHR collimators rather than VXGP.

The FWTM results also show an improvement when moving from medium to low energy collimators, however, the relative improvement is lower.

Comparing the resolution results with the distances seen in clinical imaging, the expected planar resolution for the organ of interest will range from 6.5mm to 10mm FWHM.

Normally the compromise in collimator choice is between resolution and sensitivity, however, the sensitivity is increased when using low energy collimators than compared to MEGP as less of the camera face is covered with lead. These results show a clear improvement with low energy collimators, with little difference between the LEHR and VXGP styles. A possible problem that has shown here is the varying sensitivity with distance that is seen with the two low energy collimators (Figure 2-7). This effect should not be present in parallel-hole collimators, but appears here because of septal penetration. Figure 2-7 also shows how scatter correction can reduce this effect. Scatter correction is discussed further in chapter 3. The effect of scatter correction is also seen in Figure 2-5 where the FWTM shows a decrease when scatter correction is used. This relates to a reduction in the broad tails of the line-spread function and suggests improved image contrast. This is in contrast to some results published using a different correction method⁸⁸ where the improvement in FWTM was not found to be significant. In that work only a down scatter correction was being applied, rather than full scatter correction method shown here.

The resolution performance defined here can be further used when investigating

the reconstruction techniques, where the depth dependence characteristics are modelled in the reconstruction. This will be further discussed in chapter 3.

2.3.4 Collimator comparisons – Conclusions

The two low energy collimator designs should give improved image quality in terms of reduced blurring. The LEHR is marginally superior to the VXGP collimator but both will be used for clinical imaging. The measured line spread functions can be used in the reconstruction process to try to improve reconstructed resolution (see chapter 3). The expected FWHM for the striatum, varying between 6.5mm 10mm can be used to guide the acquisition pixel sizes (see section 2.4).

2.4 Pixel size / Image Matrix

Sampling theory dictates that the highest spatial frequency data that can be defined by an imaging system will be half of the spatial sampling frequency. In the absence of noise considerations, the spatial sampling (i.e. the pixel size) should be defined based on the FWHM at either the centre of rotation or the depth of interest. In section 2.3 the limits to the imaging system were defined in terms of blurring factors at distances from the collimator. The EANM guidelines⁸⁰ suggest a pixel size between one third and half of the expected FWHM.

2.4.1 Pixel and matrix size – Results

From the resolution measurements, (section 2.3) and from examining typical orbits used clinically, the following calculation can be made:

FWHM at depth of object of interest covers the range 6.5mm to 9.0mm

FWHM at centre of rotation = 9mm

Using the EANM guidelines for the FWHM at the depth of the organ of interest - the maximum pixel size should be between 2.1 and 3.2mm. When the resolution at centre of rotation is used the maximum pixel size should be between 3.0 and 4.5mm.

The image matrix sizes available for the departmental scanners are limited to pre-defined ranges (see Table 2-3). In SPECT mode only 128*128 and 64*64 matrices are allowed. Also, the range of zoom settings is restricted to 48, 38, 30 and 25.4 cm. (zoom factors 1, 1.46, 1.85 and 2.19 respectively).

Table 2-3 Pixel Sizes for SPECT mode acquisition

	Field of View			
	48cm	38 cm	30 cm	25.4 cm
64*64 matrix	9.3mm	6.4mm	5.1mm	4.3mm
128*128 matrix	4.7mm	3.2mm	2.5mm	2.1mm

2.4.2 Pixel and Matrix size - Discussion

From the resolution measurements, a pixel size of 2.1mm is suggested. This figure is significantly smaller than the “~4.6mm” recommended by the pharmaceutical manufacturer. This discrepancy can be understood once the effects of noise are considered. By halving the length of a pixels edge, the scanning area it represents will decrease to $\frac{1}{4}$, with a corresponding drop in counts to $\frac{1}{4}$. The lowered counts may lead to noise problems in the clinical data. As a compromise, data could be acquired in 128matrix form, with a zoomed field of view. This will allow the high spatial frequency data to be acquired, but will

still allow the pixels to be combined to give 4.3mm voxels.

2.4.3 Conclusions:

The recommended matrix size should be 128*128 using a 25.4cm field of view, with the option to combine data into 64*64 data.

2.5 SPECT gantry orbits

A previous section (section 2.3) showed how the distance from the camera affects the resolution of the images. When performing SPECT acquisition, the detectors are rotated around the object of interest, maintaining the minimum distance to ensure the best resolution. There are two shapes of orbit, circular and non-circular. A circular orbit uses a fixed radius of rotation for each head, which when equal, produce a circular orbit. When images are acquired using a non-circular orbit, the camera heads will change radius of rotation as the gantry rotates. The idea of the non-circular orbit is to allow the camera movement to better reflect the shape of the object being scanned. Theoretically the non-circular orbit has the potential to give better images, as it should be able to maintain a smaller object to camera distance, however there are problems in the implementation. The non-circular orbit is defined by marking the radii to be used at 0, 90, 180 and 270 degrees. The gantry computer then uses these values to define the ellipse. It had been noted anecdotally that the resulting radii were significantly bigger than the set values. This was investigated.

2.5.1 Gantry Orbit measurements - Method

To investigate how the camera orbits are defined a series of dummy acquisitions were performed. This entailed setting up the camera for a series of acquisitions,

at different starting radii, and then monitoring the radii used in the acquisition. The starting radii were chosen based on the starting radii taken from clinical acquisitions within the department. The radii were recorded from the gantry display during the acquisition.

2.5.2 Gantry Orbit measurements - Results

Figure 2-8 shows the radii of the camera heads whilst acquiring a circular or non-circular orbit. For clarity, only one representative orbit set-up is shown. The original gantry definitions were for a 13cmx10cm radii elliptical orbit and 13cm circular orbit.

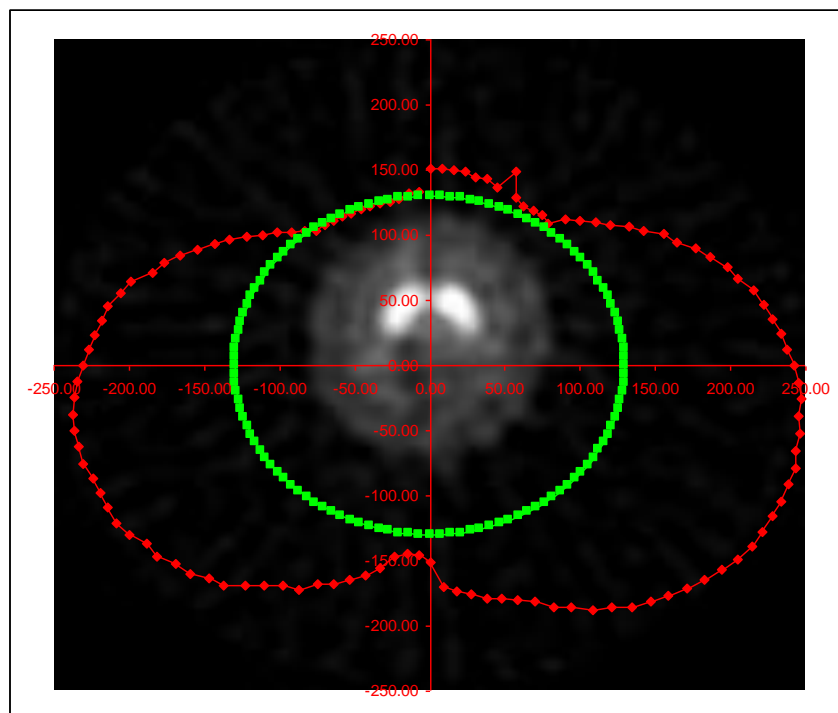


Figure 2-8. The radius of rotation as measured for a phantom experiment. The scaled phantom image is also included

Figure 2-8 shows the gamma camera detector radii as the system rotates around the phantom. A suitably scaled transverse slice from the phantom has been added to show the extent of the radius.

2.5.3 Gantry Orbit - Discussion

Figure 2-8 clearly shows a significant increase in radius of rotation when non-circular orbits are used. The gantry software automatically adjusts the orbit definition to avoid the patient couch, even when the couch is retracted away from the detectors, as is the case with brain imaging. The same phenomenon is shown on both scanners.

The increase in rotation radius when non-circular orbits are used will clearly lead to a decrease in image quality, and hence circular orbits should be used for all brain imaging until updated software has been provided and tested.

One advantage of using circular orbits occurs when the data is reconstructed. With the software available, a depth dependent resolution recovery calculation can be performed.

The camera manufacturer was contacted regarding this issue. To date no changes to the gantry control software have been made.

2.6 SPECT Gantry movement

The gamma camera is capable of acquiring data in two modes, step-and-shoot mode (SSM) and continuous step and shoot mode (CSSM). The first option only records counts when the detectors are inline with the predetermined projection angles. At this point, they are held stationary for the full time of the "time-per-azimuth" setting. In the second option, CSSM will record the counts during the movement between the projection angles together with the counts when the detectors are in position. The advantage of CSSM is a decrease in unused imaging time, at the expense of a theoretical blurring effect.

The advantages of using CSSM over SSM when fast SPECT acquisitions were performed have been described ⁹³. The increased count statistics within the same total imaging time lead to improved image quality. However, this increased imaging time is of reduced importance when the time per azimuth is much greater than the gantry rotation time. The theoretical blurring was shown in noiseless images, but was masked in realistic images when 64 views were used. They also note that the blurring effect will decrease further with increasing number of azimuths. The theoretical blurring introduced in CSSM is a function of distance from the COR.

In ¹²³I-FP- β -CIT imaging 128 views are used, and the main structure of interest lies near the COR, it is unlikely to be significantly affected. The "deadtime" associated with step and shoot acquisition is 1.5% of the imaging time. Since the exact choice was unlikely to have a significant effect on the resulting images, the continuous step-and-shoot mode was chosen as it has been suggested that it may cause less damage to the gantry.

2.7 The number of projections

A parameter related to the gantry movement is the number of projections (or azimuths) that are acquired. When too few azimuths are acquired, some artefacts are seen in the reconstructed slices. The effect gets progressively worse as the distance from the COR increases. A number of different approaches have been used to define the ideal number of azimuths⁹⁴. The spacing between line integrals should be substantially less than the resolution of the final image. Values of 0.4 to 0.7 times the resolution have been quoted. For the geometry used here this gives an arc of 0.54cm and hence a minimum of 128 azimuths.

The choice of number of projections will also have an effect on reconstruction speed, particularly if iterative methods are employed. An increase in the number of azimuths will cause an increase in reconstruction time. However, most reconstruction algorithms, including the ones available for this work, are performed using ordered subsets methods that will limit this effect. A further practical implication of the number of azimuths is the associated increase in imaging acquisition time.

2.8 Time per projection

Noise has a large effect in nuclear medicine imaging. The random nature of radioactive decay, and interactions of photons with matter, coupled with the low sensitivity of a gamma camera, means noisy images. The maximum activity that can be used is given in national guidelines ⁹⁵ so the only way to increase the number of counts is to increase the acquisition time. Unfortunately, it is impossible to ensure patients will not move for long periods of time. Even when patients are held still using immobilisation devices such as chin and head straps, and leg supports, there is a limit to the length of time they can be expected to be scanned for. This limit is normally set at around 45 minutes. Movement can be a particular problem when a patient has movement disorders and/or dementias.

2.9 Conclusions

In our department the imaging is now performed with the following parameters. LEHR or VXGP collimators, 128x128 matrix with a 25.4cm field of view. Circular CSSM orbits of 128 azimuths. Radius of rotation should be set as small as possible, and should be less than 13cm. This measurement should be recorded

on the patient notes for use in image reconstruction. Energy windows should be defined as: Photo-peak - $158 \pm 10\%$ Low scatter - $137 \pm 3.5\%$, High scatter - $178 \pm 3.5\%$.

Table 2-4 Comparisons of local settings with published guidelines

Parameter	Local setting	EANM 2010	GE
Delay to imaging	min 3 hours	3-6hours	3-6hours
Imaging device	2 detectors	Minimum 2 detectors	2 detectors
Collimators	LEHR or VXGP	LEHR or LEUHR Fan beam preferred	VXGP / LEHR
Orbit style	Circular	Not stated	Circular
Orbit mode	Continuous step and shoot	Step and shoot, (but continuous may reduce time and reduce mechanical wear)	Not specified
Radius of rotation	Smallest possible <13cm	Smallest possible	Smallest possible, 11-13cm
Matrix	128*128	128*128	128*128
Angular sampling	2.8degrees	3degrees	3degrees
Pixel size / Zoom	2.1mm	Between $\frac{1}{2}$ and $\frac{1}{3}$ expected resolution	No zoom, ~4.6mm
Total counts	As acquired	>3million	>1.5million
Total time	45mins	30mins typical for triple headed system	30sec/projection (30mins total)

Table 2-4 shows a comparison of the new local acquisition settings to published recommendations. These settings are similar to those used elsewhere and in the European guidelines⁸⁰ but with slightly reduced pixel sizes and the addition of scatter correction. Summing the pixels to give 64*64 matrix data will give similar data to the manufacturers guidelines.

3 Image reconstruction and correction techniques

The previous chapter described the various parameters and options available for the acquisition of images from a patient undergoing brain-SPECT imaging. The resulting images are known as projection images. These projection images cannot be analysed in this form and must be combined to produce images that are suitable for clinical reporting and analysis. This process is known as tomographic reconstruction, the output of which is a series of transverse slices. Reconstruction can be performed using a number of algorithms, most commonly Filtered Back Projection (FBP) or some form of iterative algorithm, (e.g. Ordered-Subset Expectation-Maximization, O.S.E.M.). For both techniques, corrections can be applied for attenuation, scatter and partial volume effect. Some post reconstruction filtering can also be applied to remove excess noise.

Although acquired as a series of 2D projection images, the reconstruction models described below work on smaller groups of this data described using sinograms. Sinograms are simply the projection data re-binned into a different form. Figure 3-1 shows the projection and sinogram data from a scan of a brain phantom. The images on the left show four projections at (from top to bottom) 270, 180, 90 and 0 degrees, relating to views from the left, posterior, right and anterior of the phantom. Each projection has the same central row of pixels highlighted. Each projection image is orientated so the horizontal axis is 'x' and the vertical axis 'z'. The right hand image shows the sinogram for the row of pixels highlighted in the projections. The sinogram is produced by stacking the same row of pixels from each of the projection angles, so the horizontal axis is also 'x', but the vertical axis is the projection number/angle. In this example, there are 128 projections of

128x60 pixels, which can also be viewed as 60 sinograms each measuring 128 x128.

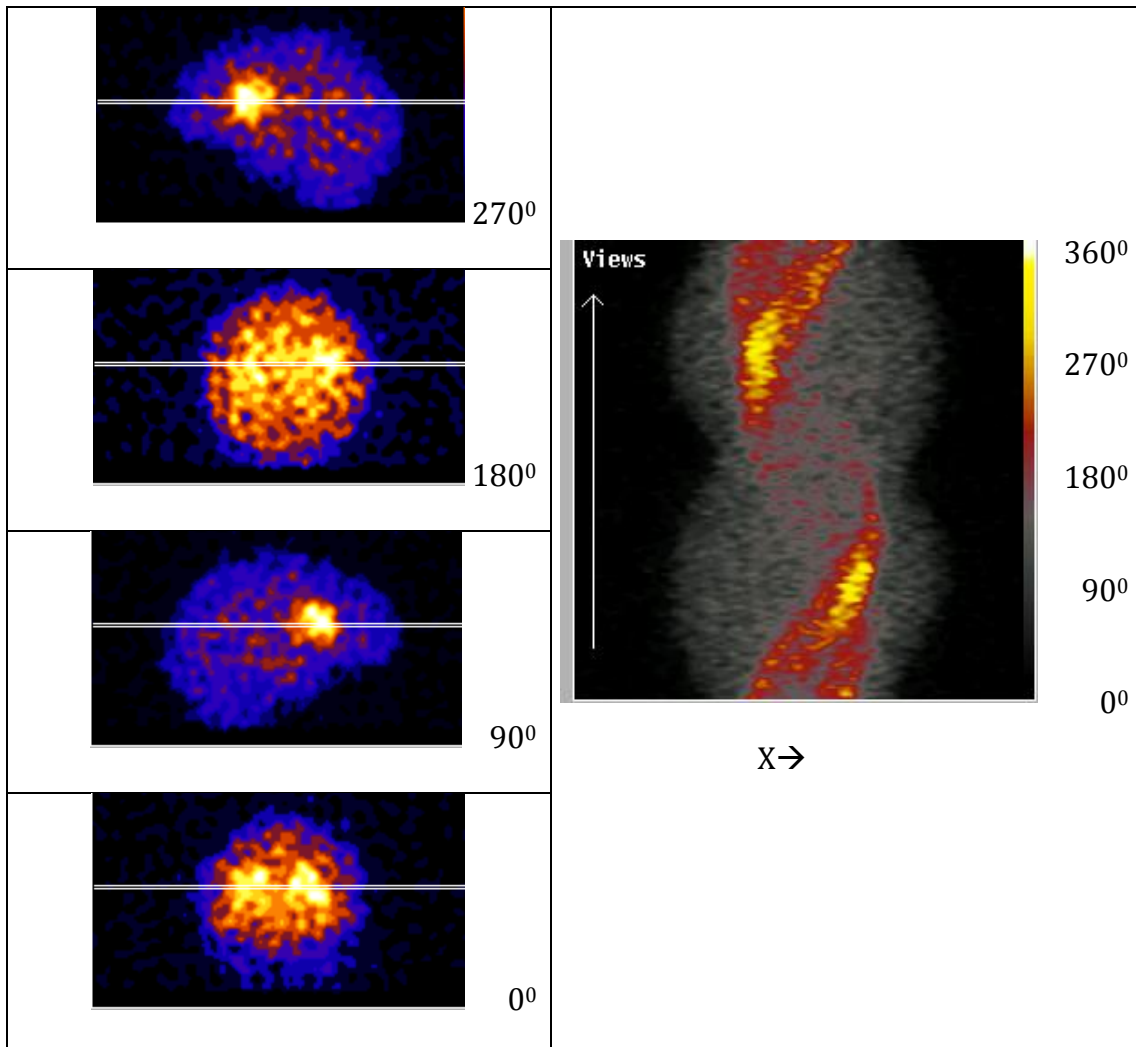


Figure 3-1 Projection images and sinogram data

Each transverse slice is therefore described by a single sinogram. The following descriptions of reconstruction models describe how this sinogram data is used to produce a transverse slice.

3.1 Filtered Back Projection

All gamma camera systems are capable of using FBP, and as such it is often the default reconstruction to use. Because of its linear form, it should give consistent performance with regard to quantitation.

There are two filters used in the FBP reconstruction process. The ramp filter is required to counteract the smoothing function inherent in the back projection process. For a description of this effect the reader is directed to text books on this subject⁹⁶ a full proof of this can also be found⁷². The amplification power of the ramp filter is proportional to the frequency and hence the ramp filter amplifies high frequency data. In theory, the ramp filter will apply exactly the right level of filtering for the reconstruction process. In practice the high frequency data that the ramp filter amplifies the most will mainly be noise. To prevent the amplification of noise the behaviour at higher frequencies needs to be modified. A sharp cut-off in the frequency response will lead to ringing artefacts in the image, so a range of filters have been implemented that act as low-pass filters with smoothly changing frequency responses. These filters are often known as windowing filters, as their action is described as a modification of the normal ramp filter. In practice, these two filters are normally applied separately during the reconstruction process.

The Butterworth filter is commonly used for clinical SPECT processing. This filter has two variables, the cut-off frequency and the order or power. The equation used to describe this filter has a number of different forms, with slight variations between software companies. The FBP software used in this work uses the form shown in Equation 3-1.

$$\text{Amplitude} = \frac{1}{1 + \left(\frac{f}{f_{co}}\right)^{2N}}$$

Equation 3-1

The cut-off frequency f_{co} defines at which frequency the signal is attenuated to 50%. The filter order defines how steeply the filter rolls off with increasing frequency. Figure 3-2 shows the frequency response for four Butterworth filters. Three filters are shown with filter order 10, and $f_{co} = 0.8, 1.5$ and 2.0 cycles/cm. A fourth filter is shown using $f_{co} = 1.5$ cycles/cm and order 5. The Nyquist frequency (2.4 cycles/cm) and the frequencies related to 6.5mm and 9.0mm features are also shown.

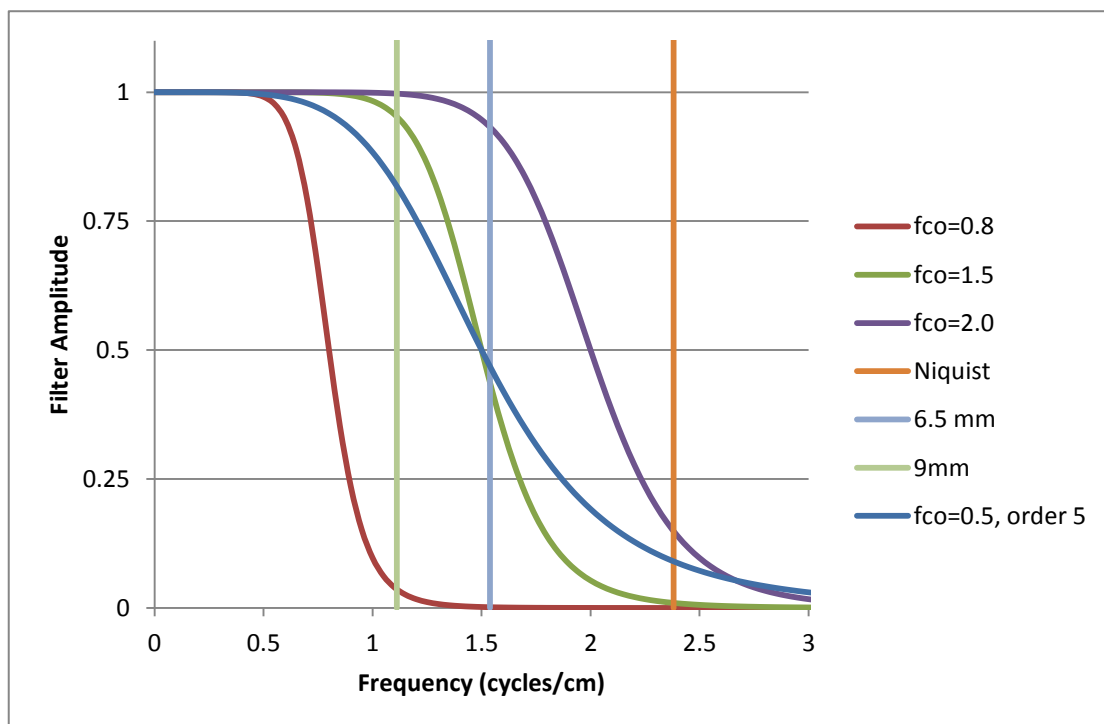


Figure 3-2 Example Butterworth filters

The Butterworth filter is also used for post-reconstruction smoothing where it can be applied to images from both FBP and iterative reconstructions.

3.2 Iterative reconstruction

Iterative reconstruction involves starting with an estimate of the activity distribution. The projection images that would be acquired from such a distribution are then calculated. A comparison of the measured and predicted projection data is made, which can be back-projected onto the distribution estimate and used to improve the estimate. Iterative reconstruction has a number of features that may give it advantages over filtered back projection.

The forward projection step, where estimated projections are calculated from the activity distribution estimate, allows mathematical models of the imaging process to be incorporated into the algorithm (see section 3.3). In the ML-EM iterative model, the measured sinogram is divided by the predicted sinograms to give the error estimate. As the noise in the projection data can be included explicitly in the reconstruction algorithm, iterative techniques tend to deal with low count data better than FBP.

As further iterations are performed, the estimate should converge towards the measured projection data. The convergence of the image with the true activity distribution is not uniform. The local convergence will be dependent on size, location and surroundings of an object. In addition, the convergence is towards the measured projection data, rather than the true activity distribution. Large numbers of iterations will eventually lead to excessive noise levels in the reconstructed image. Figure 3-3 shows a transverse slice through a phantom, reconstructed using a range of numbers of iterations. In practice, a pre-defined number of iterations are performed. This can be followed by the application of some form of post-reconstruction filtering.

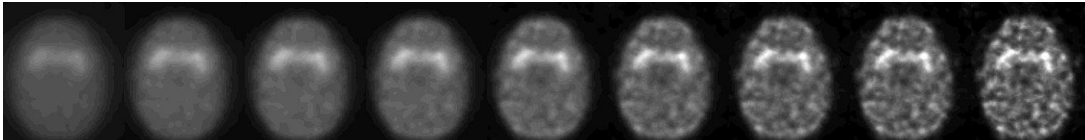


Figure 3-3. Example central transverse slice reconstructed using OSEM with depth dependent resolution recovery and a range of iterations and subsets. The numbers of equivalent iterations (iterations*subsets) are 10, 16, 20, 24, 40, 64, 96, 128 and 200 from left to right.

An accelerated form of ML-EM, known as Ordered Subset Expectation Maximisation (OSEM) ⁹⁷ is used in most commercially available software. Instead of calculating all the projections for every iteration, only a subset of the projections is calculated. The next sub-iteration calculates another subset of the projections and so forth until all the subsets have been used once. In this way, the convergence of the model is speeded up dramatically. The increase in speed is approximately equal to the number of subsets used. The number of iterations multiplied by the number of subsets gives the number of image refreshes or equivalent iterations. A reconstruction using four iterations and four subsets is almost identical to a reconstruction produced using two iterations of eight subsets, or 16 iterations of ML-EM^{72,97}.

For both iterative and FBP reconstructions there are a wide range of parameters that will affect the quality of the images produced. For most of these parameters, their effect on the end result will be dependent on other parameter choices. For example, the optimum number of iterations and post reconstruction filter are heavily interdependent, and related to the use of resolution modelling. With this in mind, the obvious solution is to compare all combinations of parameters; however, this would lead to so many permutations as to be impossible. As such, a step-by-step approach must be made comparing ranges of groups of parameters that are known to heavily influence each other.

3.3 Image corrections

A number of physical interactions can occur between the emission of the gamma ray and the formation of the image. These interactions, along with the properties of the imaging device, will affect the quality and quantitative accuracy of the image produced. By examining these limitations some corrections can be made to compensate for them.

In general, gamma rays can undergo four types of interaction in tissue, Rayleigh Scattering (RS), Compton Scattering (CS), photoelectric absorption (PE) and pair production (PP). Rayleigh scattering is only of importance at energies below 50keV, and pair production can only occur at energies above the rest mass energy of the particle pair ($2 \times 511 \text{keV}$). Hence, for most SPECT imaging including that performed with ^{123}I , the Compton and Photoelectric interactions are of most importance. These interactions are dependent on the energy of the photon, the density of the material and the atomic number of the atoms in the material.

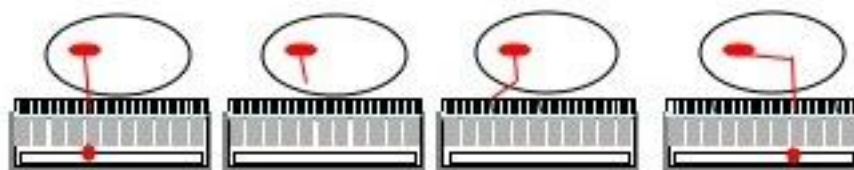


Figure 3-4 Four possible photon paths

PE interactions involve the complete absorption of the photon energy and creation of a photoelectrons and Auger electrons. Compton Scattering results in a photon being deflected with some of its energy being transferred to a scattered electron. Figure 3-4 shows four example photons paths that can occur. From left to right:

- A non-scattered photon hits the detector. It is imaged in the correct position
- The photon undergoes PE in the body and is not detected
- The photon undergoes CS inside the body and is scattered out of the collimator acceptance angle – the photon is not detected
- A photon undergoes CS inside the body and is then imaged in the incorrect position

3.3.1 Attenuation

Attenuation of the signal occurs when photons with the correct trajectory are scattered away from normal incidence with the detector or undergo PE absorption. The effect of attenuation is to decrease the number of photons reaching the detector. The effect is non-stationary and depends on the density and composition of the material. In general, the signal from areas deep within the body will be attenuated more than the signal from superficial structures.

The amount of attenuation in different parts of the object, that is the attenuation map, can be either measured or estimated. Measured attenuation maps can be acquired using transmission-CT with either x-ray or isotope sources, and a suitable detector. Estimated attenuation maps can be produced by defining the object outline and then making assumptions about the attenuation distribution, (e.g. uniform attenuation) within this outline ⁹⁸.

This attenuation map can be used to correct the reconstructed image either post-reconstruction, or in some iterative models, as part of the system model in the forward projector.

For patient data, a separate CT acquisition is rarely performed and even in departments that have combined SPECT-CT systems, the extra radiation exposure from the CT is generally not justified. Within this work, even for the few patients where CT data has been acquired, the software available does not allow the use of CT based attenuation corrections to be included in the reconstruction.

The correction techniques available use a uniform attenuation factor, which can be specified by the user. For FBP based reconstructions a basic uniform ⁹⁸ correction is available. For the OSEM iterative model two implementations of the uniform correction is applied. The first method, (labelled as 'fast correction') follows the FBP implementation, applying the scaling factor after the reconstruction has finished. The second method uses the uniform attenuation map as part of the forward projector within the reconstruction model.

3.3.2 Scatter

Compton scattering has the effect of adding misplaced counts into the image. This leads to a decrease in contrast and the addition of noise. The effect is non-stationary depending on both the density and composition of material in the scattering medium.

In general, the effect of scattering will be to slightly increase the number of counts imaged near the centre of the object. The simplest scatter compensation is to use a reduced attenuation coefficient for the attenuation correction.

Other scatter compensation methods use a measured estimate of the scattered photons. These include Dual and Triple Energy Window techniques (DEW and TEW ^{99,100}). The number of scattered photons being recorded in the photopeak

window is estimated using one or two scatter windows adjacent to the photopeak window. The estimated scatter component can then be subtracted from the projection data prior to reconstruction, or be introduced within the iterative calculation. The latter is preferable in iterative methods, as the subtraction prior to reconstruction will affect the noise properties of the projection data and hence the Poisson noise assumption in ML-EM will be incorrect.

DEW and TEW compensations can lead to an increase in the noise of the reconstructed data depending on their implementation.

The software available in this work allows pre-reconstruction scatter correction in the form of Dual or Triple Energy Window (DEW or TEW) methods. In this form, the scaled counts in the scatter windows are used to estimate the scatter in the photopeak window, which are then subtracted to produce a new set of projection data.

3.3.3 Septal penetration

Some level of septal penetration is always present in gamma camera imaging, however, when higher energy emissions are present, as is the case for 123-Iodine, a greater level of septal penetration will be present. Thus, the choice of collimators is not straightforward ⁷¹. Estimates of septal penetration can be calculated following similar models to the scatter correction estimates. The TEW scatter correction method measures the number of higher energy photons reaching the detector, and as such will include some septal penetration within its correction.

3.3.4 Collimator blurring

The spatial resolution of a gamma camera image is dependent on the collimator properties, and the distance of the object from the collimator. In section 2.3 these depth dependent resolution effects were described and measured. Since these effects are known and predictable, it is possible to incorporate a description of these blurring effects into the reconstruction program.

In the software used in this work, the Gaussian curve is described by (Equation 3-2) where σ is the standard deviation of the curve in direction x.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x)^2}{2\sigma^2}}$$

Equation 3-2

This blurring effect increases linearly with distance from the crystal surface (z) (Equation 3-3).

$$\sigma_Z = \sigma_0 + Z\sigma_{scale}$$

Equation 3-3

σ_0 and σ_Z are the Point spread functions at the collimator surface and at distance z from the collimator surface respectively. σ_{scale} describes the rate at which the blurring increases with distance, z. Default values for these factors are $\sigma_0 = 1.5\text{mm}$, and $\sigma_{scale} = 0.015$.

In practical terms, the blurring function is normally described at specific distances by the Full Width Half Maximum (FWHM), which is linearly related to σ (Equation 3-4).

$$\sigma = \frac{FWHM}{2\sqrt{2\ln(2)}} \approx \frac{FWHM}{2.35}$$

Equation 3-4

Thus, two or more FWHM measurements can be used to describe the blurring effect for the reconstruction program.

The reconstruction program uses either distance from collimator or crystal surface – this has to match the radius of rotation (ROR) measurement given in the projection image parameter details, or entered manually during reconstruction.

The FWHM data described in section 2.3 was scaled by 2.35 to convert to σ and then fitted with a straight line using a least squares fit, to give values for the collimator constant and the scale (Table 3-1). These numbers are all related - the offset for with and without collimator is linear, and the conversion from σ to FWHM is just a scaling factor.

Table 3-1 Resolution characteristics for reconstruction software

		FWHM	σ
Measured from crystal	VXGP	0.0441x+0.9612	0.0187x+0.4082
	LEHR	0.0407x+1.3196	0.0173x+0.5604
Measured from collimator	VXGP	0.0441x+3.5174	0.0187x+1.4937
	LEHR	0.0407+3.0709	0.0173x+1.3041

The values for σ_0 and σ_{scale} for these collimators are very similar to the suggested typical values given in the software manufacturer’s User Manual.

With a depth dependent resolution model, in the absence of noise and with high spatial sampling rates, the resulting images should produce images with uniform

resolution. It is also possible to use a fixed blurring function that is not depth dependent. By using a fixed blurring function, this uniform resolution is not ensured. At a practical level, when noise is included, and discrete sampling is used, this compromise may not be important.

The inclusion of resolution modelling (either fixed or depth dependent), introduces a level of smoothing to the reconstructed images. This smoothing causes the reconstructed image slice estimates to converge more slowly, but can potentially improve the resulting image quality.

In both the fixed and depth dependent resolution recovery models, the forward projection step from activity distribution estimate to calculated projection involves a convolution with a Gaussian blurring profile. Although a 1D Gaussian profile extends infinitely in each direction, for the blurring functions seen in medical imaging the bulk of the counts are seen in the first few pixels. The contribution of counts from a central pixel to neighbouring pixels trails off over a few voxels and becomes small over a short distance. It can therefore be proposed that including just the first few pixels of a blurring function may give similar results to a convolution with the full curve. In the software used in this work, the width to which the blur is calculated is defined in terms of pixels in the X and Z-axis, with Z defining the number of transverse slices the blurring function extends over.

3.3.5 Partial Volume Effect

The Partial Volume Effect (PVE) is present in all discrete imaging modalities, but is particularly prominent when the resolution of the imaging system is of a similar magnitude to the object being imaged. Since the imaging process

involves a certain level of blurring (the system resolution) the counts from a single point source will be spread out over a region. Thus, a region of interest focused on the object will not include some of the counts originating from the object as they will have been imaged outside of the object. This geometric effect is dependent on the resolution of the imaging system and the size of the object being imaged. As the size of an object decreases, the apparent activity concentration will reduce.

Since it is a predictable geometric effect, there are corrections that can be applied to image data. Unfortunately, these corrections require an estimate of the object size. Some work has been done using CT and MRI data to correct SPECT data, however there is an assumption that the MRI or CT volumes are the same as the receptor imaging studies. In dopamine transporter imaging studies, CT and MR data is not generally available. In cases of Parkinsonsonian syndromes, not all of the striatum will actually take up the tracer. In such circumstances, a CT or MR could over estimate the size of the functioning volume, and hence under estimate the correction for some patients.

PVE has been applied in some work to SPECT imaging data, where MRI scans were also available. Soret et al.⁸⁴ found that not including PVE could lead to a bias in uptake ratios of up to 50%. MRI based PVE was also used in a study comparing AD and DLB patients, along with the application of scatter and attenuation correction¹⁰¹. Due to the patient group used, perfect separation between the two groups was seen both with and without PVE. A second set of data that had been simulated using MRI data did suggest that this form of PVE might be of use for diagnostic classification. However, this modelling was based

on the various uptake levels in the simulated data being the same size and distribution as the MRI data that was used to correct it.

A similar implementation of this form of PVE correction was used as part of the BasGan software¹⁰². In this work, the patient data is automatically aligned to a template. The PVE correction factor is calculated for this template and then applied to the patient data set. Although such a technique has been shown to produce more accurate uptake values when used with phantom data, an improvement of clinical diagnostic power was not tested.

The scale of the PV effect is dependent on size of object and relative count density in the two objects. The relative density of counts will affect how many counts “bleed” across the boundary. For high contrast data the “bleeding” will appear more severe, as less counts will be bleeding into the caudate region from the background. However, this effect is a straightforward function of the relative ratios, and hence an uptake figure would be multiplied by some factor to give the true uptake. This is just a linear transformation of the results – a universal scaling factor.

For most forms of Parkinsonism, the disease progression shows both a reduction in the number of DAT sites as well as a reduction in the volume of brain in which these sites are prevalent – relating to the loss of neurons in the putamen prior to the caudate loss. The partial volume effect will actually work to amplify these two effects - with a loss in size of active uptake volume resulting in an apparent loss of uptake counts.

3.4 Image quality metrics

What makes a ‘good’ clinical image? Which parameters should be optimised? A large number of image quality metrics have been used for assessing the quality of medical images. The choice of metric will determine the image characteristic that is optimised. They range in specificity from absolute image qualities, such as measurements of pixels values and noise, through to applied measurements such as the ability of a particular image group to allow clinical disease differentiation. In much of the work in developing reconstruction algorithms, reconstructed slices are often assessed for noise, resolution and contrast. A good algorithm is able to reproduce the contrast in the object, by improving resolution, but still controlling noise. In phantom based studies the reconstructed slices can also be compared to the known activity distribution. A number of intermediate end-points are also commonly used, such as the striatal uptake measurements. In the following descriptions, the image metrics are listed in order of specificity to the imaging task.

3.4.1 Pixel by pixel similarity measurements

When using phantom data, the exact activity distribution is known, and hence the reconstructed images can be compared directly with this distribution. A common measure used in image optimisation is the Normalised Mean Square Error (NMSE) ¹⁰³. A number of definitions for this metric are used using different normalisation factors, but all take the same general form (Equation 3-5).

$$NMSE = \frac{1}{N} \frac{\sum_{i=1}^N (Meas_i - True_i)^2}{\sum_{i=1}^N True_i^2}$$

Equation 3-5

The difference between the measured pixel values (Meas) and expected values (True) are squared and summed over the range of voxels, N. The range of voxels for which this error is defined can affect this optimisation. One choice of voxel range is to include all voxels within the image. However, this may lead to the optimisation procedure being biased by voxels values outside the body. A common choice is to use all voxels that fall within the body outline, although it would be possible to define a subgroup of voxels. One such method uses just the pixels lying along a line profile through the image. This will bias the optimisation towards the exact area contained within the line profile. Such a bias may be desirable, or even be necessary, if the optimisation is to be moved away from unimportant features, but must of course be used carefully.

When comparing reconstruction algorithms, the parameters that produce the minimum NMSE are looked for, so the differences between the normalisation factors used are unimportant.

3.4.2 Noise levels

Noise in reconstructed images has the effect of masking real changes in pixel count density. In general, a lower noise level will mean improved image quality, but this cannot be used as a metric in isolation, since noise suppression techniques can affect image contrast. In phantom studies, it is possible to define regions/volumes where the true activity concentration is uniform. In such circumstances the coefficient of variation in counts for that region can be used as a measure of image noise (Equation 3-6), and will be directly related to the image count statistics and reconstruction performance. For patient studies, this

measurement is confounded by variations in the true activity distribution, which will increase the measured count variation.

$$noise = \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N}$$

Equation 3-6

3.4.3 Region of interest measurements

Moving to larger scales, a calculation in the same form as NMSE can be performed on the measured and expected values of groups of voxels, in Volumes of Interest (VOI) or, in 2D, regions of interest (ROI).

This technique has been applied using a variable group sizes ¹⁰⁴. By changing the size of the VOI, the image optimisation process can be customised to the imaging classification problem being studied. By definition, this approach involves averaging of voxel values, and therefore removes some dependency on noise levels within the image.

The logical progression of this is to apply a VOI which exactly match phantom volumes and compare the reconstructed regional counts with known phantom regional activities.

A simplified version of this is to look at the ratio of region count densities rather than the absolute values (Equation 3-7).

$$UptakeRatio = \frac{Target(CountDensity)}{Background(CountDensity)}$$

Equation 3-7

This comparison of target to background ratios with known target to background ratios removes absolute scaling effects and when expressed as a fraction is known as the recovery coefficient (Equation 3-8).

$$RecoveryCoefficient = \frac{UptakeRatio}{KnownActivityRatio}$$

Equation 3-8

The recovery coefficient is often used when describing the size dependency of partial volume effects in image reconstruction ⁷².

These VOI based measurements have often been used in characterising image quality in neurotransmitter studies as the activity distributions often follow specific anatomical uptake patterns.

These figures were used as a measure of test-retest variability in ¹²³I-FP- β -CIT imaging ⁴⁰. This paper quotes the variability for different uptake times (injection to scan) for ¹²³I-IBZM imaging. They also measured this variability as a function of reconstruction settings to suggest optimized model and parameters.

3.4.4 Signal to Noise Ratio

Since noise suppression methods often result in the reduction of image contrast, this trade-off is often measured directly as the signal to noise ratio. The SNR will depend on regions of voxels used to define it. The noise measurement can be problematic in clinical images as for a true measure of image noise an area of uniform uptake is needed.

3.4.5 Visual interpretation, scoring and confidence

In the vast majority of clinical imaging applications, the main image classification technique is a visual interpretation. Human observers are naturally equipped to assess and grade clinical images on a range of parameters. Visual scoring is generally performed in a qualitative way giving scores on an arbitrary scale (e.g. 1-5) or relative in nature (i.e. selecting preferred option from a presented group

of images). These are used either as an overall classification or for a set range of image quality parameters, such as noise and contrast, as well as overall indices such as image quality. Further qualitative measures can be gained by recording the diagnostic confidence given for an image set.

In practice such measures often show surprisingly good correlations with more quantitative measurements such as signal to noise ratios.

3.4.6 Correlation of uptake measures with clinical disease progression and disease group

Moving to more task specific measures, correlations between these intermediate measurements (e.g. visual scoring, or uptake ratios) and disease progression or disease group are sought. Various clinical disease indices are available which can act as surrogates for disease progression, such as the DSM scoring ¹⁰⁵. An alternative is to use long-term follow up of patients, and retrospectively analyse the images once the clinical diagnosis is defined. Better correlations between these image metrics and disease measurements would imply that the image quality is improved.

The logical conclusion to these progressions is that the images that allow the best differentiation between disease groups are the 'best' images. That is, they are most fit for purpose. However, since the optimisation of the images requires some diagnostic measurement (e.g. VOI ratios) then it is possible that there will be a different optimisation for each particular diagnostic image analysis technique. Image analysis and classification techniques are discussed and investigated in chapters 5 and 6.

3.5 The choice of images for clinical visual reporting

All the patient data used in this project was based on clinical work that was analysed retrospectively. It needed to be reported using visual assessment as part of the normal clinical process. To select the preferred reconstructions for visual reporting three consultant radiologists with specialist nuclear medicine training were asked to score a series of reconstructions.

3.5.1 The Choice of images for clinical visual reporting- Methods

Each radiologist was investigated separately with no knowledge of the others preferences or results. The images were presented in groups of three representing a broad range of noise/contrast compromises. The readers were asked to select their favourite or joint favourite images, using their own criteria for optimum image quality for reporting. They were given no guidance on what they should be optimizing. Based on their choice a further three images were presented covering the range specified. In this way through a succession of levels of choices, their preferred image choice was discovered.

This sequence was repeated for three patients chosen to be representative of the image quality seen clinically.

Using this process optimum OSEM settings were identified using no attenuation or scatter corrections. The process was repeated for FBP images. Their preferred FBP and OSEM images were then compared.

Attenuation corrected images were then produced for this preferred setting and compared with the non-attenuation corrected images.

The process was repeated after the changes were made to the acquisition parameters, as described in chapter 2. With this second set of data a further test was applied to assess scatter correction.

3.5.2 The choice of images for clinical visual reporting - results

The preferred OSEM settings for are to use between 8-12 iterations (with 8 subsets), with little preference between them. A post reconstruction filter with a cut-off frequency in the range 1.2-1.5 cycles/cm was judged to be the visually preferred setting. FBP settings were chosen to include a pre reconstruction filter of 1.3 cycles/cm.

The OSEM images were preferred visually to the FBP images. No preferences between with or without attenuation correction. The scatter correction images were deemed too noisy for visual reporting.

Following these preferences the reconstruction parameters for visual reporting involve the iterative settings without scatter correction. The data for clinical reporting was reconstructed using 12 iterations of 8 subsets, and a post-reconstruction smoothing filter Butterworth order10, cut-off 1.4 cycles/cm.

3.6 Practical investigations of the properties of reconstruction models using phantom data

This chapter focuses on two groups of experiments; the first uses images of phantoms to show the differences between reconstruction models settings and corrections techniques. The second group of experiments investigates the application of correction and compensation methods to these reconstructions.

A number of acquisitions were performed using an anthropomorphic phantom under a range of conditions and settings, chosen to mimic the range of conditions seen in clinical imaging. A large number of reconstructions were then performed on this data using different settings, processing options and correction techniques. The resulting reconstructed data were analysed using the range of image quality metrics described in section 3.4.

3.6.1 Methods – Acquisition of phantom data

An anthropomorphic brain phantom (ECT Striatum phantom, RSD-Alderson, CA, USA) containing five fillable volumes (left/right caudate and putamen and brain background) was used to simulate clinical dopamine transporter imaging. All data were acquired using the VXGP collimator. A series of acquisitions at varying Radius of Rotation (ROR) were acquired using circular orbits of 13, 14, 16, 20 and 25cm. Each series was repeated with different activity concentrations. The time per acquisition was adjusted to achieve total counts close to 2million (Table 3-2).

Table 3-2 Summary of phantom filling parameters and acquisition times for anthropomorphic phantom investigations

Phantom run number	Ratio (target: background)	Activity concentrations (first acquisition) (kBq/ml)	Time per azimuth	Radius of rotation	Total counts acquired (first acquisition)
Run-1	6.9:1	58.6 : 8.5	40s	13,14,16,20	2.25x10 ⁶
Run-2	3.91:1	103 : 34.4	15s	13,14,16,20,25	2.47x10 ⁶
Run-3	2.6:1	93.4 : 35.9	10s	13,14,16,20,25	2.17x10 ⁶
Run-4	1.84:1	87.2 : 47.4	7s	13,14,16,20,25	2.00x10 ⁶
Run-5	1.28:1	82.5 : 64.5	5s	13,14,16,20,25	1.91x10 ⁶

In total, the phantom was filled to five clinically appropriate contrasts, each scanned at five ROR distances using the clinical acquisition parameters as defined in chapter 2 (except the time per projection and ROR). Phantom parameters are detailed in Table 3-2.

After scanning, 1ml samples were taken from the phantom and counted in a well type gamma counter (Wizard 1470 automatic gamma counter, Wallac Oy, Turku Finland). The well counter was calibrated, and normalised using ^{123}I emitters. The count results were corrected for dead time using the counter inbuilt correction models and manually adjusted for background radiation. Results of the gamma counter measurements were used as the true activity ratios within the phantom as shown in Table 3-2.

At a later date, the phantom was imaged using a CT scanner (Biograph64, Siemens Knoxville USA). For this acquisition the four striatal volumes were left empty and the brain background filled with water. CT data were reconstructed with a Bone kernel using a 512*512 matrix and a slice thickness matching the SPECT acquisition.

After reconstruction, the CT data of the phantom was spatially normalised to each of the five SPECT reconstruction data sets manually using a rigid body transform.

3.6.2 Methods – Reconstruction parameters investigated

3.6.2.1 Filtered Back Projection

The simplest algorithm in clinical use is Filtered Back Projection (FBP), for which the algorithms only inherent free parameter is the choice of filter. Some correction factors can also be applied before or after reconstruction but they are

separate from the actual reconstruction. As such, it should be possible to assess the majority of the possible range of images that can be produced by FBP by reconstructing with a selection of filters. Published work generally concludes that iterative models perform better than FBP once the correct parameters have been chosen. Since the range of free parameters available in iterative reconstructions far outnumber those of FBP, the simplest approach will be to find an iterative reconstruction that is better than the best FBP image. If this can be found then FBP data need not be considered further. However, what constitutes the best image will depend on the intended use of the image. For example, Figure 3-5 shows that by varying the filter used the reconstructed slices can vary from very smooth to very sharp, with corresponding changes in noise, contrast and fine detail. Hence the approach taken here is to define the properties of the FBP reconstructions using the methods described in section 3.6.3 and use them as baseline figures with which to compare the iterative models.

Filtered Back Projection reconstructions were performed using commercial reconstruction software (FBP software, version 3.7, Hermes medical, Stockholm, Sweden).

Projection data was convolved with 2D Butterworth smoothing filter, before being reconstructed with FBP and a ramp filter. The Butterworth filter cut-off (f_{co}) was varied from 0.5 to 2.0 cycles/cm, with a fixed order of 10. Figure 3-5 shows a central transverse slice of the reconstructed data from a phantom acquisition with a range of filters.

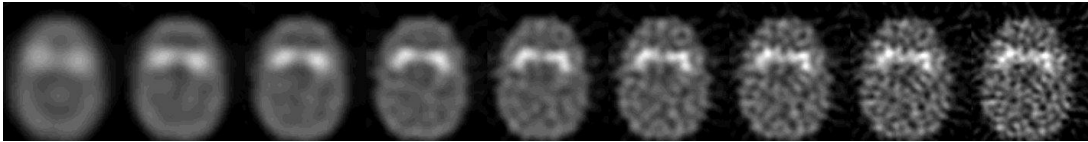


Figure 3-5 A central slice from a number of reconstructions using FBP and a range of Butterworth filters of order 10. From left to right, f_{co} = 0.5, 0.7, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8 and 2.0 cycles /cm. Data from Phantom run number 2, at 13cm ROR.

This range was chosen to extend beyond the range usually used; going from a very smooth filter, to one that approaches the Nyquist frequency of the data. Image reconstruction for visual reporting of ioflupane SPECT imaging is generally performed with $f_{co} = 1.2$ - 1.5 cycles/cm. The Nyquist frequency for the projection pixel size (0.21cm) is 2.38 cycle/cm.

The manufacturer of ioflupane suggests a filter of 0.6 of the Nyquist frequency should be applied, i.e. 1.43 cycles/cm. However, these suggestions are coupled with projection data imaged with pixel size 0.42 cm as opposed to 0.21cm used here.

This set of FBP reconstructions was then used as baselines for comparing the various implementations of the OSEM algorithm under different acquisition parameters.

3.6.2.2 OSEM reconstructions - Resolution modelling

Acquisition data were reconstructed using an implementation of the OSEM algorithm (HOSEM, version 3.7, Hermes Medical, Stockholm, Sweden) with different resolution recovery techniques. Baseline results are produced using OSEM with no resolution modelling. A second series of iterative reconstructions were produced using a fixed blurring function (PSF= 4.0mm, FWHM = 9.4mm).

The third series used a depth dependent blurring function based on resolution measurements in section 2.3.

3.6.2.3 Extent of resolution model in OSEM reconstruction

The default extent of the resolution model employed for these reconstructions was two pixels in x-plane and one in z plane. For one acquisition data set (run1, 13cm ROR) the reconstructions were repeated using a range of x and z blurring sizes. Due to differences in the levels of smoothing inherent in these reconstructions, the number of equivalent iterations was extended up to 750 equivalent iterations.

3.6.3 Methods - Image quality measurements

3.6.3.1 Specific Uptake Ratio and Noise measurement

Volumes of interest (VOI) were drawn on the spatially aligned CT data to cover the known structures in the phantom. A VOI was also placed in the posterior section of the brain. These volumes were then applied to the various SPECT data.

Count densities in the target (Str_{cnts}) and background (Bkd_{cnts}) VOIs were used to define the Striatal Uptake Ratios (SUR) for the various reconstruction parameters using Equation 3-9.

$$SUR = \frac{(Str_{cnts} - Bkd_{cnts})}{Bkd_{cnts}}$$

Equation 3-9

The noise in the image was calculated as the standard deviation of pixel values in the background region divided by the mean average counts in the same region.

3.6.3.2 Calculating NMSE, Noise and contrast

To calculate the NMSE of the reconstructed image, the actual activity distribution must be known on a pixel-by-pixel basis. To calculate this distribution the CT images of the phantom were used.

The realigned CT data of the phantom was imported into the MRICro program ¹⁰⁶. Automatic VOI drawing tools were used to define the extent of the four striatal volumes and the whole brain background. The whole brain VOI includes the regions defined in the striatal volumes. These VOIs were exported as binary image data sets, where voxel values equal to one inside the volume and zero elsewhere. Although the putamen and caudate compartments within the phantom are physically separate objects within the phantom, they touch on each side. The walls of these compartments are made of very thin plastic. When the automatic growing tools are used the Caudate and Putamen regions on each side are formed as single volumes. The phantom scans had been performed using equal activity concentrations in the four striatal chambers, so the striatal VOIs were combined to give one single VOI. The striatal VOI was subtracted from the whole brain VOI to produce a mask that represents the background activity.

The background mask was scaled to the known activity concentration in the phantom and multiplied by a sensitivity factor. Likewise, the striatal VOI was also scaled up, but with an addition scaling factor of the known target to background ratio. The striatal and background activity distributions were summed to give the known activity distribution. Figure 3-6 shows a transverse slice through the calculated activity distribution and a corresponding slice from one of the test reconstructions.

For some phantom acquisitions, the phantom had to be removed and re-fitted between scans. For this data the CT images were re-aligned to the new position following the method above.

All reconstructed phantom data were exported in native format (interfile) converted using ImageJ ¹⁰⁷ into analyse format and then imported into Matlab. The following processing steps were then performed using Matlab functions.

For each test reconstruction the whole brain VOI was used as a mask to set all voxels outside the brain equal to zero. The known activity distribution and the test reconstruction were used to calculate NMSE (as per Equation 3-5).

A measure of noise was recorded using the brain background VOI following Equation 3-6. The striatal uptake ratio was also measured using the average count densities in the striatal and background VOIs (Equation 3-7).

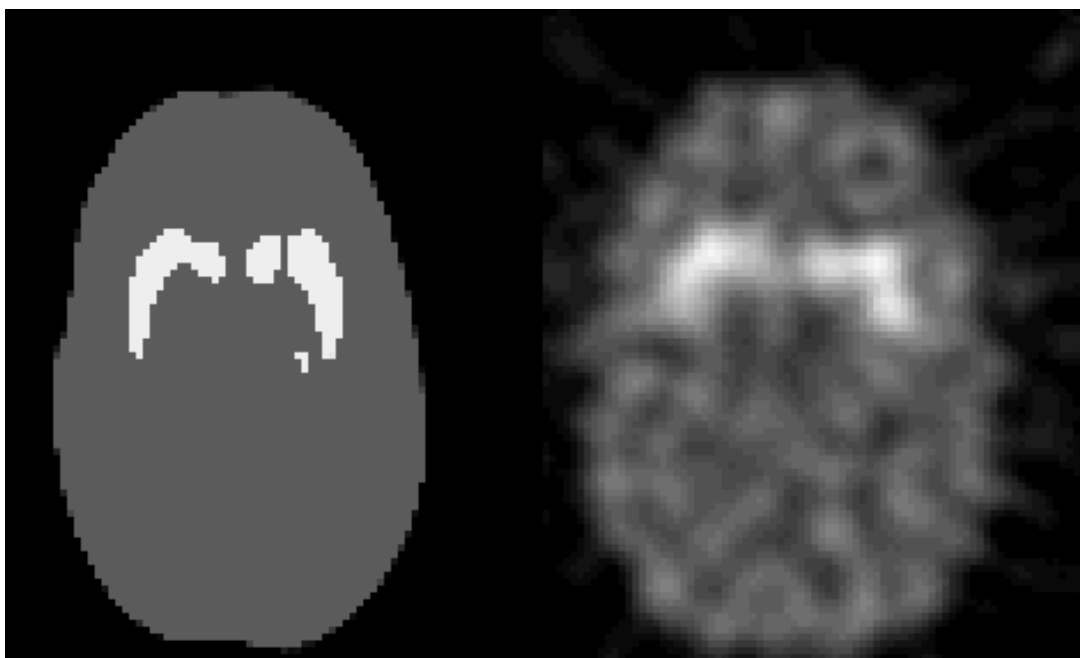


Figure 3-6 Transverse slice through the ideal count distribution (left) and an example reconstruction (right)

3.6.3.3 Calculating Profile-NMSE, noise and contrast

A line profile was drawn across the phantom data on a central transverse slice, from left to right crossing both putamen tails. This line profile was applied to both the reconstructed SPECT data and the CT data. Figure 3-7 shows this profile drawn across a transverse slice. These profiles were exported to a spreadsheet program for further analysis (Excel 2003, Microsoft). Changes in the CT number (measured in Hounsfield Units) across the profile allowed delineation of the separate brain volumes. The CT numbers were then substituted for the expected count value for those regions (defined as per section 3.6.3.2) to give the expected line profile counts. The expected and measured line profiles were used as inputs for the NMSE measurement (Equation 3-4).

As the line profile covered both striatum and background regions another measure of contrast could be made, using the region boundaries given by the CT line profile.

In this phantom study the central region of the brain between the two putamen tails should give a uniform count level. In practice, the variation in counts in this region will be a combination of noise and counts blurring in from the two putamen regions on either side. The coefficient of variation in this region was recorded as further image quality metric.

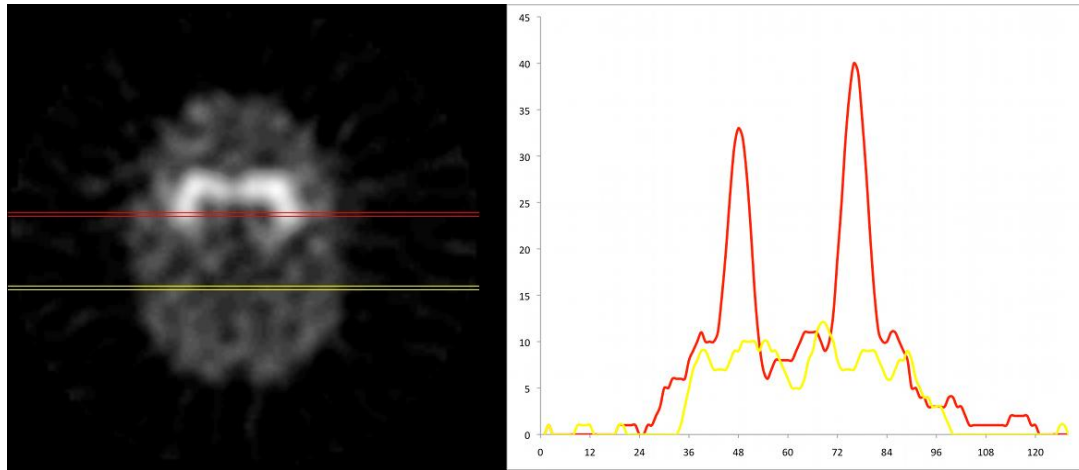


Figure 3-7 Two line profiles drawn across a transverse slice with corresponding profile chart displaying counts per pixel as a function of pixel number

3.7 Results: Phantom experiments with FBP

3.7.1 Contrast and noise - results

The first set of results is from the filtered back projection data. The contrast and noise were measured using method described in section 3.6.3.1 using VOIs defined on the CT image and transferred to the SPECT data.

Figure 3-8 shows the contrast measured as a function of f_{co} of the filter used in the reconstruction. The four contrast curves relate to the first four of the phantom acquisition runs, all acquired at 14cm Radius of rotation. For details of phantom acquisitions see section 3.6.1. The data from phantom run five could not be reliably aligned with the CT template and was therefore left out of this analysis. No corrections for attenuation or scatter were used.

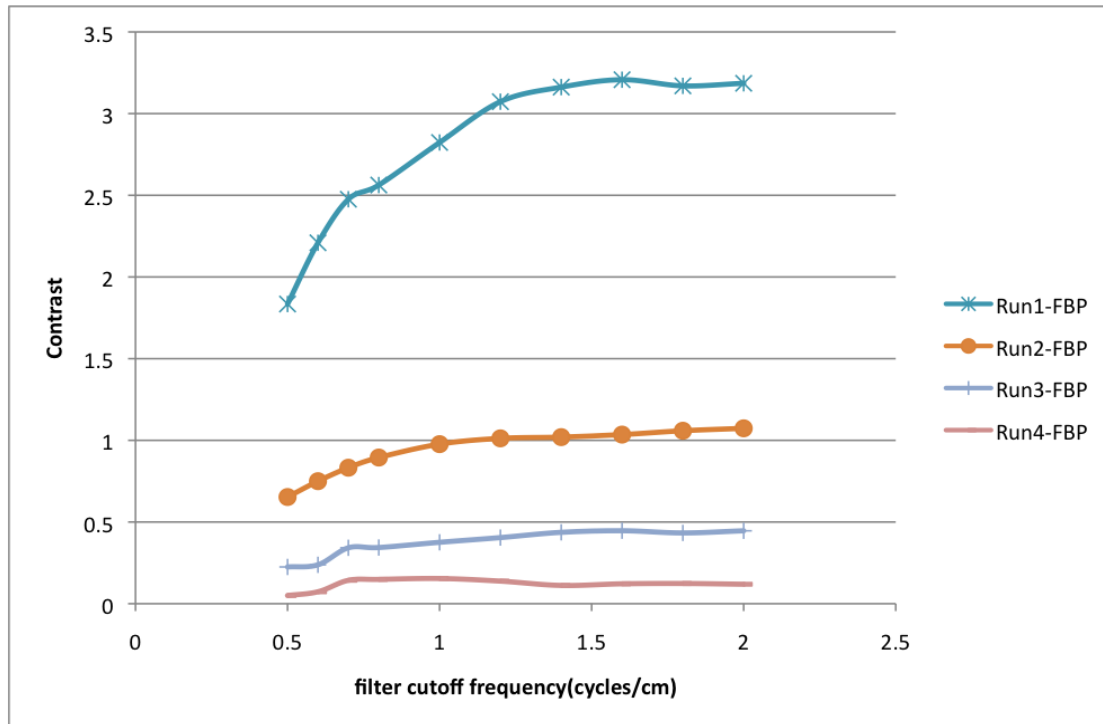


Figure 3-8 Contrast as a function of the cut-off frequency for FBP reconstructed phantom data

The general form of the curves seen in Figure 3-8 is an increase in contrast as f_{co} is increased, before levelling off to a plateau. The same data was rescaled as a fraction of the known contrast in the phantom. This is known as the recovery coefficient. The recovery coefficient is shown as a function of f_{co} in Figure 3-9. In these recovery coefficient graphs, the increase in recovery coefficient with f_{co} is more clearly shown for the lower contrast acquisitions. The highest recovery coefficient for all but one of the phantom runs is seen at the highest frequency cut-off. The maximum recovery coefficient for each run, averaged across the four runs is 0.58.

Using the coefficient of variation across the different phantom runs, the recovery coefficients are most similar at high frequency cut-offs.

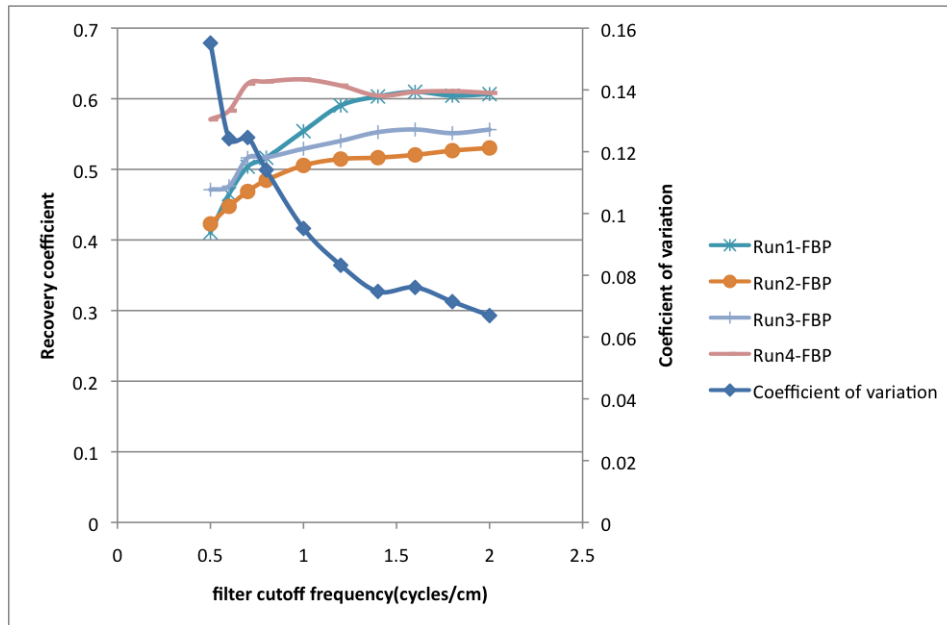


Figure 3-9 Recovery coefficient as a function of cut-off frequency. The variation between recovery coefficients is shown on the secondary axis

Figure 3-10 shows the noise levels recorded in the uniform background region. The reconstructions shown are same as those used to create Figure 3-8 and Figure 3-9.

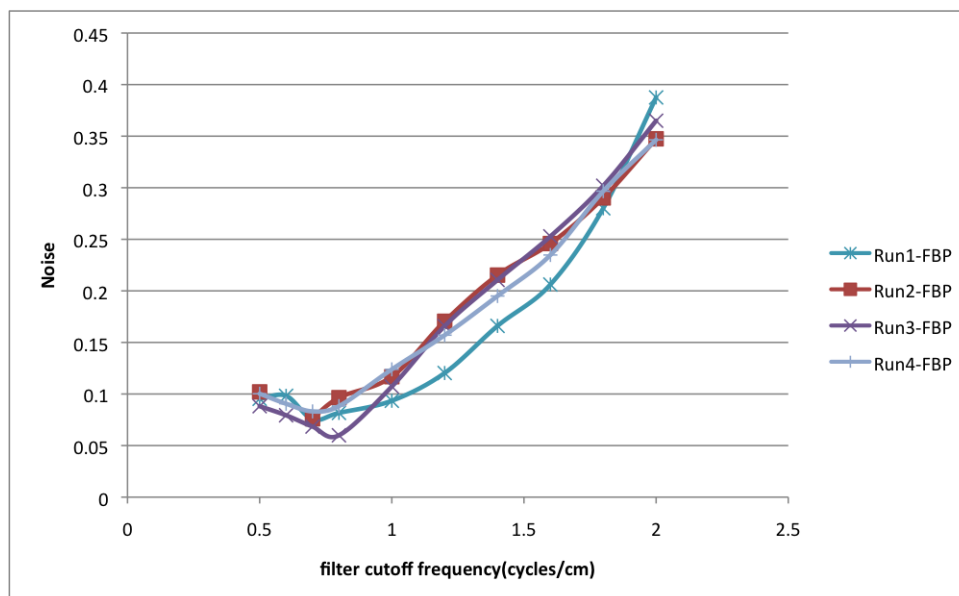


Figure 3-10 Noise in the background uniform section of the phantom, as a function of the filter cut-off frequency

The noise- to cut-off relationship is close to linear for the range 0.8 cycles/cm up to the maximum frequency. At low cut-off frequencies ($f_{co} < 0.8 \text{ cycles/cm}$) the

noise levels show a weak negative relationship with the cut-off frequency. Very similar noise levels are seen across the different phantom acquisitions.

The contrast and noise properties are combined in Figure 3-11. Figure 3-12 shows the same data but with the contrast rescaled to give recovery coefficient.

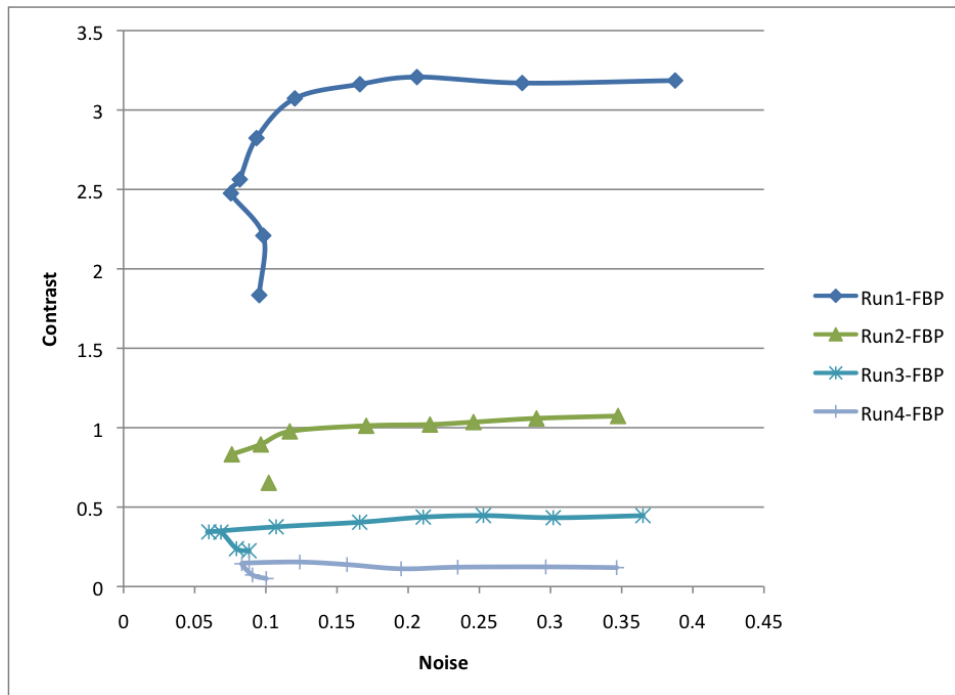


Figure 3-11 Contrast as a function of noise for FBP reconstructions

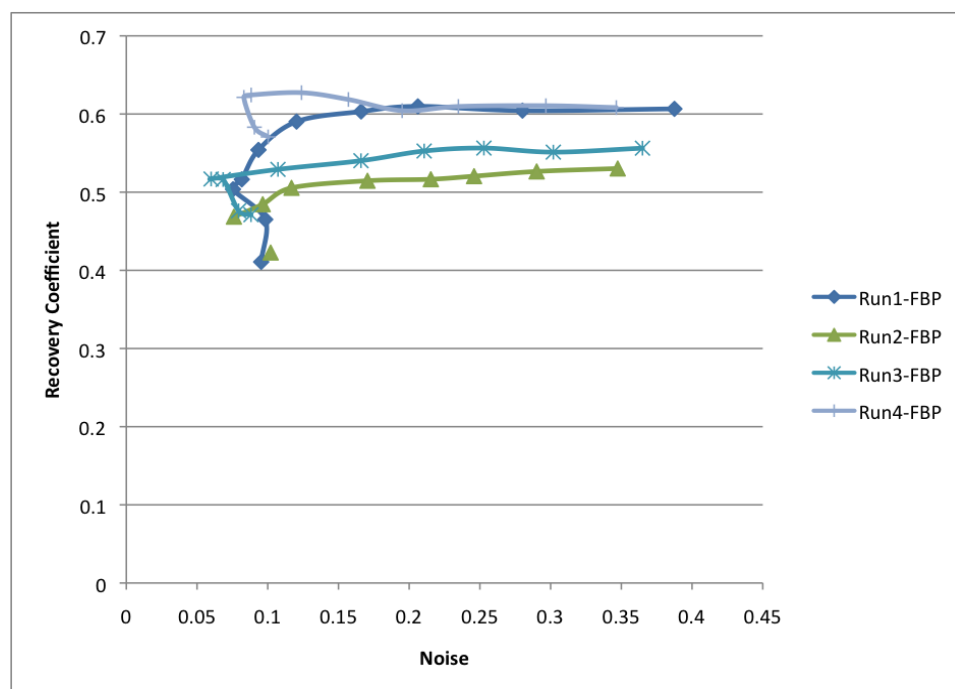


Figure 3-12 Recovery coefficient as a function of noise

The same data expressed as Signal-to-Noise Ratio (SNR) values as a function of f_{co} is shown in Figure 3-13. The peak in SNR relates to the range $f_{co} = 0.7-0.8$ cycles/cm. This frequency is the point of minimum noise in the noise-frequency graphs.

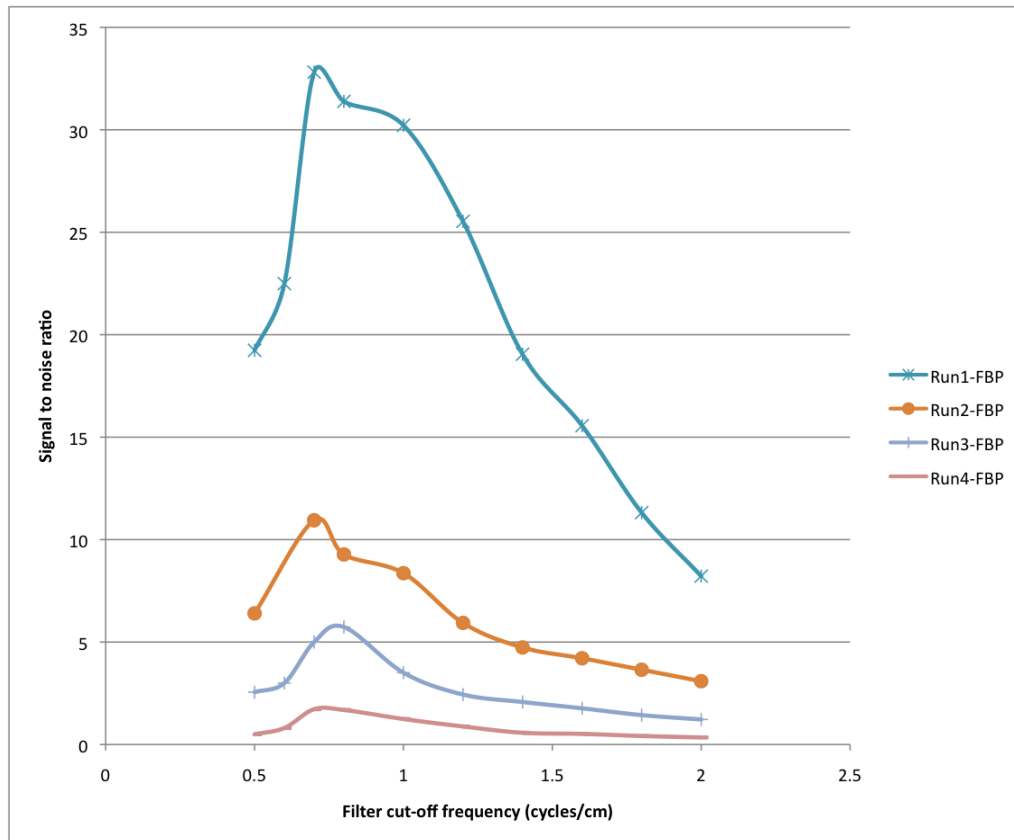


Figure 3-13 Signal to noise ratio as a function of cut-off frequency for FBP reconstructed data

3.7.2 Contrast and noise - discussion

The contrast and noise results from the FBP reconstruction show responses in fitting with a visual description of the images. With reference to the images shown in Figure 3-5 and Figure 3-14, the low frequency cut-off is seen to be smoothing out some of the detail of the size and shape of the striatal volumes. Increasing frequency cut-off allows finer details to be included in the reconstruction. This is seen in the results (Figure 3-8 and Figure 3-9) as a sharp

rise in contrast at low cut-off frequencies. A difference between the visual assessment and the VOI contrast measurement is at which point the contrast is fully recovered. In the contrast curves (Figure 3-8 and Figure 3-9) most of the contrast has been recovered once the filter is above 1.2 cycles/cm. A visual assessment of the images shows that raising this filter to over 1.6 cycles/cm results in better delineation between the caudate and putamen. This is a result of the relatively large VOI used to define the contrast.

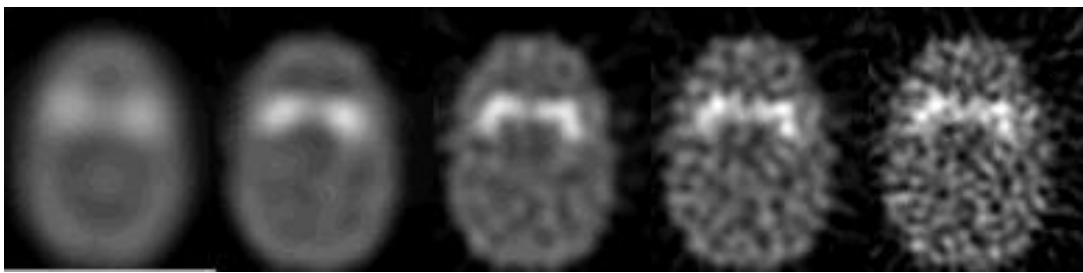


Figure 3-14 A central transverse slice reconstruction using FBP and a range of frequency cut-offs (from left to right, 0.5, 0.8, 1.2, 1.6 and 2.0 cycles /cm).

The recovery coefficient is a function of the size and shape of the object as well as the resolution of the imaging system. Since in these tests the size and shape of the object are fixed, the point at which the variation in recovery coefficient is minimal may be of importance. This minimum appears at high cut-off frequencies. There is no dependency in these results on the actual contrast seen in the phantom. This is perhaps surprising, as PVE would tend to affect low contrast data less than high contrast data. It is postulated that the differences seen in these maximum contrasts is due to the differences in positioning of the VOIs between each set of reconstructions.

The trends in the background noise are interesting as they appear to show a drop in noise over these same low frequencies – examining the transverse images (left most slice in Figure 3-14) it can be seen that the higher count

variation in the 0.5cycles/cm image is not noise but some form of structured pattern. Since the activity concentration is uniform this pattern must be some form of interference pattern.

3.7.3 NMSE - results

A series of NMSE calculations were performed on the data from the filtered back projection reconstructions following the method described in section 3.6.3.2. The use of NMSE as an image quality metric relies on the definition of the expected pixel counts. The expected counts were calculated using the known activity concentrations and sensitivity of the scanner. A large discrepancy was found between the expected counts and measured counts for all reconstructions using FBP. For example, using phantom run number 2 the background counts should average 28.6 counts/voxel. For the range of filtered back projection reconstructions (without attenuation or scatter correction) the average counts in the background region covers the range 8.66 to 8.91 counts/voxel depending on filter. When attenuation and scatter correction are applied this range increases to between 14.07 to 14.15 counts per voxel.

It has been reported that different reconstruction models can give different scaling factors for the image data ^{108,109}. The absolute scaling of these images cannot therefore be used on its own. The relatively small range of voxel intensities seen in the background regions does suggest that some universal scaling factor may be being included in the reconstruction algorithm.

The NMSE analysis was then performed using a base count concentration equal to the background count concentration. Figure 3-15 shows the NMSE score, Signal to Noise Ratio (SNR) and noise for the FBP reconstruction data.

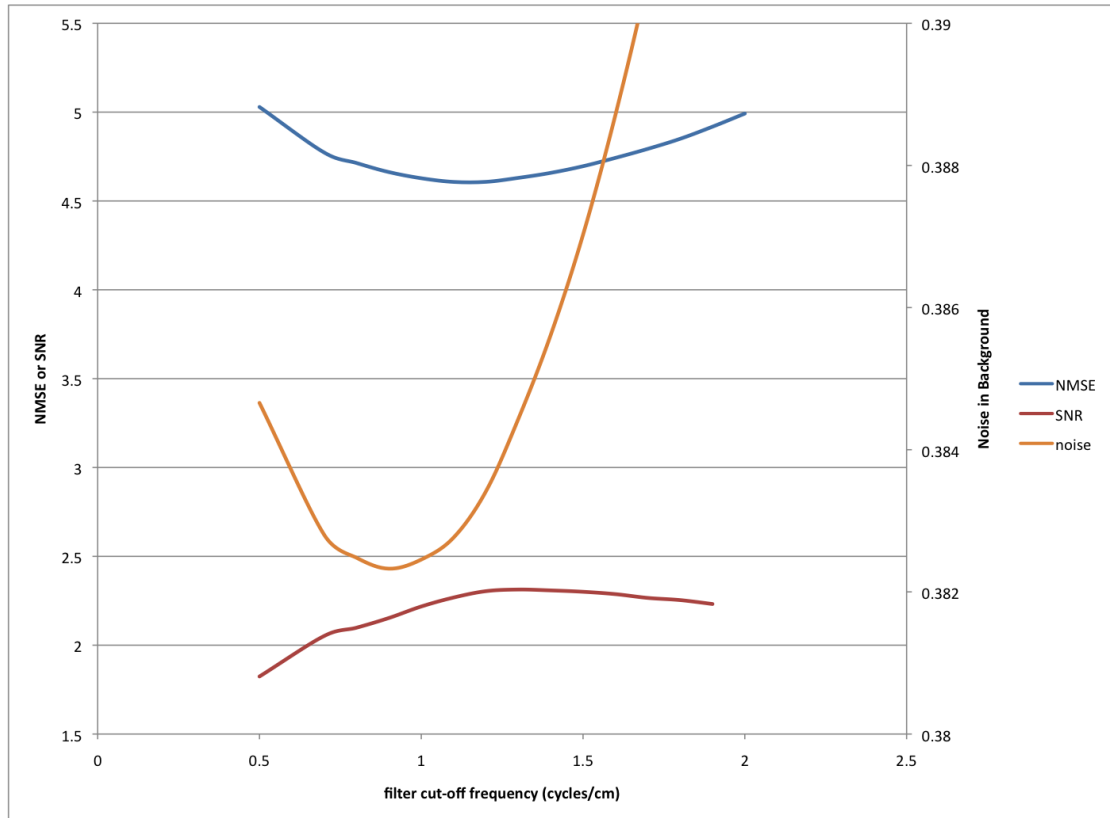


Figure 3-15 NMSE, SNR and Noise as a function of filter cut-off frequency for FBP reconstructed data

In these calculations the noise and SNR levels are measured using the whole brain volume as the background region. The noise levels are at a minimum at a cut-off frequency of 0.8 cycles/cm. The minimum NMSE score occurs at 1.1-1.2 cycles/cm and the maximum SNR occurs between 1.2 and 1.4 cycles/cm.

3.7.4 NMSE - discussion

The use of a scaling factor with the NMSE reduces confidence in its use as an absolute comparison tool. By adjusting the scaling factor the optimum reconstruction parameters will change. A second problem can occur with application of the NMSE score to the phantom data described here. The NMSE score is similar to standard deviation, but it is the variation from an expected value, rather than mean value. Since the vast majority of voxels in this study are

uniform brain background areas (97.7% of active volume), the NMSE score is heavily based on noise in background. This can be seen in Figure 3-15 where the curves for NMSE and background noise are highly correlated. The background in the phantom is a uniform volume, so a reconstruction method that optimises this feature (i.e. a uniform spread of pixel intensities) may not actually be the most appropriate.

3.7.5 Line profile NMSE - results

A line profile was drawn across a series of FBP reconstructions as described in section 3.6.3.3. To overcome the problems with the scaling factors encountered in the whole brain NMSE calculations, the expected background count was set to match the average background pixel count.

The line profiles from the reconstruction using Butterworth filters with $f_{co} = 0.5$ to 2.0 cycles/cm and order 10, are shown in Figure 3-16. The expected profile, based on the CT data and scaled to average background count, is shown as the solid black line.

As the filter cut-off frequency is increased the counts in the centre part of the profile become smaller. At the same time, the two high count regions increase. The positions of the two peaks are also seen to widen slightly with increasing cut-off frequency. The profile for the 2.0 cycles/cm reconstruction shows large fluctuations in the central part of the image. Large fluctuations are also seen outside of the brain volume.

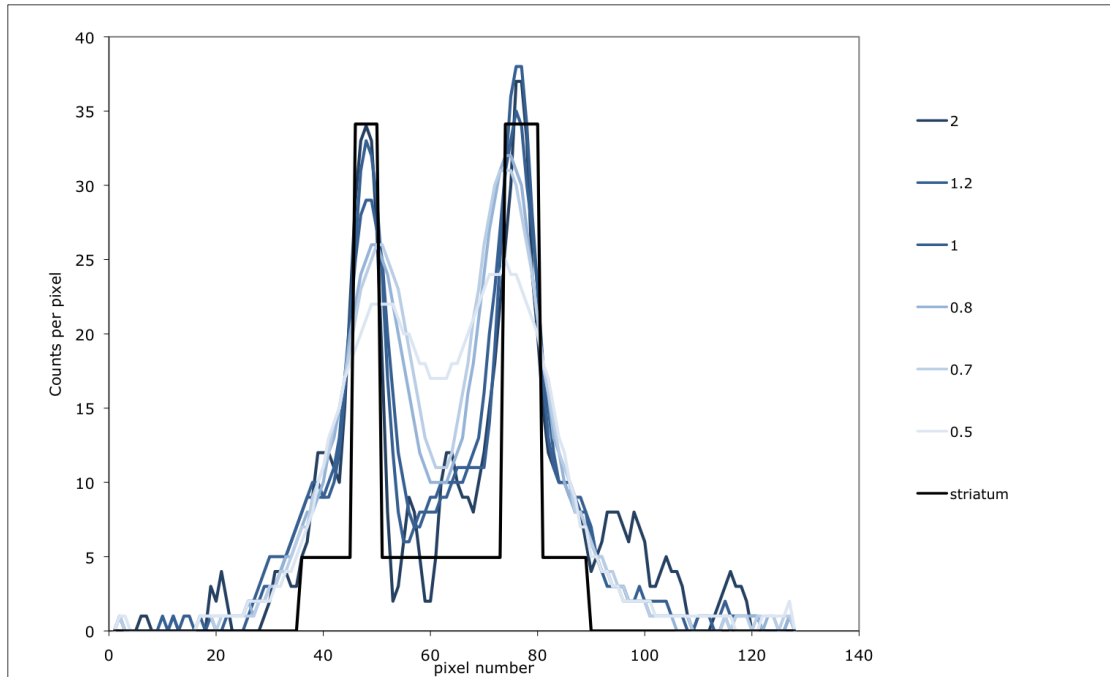


Figure 3-16 Line profiles of FBP reconstructed data using a range of filter cut-off frequencies.

These line profiles were used to define NMSE, contrast and noise as described in section 3.6.3.3. The noise was measured as the coefficient of variation for the central section of the image between the two peaks.

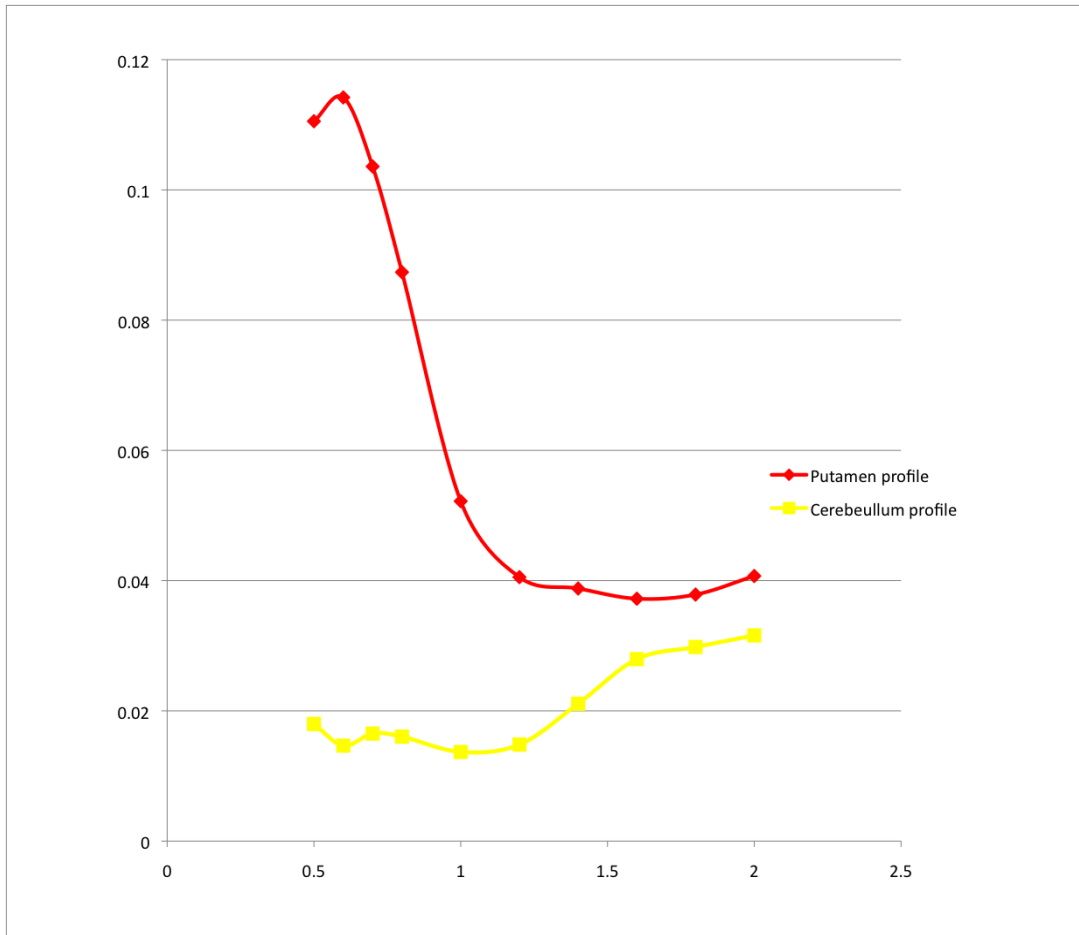


Figure 3-17 NMSE calculated for the line profiles as a function of the cut-off frequency

Figure 3-17 shows the NMSE score as a function of the cut-off frequency for the two line profiles. The NMSE score for the line profile through the putamen region shows a sharp drop for filters from 0.5 to 1.0 cycle/cm. The minimum is reached for a filter with $f_{co} = 1.6$ cycles/cm, but it is a board minimum. The NMSE for the profile drawn across the cerebral region shows an increase between 1.2 and 1.6 cycles/cm.

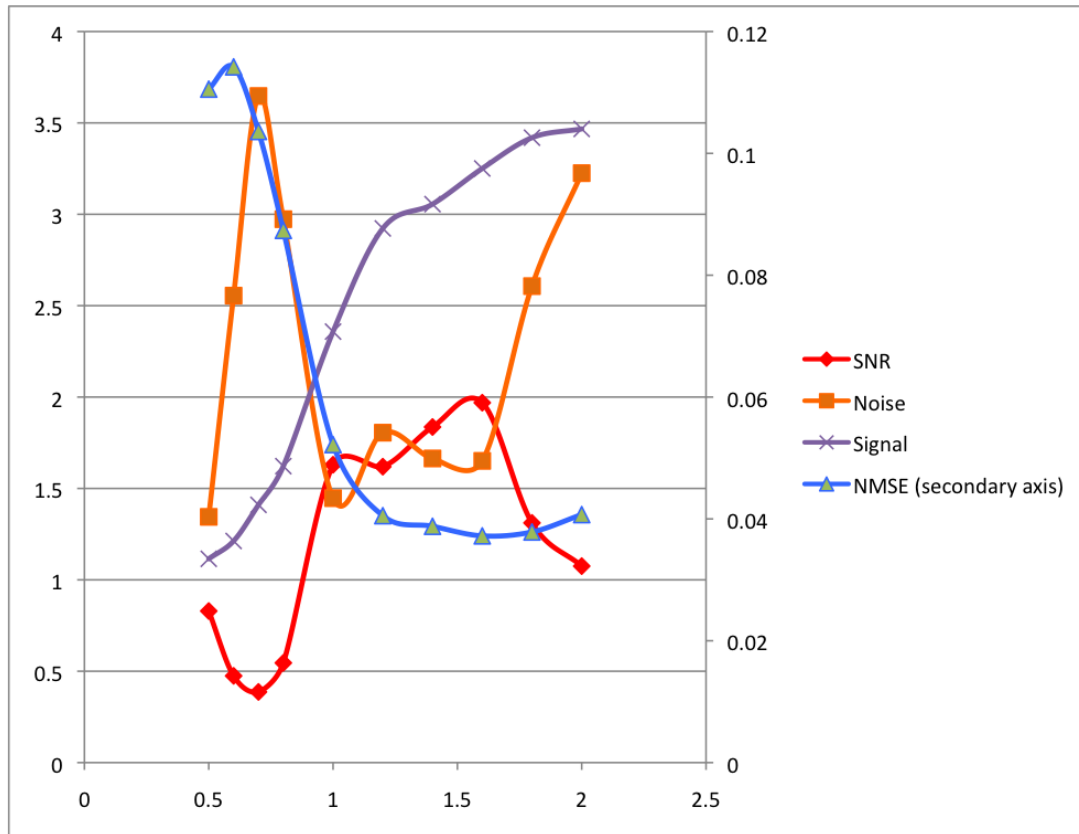


Figure 3-18 Line profile metrics as a function of reconstruction filter

Figure 3-18 shows the Noise, Signal, SNR and NMSE scores as a function of the reconstruction filter. The contrast is seen as a rapidly rising curve up to 1cycle/cm, after which it flattens out. The noise initially decreases as the image sharpens – this relates to the steepening of the boundary between the putamen and background regions. The NMSE score is the same data as depicted in Figure 3-17.

3.7.6 Line profile NMSE - discussion

The progression of line profile curves shown in Figure 3-16 shows the two related effects of increasing reconstruction filter cut-off frequency. The low frequency reconstructions show smooth transitions from putamen to background, with smooth near-uniform background regions. The curves showing the higher frequency data show a much-improved contrast between the

putamen and background regions at the expense of increased noise. The expected background region counts were taken from the average counts in this region, so the noise measurement is simply a measure of variation from this mean.

Comparing the trends in these figures with those from the whole brain NMSE results, similar trends are seen. The SNR and noise measurements show larger variations in the line profile data than in the whole brain data, most likely due to the reduced number of voxel samples being used to create the figure, and the more complicated interplay between the sharp boundaries and the noise levels. The whole brain data on the other hand is dominated by uniform brain background noise, with only a small component coming from the striatum-background boundary regions. The minimum NMSE score is found at slightly higher cut-off frequencies in the line profile data than in the whole brain data. The NMSE score for entire image combines the information from the two line profiles with many other profiles of similar form. However, the whole brain NMSE behaviour will be heavily weighted towards the cerebellum line.

3.7.7 Varying Radius of Acquisition - results

The contrast results were compared for a single phantom run, repeated at different acquisition radii of rotation. The results are shown in Figure 3-19.

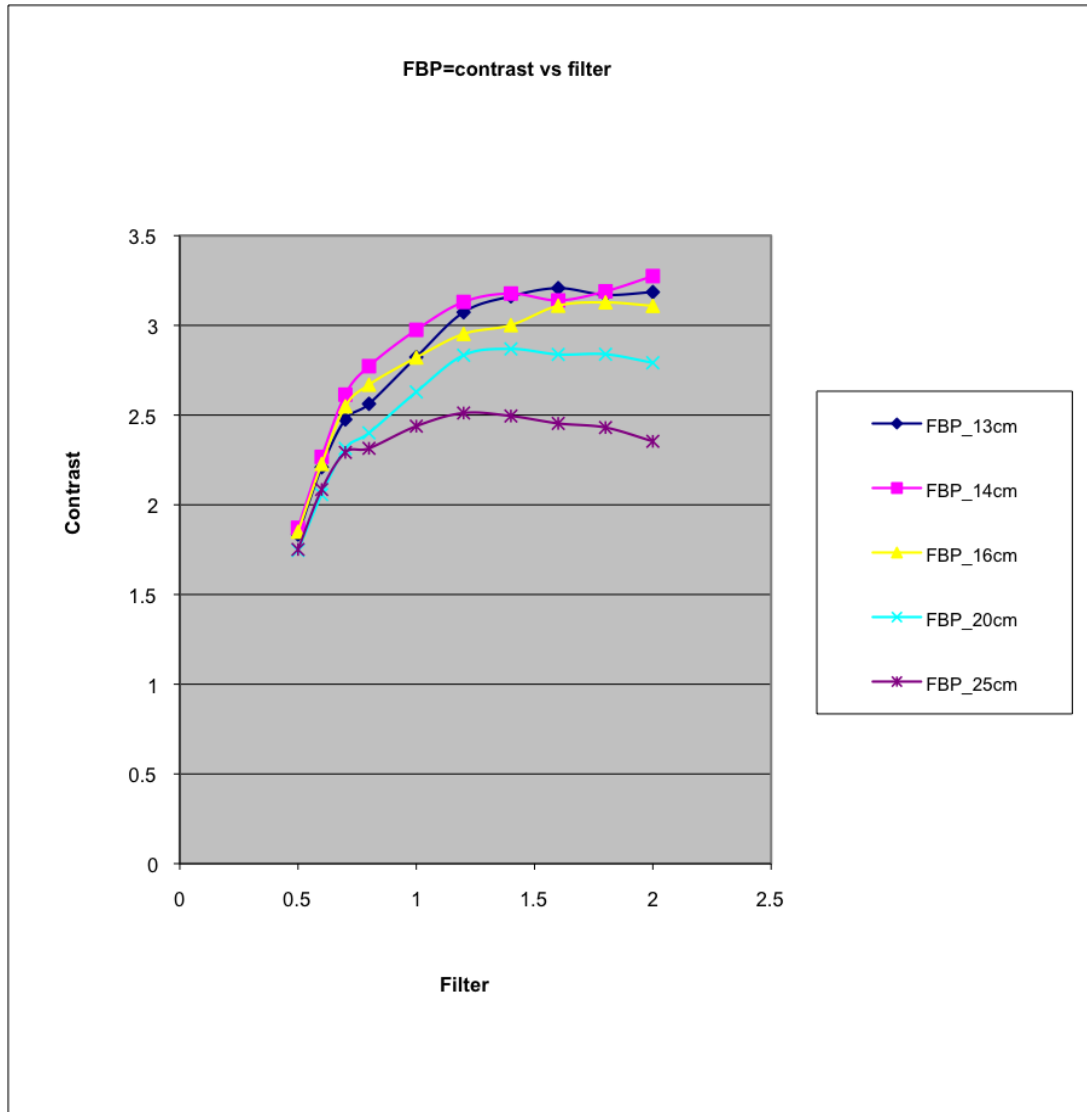


Figure 3-19 FBP contrast as a function of acquisition radius of rotation

The graph shows that as the radius of rotation is increased, the image contrast goes down.

3.7.8 Varying Radius of Acquisition - discussion

The radius of rotation seen clinically is typically 13-15cm. For patients with mobility/flexibility issues this must sometimes be increased to 16 or 17cm. The results show that a similar contrast is seen for the 13-16cm acquisitions. However, when the ROR is increased to 20cm or greater then there is a significant loss in contrast (30% drop for the $f_{co}=2.0$ cycle/cm).

Looking at the scatter free line source resolution measurements described in section 2.3 for the collimator source distances of 13,16 and 20cm, the planar resolution (as FWHM) is 9.25mm, 10.57mm and 12.34mm respectively. This equates to an increase in blurring width of 14% and 33% when moving from 13 cm up to 16cm and 20cm. When the ROR is extended to 25cm, the blurring increase is over 57%.

3.7.9 Conclusions on Filtered back projection images

Although the results of the various measurements performed on the filtered back projection images show a wide range of results in terms of the exact filter that is ideal for each test, a broad consensus between these measurements is seen and so some conclusions can be drawn.

The sharpest images, those with the highest cut-off frequencies for the filter, give the highest contrast. These images show the smallest differences in recovery coefficient between acquisitions. These two properties would suggest that it could be the best for VOI based analysis programs. However, these images show very high levels of noise, and are visually very hard to interpret. Automatic image registration does not work very well with excessively noisy images, so alignment of VOIs may be problematic unless a surrogate image is used for the alignment.

At the other extreme, the very smooth images show low levels of noise, but lack the detail in the striatal activity.

The images preferred for visual reporting are near the centre of the filter ranges at 1.2-1.3 cycles /cm. These images are slightly sharper than those with the

highest SNR as defined using the larger VOIs. They are very close to those with the highest SNR as measured using the line profiles.

Comparing these results with settings used in published data, a wide range of settings have been used. To date no work has been published optimising the settings for FBP in DAT imaging, however a number of settings have been used. In their work showing the usefulness of a large VOI based analysis package, Tossici-bolt et al. ¹¹⁰ used a relatively smooth filter (0.6 cycle/cm order 10). The argument for their specific analysis package was to create a resolution independent system, which incorporates large VOIs. A similar filter was used by Badiavas et al. ⁷⁷ for their implementation of the same uptake measurement. Work utilising more tightly defined VOIs have used a range of filters (0.5 cycles/cm up to 2.2 cycles/cm) ^{49,111-115}. The use of smoother filters would allow visual reporting, but at the expense of some contrast. Dickson et al. ¹¹⁵ noted that although they had suggested the use of very sharply reconstructed images, that smoothed options might be required for visual reporting. A note of caution should be taken when comparing the filter settings from different published data due to the variation in the definition of the filter (see section 3.1).

3.8 Results: Phantom experiments with OSEM

3.8.1 Contrast and noise for three iterative reconstruction models

The contrast and noise were measured using method described in section 3.6.3.1 using VOIs defined on the CT image and transferred to the SPECT data as per the FBP data in section 3.7.1. No attenuation or scatter compensation techniques

were included in these reconstructions and no post reconstruction smoothing was applied.

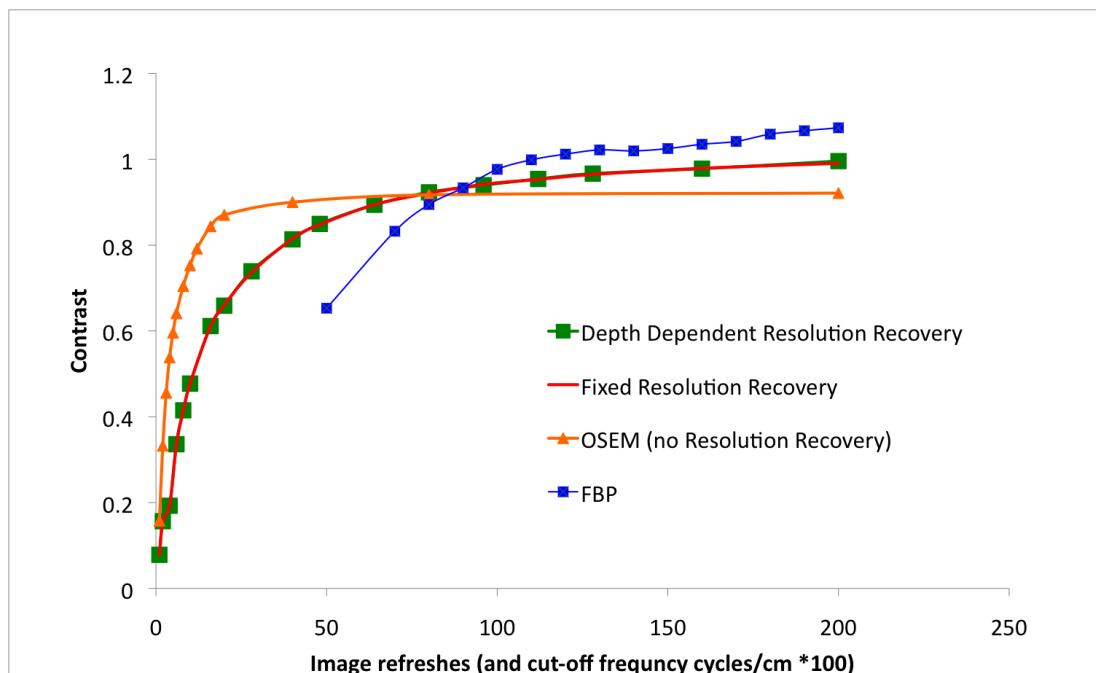


Figure 3-20 Contrast as a function of image refreshes for different OSEM based reconstructions

Figure 3-20 shows the results from phantom Run 2, using a radius of rotation of 14cm. The Gaussian blurring function is extended to +/- 2 pixels in the x-axis and +/-1 pixel in the Z-axis. The FBP data is also shown for comparison, as a function of cut-off frequency. The cut-off frequency (f_{co}) was scaled to allow visualisation on the same chart by multiplying by 100.

Figure 3-20 shows the data from three iterative models, two of which include resolution recovery. The Depth Dependent Resolution Recovery (DDRR-OSEM) and Fixed Resolution Recovery (FRR-OSEM) give almost identical results - as shown by the overlapping data on the graph. The basic OSEM model shows faster convergence, but lower contrast. The inherent filtering in the resolution recovery based models means they converge more slowly, but result in higher image contrast. This trend is seen in all datasets: the resolution recovery models

give increased contrast when compared to OSEM. The FBP reconstructions are seen to give higher contrast in these measurements than those achievable by OSEM.

The noise in this same set of reconstructions was measured using the same region as per section 3.7.1. This noise data was combined with the contrast data shown in Figure 3-20 to give Figure 3-21. The two curves for iterative reconstruction are produced by varying the number of iterations/image refreshes in the reconstruction. This variable produces changes in both the noise and contrast in the images.

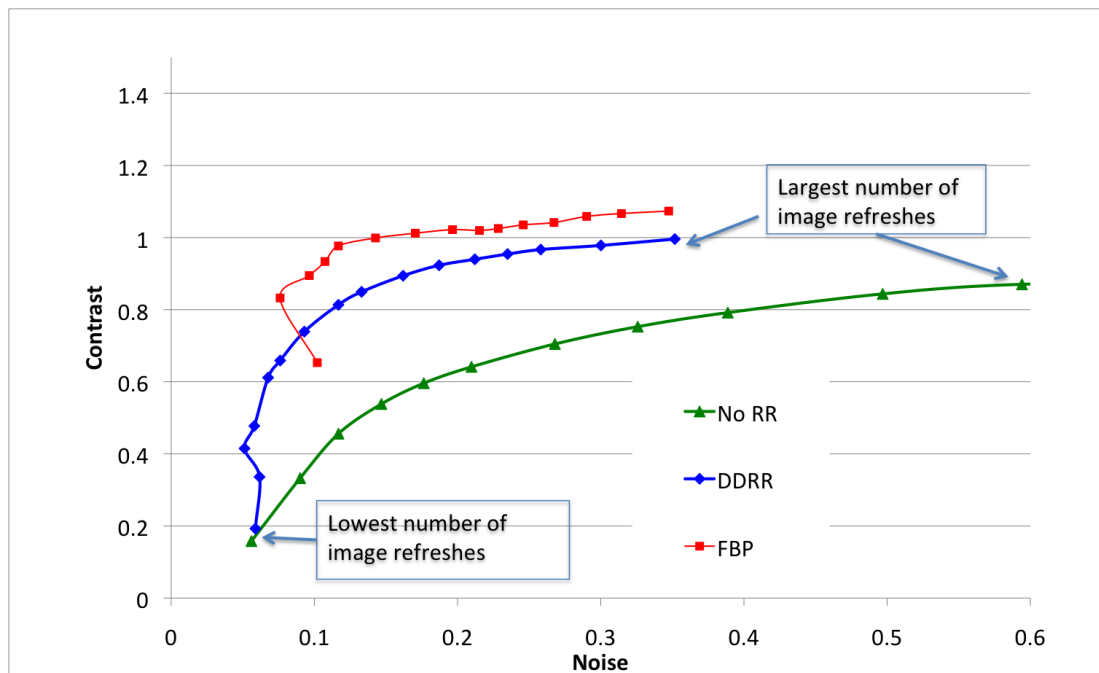


Figure 3-21 Signal as a function of Noise for different Iterative reconstructions

These signal-noise graphs show that at any specific noise level, the DDRR-OSEM will give better contrast than OSEM. The filtering inherent in the DDRR-OSEM and FRR-OSEM mean the noise levels are controlled better.

Some researchers (e.g. Hutton et al. ¹¹⁶) have suggested that a post-reconstruction filter can improve noise characteristics. Due to the inherent

smoothing in the resolution recovery models, their improvement with post filtering was reported to be smaller.

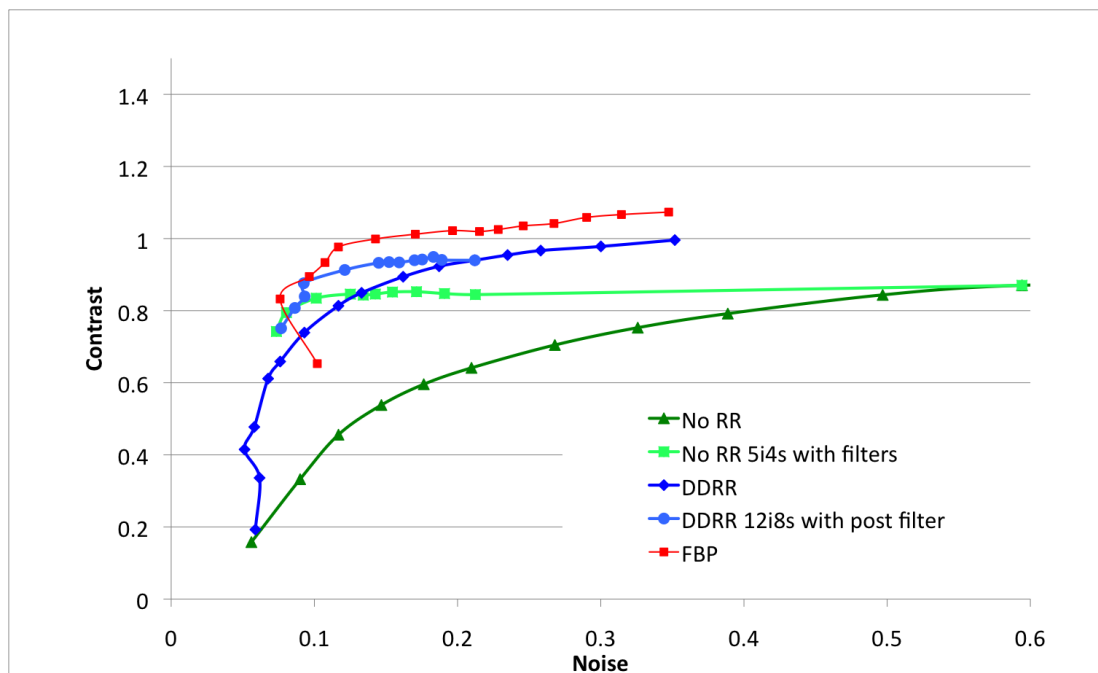


Figure 3-22 Contrast as a function of noise for different OSEM based reconstructions with post-reconstruction filter

Figure 3-22 shows the effect on noise and contrast when a post-reconstruction smoothing filter is applied to the data in Figure 3-21. Two representative reconstructions are shown with post-reconstruction smoothing. The data from an OSEM reconstruction with 20 image refreshes (5 iterations of 4 subsets) was smoothed using a range of Butterworth filters. The same series of Butterworth filters was applied to the Depth-Dependent resolution recovery image reconstructed using 12 iterations of 8 subsets (96 image refreshes).

A dramatic reduction in noise is seen in the OSEM data, with contrast loss only when very smooth Butterworth filters are applied. The reduction in noise for the DD-RR OSEM data is less pronounced, but higher contrast is still seen for almost all noise levels.

The shoulder in these graphs, where the smoother filter starts to have a dramatic effect on the contrast is seen at filters around 1.4 cycles/cm.

The ratio of signal to noise (SNR) can also be used to describe this trade off as is seen in Figure 3-23.

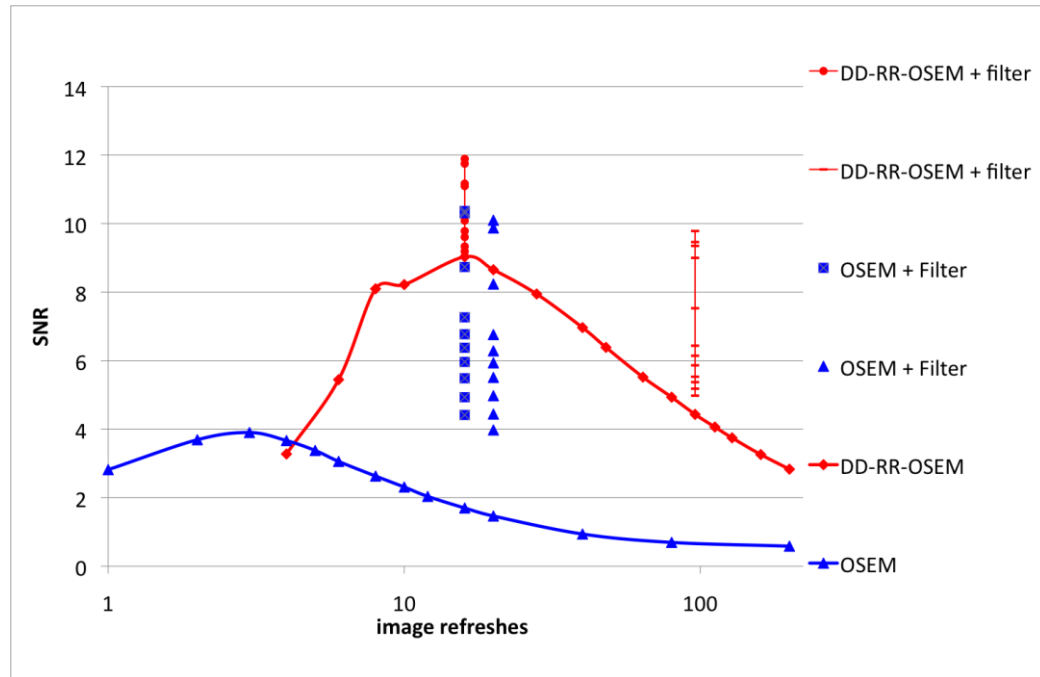


Figure 3-23 SNR as a function of image refreshes

The curves for SNR for OSEM, FRR-OSEM and DDRR-OSEM show peaks at 4 and 16 refreshes. The SNR curves from FRR-OSEM and DDRR-OSEM are superimposed.

Different levels of post-reconstruction filtering are shown as the vertical lines coming up from the SNR curves. Filtering was applied to specific reconstruction data sets:

- The data that gave Peak SNR in non-filtered data,
- At a level where 95% of maximum contrast is seen, i.e. at 16 or 20 image refreshes for OSEM, and 96 image refreshes for DDRR-OSEM.

The peak SNR occurs when relatively smooth post-reconstruction filters are applied. For both sets of OSEM models the peak SNR is at post reconstruction filters of 0.8 cycles/cm.

3.8.2 Contrast and noise for three iterative reconstruction models – discussion

The contrast and noise results for the iterative reconstruction processes show complicated dependence on the number of image refreshes and the use of post reconstruction smoothing.

When resolution modelling is included, the images will converge at a slower rate than for the basic OSEM reconstruction. However, if sufficient image refreshes are performed, and the post-reconstruction filter is chosen appropriately, the DDRR-OSEM can always produce an image data that is either higher contrast for the same noise, or lower noise for the same contrast.

No discernable differences are seen between DDRR and FRR. The depth-dependent model calculates the blurring function for the distance between the object and the detector. However, for the contrast measurements here, the range of depths of interest is very small, only 6.5 to 10.5 cm. At these distances the calculated blurring function will give a FWHM of 7 to 9.25 mm, which is very similar to the fixed filter model of 9mm.

In two papers investigating the use of OSEM models in DAT SPECT imaging,^{115,117} a depth independent PSF was used with a value of 4.0mm (equating to 9.4mm FWHM). This figure was adopted as a compromise setting for data from a number of different cameras and collimator types.

A post-reconstruction smoothing filter can improve the noise properties of OSEM data. A filter can reduce the noise levels in the image, but if the filter is too smooth, the contrast will decrease. For a range of filters, only very small decreases in contrast are seen. For both OSEM and the resolution recovery OSEM models filter cut-off frequencies below 1.4 cycles/cm have an adverse effect on image contrast. However, the peak SNR as measured using this system occurs at much smoother filters of 0.8 cycles/cm.

In a very recent work, Warwick et al.¹¹⁸ compared resolution recovery with fixed or depth dependent resolution model to reconstructions without resolution modelling. They found that post reconstruction smoothing was required when no resolution recovery was included. They also found a small improvement with depth dependent resolution model, over fixed blurring.

3.8.3 Direct comparison of FBP and DDRR-OSEM

This section compares SUV uptake data for the FBP and DDRR-OSEM reconstructions. Using a specific ROR (13cm), and the four different phantom contrasts, data was reconstructed using FBP and DDRR-OSEM using a range of filters and iteration numbers respectively. No post-reconstruction filtering was used in these measurements.

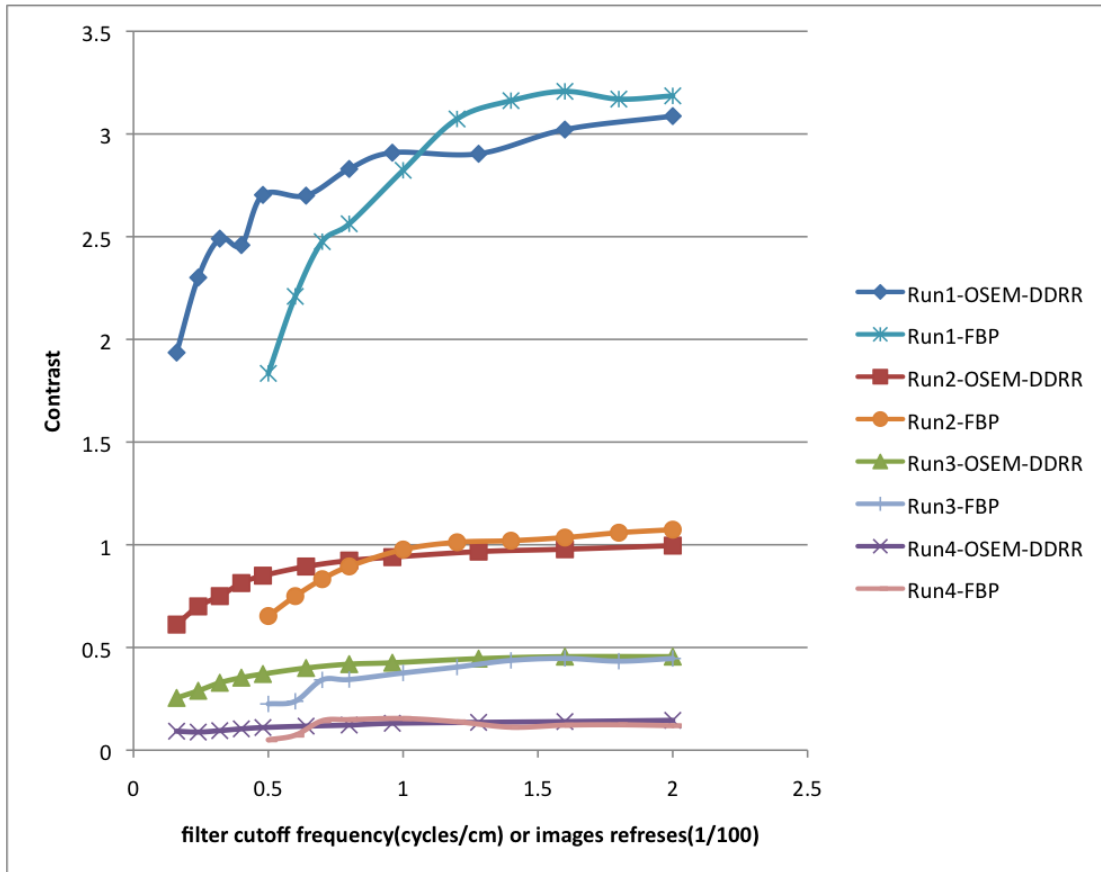


Figure 3-24 contrast for different phantom runs using FBP and OSEM

Filter cut-off (in FBP) and number of iterations (in OSEM) affect the contrast and noise levels of the resulting images, however they are of course completely separate variables so they should not be compared directly. That is, the contrast for a particular filter should not be compared with the contrast from that number of iterations. However, with this proviso, the data from the two groups of reconstruction can be shown on the same graph. The x-axis shows both filter cut-off frequency (cycles/cm) and iterations (measured in hundreds of iterations).

This data shows clear dependencies on number of iterations, and filter levels for the two groups of reconstructions.

When the data is normalised to the actual phantom concentration ratios, i.e. measured as recovery coefficients, the variation for low contrast data is more clearly seen (Figure 3-25).

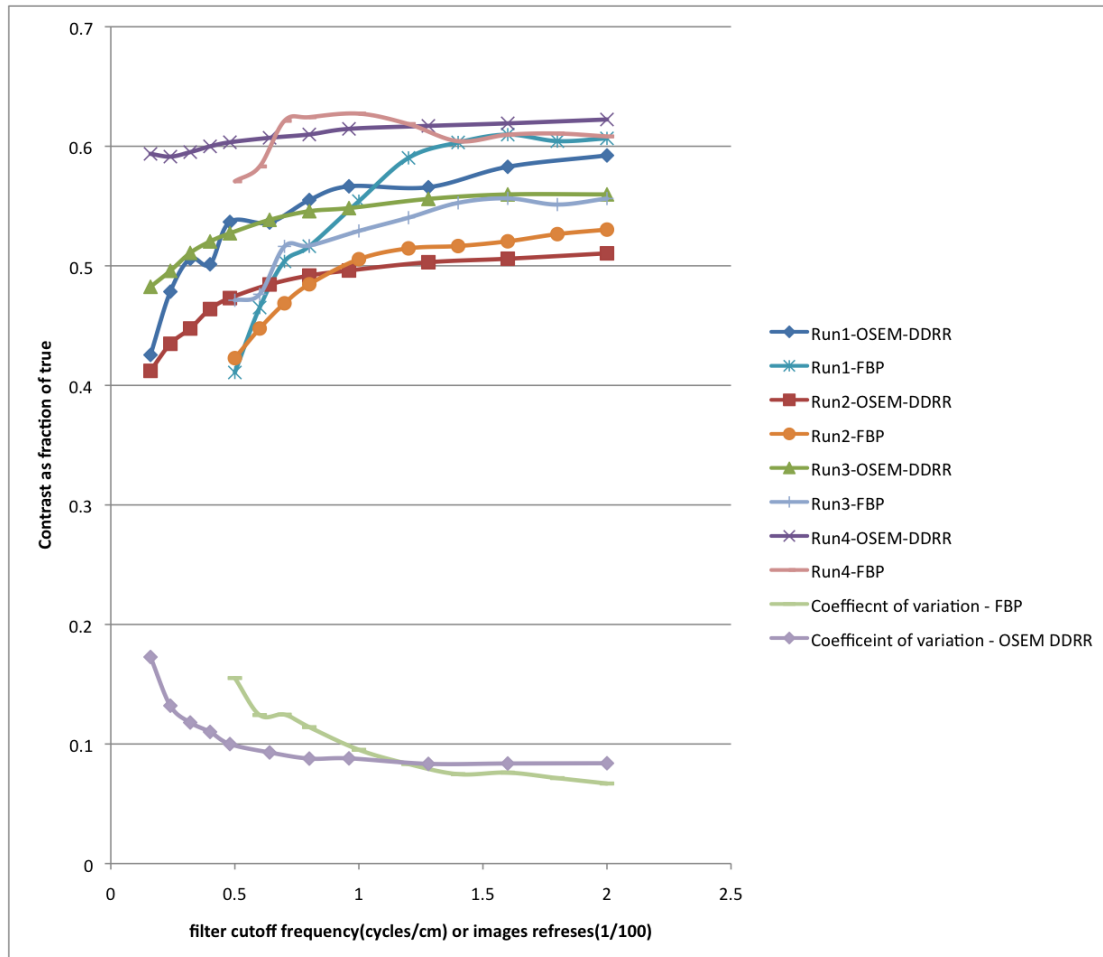


Figure 3-25 recovery coefficient as a function of image refreshes or cut-off frequency

This chart shows mainly smooth variation of contrast with number of iterations used. Only very small changes in image contrast are seen at over 100 iterations. The FBP shows abrupt changes in measured contrast for small changes in filter settings. However, these variations are only seen when very smooth filters are used.

The peak contrasts are seen when large numbers of iterations are used, or very sharp filters are applied. For both reconstruction types there is a marked dependency on phantom contrast.

Using just the highest contrast data available, the DRR-OSEM results are on average marginally lower than the FBP data. This result is in fitting with the results published by Koch et al. ¹¹⁷. They used similar software, but fixed the iterative reconstruction to 3 iterations of 8 subsets and had a fixed resolution function of 4mm. The extent of the model was not quoted. Their FBP reconstruction used only a ramp filter coupled with a specific post-reconstruction filter. They measured an average 5% lower binding ratio with OSEM compared to FBP.

3.8.4 Size of Blurring Function

In section 3.8.2 the speed at which the contrast in the images converges on the final value was found to be dependent on the use of resolution recovery. The extent of this resolution was investigated next.

The extent of the resolution model is measured in the number of voxels in the x and z directions (respectively, across the transverse slice and between slices).

The software manual suggests a starting point of only a small number of pixels (x1,z0,) but does not suggest or put any limits on how far this can be extended.

Routine clinical practice has been to employ 2x-z1.

The influence of these settings was investigated using data from Acquisition 1, at 13cm radius of rotation, with a range of iterations and extent of resolution model (x's and z's). The contrast was measured using an ROI based on CT data as per section 3.6.3.1. Results are shown in Figure 3-26.

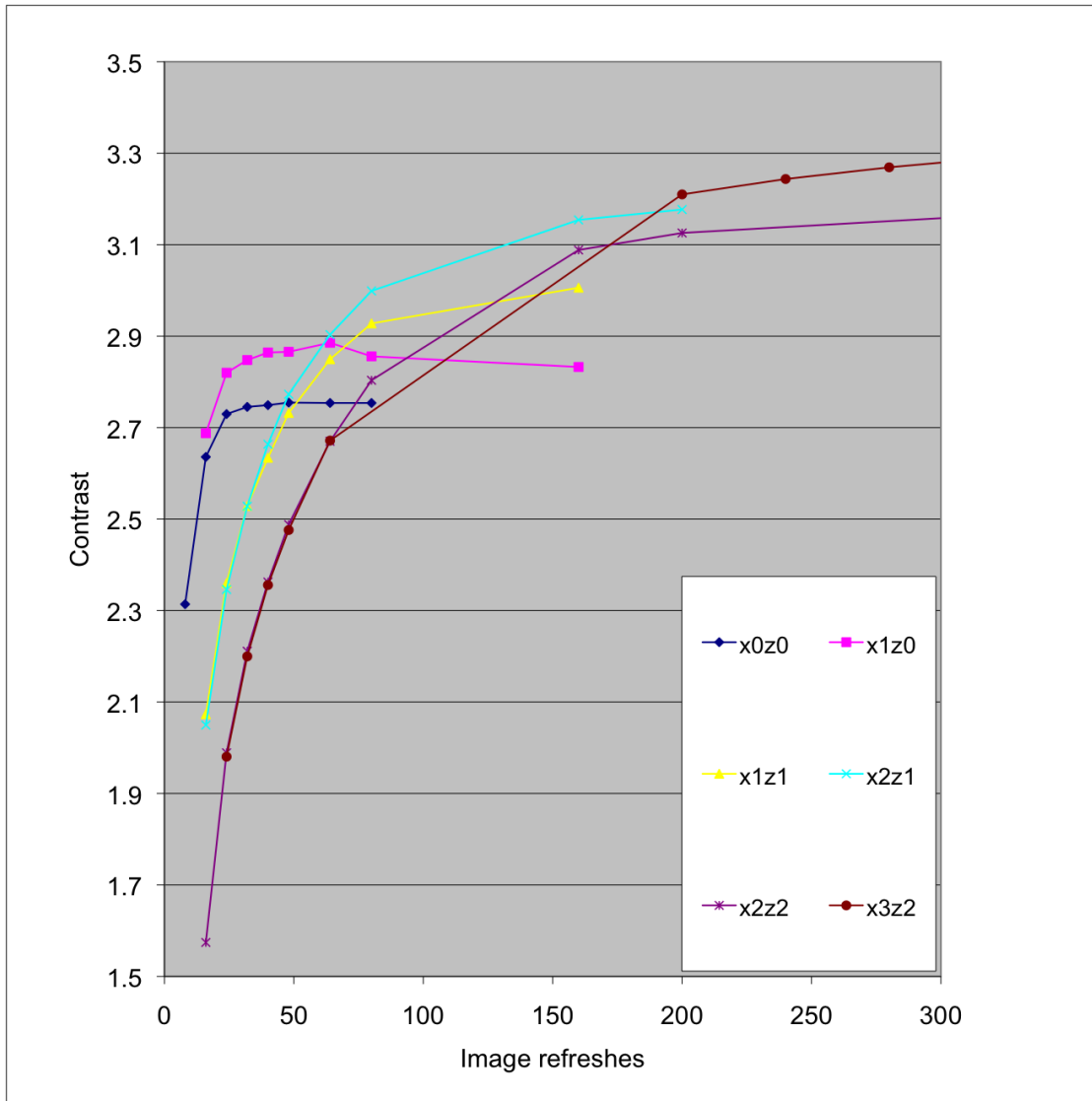


Figure 3-26 Contrast as a function of image refreshes for varying extent of resolution model

The results show that as the Gaussian profile is modelled to a larger size, the slower the reconstruction. The x3,z2 data converges at over 700 image refreshes. An increase in the x-width appears to increase the contrast, whereas for this measure of contrast at least, the widening of the z-model slows the convergence.

The dark and light blue lines relate to the normal OSEM (x=z=0) and DDRR-OSEM (with x=2,z=1) used elsewhere in this work.

These contrast curves would appear to recommend the use of wider modelling with more and more iterations, however, the visual appearance of these images show an unwanted side effect. Figure 3-27 shows a central transverse slice from a reconstruction using a broad collimator model and lots of iterations (x3,z2, 75 iterations of 8 subsets). The figure on the right is the same data as the left, but with a post-reconstruction Butterworth filter ($f_{co} = 1.5\text{cycle/cm}$, order 10). There are obvious non-uniformities in the background regions. These are reduced in the smoothed data, but are still present. There are also two very low count regions in the arch of the two striata – appearing as two black circles medial to the putamen and posterior to the caudate heads. On the non-smoothed imaged there are distinct detail along the putamen that is artefactual – they appear as two separate sections – although this is removed when filtering is applied. It would appear the images have been over corrected. A similar affect could be achieved using an edge enhancement filter, or a non-uniform smoothing filter.

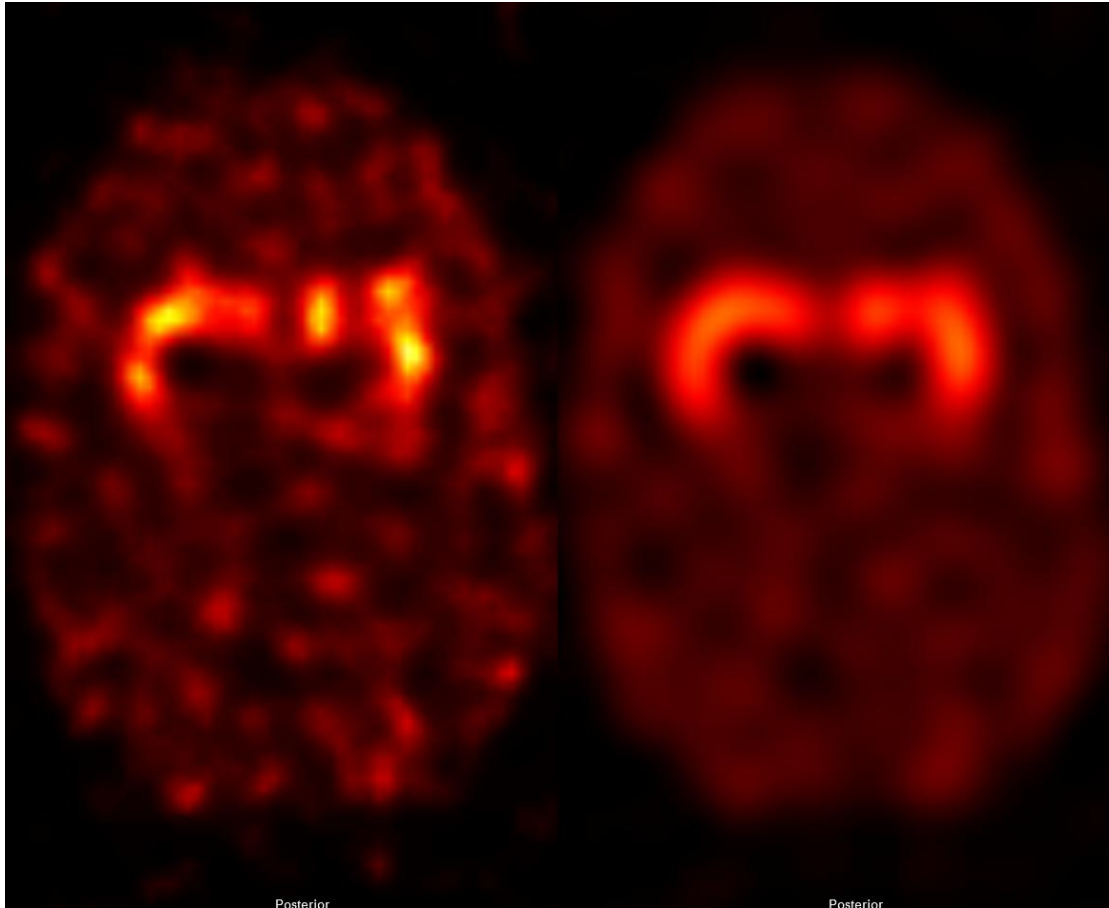


Figure 3-27 Central transverse slice from DDRR-OSEM reconstructed volume using wide Gaussian modelling (x3z2). The image on the right has been smoothed post-reconstruction Butterworth filter

3.8.5 Size of Blurring function line profile NMSE - results

To investigate the formation of this structured background, the line profile techniques described in section 3.6.3.3 and used in section 3.7.5 were employed. Some post reconstruction filtering was also performed to investigate the differences between structured and random noise in the images. A post reconstruction filter of 1.5cycles/cm was applied to the reconstructed data. This filter was chosen to be sharper than the point at which contrast is lost, as based on the curves in section 3.8.2.

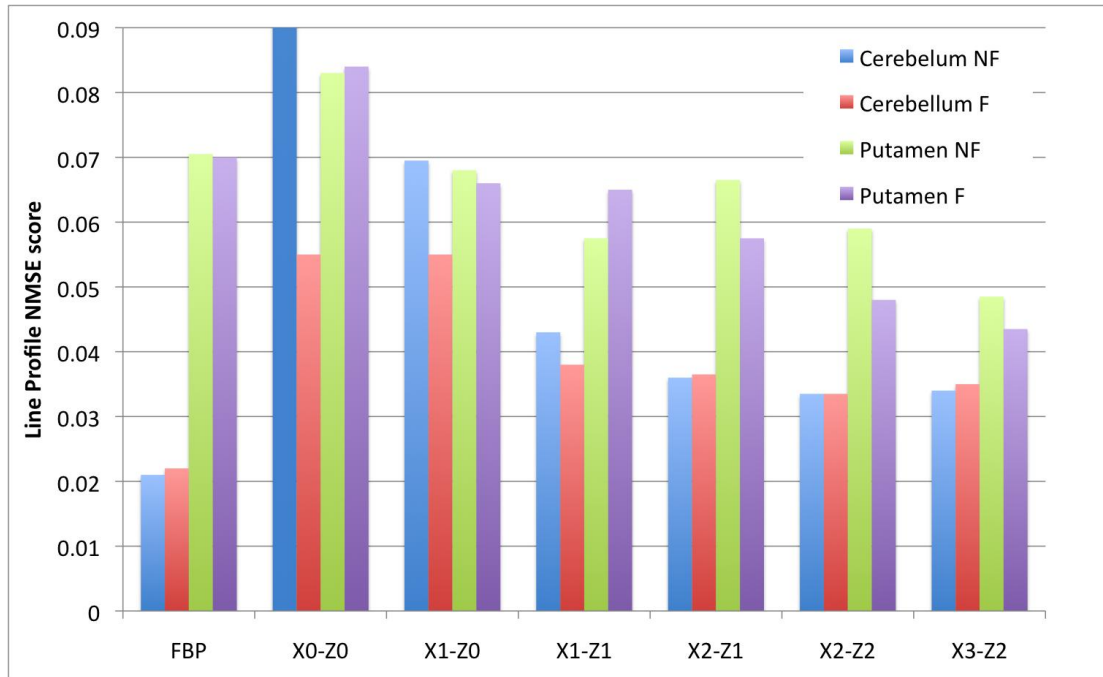


Figure 3-28 Line Profile NMSE minimums for a range of iterative reconstructions and FBP with and without post-reconstruction filtering (F /NF)

Figure 3-28 shows the minimum NMSE score for the putamen and background regions for a range of OSEM reconstructions using different Gaussian blurring widths, with and without filtering.

The minimum NMSE scores are generally seen for data with a post reconstruction filter. The FBP has the lowest background NMSE score, while the wide (x3z2) DRR-OSEM reconstruction has the lowest putamen region score. The wide-Gaussian data is capable of giving both the minimum putamen score, and a low cerebellum score. However, these two minimum values are produced from different reconstructions. Figure 3-29 shows the trends in the x3z2 data along with the FBP data for comparison. The minimum NMSE scores for the putamen and background regions occur at 64 and 400 iterations respectively.

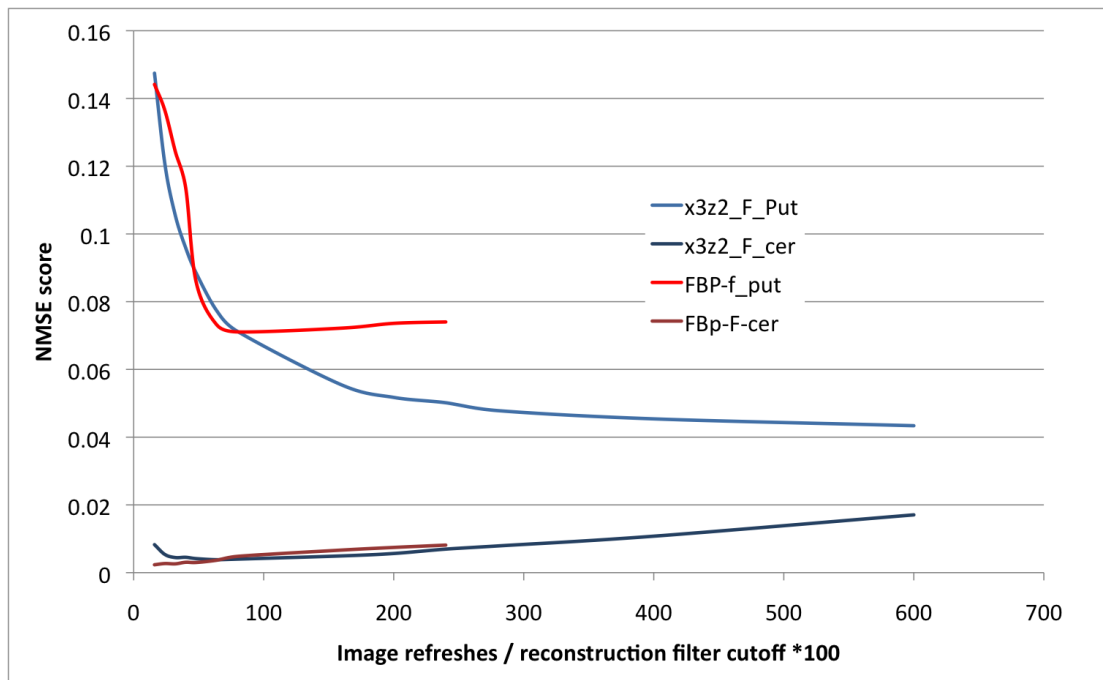


Figure 3-29 Line profile NMSE scores for OSEM and FBP reconstructions

In Figure 3-30 the NMSE score as a function of image refreshes is shown for the putamen line profile for the various levels of resolution modelling. The lower chart is based on the same reconstructions but utilising a 3D post-reconstruction smoothing filter (Butterworth 1.5cycles/cm, order 10). Using different widths of Gaussian resolution model will change the number of iterations needed to make NMSE reach a minimum. Once this minimum is reached, adding further iterations will generally increase the NMSE score as more noise is added to the image. Adding a post reconstruction filter removes this extra noise and so the NMSE score does not rise with further iterations.

The curves for the background region are shown in the two charts in Figure 3-31.

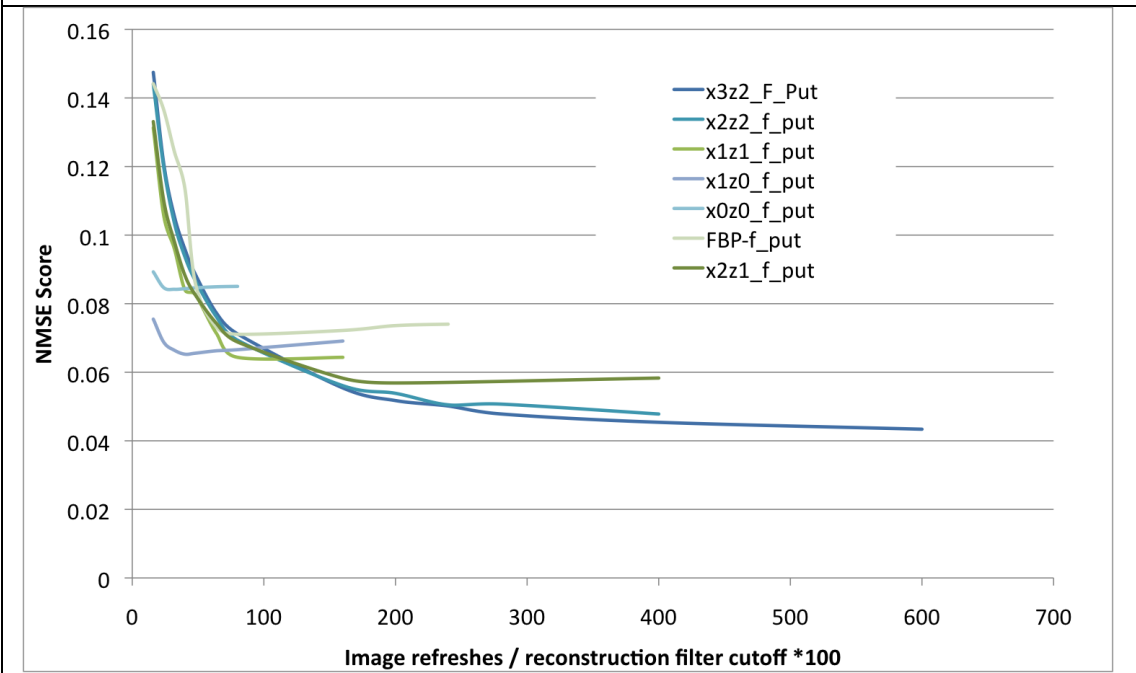
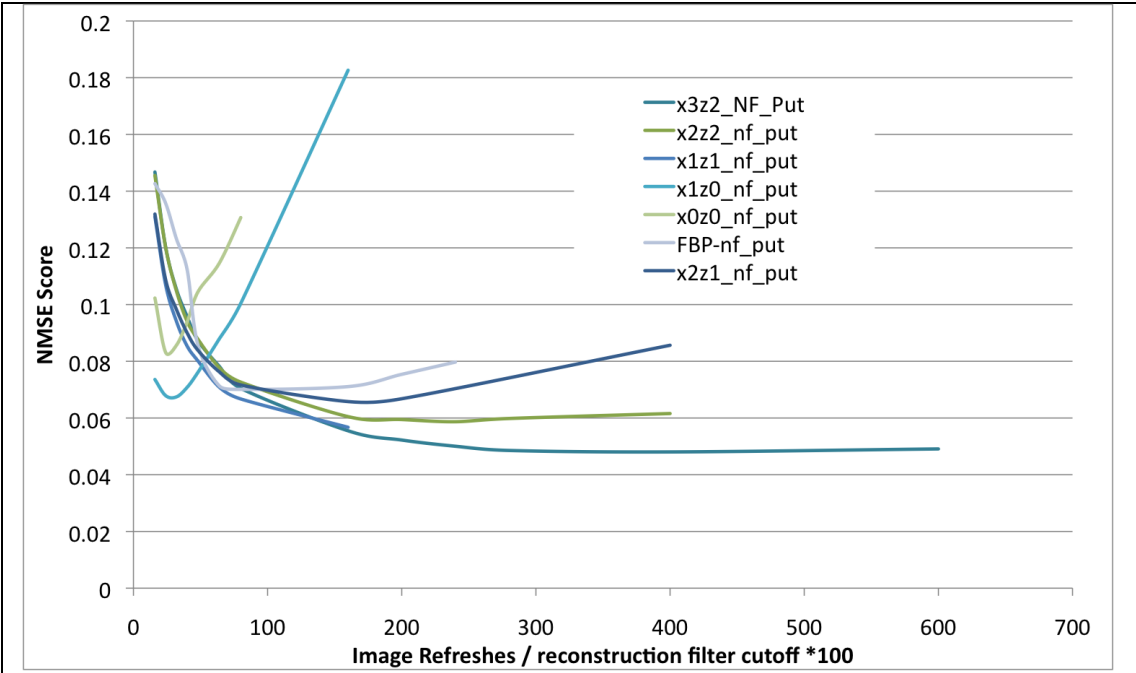


Figure 3-30 Line NMSE score for the Putamen region for iterative reconstructions using different sizes of resolution model. The lower chart includes a post-reconstruction smoothing filter

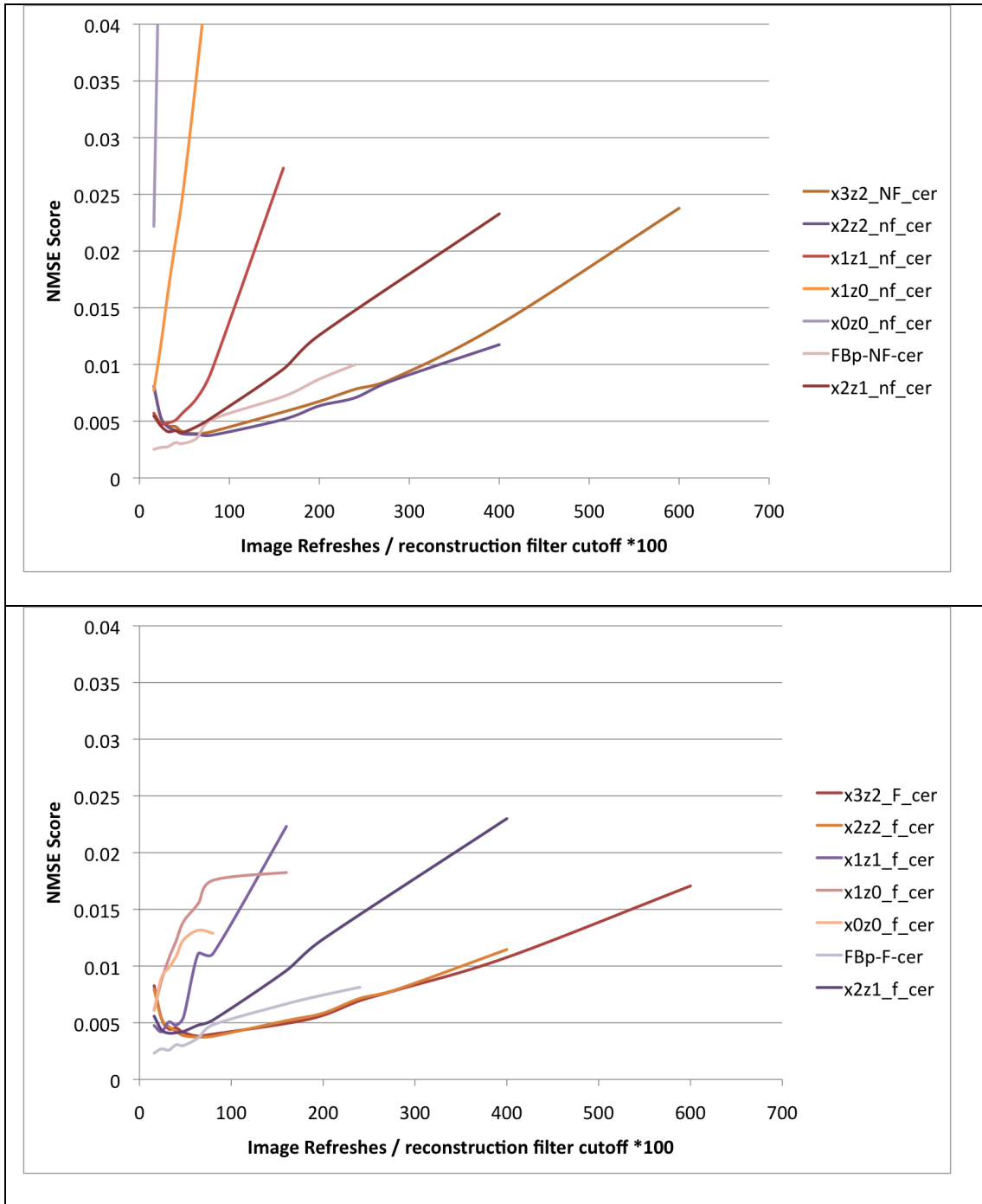


Figure 3-31 Line NMSE score for the background region for iterative reconstructions using different sizes of resolution model. The lower chart includes a post-reconstruction filter. The background region charts (Figure 3-31) appear to show quite variable results. There is an obvious trend for the data from the larger blurring model to give better results at high numbers of iterations. When a post reconstruction filter is added the curves follow the same trend, but are more closely matched.

3.8.6 Size of blurring function – discussions

For both the filtered back-projection and the OSEM NMSE line profile results, there are opposing trends going on. For the putamen region line profile, the best results are occurring when higher frequency data is included in the image: That is, when either high cut-off frequency filters are used in the FBP reconstruction, or large numbers of iterations and wide Gaussian modelling are used within the OSEM reconstruction. The opposite is true for the background regions, where the best results are seen with smoother reconstructions. For all of the images a post-reconstruction smoothing filter improves this image quality metric.

Relating these NMSE scores to the image quality in terms of classification is not straightforward. For the particular imaging task being investigated here, where the importance is in the size and intensity of uptake of a particular organ/sub-organ, the ability of a reconstruction system to give smooth background regions is less important than its ability to clearly define the extent of an active region. As such, the line NMSE score for the putamen region may be a useful indicator of reconstruction performance, whereas the cerebellum region is less important.

3.8.7 Varying Radius of acquisition – iterative results

The depth-dependent resolution recovery model is described as using a different blurring function depending on the radius of rotation for the acquisition. It is possible that the DRRR-OSEM will outperform FBP when patient positioning requires a larger ROR.

The same projection data used in section 3.7.7 was reconstructed using DRRR-OSEM using the known Radius of rotation. Figure 3-32 shows the contrast measured for phantom run1 using different acquisition ROR. The left figure

shows the FBP results, and the right DDRR-OSEM model. Using the variation in contrast at the highest contrast setting used, the DDRR-OSEM model shows a smaller variation in the measured contrast across the ROR used compared with FBP (7% vs. 10%). However, when just the 13-16cm data is used, the FBP has a smaller spread of results (2.4% vs. 3.8%).

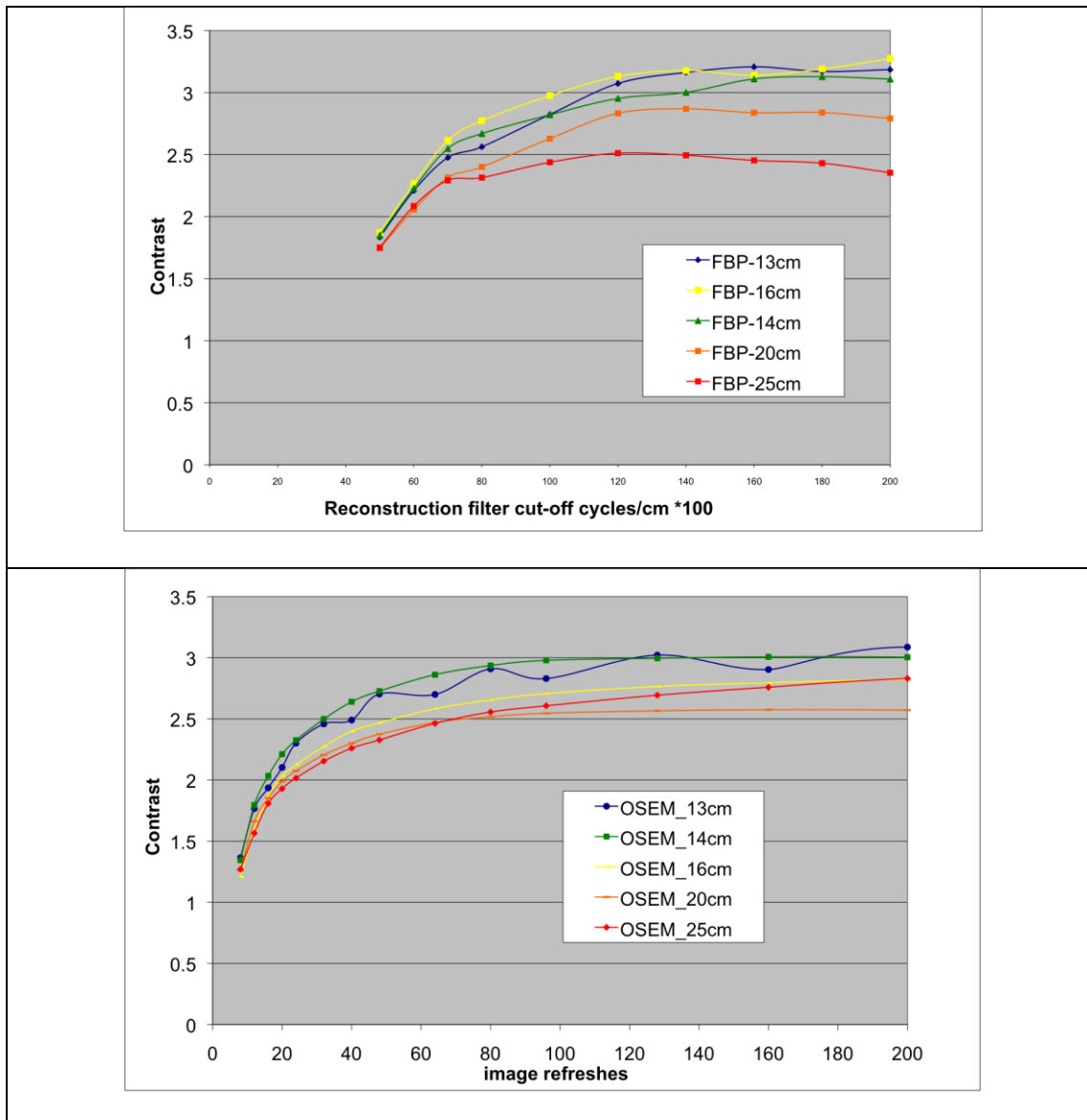


Figure 3-32 Contrast measurements using different acquisition radii, as a function of cut-off filter used for FBP (top) and image refreshes for DDRR-OSEM (bottom)

3.8.8 Varying Radius of acquisition – discussion

The depth dependent resolution model appears to make significant improvements when the phantoms are imaged over a wide range of ROR. However, such wide variations are rarely seen in clinical brain imaging. When the more clinically relevant ROR range is used, the FBP data shows a smaller variation. As such, the results suggest the depth dependent model does not reduce the negative effects of increased ROR over the typical range of brain imaging distances. Such a result may be important for SPECT imaging of the body.

3.8.9 The effect of errors in ROR measurement - results

The use of depth dependent Resolution recovery requires that the ROR be input into the reconstruction program. On the scanners used in this work, this data is not stored within the image information file (the header file). The Radius of Rotation (ROR) has to be recorded manually. If the radius were not recorded properly, then an estimate would have to be made.

A test was performed using the 13cm ROR data, with a range of incorrect radii input into the reconstruction program. The input range was 175mm-205mm Centre of rotation-crystal distances, relating to 13-16cm ROR. The contrast was measured using the CT derived ROIs.

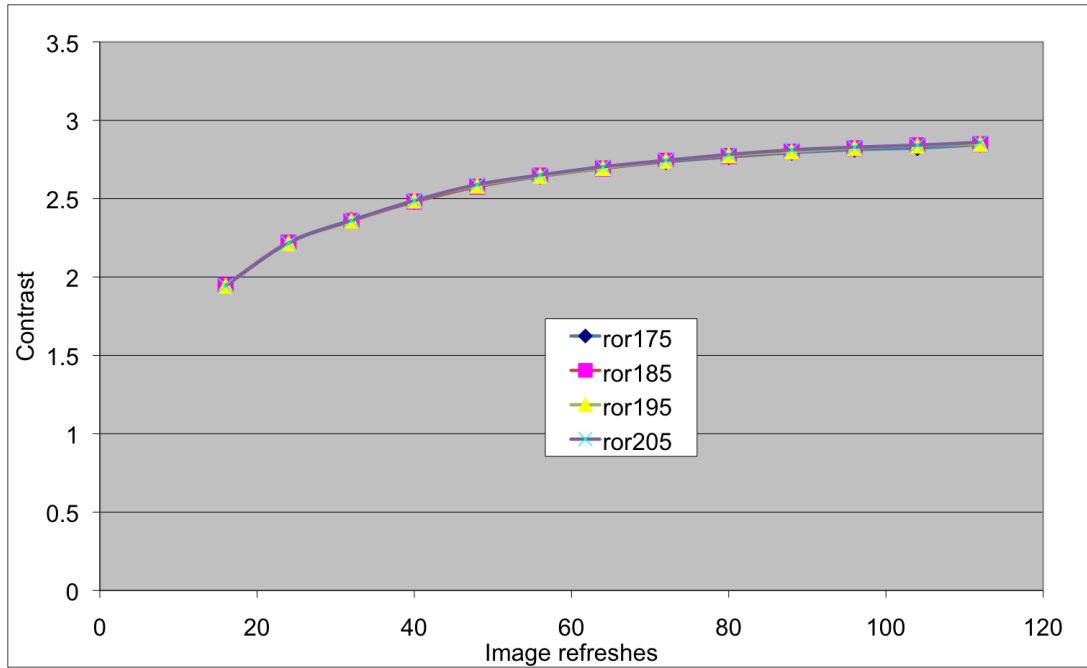


Figure 3-33 Contrast as a function of image refreshes when different ROR distances are used

Figure 3-33 shows the contrast measured for Run1 data acquired at 13cm, using different ROR input data. The coefficient of variation for the contrast measurements at maximum contrast is 0.2%. The contrast in the reconstructed images is not dependent on the ROR used in the program over the range of distances that are likely to be used.

3.8.10 The effect of errors in ROR measurement – discussion

The collimator blurring function used in this program is a function of distance as given by the experiments in chapter 2. The blurring function sizes are shown in Table 3-3. For this range of radii, there are only minimal differences between the blurring function used, and the correct blurring function.

Table 3-3 Calculated blurring function at different distances from collimator

Collimator source distance (mm)	FWHM (mm)	Percentage change from 130mm ROR
130	9.25	0%
140	9.69	4.77%
150	10.13	9.53%
160	10.57	14.3%

3.9 Compensation methods for attenuation

The effect of attenuation is to reduce the apparent levels of activity within the reconstructed image. In particular, regions central to the body, i.e. surrounded by attenuating material will be affected the most. The counts in a central region of the brain will be reduced by a factor of around 60%. However, what matters is not the size of bias or even the error in the uptake measurements, but the relative differences between patients. The composition of patient's heads is very uniform between patients, regardless of body habitus. In fact, as is shown here, patient head sizes do not vary considerably, and so the range of the attenuation effect in the very centre of the brain will typically vary from 58% to 64%.

Patients with larger than average heads will exhibit increased attenuation effects and hence reduced apparent uptake within the brain. There are however a number of factors that will change the absolute level of uptake within the brain, and these are accommodated for by scaling the images using a background or non-specific binding region (see chapters 5 and 6). When the attenuation effect is combined with this scaling effect, some of the attenuation effects are cancelled

out. What remain are the differences in relative attenuation effect to the target and background regions due to the head size. Depending on the background region being used, it is possible for the target region and background region to be at the same effective depth and in this specific case the size of the patients head will have no bearing on the measured up take in the image.

This effect can be calculated theoretically for specific regions of the brain and for pre-determined head sizes. This calculation will not give accurate attenuation measurements for specific patients, but will give a good estimate of the likely magnitude of differences between patients, and hence a good indication of the likely variation in uptake measurements due to patient head size.

Following Chang's description of an attenuation correction technique ⁹⁸, and using a simple ellipse as the head outline, the attenuation effect at different positions can be calculated.

3.9.1 Attenuation compensation calculations - Methods

A series of 21 patients attending for PET-CT brain investigations were chosen at random. After appropriate manual realignment to correct for head position, a central slice covering the striatum was selected on the CT image volume. The width (left-right) and depth (ant-post) distances were measured manually using clinical CT reporting software.

Various head sizes were modelled using a simple spreadsheet program as an ellipse following Equation 3-10. The ellipse is centred on the coordinate origin, and the half lengths (or radii) are 'a' and 'b'. The coordinates x_{θ} and y_{θ} are the points on the ellipse that are found at angle θ from the point of interest X_0, Y_0 .

$$\text{for } -\frac{\pi}{2} < \theta < \frac{\pi}{2}, x_{\theta} = -\frac{ab\sqrt{a^2 m^2 - c^2 + b^2} + a^2 c m}{a^2 m^2 + b^2}, y_{\theta} = -\frac{abm\sqrt{a^2 m^2 - c^2 + b^2} - b^2 c}{a^2 m^2 + b^2}$$

$$\text{for } \theta > \frac{\pi}{2} \text{ or } \theta < -\frac{\pi}{2}, x_{\theta} = \frac{ab\sqrt{a^2 m^2 - c^2 + b^2} - a^2 c m}{a^2 m^2 + b^2}, y_{\theta} = \frac{abm\sqrt{a^2 m^2 - c^2 + b^2} + b^2 c}{a^2 m^2 + b^2}$$

$$m = \tan \theta, c = Y_0 - mX_0$$

Equation 3-10

The thickness of tissue (l_{θ}) from the point of interest (X_0, Y_0) for a particular projection angle θ is given by Equation 3-11.

$$l_{\theta} = \sqrt{(x_{\theta} - X_0)^2 + (y_{\theta} - Y_0)^2}$$

Equation 3-11

The attenuation effect at that point is then calculated by averaging the attenuation effects over all projection angles (M) (Equation 3-12) (adapted from 98).

$$\text{Att}_{X_0, Y_0} = \frac{\sum_{i=1}^M e^{-\mu l_{\theta_i}}}{M}$$

Equation 3-12

To measure the effect of head size on DAT SPECT imaging, the attenuation effect at a series of specific points of interest was calculated. These points of interest are shown in Figure 3-34 and relate to the striatum and two series of background non-specific binding areas. The attenuation effect was calculated for each point and averaged for each of five head sizes.

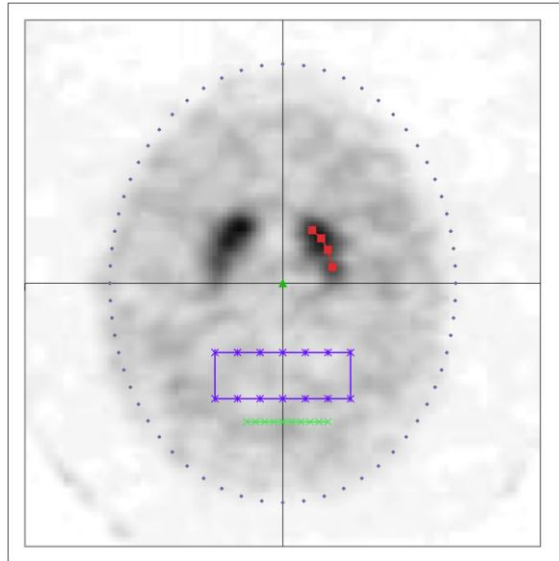


Figure 3-34 The position of points of interest within a brain outline used for attenuation effect calculations a string of points are shown in the left striatum along with two separate regions used for non-specific binding estimates

3.9.2 Attenuation compensation calculations - Results

A summary of the head sizes measured for the 21 subjects is shown in Table 3-4. The minimum and maximum head depths were not seen on the same patients as the minimum and maximum head widths i.e. the ratio of width to depth was not constant. The data in Table 3-4 was used to model the range of head sizes to be seen in clinical imaging. The mean depth and mean width were used to create the “average” head size.

Table 3-4 Head sizes for a test population

	Depth (ant-post)	Width (left-right)	Ratio Depth/Width
Minimum	17.8cm	14.0cm	0.745
Mean	19.0cm	15.3cm	0.801
(Standard deviation)	(0.8)	(0.7)	(0.044)
Maximum	20.6cm	16.7cm	0.908

The maximum and minimum head sizes were calculated using the maximum depths and widths recorded. These do not relate to specific patients and are likely to show a larger range than is present in the population. The measurements for the patients with the maximum and minimum depth-width ratios were used directly and labelled “maximum ratio” and “minimum ratio”. These five head sizes are shown in Figure 3-35.

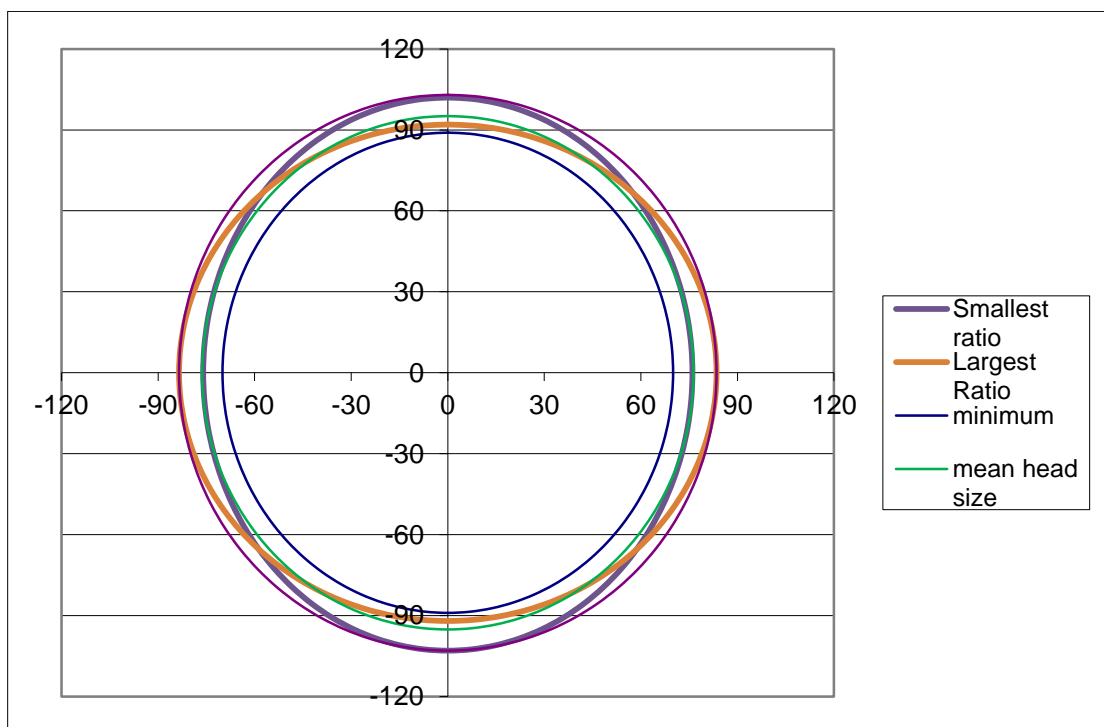


Figure 3-35 Elliptical outlines based on head size measurements

In Figure 3-36 a series of three points of interest are evaluated. The points marked on the perimeter of the head show the position of the points for which the attenuation depth is measured for a specific point of interest – marked with a green triangle.

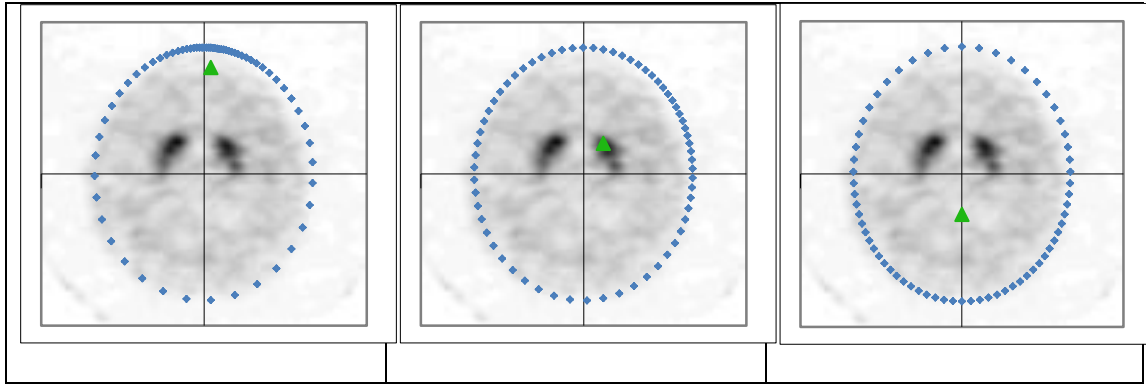


Figure 3-36 Calculation of the attenuation effect at three points

Table 3-5 Calculated attenuation affects in different positions within different sized heads

	Mean Brain size	Minimu m brain size	Maximu m head size	Maximu m ratio	Minimu m ratio	Min ratio from mean	Max ratio from mean
Target average	0.410	0.440	0.377	0.396	0.399	-8.0%	7.4%
Background average 1	0.482	0.522	0.440	0.476	0.462	-8.8%	8.1%
Background average 2	0.444	0.479	0.407	0.433	0.425	-8.4%	7.8%
Target – background ratio 1	1.177	1.185	1.167	1.203	1.157	-1.7%	2.1%
Target to Background ratio 2	1.084	1.087	1.079	1.094	1.065	-1.8%	0.9%

The calculated attenuation effect for the various head sizes and shapes are shown in Table 3-5. The ratios of the attenuation effects between the two background regions and the target region is also shown. The range of attenuation effects was calculated using the minimum and maximum values of the head size assessed. This was expressed as a fraction of the mean head size. The maximum differences in attenuation effects are seen for background region

1, and ranges from -9% + 8%. However, the ranges for the two ratio calculations is much lower at -2% to +2% of the mean value.

3.9.3 Attenuation compensation calculations – Discussion

While the bias in absolute activity due to attenuation is a large factor, the effect is relatively uniform between different head sizes with less than 10% deviation from the average. When the scaling effects of a non-specific binding region are included the differences come down to only +/- 2%. These results suggest that while the inclusion of attenuation correction may have a significant effect on the absolute quantitation, it is unlikely to have a large impact on classification systems. This error can be compared with the test-retest variability for this pharmaceutical published by Booij et al. ⁴⁸. They scanned 6 healthy volunteers, and 10 patients with Parkinson's disease twice each, with repeat scans occurring at 3-6 weeks after the first. With an identical scanner, and acquisition, processing and analysis techniques, they showed a test-retest variability of 7.47% (+/- 6.36%) and 7.36 (+/- 6.16) for the health volunteers and PD groups respectively.

3.10 Compensation methods for scatter - discussion

In section 2.2 Triple Energy Window scatter corrections were investigated with respect to the collimator sensitivity with distance. In that section it was found the TEW could compensate for these septal penetration effects thus creating a more stable sensitivity value. In section 2.3 this same TEW was found to make improvements to the FWHM and FWTM resolution measurements. The former

effect is likely to improve quantitation in SPECT calculations, and the second effect may improve the reproduction of fine detail in SPECT images.

For the first of these effects, the improvement is likely to be head size dependent scaling factor, which, like the attenuation effects in the last section may not have a large effect on quantitation. The second effect - improvement in resolution - may allow improved contrast.

The main use for TEW corrections is to compensate for scattered photons being included within the photopeak window. The amount of scatter depends in part to the size of the head and distribution of activity within it. The distribution of activity in other parts of the body may also be important. For many patients there is significant uptake in the salivary and thyroid glands. These tend to be on the edge of the field of view, or just outside it. There is also uptake in other organs outside the field of view that will be variable between patients.

Because of these out of field effects, and the variable glandular uptake, modelling with phantoms may not provide results representative of real patient data.

The downside of this form of scatter correction is in the increases in image noise. Again, representative measurements of this cannot be made using phantom data as it does not truly represent the noise properties of clinical data, so it can only really be assessed using clinical patient images.

Published work ¹¹⁹ has shown that these corrections can improve the accuracy of absolute uptake measurements in patients. No studies have shown improvements in diagnostic accuracy when these corrections are included.

In their work with cross-site camera comparisons using phantom data Dickson et al. ¹¹⁵ found problems with OSEM reconstructions in data with low pixel counts. They related this to the non-negativity constraint in OSEM. Such a constraint

could be a problem when using TEW corrected projection data due to the low background count levels and increased noise. A related issue when dealing with low count statistics is the use of integer values in image formats. For data with low counts the rounding off to the nearest integer value could affect the resultant images.

3.11 Overview of reconstruction results

Looking at the general results seen across the range of parameters for the noise and contrast measurements the OSEM models that include resolution modelling perform better than the basic OSEM model. However, there are no significant differences between the fixed and depth dependent methods for the SUR measurements used here.

The DD-RR and fixed-RR OSEM models cannot be separated using the measured described here, and it may be that the full advantage of a depth dependent model only becomes noticeable over larger distances. Since there were no drawbacks from using the more complex model this system was used.

The FBP and DDDR-OSEM models can be optimised to give very similar performance for the large VOI based uptake measurements. The size of the VOIs used to measure the contrast using striatal uptake ratios are quite large – they cover the entire striatum on one side. It is conceivable that the DDDR-OSEM outperforms the FRR-OSEM model and FBP for smaller details. To measure such changes a large range of different sized ROIs and VOIs would have to be employed. The exact choice of these VOIs would affect the outcome of such a test, and without a clear anatomical reason for choosing the VOIs such a choice may introduce bias.

Another way of measuring these fine details is to use the line profile NMSE tests. There will of course be some variability in these results based on where exactly the line profile is drawn, but there are differences seen between the FBP and resolution recovery based OSEM results. The DRRR-OSEM model can give superior results in this measure of fine detail. The improved score for this measure suggests an ability to separate out the edges of the striatum. This difference correlates with the preference of the expert observers based on visual assessment.

The extent of the collimator blurring function was shown to have some unpredictable effects on the resultant images. Although giving better results in both the VOI uptake measures and the line-profile NMSE score, the very wide models gave significant artefacts within the images. The middling results of (x1z1) and (x2z1) gave a boost in both these metrics without obvious artefacts.

The number of iterations used in the reconstruction must be sufficient to ensure adequate convergence of the images. After 96 equivalent iterations (12i 8subsets) 95% of the maximum contrast (as measured at 200 equivalent iterations) has been recovered. This figure agrees well with the phantom work done by Dickson et al. ¹¹⁵ who suggested 100 iterations should be performed.

The levels of noise and contrast are tuneable using different levels of post-reconstruction smoothing. For automatic or semi-automatic quantification systems that use VOI Ratios, it is likely that reconstruction methods that recover high levels of contrast in the images would be beneficial. The optimum level of image smoothing may be dependent on the particular classification process being used. Techniques using large VOIs are unlikely to be sensitive to filtering levels, whereas techniques using tightly constrained VOIs may be heavily

affected. The new techniques proposed using machine learning are also likely to be dependent on the level of smoothing.

As discussed in section 3.4 the ultimate test of the quality of an image is that it allows the correct diagnosis to be given. To this end, a set of patient data will be reconstructed using a set reconstruction model and a range of post-reconstruction filters. Each subjects scan data will be reconstructed using the DRR-OSEM iterative reconstruction algorithm with 12iterations of 8subsets. The resolution will be modelled to a width of +/- 2 pixels in the x axis and +/- 1pixel in the Z axis. The post-reconstruction filtering will be performed using a Butterworth filter with settings ranging from $f_{co}=0.5\text{cycles/cm}$ up to 2.0cycles/cm . A non-filtered dataset will also be used in the analysis.

4 Clinical Data

Clinical procedures for scanning and processing DAT SPECT scans within Imperial College Healthcare NHS Trust are as follows.

Patient preparation involves thyroid blocking using tablets. These are taken twice per day for three days starting the morning of the day before the scan. A total of six doses of 60mg of potassium iodide are taken. The thyroid blocking procedures are in place to minimise uptake of free ^{123}I iodine within the thyroid, hence minimising the radiation dose to the patient.

Some medications can affect the way in which ^{123}I -FP- β -CIT is taken up by the brain. Within the drug safety leaflet for ^{123}I -FP- β -CIT a number of such medications are listed, however the advice is that the prescribing clinician may or may not restrict these medications. The recommendations are that such medication may affect absolute quantification methods, but are unlikely to cause changes in visual appearance of the tracer distribution. In this institution no restrictions are placed on these medications. This choice was made out of simplicity for the patients. In a prospective study this choice may be made differently, but for this retrospective analysis the data collection has had to follow the clinical routine. In fact, the purpose of this work is to investigate the use of classification methods that can be used in clinical practice and as such making restrictions on these drugs may bias the results.

On the day of the scan the patients is allowed to eat and drink normally. They are requested to stay well hydrated throughout the day and to void their bladder as often as necessary – this too reduces the patient's dose.

The patient is injected intravenously with 185MBq ^{123}I -FP- β -CIT. A delay of between 3 and 5 hours is then needed for the tracer to be taken up by the dopaminergic system in the brain and reach a dynamic equilibrium with the non-specific binding. Immediately prior to the scan the patient is asked to void their bladder. They are then positioned supine on the scanning couch with their head positioned in a headrest that extends out of the patient couch. See figure 1.7, 1.8. The patients head is held still using chin and forehead straps.

The SPECT acquisition is performed using the parameters defined in chapter 2 using Forte gamma cameras (Forte, ADAC Laboratories, California, USA) and LEHR and VXGP collimators. The acquired data is then reconstructed using the processing steps selected for visual reporting.

All images are then realigned manually to compensate for patient positioning. The realigned studies are then reported by experienced clinicians using dedicated nuclear medicine reporting software.

Following approval from the Clinical Governance Department of Charing Cross Hospital an anonymised patient database was created for auditing the value of ^{123}I -FP- β -CIT SPECT imaging for the management of movement disorder patients. The database includes salient clinical details and the (clinical) reported results of ^{123}I -FP- β -CIT SPECT imaging. Where possible, a definitive diagnosis was given for each patient by a consultant neurologist who was responsible for his or her care. This diagnosis was given at a later date using clinical history, scan results and long term follow up of the patients. No specific minimum follow up time was used for this definitive diagnosis.

The audit database included only anonymised data, and all image data was

anonymised before being copied into the study database. The audit database only contained data from a single referrer and his registrars. They referred patients from their movement disorders clinic and hence the vast majority of the patients in this database have movement based disorders, as opposed to dementia type disorders.

Image data from patients within this database were reconstructed with the range of test reconstructions defined in chapter 3 to allow development and testing of automatic classification systems.

Over the course of this work, the details of 299 patients were included onto this database. Two entries were duplicates and were removed from the database. Six patients were lost from follow up, did not have a scan or requested to be removed from the database, leaving 291 records.

Patients for whom a definitive diagnosis could not be given were excluded from this study. These included patients for whom the clinical symptoms, disease progression and scan results were inconclusive at the time of this study. This resulted in 74 patients being excluded from this study.

Patients whose diagnosis involved both small vessel cerebrovascular disease (SVCD) and a non-PS disease were excluded from this study. SVCD can produce localised reduction in DAT that may interfere with the normal presentation of these scans. Patients with a mixture of SVCD and a Parkinsonian syndrome were kept within the database.

Following these exclusions, 194 patients remained in the database including 135 with a definitive diagnosis of a Parkinsonian Syndrome (PS), and 59 with a non-Parkinson's syndrome (Non-PS). Of the 194, a group of 116 patients were

scanned before changes in the acquisition techniques (as described in chapter 2) were implemented. The demographics of these patients are shown in Table 4-1. A further 78 were included while the newer acquisition protocols were being used. The demographics for this second group are shown in Table 4-2. A number of patients reported their “handedness” as mixed, or variable. These were included as “other” in the table of demographics. Statistical analysis of these demographics was performed using SPSS software (SPSS statistics, v19.0, IBM, USA) using a significance value of P=0.05.

Table 4-1 Demographics of patients scanned using original parameters

Diagnosis	Number of patients	Sex (M/F)	Age (years) Mean (range)	Handedness (R, L mix/other)	Length of follow up mean (range)
PS	79	52 / 31	60.2 (29-85)	51/28/7	32.28 (0-108)
Non-PS	37	22 / 24	64.7 (38-81)	31/3/3	24.28 (0-60)

Table 4-2 Demographics of patients scanned using optimised parameters

Diagnosis	Number of patients	Sex (M/F)	Age (years) Mean (range)	Handedness (R, L mix/other)	Length of follow up mean (range)
PS	56	43 / 21	62.89 (33-84)	61/2/1	6.63 / 13 (0-60)
Non-PS	22	15 / 15	62.73 (34-83)	29/1/0	5.1 / 9.3 (0-36)

These divisions were not significantly different for the two groups. Ages compared using Mann-Whitney U test. Sex and handedness compared using χ^2 test.

Table 4-3 Breakdown of the diagnosis subtypes for the two datasets

	Diagnosis*	Using Original acquisition parameters	Using optimised acquisition parameters	Total
Parkinsonian syndromes	Parkinson's Disease	49	39	88
	PD & CVPS / SVCD	7	3	10
	DLB	4	1	5
	PD and ET	10	4	14
	MSA	2	1	3
	PD & ET & SVCD	1	1	2
	PSP	3	5	8
	PD & MS	1	0	1
	Stroke & secondary PD	1	0	1
	Young onset PD	1	2	3
Non-Parkinsonian syndromes	ET	20	11	31
	Dystonia	9	4	13
	Drug Induced	4	3	7
	OT	2	2	4
	Psychogenic/ stress/ Malingering	0	2	2
	FXTAS	1	0	1
	MS	1	0	1

*For definitions of abbreviations and disease groups please see chapter 1.

The patient data described in **Table 4-1** (using the original acquisition techniques) was used in work comparing different automatic classification systems ¹. The breakdown for the various subtypes of disease is shown in Table 4-3. The two groups of data were used to test various classification techniques in chapters 5 and 6.

5 Development of a PCA based classification system

5.1 Feature Extraction and machine learning

There are a range of matrix transforms that can be used to summarize large data sets. The transforms are simply ways of describing the same data using different measures. One example of a simple transformation is the Fourier Transform (FT). This describes an image as a collection of sine waves, with frequencies and amplitudes. When in this form the data is said to be in frequency space, or 'k' space. Examining the data in the frequency domain shows how much of the information in the image is coming from different frequency data. This can be useful in image filtering. FTs and inverse FTs allow the transfer of information between real space and frequency space and assuming a sufficient range of sine wave frequencies is used, the process is reversible with no loss of data.

Other transforms can be applied that extract different sorts of information from the images. The transforms take the general form of a series of components and weights, analogous to the sine waves and amplitudes in FT. The different transforms optimise different properties of these components.

When applied to image analysis, the input matrix will typically be multiple images from a time series, from repeat scans or from groups of patients.

Some of the transformations that have been applied to medical imaging tasks are described below.

5.1.1 Principal Component Analysis and Singular Value Decomposition

Principal Component Analysis (PCA) allows data to be represented by a set of orthogonal eigenvectors and corresponding eigenvalues or weightings. The eigenvectors are common for all subjects whilst the weightings are unique to each subject. With reference to FT, the eigenvectors are equivalent to the sine waves, and the weights equivalent to the amplitudes. The weighting factors show how much a particular component needs to be scaled by to give that subjects image. PCA finds the uncorrelated components that describe the input data in the most economical way. That is, the variance in the components is maximised. There are a number of computational processes that can be used to perform PCA including Singular Value Decomposition (SVD).

The main applications for PCA in nuclear medicine imaging have been in characterising regional Cerebral Blood Flow (rCBF) studies and as a starting point to factor analysis.

For rCBF studies, Houston ¹²⁰ defined a brain template using a set of normal patient scans. The range of component weights seen in template data was used to limit the allowed weights for subsequent test data. They calculated the residual information: the information in a test patient that could not be described by the principal components using weightings up to three standard deviations.

Instead of looking for separate clusters of points, some groups have looked at the correlation between the weighting for a particular component with a clinical grading. This technique was applied to the ¹⁸F-FluoroDeoxyGlucose (FDG) PET

studies performed on patients with idiopathic PD and age-matched controls ¹²¹. They found two components that correlated to motor dysfunction scores and executive process (such as working memory and planning tasks scores).

Of course, the neuropsychological results can also be included in such an analysis. One example of this approach ¹²² combined the test results with ROI scores from SPECT data. Their analysis identified specific correlations between different groups of test results.

5.1.2 Factor Analysis

Factor Analysis (FA) is a group of rotations/transformations that use PCA as a starting point. They will typically employ prior knowledge of the data to constrain the components. One example is the non-negativity constraint that has been employed in FA of dynamic nuclear medicine data. The studies of dynamic data have generally involved analysing the components as these can relate to bio kinetics of the structures being imaged ¹²³⁻¹²⁶.

Another variation of FA is Non-negative factorisation, which as the name implies, applies a transformation that results in purely additive components. This was applied to rCBF data and when coupled with a support vector machine classification process, successfully classified groups of AD and non-AD patients ¹²⁷.

5.1.3 Independent component analysis

Independent component analysis (ICA) is different from PCA in that the components are chosen to be independent rather than merely uncorrelated. ICA has been used predominantly in blind source separation problems. The focus of

ICA work is in identifying the components themselves, often as a method for removal of extraneous data and noise.

5.1.4 Logistic Discriminate Analysis

Logistic discriminate analysis (LDA) performs a transformation that maximises the correlation between the classification groups and the components. When applied to the classification of images into two groups, a discriminator is defined for each voxel. The clustering of these discriminating voxels allows a simplification of the process using methods related to those in SPM. This process has been applied to DAT imaging with ^{99m}Tc -TRODAT1 ¹²⁸.

5.1.5 Projection to Latent Structures or Partial Least Squares

The components are chosen to maximise the separation between training groups using brain perfusion images to differentiate between AD and frontotemporal dementia ¹²⁹.

5.2 Machine learning and classification methods

The transforms described above take groups of patient data and map them into some form of feature space. The exact dimensions of the feature space being related to the form of the transform. The weightings for a set of patient images can be plotted as vectors in this p-dimensional 'feature' space. Patients of similar disease state should display similar tracer distributions, and hence their weighting vectors should cluster in feature space. A population of patient images with known disease state(s) can be used as a training set to define the feature space.

The image volume from a test patient can then be mapped into this feature space. A measure of the proximity of this vector to either the healthy group, or to two or more groups representing particular patient disease groups may provide a powerful tool for classification of the patient. What is needed is a way of defining the boundaries in this feature space that relate to specific disease groups.

There are a number of approaches that may be used to define clusters of points in one or more dimensions. To describe these systems some example data was created using normally distributed random numbers, defined in two dimensions.

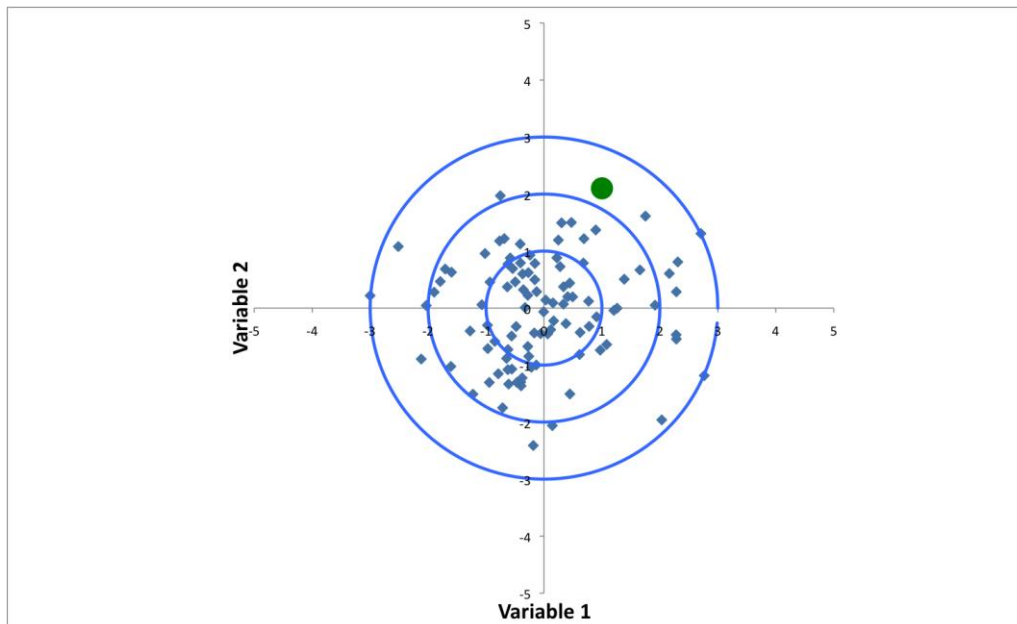


Figure 5-1 Example data for normally distributed population of 100 subjects, with mean of 0 and standard deviation of 1 for both axes. The circles represent 1, 2 and 3 standard deviations from the mean.

For the basic single class classification task, the standard deviation of the training set can be used to define the probability that a patient is from that group. This is shown in Figure 5-1 for data defined using two variables. A test patient is shown as a green circle. A score can be defined in terms of the number of standard deviations away from the group centre, which can then be translated into a probability. In this example the test patient would be given a score of 2.3

standard deviations, meaning that a patient from the healthy group would have a 2.1% chance of being at least this far from the cluster centre. To use this figure for binary classification a cut off would need to be defined depending on the preferred trade-off between sensitivity and specificity.

5.2.1 The group prototype method

Moving on to two-class systems, the task is to define which group the test patient is more likely to belong to. One conceptually simple technique defines a Group Prototype (GP) as the centre of each cluster by calculating the average vector for each classification group. The distance from test patient to the Group Prototypes can then be calculated and the relative distances compared to give the diagnosis

3.

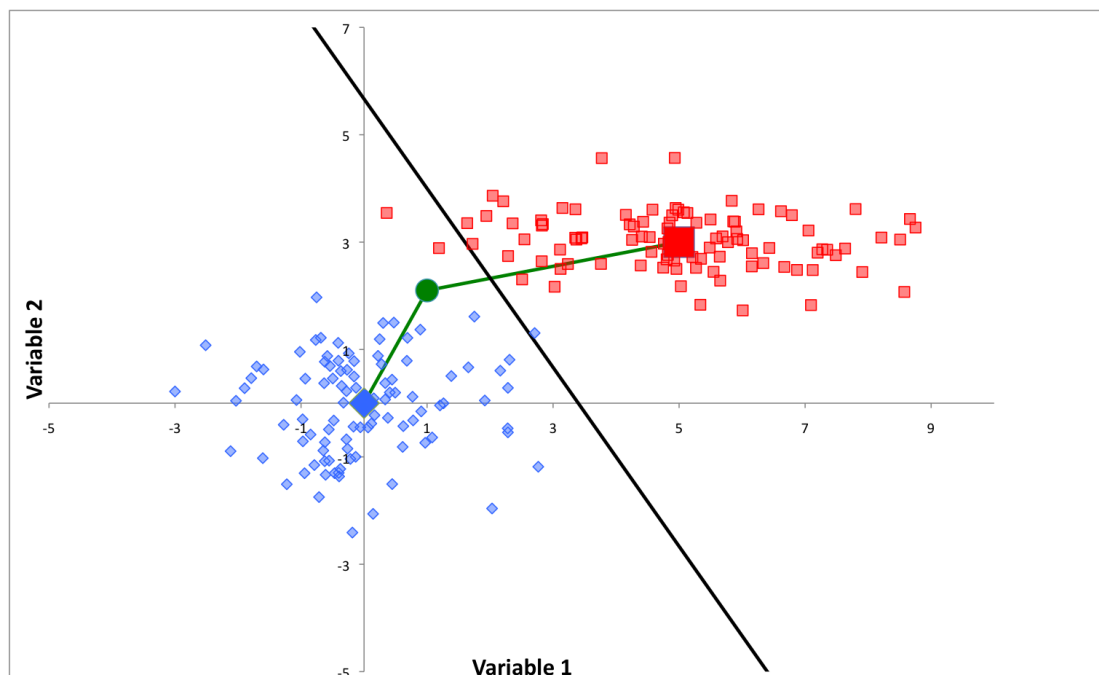


Figure 5-2. Two example data sets are shown along with a test patient in green. The group-prototype method is used here to classify the test patient in the normal (blue) group. The decision line for the Group prototype method is shown and is defined as the being equidistant from the two group centres.

Figure 5-2 shows the same “healthy” data from Figure 5-1, but with the addition of a cluster of non-healthy patients, shown in red. The “group prototypes” for healthy and non-healthy are shown at the centre of each cluster. The vectors from the test patient to the two group prototypes are shown in green. Their relative lengths give the classification. The decision line, is a straight line, and is shown in black. In this example the test patient would be classed as normal.

In some approaches to transforming data into feature space, the data is rounded and centred. Centring moves the data so that the mean vector is positioned at the origin. The rounding is a rescaling of the axes according to the variation seen in that axes. This second effect could theoretically make a significant difference to the above classification.

5.2.2 The Mahalanobis Distance

The amount of variation contained within each component may be different for different groups of patients. For example, if one component were to relate to disease severity of progression, it would be expected that the normal data set would bunch closely together, whilst the diseased state would spread out in a continuum from near the non-diseased state through to very advanced disease. These differences can be included by scaling the group prototype measurement by the variation of that specific group in that specific direction. This is known as the Mahalanobis distance.

This effect can be described using the same data as in the previous example and marking on the extent of the 1st 2nd and 3rd standard deviations for the two groups. See Figure 5-3.

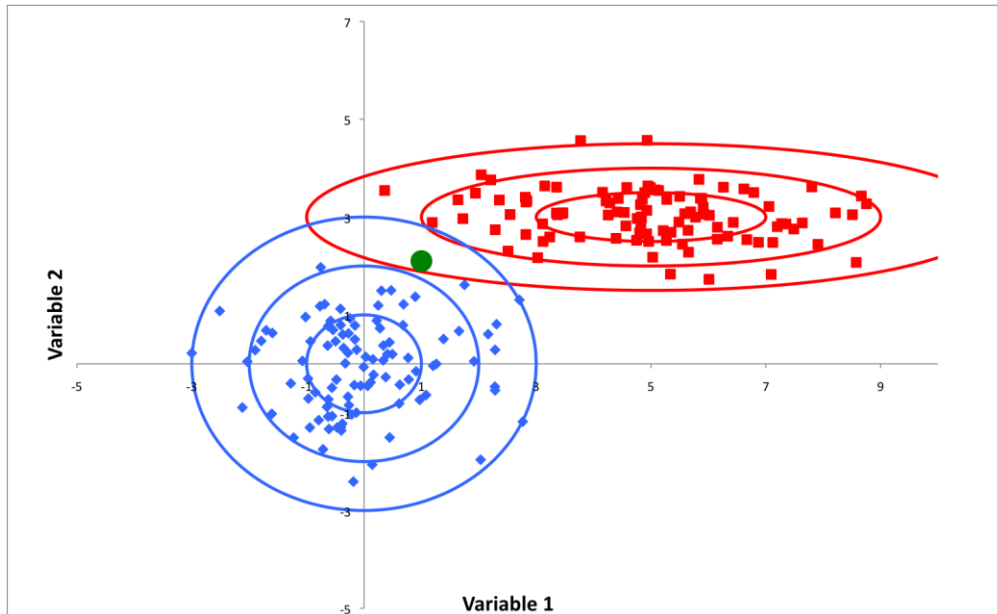


Figure 5-3. Two example data sets are shown with dissimilar amount of variation in the two components. The concentric ellipses represent the extent of the 1st, 2nd and 3rd standard deviations for the two groups.

5.2.3 The Bayes Classifiers

These Mahalanobis distances can be used as the probability measurement for a Bayesian classification approach. This is calculated by modelling the distributions seen in the training set and combining it with prior knowledge. The probability that a test patient belongs to a particular group (Y), based on their component weightings (W) is given by Equation 5-1.

$$P(\text{Group}_Y | W) = \frac{P(W | \text{Group}_Y)P(\text{Group}_Y)}{P(W)} \quad \text{Equation 5-1}$$

The test patient will be assigned to the group with the highest probability. Since $P(W)$ is constant, only the prior probability $P(\text{Group}_Y)$ and the function $P(W|\text{Group}_Y)$ need to be defined. Typically the prior will be neutral or will be equal to the ratios of classifications in the training data set. All patients in each training group are used to define the probability functions. The distributions are assumed to be normal distributions (Gaussian), with means and standard

deviations calculated for each group and component. Such an approach was used in classifying images of Alzheimer's Disease following PCA ¹³⁰.

With no prior probability, the decision line falls where the Mahalanobis distance is equal for the two classification groups. However, adding a value for the prior probability can tune the performance of such a classifier to improve either sensitivity or specificity. Figure 5-4 shows the decision lines (in black) using a prior equal to 1/3, 1/2 (i.e. equal probability) or 2/3. Note that the example test patient will be classified differently when different prior probabilities are used.

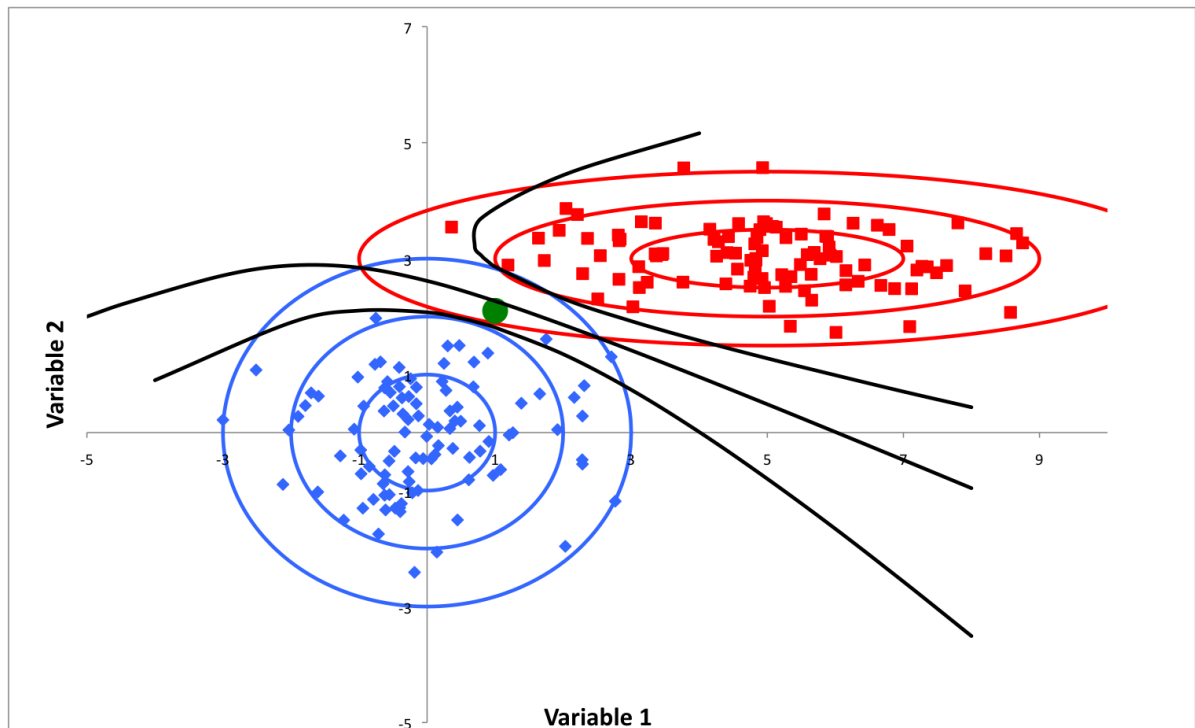


Figure 5-4. The decision lines when using different prior probabilities in a Bayesian Classifier

5.2.4 Support Vector Machines

Support Vector Machines (SVMs) are supervised learning techniques that have recently been proposed for classification problems ¹³¹. The aim is to define a multi-dimensional plane that best separates the two training groups. In Figure 5-5 an example is given using two groups of data in two dimensions. All three of

the dividing planes shown give perfect separation between the two groups, however, visually it can be seen that the middle line may be the best separator.

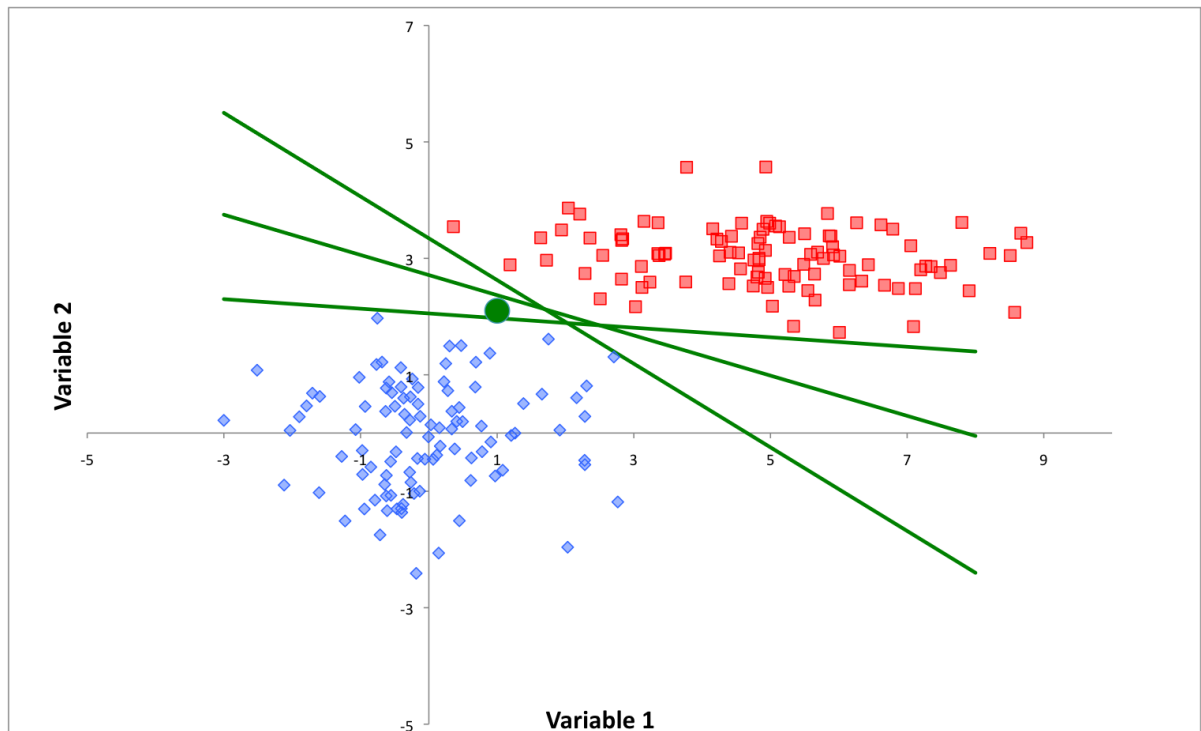


Figure 5-5 Two groups being separated by three straight decision lines

In SVM, the optimised plane is defined by a subset of the training data, the support vectors. These vectors lie along the boundary between the two classes. The plane is defined to maximise the distance between the plane and the support vectors on either side. In Figure 5-6 the central division line from Figure 5-5 is shown along with the “fat margins” and the distance B , which is maximised in the SVM process. The plane can then be used to automatically classify other data points as they are mapped into the same feature space. It should be noted that the classification does not give a degree of severity, or a probability of disease. The result of this classification is binary.

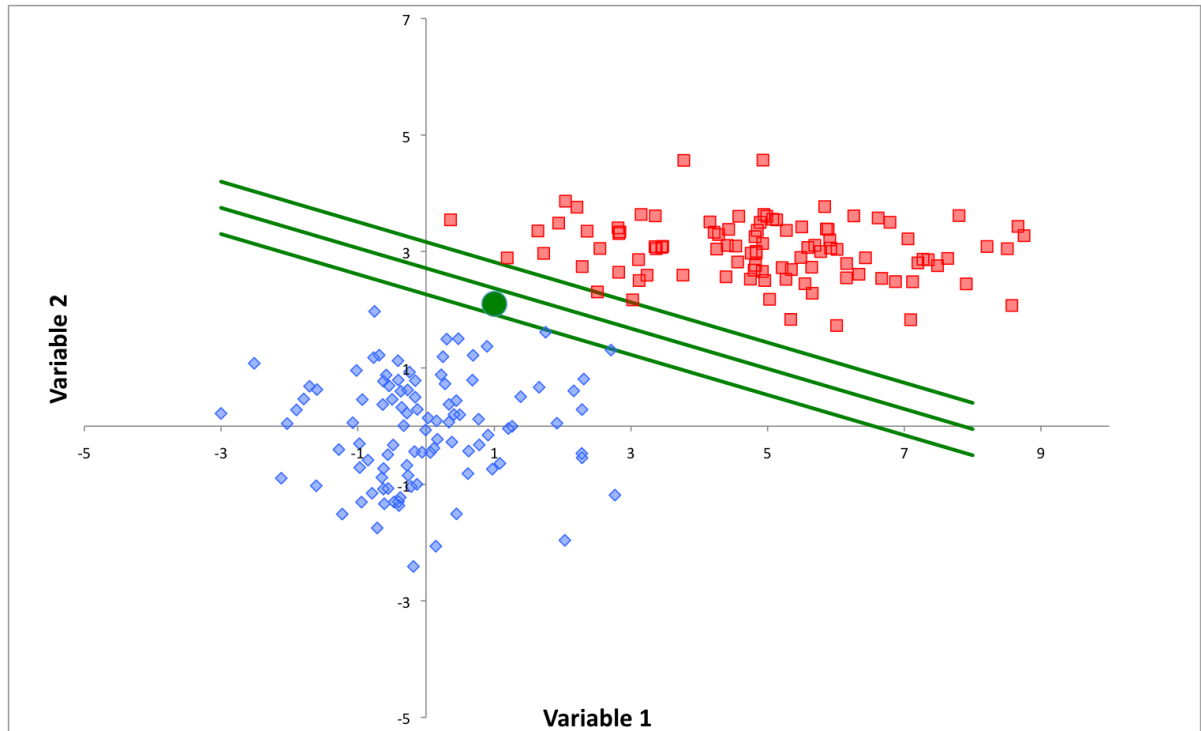


Figure 5-6 The optimum decision line as picked by SVM the parallel lines either side show the so-called 'fat margin' the width of which is maximised in the SVM process.

Using a linear decision plane, the computational procedure is to maximise the thickness of the fat margins. The identification of this maximum is found using standard programming solutions. A powerful extension of this process is to map the data into further dimensions using non-linear mappings. The decision plane will still be defined using the solutions to the same standard solutions, but the affect is to apply non-linear boundaries in the original coordinate system. This is known in programming parlance as the 'kernel trick'.

5.3 Image processing and preparation

It is proposed that an automatic classification tool can be created using a combination of PCA and machine learning tools. Before being used in such a system, the reconstructed data in the form of transverse slices must be put through a number of processing steps.

5.3.1 Image registration / spatial normalisation

For meaningful comparison of images from different patients, a method of identifying equivalent areas on multiple images must be found. For the manual ROI/VOI methods this involves the user marking specific regions of the brains by placing ROIs/VOIs on known areas following visual interpretation. For the automatic ROI/VOI options and the methods using direct pixel comparisons, the image volumes are spatially aligned to each other. For further interpretation, it is useful if this alignment can be in a standardised space so that assessment can be related to known functional areas of the brain.

Spatial normalisation, involves realigning an image so that the voxel information at a particular coordinates on one image refers to the same anatomical region on another image.

There are three parts to any image registration process:

- The image transformations
- The measurements of how well the images are aligned
- The re-sampling of the image data

There are many options for these three parts and the ideal combinations will depend heavily on the images that are to be brought into registration. The more prominent options are discussed below. Since this work is based on three dimensional image volumes, the discussion will be focused on 3D-3D registrations. However, similar methodologies can be applied to higher or lower dimensional data.

5.3.1.1 Image spatial transformations

The simplest transformations involve rigid body movements. These transformations include moving (translations) in the three dimensions and three rotations, thus six degrees of freedom are available. These transformations are often employed when the image data is from the same subject, which has been imaged in the same imaging plane and with known pixel sizes.

Problems can arise when pixel size data is missing or incorrect. In these circumstances three more degrees of freedom are required – to scale along each dimension. Also if the data has been acquired with a gantry tilt, as can be the case in CT scans of the head, this tilt must be reflected in the transformation in the form of image skew. Transformations that include translation, rotation, scaling and skews are called affine transformations and for 3D data will include 12 degrees of freedom.

These transformations should allow for alignment of data from an individual subject being repeatedly imaged. However, some problems arise when either the shape of an organ or tissue changes, or multiple patients are included.

Non-linear warping allows many more degrees of freedom. In more complex warping fields the fields can be non-stationary, with data in different areas being moved by different amounts and in different directions. These types of fittings have been found to be very useful in MRI based studies, where not only are there inter-patient differences which require some warping to allow alignment, but also MRI data itself can contain some image warping effects.

The reason for image registration is to allow the equivalent parts of two or more images to be positioned at the same coordinates. If the warping fields are left unconstrained then the alignment may optimise to a position where the tissues

occupying the same location do not correspond to the same structures in the various patients.

In the case of dopamine transporter imaging, the disease progression is often seen as a change in shape of the activity distribution. It is possible for this change in shape to be masked if image warping is used, however most of the published work ^{56,57,83,132} using the SPM package has included some degree of warping in the spatial normalisation stage. Others ¹³³ have avoided this warping stage.

5.3.1.2 Image alignment metrics

The image landmark registration is performed by identifying reference points that occur on both images and then applying a transform that minimises these distances. Such reference points are generally not available in DAT and D2 imaging unless x-ray CT data has also been acquired, or fiducial markers are used.

Count difference (CD) is simply the sum of the absolute differences between the source and target images ¹³⁴. This is most appropriate when the pixel values are of similar magnitude and spatial distribution.

Mutual information (MI) has been proposed ¹³⁵ as an alternative to count differences as it can be used in applications where pixel values are likely to be very different, e.g. in images of different modalities or pharmaceuticals.

Kas and co-workers ¹³⁶ investigated the use of different normal templates in SPM for ¹²³I-FP- β -CIT studies, and found that quantification was insensitive to the template used for the spatial normalisation. They used the SPM implementation of the MI algorithm and found minimal errors.

Van Laere et al. ¹³⁷ used data from a range of brain SPECT tracers to investigate different automated spatial normalisation techniques, including CD, MI and a system called Uniformity index (UI). They concluded that the addition of CT data allowed the best realignment especially for receptor/transporter imaging. A secondary conclusion was that CD was superior to MI.

Of the commercial image analysis packages designed for DAT SPECT imaging, only the BRASS system from Hermes uses automatic image alignment. Their package includes a number of methodologies, but the recommended model is an implementation of the normalised mutual information technique.

5.3.1.3 Re-sampling methods

Once the alignment metric has been optimised and the best transformation identified, the source image must then be transformed to the target location. The transform is highly unlikely to represent movement of an integer number of pixels – and so the voxel values in the transformed image will be formed from an interpolation of a group of neighbouring pixels. Since interpolation is being applied, there will normally be a small level of smoothing applied to the data.

The most common form of re-sampling used is based on tri-linear interpolation. Tri-linear interpolation takes a weighted average of the eight voxels closest to the centre of the new voxel location.

5.3.2 Count normalisation

There are many factors that will affect the absolute levels of tracer uptake within the brain that are not related to the disease process being studied. These include differences in injected activity, radioactive decay between injection and scan, the

distribution volume, the size of the organs, kidney function, and hydration levels amongst others. Hence, the absolute uptake in the regions of the brain is of little direct clinical use. A more useful index is the amount of tracer that binds to the organ as a fraction of the amount of tracer that reaches the organ. A full investigation into amount of receptor sites would involve following the tracer uptake dynamically over a period of many hours. As ^{123}I -FP- β -CIT reaches a prolonged level of binding equilibrium ¹³⁸, a simpler definition of the binding potential can be employed using a measure of non-specific binding. The regional distribution of dopamine uptake was measured in a post-mortem uptake study by De Keyser et al. ¹³⁹. They showed that a range of areas in the cerebral cortex, including the occipital cortex, displayed no specific dopamine binding. It is therefore appropriate to use either the occipital cortex or the whole cerebellum to normalise the pixel intensities within DAT images.

A volume centred on the occipital cortex was identified and used as an estimate of non-specific binding in counts/voxel (C_{occ}). The volume was defined on the template image using MRlcro and then converted into a binary mask. This volume was then used to scale the counts in each voxel (C_{voi}) to give the binding potential for that voxel, BP_{voi} .

$$BP_{voi} = \frac{(C_{voi} - C_{occ})}{C_{occ}} \quad \text{Equation 5-2}$$

5.3.3 Image filtering

ROI/VOI based classifications combine the information held in a large number of voxels into just two or four uptake ratios. In this way the techniques are unlikely to be susceptible to pixel level noise as the uptake figure sums across a large number of voxels - effectively averaging the data. The pixel-pixel methods such

as SPM and the PCA methods proposed here use the similarities or differences between voxels in the scans of multiple patients. Hence, localised noise will increase the patient-to-patient differences and possibly mask any true differences between disease states. As such these measures are likely to be very susceptible to pixel noise. Some inter-patient differences will occur due to differences between patients that are unrelated to the disease. Therefore, some level of image smoothing will improve these techniques. However, the level of smoothing applied will have to be optimised for a specific imaging task. As the level of smoothing is altered the classification tool will become less sensitive, but more specific to real group-group differences. The exact level of smoothing will be a trade-off that allows removal of the influence of trivial patient-patient differences, whilst retaining the salient disease features. The exact level of smoothing will be a function of the inherent patient-patient differences, the resolution of the images being used, the size of the real differences between the disease states and the number of patients used within the study.

There is a balance to be achieved for optimum smoothing of datasets prior to analysis with SPM. The filter should smooth/combine data from individual structures, allowing for inter-patient and intra-regional variability, but should not be so broad as to significantly reduce contrast in real defects/activations.

These factors are necessarily going to be dependent on the resolution of imaging system, the noise levels in the reconstructed images, and the size of the defects seen in the study.

Van-Laere et al. ¹⁴⁰ investigated the effect of filtering when using SPM in SPECT imaging. Their phantom based study used a ^{99m}Tc based phantom with a range of activation foci. The phantom consisted of a series of 4mm thick plastic sheets,

interspaced with paper sheet printed with activity distributions. They showed the trade-off between the numbers of subjects (or study pairs) needed in the study design, with the size of the activation effect. They also investigated the ideal levels of smoothing required using Gaussian filters. In general they found that the group comparison gave highest significance when the filter was around $2 \times \text{FWHM}$ of the system, but were relatively insensitive over the range $1.5\text{-}2.5\text{FWHM}$. They found that the ideal level of smoothing is dependent on the size of the activation as well as the position/surroundings of the defect resolution of the imaging system.

In other works using SPM and SPECT imaging a range of filter sizes have been implemented. In perfusion imaging a Gaussian filter with a FWHM of 10mm⁸³ has been quoted. In DAT imaging, researchers from Innsbruck⁵⁶ used a 6mm Gaussian filter for their work with $^{123}\text{I}\text{-}\beta\text{-CIT}$. For $^{123}\text{I}\text{-FP-}\beta\text{-CIT}$ imaging both 8mm¹³² and 10mm¹³³ filters have been used.

Transferring these ideas of noise levels to PCA based techniques, it should be remembered that the PCA transform finds the components that encapsulate the largest fraction of the variance in the training data set. If pixel noise is a large component of this variance then the transform will produce components that describe the noise, rather than the differences between subject groups. It is apparent then that the level of smoothing within the image data is likely to affect the usefulness of PCA as an image classification tool.

5.3.4 Applying PCA to Groups of SPECT volumes

The range of feature extraction processes described above (in section 5.1) either start with PCA or, another transform which works in an equivalent manner. As

such, the initial stages of a machine-learning model will start with similar data processing. To apply PCA to groups of SPECT images, the volumes of patient data would first need to be spatially aligned, and count normalised.

SPECT image data describe activity distributions in a three dimensional matrix $a_{x,y,z}$. This can be reshaped into a one-dimensional vector:

$$a_{x,y,z} \rightarrow a_i \quad i = 1 \rightarrow m \quad \text{Equation 5-3}$$

Where m equals the total number of voxels in the SPECT volume i.e. the volume dimensions $x*y*z$. Data from a group of n patients can then be combined to form a 2 dimensional matrix, \mathbf{A} ;

$$A = a_{i,j} \quad i = 1 \rightarrow m \quad j = 1 \rightarrow n \quad \text{Equation 5-4}$$

Where $\mathbf{A}_{i,j}$ is the i^{th} voxel of the j^{th} patient. Using Singular Value Decomposition (SVD), \mathbf{A} can be transformed to produce three matrixes \mathbf{U} , \mathbf{S} and \mathbf{V} ,

$$A = USV^T \quad \text{Equation 5-5}$$

\mathbf{U} is the left-eigenvectors, defined as the eigenvectors of \mathbf{AA}^T matrix, \mathbf{S} is a diagonal matrix, whose main diagonal is populated with the singular values of \mathbf{A} in decreasing order, \mathbf{V}^T is the transpose of \mathbf{V} and is the right-eigenvectors, defined as the eigenvectors of $\mathbf{A}^T\mathbf{A}$.

Since the original data was built up from reconstructed volumes, the left eigenvectors, \mathbf{U} , are referred to as EigenVolumes. The full transform would produce m EigenVolumes of m voxels. As $m \gg n$, a restricted transform is performed to give just n EigenVolumes each containing m data points. The product \mathbf{SV}^T is the component weights, sometimes referred to as loadings, or component weightings \mathbf{W} .

$$W = SV^T \quad \text{Equation 5-6}$$

It is the contribution of each of the EigenVolumes needed to recreate each patient's image volume. For the full transformation, \mathbf{SV}^T is an $n*m$ matrix. Each row relates to an individual patient in the original dataset.

A characteristic of this transform is that the EigenVolumes are ordered in terms of their contribution to the variance in the original data. Most of the original image data is contained in the first few EigenVolumes. The first EigenVolume is defined so that combined with its weighting gives the best approximation of the original matrix. The second EigenVolume is the optimal choice to represent the data not covered by the first EigenVolume. In this way, the EigenVolumes are defined in order, to give the original matrix in a mathematically economical way. We expect only the first few EigenVolumes to have significant weightings. The entire cohort of patients can be defined in terms of p -EigenVolumes ($p < n$) common to all patients, along with defining weightings individual to each patient.

The SPECT data for an individual patient j , as given by summing the weighted components.

$$A_j = \sum_{k=1}^p U_k W_{j,k} \quad \text{Equation 5-7}$$

In Principal Component Analysis (PCA) the significant EigenVolumes are called the principal components. They are the main features seen in the columns of the input matrix \mathbf{A} . Although defined here from SVD, principal components were originally defined using the covariance matrix of \mathbf{A} ^{141,142}.

By applying this transform to the input matrix \mathbf{A} , patterns that are common across the columns of \mathbf{A} are defined, along with the spread of weightings seen across the columns of \mathbf{SV}^T .

It is common to subtract the mean patient image (sometimes referred to as Component 0 or PC0) from each patient image prior to performing the SVD. PC0 must also be subtracted from test patients prior to analysis.

5.4 Classification models

5.4.1 Single, dual and multi-class classification tasks and decision trees

Classification tasks can be defined based on the number of groups that are identified. In a simple one-class problem, a definition of a group/state is given, and the task is to state if a test subject belongs to that class or not. When applied to the ioflupane imaging used here, a single-class task would define the health/normal state and its ranges, and then calculate if a test image falls within these ranges and is therefore normal, or not normal.

A dual-class model is created by using definitions and ranges of both the diseased and non-diseased states. Further groups or subgroups can be added if they can be suitably differentiated. When three groups are defined, the classification tasks become a three-class problem, or it can be split into multiple two-class problems. This hierarchy system is particularly useful where subgroups are defined. In the classification of ioflupane images, this could mean first characterising the data as normal or non-normal, and then characterising the non-normal data into more specific disease types. A two-stage approach was used by Hamilton et al. ¹⁴³ to differentiate between different stages of Parkinson's disease and non-Parkinson's patients using ioflupane imaging.

In most of the published work using ioflupane, the classification normally takes the form of a single layer, a one-class problem. A normal range is defined for uptake measurement (e.g. VOI uptake as striatum to background ratio or putamen to background ratio) and patients whose image data falls outside these limits are given the diagnosis of Parkinsonian syndrome.

There is some evidence to suggest that DAT imaging can be used to distinguish between MSA-P and PD types ⁵⁶. A similar problem was also described ⁵⁷ as a three-class problem, defined as two two-class problems; the first stage discriminating between normal and abnormal, and the second stage discriminating between PD and MSA-P.

The distinction between IPD and atypical Parkinson's (including MSA, PSP and CBD) was shown using visual reporting ¹⁴⁴. In the case of D2 receptor imaging, the differences in disease progression between Parkinsons and Parkinsons plus have been shown to allow their differentiation. Kim et al. employed a multi-class system with good results ¹⁴⁵. A natural development of these results would be combining both pre- and post- synaptic imaging in the diagnostic pathway. This has been investigated ^{146,147} but these techniques have not crossed into clinical routine use.

5.4.2 The problem of unilateral disease

For many patients, the onset and progression of PD will be unilateral, with symptoms and imaging studies showing clear one sided patterns. This can make visual reporting easier as images that are clearly non-symmetrical are a good indicator of disease. When applied to ROI based approaches, the uptake in either hemisphere can give an abnormal classification even if the contra-lateral

hemisphere shows no signs of dopaminergic loss. This unilaterality has been allowed for in VOI/ROI based systems by classifying both side individually, and then assigning as normal (or non -PS) if both sides are normal, and abnormal (or PS) if one or more sides is abnormal. This is the approach taken in published ROI/VOI based work ^{78,148}.

In SPM based approaches the whole brain is used in a single-class classification system. The variation from a normal template is calculated as a function of the normal variance on a voxel-by-voxel basis. Voxels that are significantly different to the normal template are identified and if they form a cluster larger than a pre-determined size, then a positive diagnosis is given. In such a method, unilateral disease would identify a pathological state when a particular region showed disease. Normal uptake on the contra-lateral side would not affect the diagnosis. If a PCA approach is used, and unilateral disease has not been allowed for, the best fit will involve an amount of averaging of the two hemispheres. The normal training set would be symmetrical, and if an abnormal training set were defined, this would be a combination of one-sided and dual-sided disease and hence would also be symmetrical.

In previous work ¹ unilateral disease was accommodated for by realigning the brains based on gross striatal uptake. Image volumes were mirrored along the central line so that the hemispheres with the maximum striatal uptake were aligned. This approach replaces the assumption that dopaminergic loss is symmetrical, with assumptions about the level of non-symmetry.

Another approach would be to treat each half brain as a separate example of either disease or non-disease. Both halves of the non-PS patients would give examples of non-PS hemispheres. The PS data would have to be split since it

would only be known that at least one half of each PS brain is showing signs of disease. One approach would be to include just the side with more advanced disease. A second approach would be to have some classification that would allow one or both sides to be included providing they both show disease. To identify the more diseased hemisphere of the PS patients, a single class classification technique will have to be performed.

Other analysis techniques use either a basic decision tree, (e.g. in the case of ROI based data, the patient is classified as abnormal if either side is below a particular threshold) or the classification relies on an abnormal volume of interest being identified (as in the SPM analysis).

The whole brain analysis discussed so far has taken the whole brain as the input data, and the analysis identifies disease state by looking at the weighting factors for whole brain components. As such a diseased brain with a one sided degeneration will be given weightings that would best approximate to the activity distribution i.e. it will identify an average weighting between the diseased and non-diseased halves.

So, how can this be incorporated into the PCA / cluster analysis? As the averaging comes from the application of PCA/fitting to the weights, the data from the two sides of the brain has to be split before this stage. The two approaches are:

1. Split the training data, do a left-right mirror of one side and then input all the brain halves into the analysis,
2. Split the data and then apply PCA to both data sets independently. Test patients could then be split and fitted separately to the two sets of component weights.

The first approach has the advantage of using all the available data to define the D-Space / feature space. This should work well for the non-PS group of data as it increases the number of data sets. However, if both sides of the PS patients are used to define the PS cluster, this could be contaminated with non-diseased brain halves. The second approach allows the separate classification for the two halves of a test patient, but the training sets for the PS data will both contain a mix of diseased levels, and hence may restrict the sensitivity of the test.

5.4.3 The training dataset and feature space:

The SVD transform aims to describe the variance contained in a dataset. If the training set is formed of data from a single disease group, the variance will relate to normal variability in that group. If the training group contains a mixture of two groups of images, then the variance identified will be a combination of the inter- and intra-group variance. Assuming the intra-group variance is the same for the two groups, the inter-group variance will be maximised if the training set is a 50-50 mixture of the two groups. In such a situation the variation in the weightings between the two groups will be maximised.

The clearest distinction between the two groups would be given if the two groups were extreme cases. That is, a group of perfectly normal data and another containing only advanced stage disease. However, such a split may not provide the best classification system when patients of mild disease are presented. The best solution may in fact be to define the feature space using closely grouped extreme data, but to use data from a full spread of disease states to training the machine learning classification tool.

Such adjustments or biasing of the PCA results would represent a form of factor analysis. Taking this idea further, if the classification groups are known during the transformation stage, the components can be adjusted to maximise the differences between the two groups. This is formally done using the PLS approach.

5.5 Confounding features

A number of features have been shown to affect the uptake of ioflupane. In particular, DAT uptake in healthy volunteers has been shown to reduce with age¹⁴⁹. Their large cohort study showed a linear 6.6% per decade decrease in striatal DAT availability. The reduction was similar in both caudate and putamen. The same study described small hemispheric asymmetries, but it was shown that these were not related to left- or right-handedness.

These losses can be compared with the losses shown in Parkinson's Disease at around 8%³⁷, 5%⁵⁶ or 6-7%¹¹² per year. These rates show greater variability in the literature, which may be explained by the underlying disease progression being non-linear.

A smaller dependence has been shown on gender in healthy patients¹⁵⁰. However, this small dependence was not confirmed in a separate study investigating the uptake variations due to hormonal changes during menstrual cycle, for which no dependency was found¹⁵¹.

These changes suggest that inclusion of the subjects' age, and possibly gender, may improve disease classification.

5.6 Developing and Testing the PCA based classification technique

In this work, we use SPECT images of patients with known disease state to define a feature space using SVD; we then define a number of automatic classification techniques. The classification models used include Group Prototype, Support Vector Machines and Naïve Bayes. Other patients of known disease state are then mapped to this D-space and their diagnosis determined automatically given the classifier. To allow comparisons with other published data, four commercially available classification models are tested using the same patient database (see chapter 6).

A SVD based processing tool was written using the MATLAB toolkit. After prototyping, bug fixing and optimising the code, the system was tested under a number of configurations using the clinical data described in chapter 4.

The two datasets described in chapter 4 were used to train and test variations on the SVD and machine learning classification system. The first sets of measurements were performed using images acquired using the original clinical data format (dataset A). These results have been previously published ¹. The descriptions below include results and some text from that publication with permission from the publisher.

The second set of measurements investigates the use of SVD for the clinical data acquired using the optimised clinical settings as defined in chapters 2 and 3 (clinical dataset B).

5.6.1 Testing the SVD/PCA based system using clinical dataset A

- method

For this work an image volume created from ^{18}F -FDOPA PET data and aligned to the standard Talairach space was used as a template. All the ^{123}I -FP- β -CIT SPECT data was spatially normalised to this template using a routine contained in the SPM software package (SPM 8, Wellcome Trust Centre for Neurology, London, UK). This automatic registration works in two stages; the first is a linear affine transformation and is followed by non-linear warping using basis functions. After automatic spatial normalisation, the images were checked visually to ensure correct alignment. The following image processing and statistical analysis were performed using the MATLAB software package (Mathworks, Massachusetts, USA) using a combination of existing and in-house written functions.

A volume centred on the occipital cortex was identified and used as an estimate of nonspecific binding in counts/ voxel (C_{ns}). The volume was defined on the template image using MRIcro ¹⁰⁶ and was then converted into a binary mask. This volume was then used to scale the counts in each voxel (C_i) to give the binding potential for that voxel, BP_i (Equation 5-8).

$$BP_i = \frac{(C_i - C_{ns})}{C_{ns}}$$

Equation 5-8

Using this system the data from every patient in the training set was rescaled to binding potential in each voxel.

A second mask was defined on the template image marking the external boundary of the brain and brain stem. This second mask, also defined using

MRIcro, allowed the counts in the voxels outside the brain and the brain stem to be set to zero in the training and test patients. In this way, variations in the skin uptake and counts outside the body are removed to prevent them influencing the processing.

The problem of unilateral disease was assessed in this dataset by mirroring some of the patient data so that the highest uptake from each patient is on the same side. This was performed by using two more VOI masks drawn by eye on an MRI template. These masks define the size of the left and right striatum. The masks were used to identify the hemisphere containing the higher uptake (using total counts in that VOI). Image volumes were then mirrored left to right where needed to ensure the higher uptake was seen in the right-hand side. In total, 116 patient images were used in this analysis. See chapter 4 for a full description.

The 116 patient image volumes were used in a leave-one-out cross validation procedure. Each training run uses 115 patient image volumes to define the feature space using SVD. Prior to the SVD transformation being applied, the average image volume (PC0) across all 115 patients is calculated and then subtracted from each image volume. A number of automatic classifiers were then trained using the known disease classifications of these 115 patients. The average image volume (PC0) is then subtracted from the test patient image volume. This volume is then mapped into the feature space and the classification tools used to give a diagnosis. The process was repeated a further 115 times leaving out each patient in turn.

The number of components used in the classification step was varied between 1 and 5. The automatic classification tool measured included a group-prototype

model (SVD-GP), and two naive Bayes models (with and without the training group ratios as a prior, SVD-NB and SVD-NBP).

5.6.2 SVD/PCA based system using clinical dataset A-results

The leave-one out cross validation uses 116 separate training and testing runs to give the final results in terms of specificity and sensitivity. The intermediate steps created during this process are described for one of these runs (using 115 image volumes, and one test volume).

The feature space is defined using between one and five components – these components are in the form of image volumes that can be combined (using specific weighting factors) to reproduce (to a certain level of accuracy) any of the patient data. The first four of these image volumes are shown in Figure 5-7. The images show six central slices from these volumes.

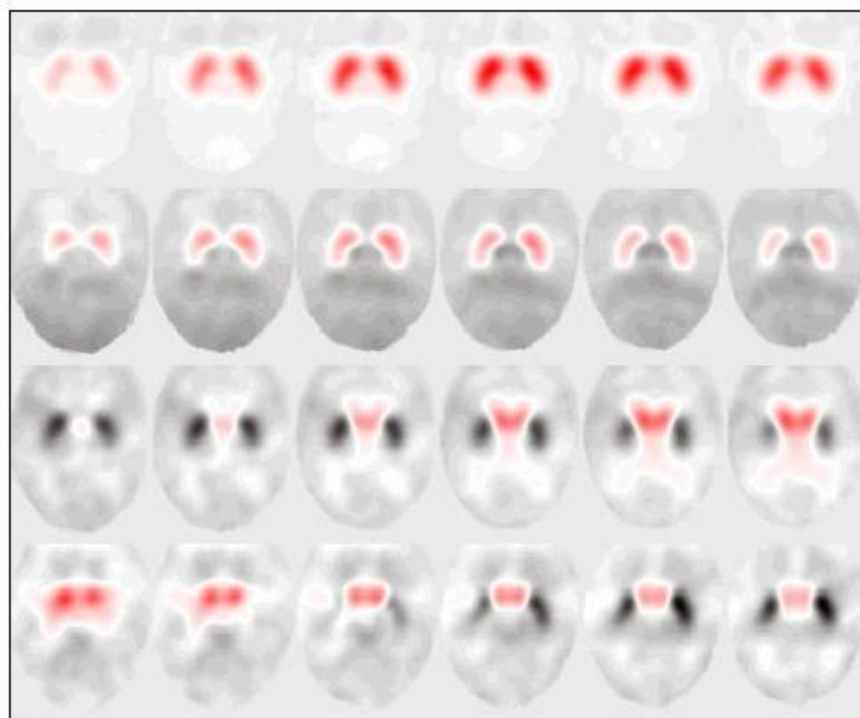


Figure 5-7 Six example slices from the first four components

The weighting factors for each of the 115 training volumes for these first five components are described as vectors in a five dimensional feature space. These vectors are plotted in Figure 5-8. For simplicity only the first three components have been shown. The 116th patient, the test patient, is also shown mapped into this feature space.

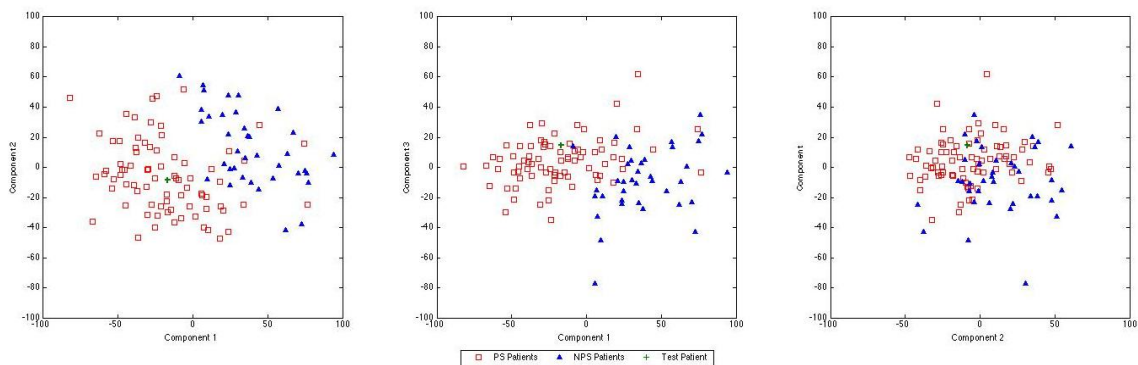


Figure 5-8 The component weights for 115 patients for the first three components. A test patient projected into this feature space is also shown. PS is Parkinsonian syndrome and NPS is non-Parkinsonian

The main result of the classification systems is a binary diagnosis. This was compared with the known diagnosis to give sensitivity, specificity and overall accuracy results.

The accuracy of each system in terms of the Area Under the Curve (AUC) as a function of the number of components is shown in Figure 5-9. The maximum accuracy is reached once the first three components have been included in the classification. The Naive Bayes and Group prototype classifications also give the relative probabilities for the two test groups, and absolute distance to GP centres respectively. These two variables were used as the input to an ROC analysis (Figure 5-10).

To visualise the two groups, the average PS and non-PS component weights for the first three components were used to scale the component volumes. The average training volume PC0 was then added to each of these image volumes. The two image volumes represent the group prototype images for the two disease states. Six central transverse slices of these image volumes are shown in Figure 5-11.

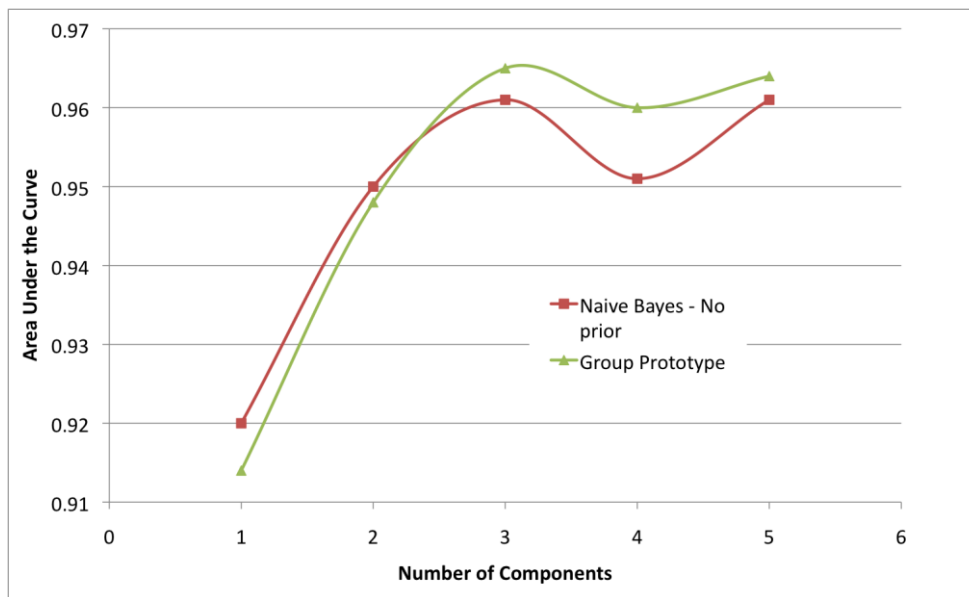


Figure 5-9 The Area Under Curve from ROC analysis SVD based classification

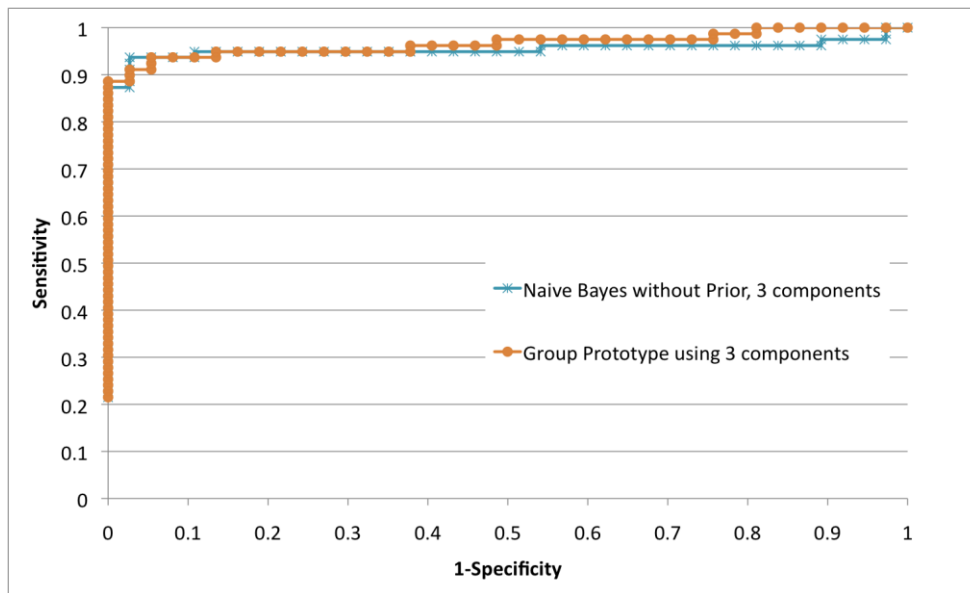


Figure 5-10 ROC analysis for Naive Bayes and Group prototype classifiers, using 3 components

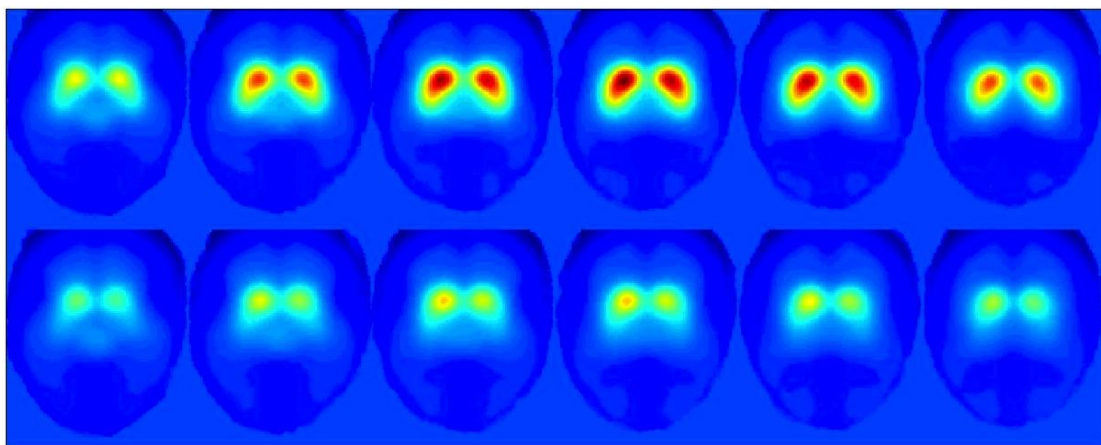


Figure 5-11 Six Transverse slices through averaged image volumes of non-PS patients (top row) and PS patients (bottom row)

5.6.3 SVD/PCA based system using clinical dataset A-discussion

The choice of feature extraction and classification model is not unique and can be optimized for the clinical situation involved. Alternative feature extraction systems have been proposed in other imaging investigations; these range in size and specificity from the raw voxel counts up to just a few VOI values. In this clinical situation, disease progression will generally follow a familiar pattern

(starting with the putamen and progressing to the caudate). Therefore, a feature extraction technique was chosen that would identify patterns in uptake across different groups of patients.

Although this technique identifies the most mathematically economical components, the first four components show obvious similarities with anatomical features (see Figure 5-7). The components relate to the differences between patients and the mean patient image. The first component shows general uptake in the caudate and putamen. The second component shows some asymmetrical changes from the mean image. This asymmetry can only be seen because some of the patient images were flipped left to right. The third and fourth components show changes in the putamen uptake. However, care should be taken when examining the components, as they are defined by the variations in the training set rather than by specific classification tasks.

These results show that a SVD based classification technique can be used as an automatic classification tool with a high level of accuracy when compared to the gold standard diagnosis. The performance was shown to be a function of both the number of components used and of the type of classification model used.

5.7 Developing and optimising the SVD/PCA based system using clinical dataset B - methods

Following the changes made to the clinical acquisition protocol (see chapter 2) and recommended reconstruction changes (chapter 3) a second clinical data set was produced (dataset B). The full breakdown of this is described in chapter 4.

This dataset was used to study the characteristics of the SVD based classification tool. As with the reconstruction optimisation work in chapter 3, the

characteristics of the classification model rely on a number of variables that must all be taken into consideration. These include:

- Features of the image quality
 - Including image filtering
- Pre-processing steps on the image data
 - Spatial registration
 - Count normalisation
 - Masking for extra-cerebellum count data
 - Non-symmetrical uptake effects – left/right alignment of maximum hemispheres/ used of half brains
- Data extraction steps:
 - The number of components used from the transformation
- Classification options
 - Classification model
 - Trade-off between sensitivity and specificity

5.7.1 Pre-processing step: spatial registration- methods

The data for the SVD analysis techniques were exported in DICOM format, and then converted to nifti format using imageJ ¹⁰⁷. The SPM spatial normalisation procedures were then used to fit the data to a template. The SPM spatial alignment method uses a version of the “count difference” method for alignment scoring. It is therefore important to use a template image of similar intensity and distribution. Spatial alignment was then performed using a ¹²³I-FP- β -CIT based template. This template originated in the IBZMtool kit ¹⁵². This template was

used to create a new template based on the local data using non-Parkinsonian subjects. The new local template could then be used for further work.

The SPM on-line help files suggest using some smoothing for both the template (target) and source images, and that the smoothing levels should be similar. By visual assessment the central range of filters (1.0-1.4 cycles/cm) showed similar smoothness to the template image. Since no other modality image data was available for these images, the only ways to assess the alignment is to use the spatial normalisation metrics, or by visual assessment. Through trial and error it was observed that most of the filtering levels within this central region gave visually acceptable results.

The non-Parkinson's patient images were first spatially aligned to the ^{123}I -FP- β -CIT template. This was performed using the data with a post reconstruction filter of 1.0 cycles/cm. Combinations of affine and non-linear warping steps were included in this transformation. The resultant transformation data was then applied to the other noise realisation for each patient, including the data to which no post-reconstruction filter had been applied. These non-filtered data were then summed together and the resulting template counts divided by 21 to give the average non-Parkinson's patient.

5.7.1.1 Pre-processing Step: Spatial Registration- Results

Central slices of this template are shown in Figure 5-12. Visual checks were then made using the MRI template to check the alignment of the local template to the standard MNI stereotactic space. Figure 5-13 shows a central transverse slice from a T1 weighted MRI template and the local DAT template.

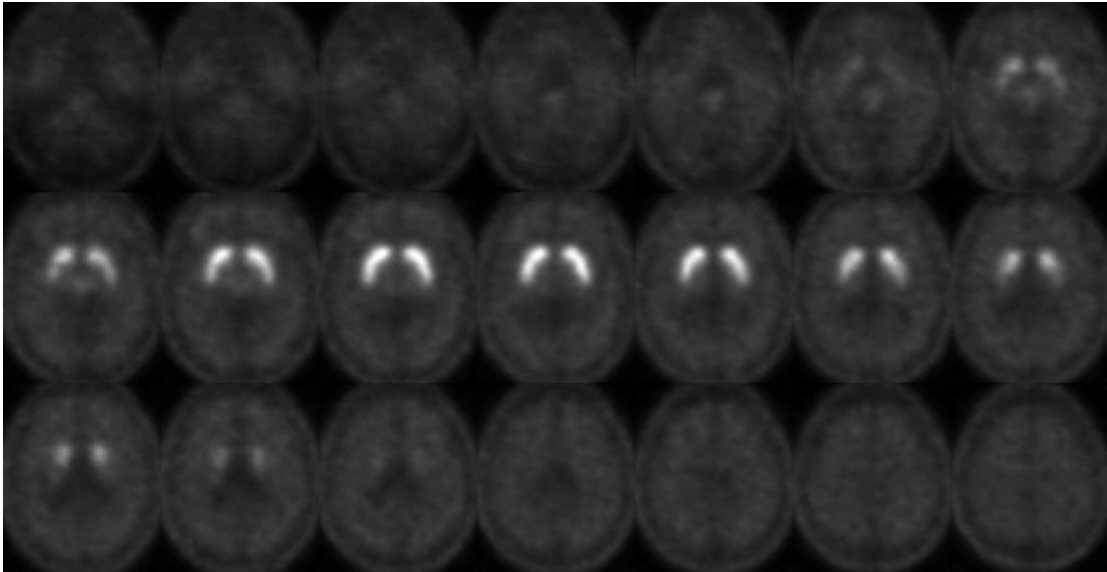


Figure 5-12 The central slices of the locally derived ^{123}I -FP- β -CIT template

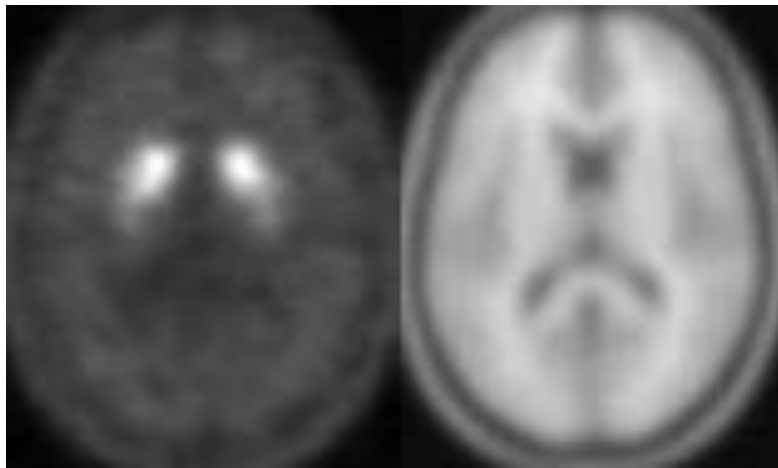


Figure 5-13 central slice from local template and T1 MR Image

This new locally defined template was used for further spatial normalisations. It was defined without any post-reconstruction filtering or filtering at the combining stage to simplify the use of any further smoothing that may be necessary.

The entire dataset was then spatially normalised to this local template using just linear transformations. No warping effects were included in this process. There is some debate over the use of non-linear warping factors in this alignment

process. In this work, the decision was made to not use warping as this was thought to mask changes in underlying shape of uptake.

During this process both source and target images were smoothed using a Gaussian filter with 8mm FWHM. This filter was applied only for the alignment process. The resulting re-aligned images had their original levels of smoothing. For each subject the non-post-filtered data was used to define the re-alignment transformation. This transformation was then applied to all other image files for that patient.

5.7.1.2 Pre-processing Step: Spatial Registration- Discussion

The use of a local template compiled with the radiopharmaceutical being investigated is generally recommended. However, visual comparisons of the spatial normalisation to an ^{18}F -F-DOPA template (as used in the study with clinical dataset-A) show that good alignment can be made with either template. A full analysis of the preferred template image was not considered. The use of different templates for spatial normalisation was investigated by Kas et al. ¹³⁶. They used four different ^{123}I -FP- β -CIT templates to spatially normalise a set of ^{123}I -FP- β -CIT data from a single site. They found no difference in analysis results when the template was changed to that of another centre. They did not investigate the use of template constructed from other radiopharmaceuticals. Colloby et al. ¹³³ followed a procedure similar to that described here except that the template used to register the ^{123}I -FP- β -CIT images was a T1 weighted MR volume. They reported good registration even with this image from a completely different image modality.

5.7.2 Number of Components Preliminary Investigation - Methods

In the previous section (5.6.2) it was shown that the accuracy of the SVD based system is a function of the number of components that are used. Dataset-B contains 78 patients, so the full deconvolution using the whole brains as input volumes will produce 78 components. Performing SVD on the full dataset of 78 patients (after subtraction of the average volume PC0), produces the three matrixes \mathbf{U} , \mathbf{S} and \mathbf{V}^T . \mathbf{S} is a diagonal matrix containing the singular values of the input matrix in size order. The values show the amount of variance in the original matrix that is described by each component.

5.7.2.1 Number of Components Preliminary Investigation - Results

The variance described by each component (as a fraction of the total variance) is plotted as a function of component number in Figure 5-14. The three curves have been created for the full dataset of 78 patients volumes following spatial normalisation, using the whole image volume (i.e. using both hemispheres without any left-right mirroring). The first curve uses just spatially aligned data. The second is created after the data has been rescaled to give binding potential. The third is from images where the data has been masked to remove extra-cerebral activity. The graph is truncated to show just the results from the first 20 components. All three curves follow a very smooth gradual decline in fractional variance all the way to 77 components. The 77th component weighting is 0.0035, 0.0040 and 0.0035 for curves 1, 2 and 3 respectively. The 78th component $<1 \times 10^{-9}$ for all three slopes.

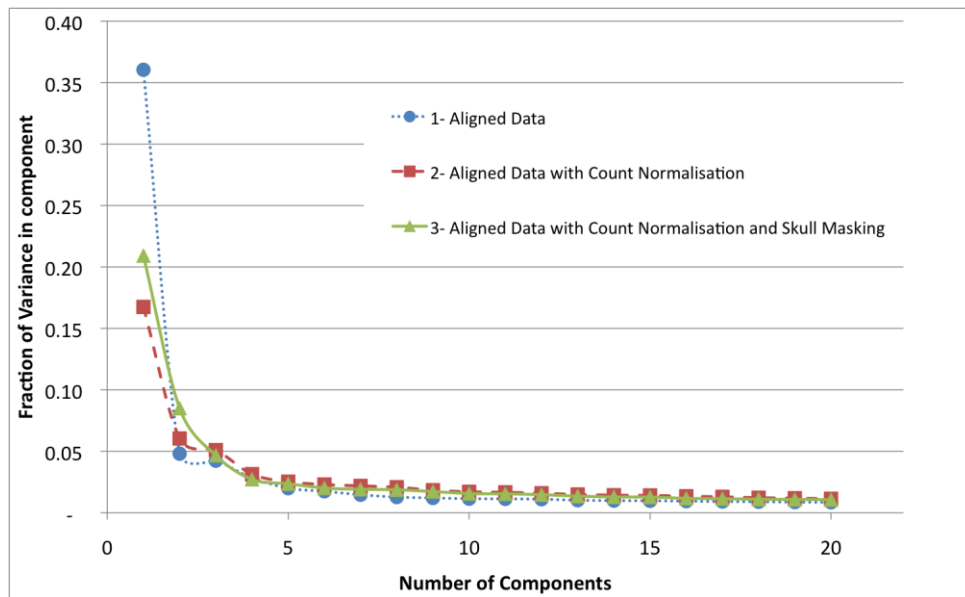


Figure 5-14 fraction of total variance described in each component for three types of image pre-processing

5.7.2.2 Number of Components Preliminary Investigation - Discussion

When assessing the variance curves the common approach is to identify a ‘shoulder’ in the data. The first two curves appear to show an abrupt change at 2 components, whereas the third shows a very gradual change with the shoulder appearing at between 5 and 10 components.

The curves only show the fraction of the total variation that is reproduced by each component, they do not show if this variance useful for the diagnosis. The first curve shows data that has not been count normalised; the variance in the components will therefore have to include this count scaling information, and since this is likely to be an image-wide scaling factor, the first component will be strongly influenced by this information.

These curves suggest that for the skull-stripped normalised data at least 5 components may be necessary. The classification procedures were repeated

with between 1 and 10 components for most of the classification tests. For one image quality (the non-filtered data) the number of components was extended to cover the first 20 components.

5.7.3 Pre-processing steps: Count normalisation and extra-cerebral data -methods

There are many variables that can affect the number of counts within the images that are not related to the underlying disease status. The data was assessed with and without scaling to binding potential (following Equation 5-8) using the same VOI for non-specific binding that was used in section 5.6.

This rescaled data was assessed with and without the masking of the extra-cerebral count data. This was performed using the same volume mask as 5.6.

5.7.4 Pre-processing steps: Count normalisation and extra-cerebral data -results

The classification accuracy for the naïve Bayes classifier using different levels of pre-processing is shown in Figure 5-15 as a function of the number of components. This data uses the whole brain data that has not had any post-reconstruction filtering. The accuracy is quoted for the naïve Bayes classifier with no prior and uses data with no left-right flipping of hemispheres.

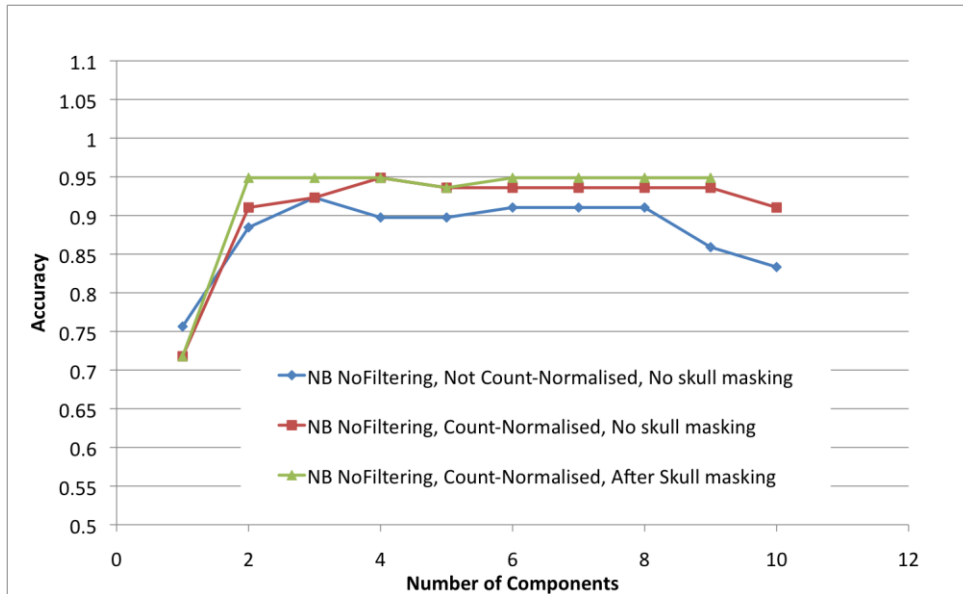


Figure 5-15 NB classifier accuracy for different levels of pre-processing

This data shows that the pre-processing steps of count normalisation and skull masking can improve the accuracy of this classification system. Similar results are seen for the SVM and GP classifiers. The accuracy level is 95% and is constant with the number of components for this system (except when only the first component is used). These results are broadly similar to the previous results (section 5.6). All further calculations were performed using the count normalised and skull masked data.

5.7.5 Image Quality Effects - Methods

In chapter 3, the image quality produced by different reconstructions was described. It was postulated that the level of smoothing used in the reconstruction would influence the accuracy of any classification model.

The clinical data was reconstructed using the depth dependent resolution recovery model, using a blurring model width of x2-z1 with 12 iterations of 8 subsets. The resulting images were saved with varying levels of post-reconstruction filtering. All were Butterworth filters of order 10, with a range of

cut-off frequencies (0.5, 0.7, 0.9, 1.0, 1.2, 1.4, 1.6, 1.8, and 2.0 cycles/cm). A non-filtered image set was also saved giving a total of 10 different levels of image noise. The data was pre-processed by spatial registration, count normalisation and skull masking.

5.7.6 Image Quality Effects - Results

The classification accuracy for the naïve Bayes classifier is shown in Figure 5-16 as a function of image filtering. The results for 1 to 10 components were included. This data uses the whole brain data that has been count normalised and been skull masked, with no left-right flipping of hemispheres.

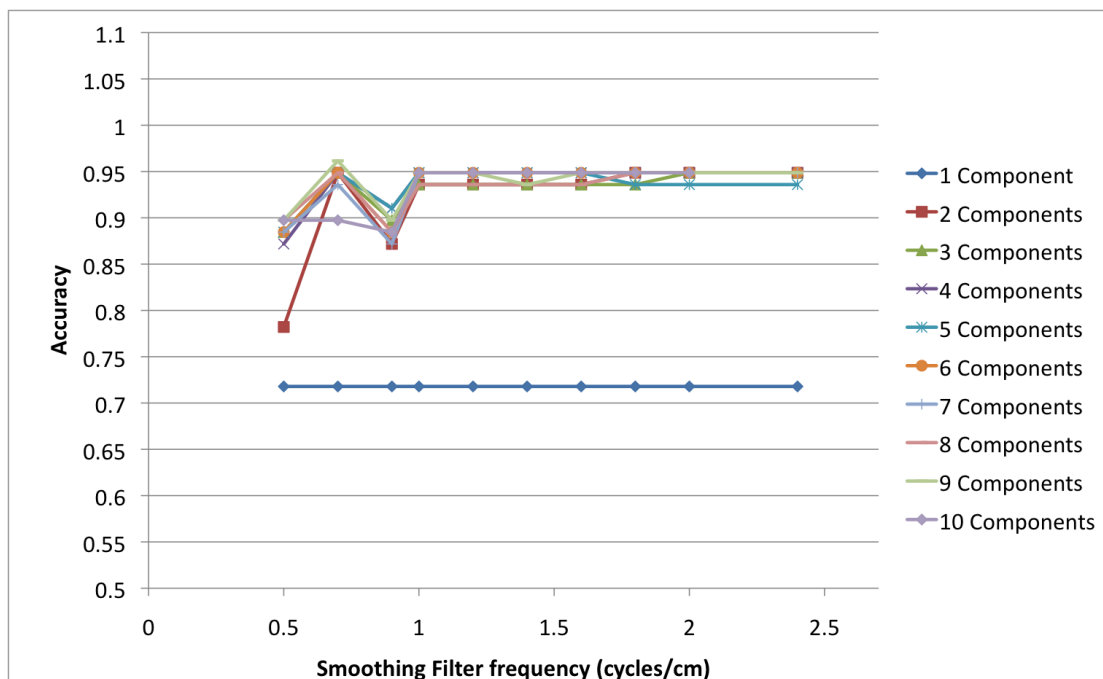


Figure 5-16 Naive Bayes classifier accuracy as a function of image filtering levels

The accuracy is seen to be highly variable for the data that has been smoothed the most. Filters with frequencies lower than 1 cycle/cm give the poorest results. The results calculated using just the first component are very poor for all levels of image smoothing. Using data with no post reconstruction filtering will give good accuracy results for this classifier.

5.7.7 Image Quality Effects – Discussion

The apparent non-dependence on image filtering is perhaps surprising when considered against the published work on SPM analysis. However, the approach here does not rely on the GLM model to give a statistical basis to the analysis. Looking back to the reconstruction work in chapter 3, the data considered here have been produced using the OSEM method including resolution modelling. This was shown in chapter 3 to impart a certain level of image smoothing in the reconstruction process. A further smoothing factor is the poor resolution of the SPECT imaging system itself. The typical resolution ranges for these images will be in the region of 6-10mm FWHM. Between them all these factors will provide a level of image smoothing before any post reconstruction filtering is applied.

The following tests will all be performed using data that has not had a post reconstruction filter applied.

5.7.8 Testing the Feature Classification Model - Methods

Four types of classification models were tested for the various combinations of data types, pre-processing steps, and feature extraction results described above.

- The group prototype method uses the distance from the test patient to the mean vectors for the two disease groups and as described in section 5.2.1. The classification given by the group with the smallest distance.
- For the single class method using individual hemispheres the Mahalanobis distance was used.
- A support vector machine was tested using a linear decision plane.
- Two forms of the Naïve Bayes model were used, with and without the prior disease split to weight the classifications (see section 5.2.3). The

classification was given by the group with the highest probability. The data was further investigated by using the relative probabilities as the input to an ROC analysis – thus allowing a maximisation of the accuracy by adjusting the sensitivity and specificity.

Due to the way these classification models use the feature space vectors to give the classification, their dependence on the number of components may be variable and so all must be recorded for a range of numbers of components. The image data used in this section has had no post-reconstruction filtering. No allowance for unilateral disease was made at this point.

5.7.9 Testing the Feature Classification Model - Results

The total accuracy for each of the classification methods is shown in Figure 5-17 as a function of the number of components. The three classification models all give good diagnostic accuracy for 2 or more components. The NB classifier gives consistent accuracy of 96% for 2 to 11 components. The group prototype results are not as good, mainly being at 91%. The SVM classifier gives the best performance (97%) but it needs a high number of components to give this accuracy. Of note there is a small improvement in both the SVM and GP classifiers when the 13th component is included.

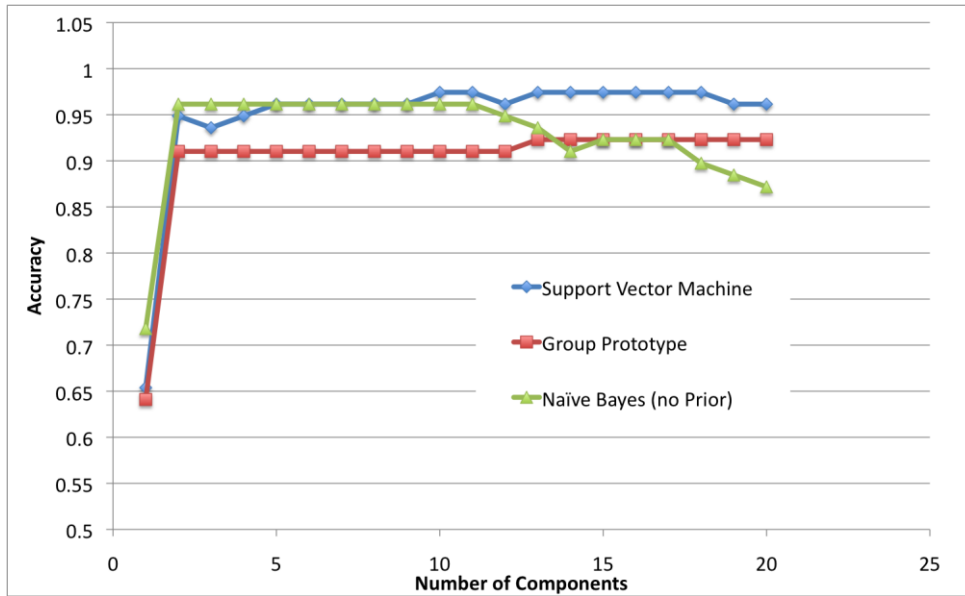


Figure 5-17 Accuracy as a function of number of components

5.7.10 Testing of the Feature Classification Model - Discussion

The way in which these classifiers form decisions planes within the feature space is quite different. The Group prototype uses the ‘distances’ in feature space for the test subject vector to the average positions in the two classification groups. Because of the way the feature space is calculated, the components which contribute more of the variance in the training set will be weighted more heavily than the lower order components. Assuming the required data is held in the first few components, it would be expected that the accuracy of this classification system would peak at low component numbers and then not change much as further components are added. This expected performance is due to the scaling factors in the feature space dimension related to \mathbf{S} . For the data discussed here, and taking the scree plots in section 5.7.2.1 into consideration, the first few components do contain a lot of the information, but the scree slope also shows

that most of the other components (above 5) have almost the same contribution to S and hence the group prototype can still make small changes in its accuracy as the higher order components are added.

The group prototype takes the entire dataset into consideration as it creates the decision boundary. However it does not take into consideration the spread of data within the training groups.

The Naïve Bayes classifier takes the spread of each group into consideration. Like the GP system the NB calculation uses all of the data in the training data to define the decision boundary. An increase in spread of the data for training group along a specific components will have the effect of reducing the importance of a variation along that component for a test patient. The NB system therefore weights the importance of each component by the spread of scores for that components within the training groups.

Because the GP and NB training system use the entire training sets they are less susceptible to problems arising from individual training subjects especially when the training group is large. An incorrectly classified data point within these two datasets would be unlikely to have a significant effect on the classifiers performance.

The SVM classifier is very different from the GP and NB systems as it uses a small subset of the data to form the decision plane. The variation of component weightings within a training group has little impact on the decision boundary. Because the classifier uses just the subset of training points near the decision boundary this system is either sensitive or completely insensitive to misplacement data point depending on its proximity to the boundary. A

mislabelled data point is likely to have an impact on the decision plane, as it will be selected as one of the support vectors.

The emphasis placed on specific components by SVM depends on the usefulness for classification. Taking the weightings for the support vectors alone in the two training groups, if there is no correlation between the component and the training group then the boundary will not be dependent on that specific component. i.e. the boundary will run normal to that component axis. If such a correlation occurs by chance for a component then the resulting decision plane may perform badly. It is these random correlations skewing the SVM decision plane that will lead to the phenomenon known as over fitting. With any noisy training set such correlations will occur and hence the ideal number of components is likely to be considerably less than the full range of components available for training.

5.7.11 Non-symmetrical uptake effects -methods

The simplest approach uses the whole brains for each patient as an input function. This was used for the baseline calculations.

A second approach is to split each brain into two hemispheres. The left hand hemispheres were mirrored left to right. The training set was then composed of 156 hemispheres. Of these 44 represented Non-PS hemispheres and 112 PS hemispheres. The classification models were then trained using 154 hemispheres, with the other 2 belonging to the test patient. Both of the test hemispheres were then classified and if both hemispheres were classed as normal than the test patient was classed as NPS. If one or both hemispheres were classed as PS then the test patient was classed as PS.

A third approach used all the hemispheres to define the feature space. The NPS hemispheres were then used to define a normal group. This normal group would then include either 42 or 44 hemispheres depending on the classification of the test patient. Test hemispheres were then classified using the Mahalanobis distance from the NPS group. The test hemisphere with the larger distance is assumed to be the least normal and was hence used for the classification. This single class problem was used to test if the presence of unaffected hemispheres in the PS group would affect the results.

A fourth approach was used which re-aligned the brains so that the striata with the highest uptake was on the same side. This was performed using the same striatum VOI as used in section 5.6.

5.7.12 Non-symmetrical uptake effects -results

The results for the data classified using just the NPS half brains was in the form of Mahalanobis distances. These were used as an input into a ROC analysis to give accuracy figures. The other three systems used the Naive Bayes classifier. The input training data was from the Non filtered images and calculations were performed for between 1 and 20 components. The accuracy of each system is shown in Figure 5-18 as a function of number of components. Using this NB classifier the base line figures are better than the various non-symmetrical classification approaches. The other classification tools were also assessed with the non-symmetrical training sets. The accuracies for the NB, GP and SVM classifiers tested with Maximum-hemisphere alignment system are shown in Figure 5-19. A broad maximum accuracy is seen for the SVM classifier.

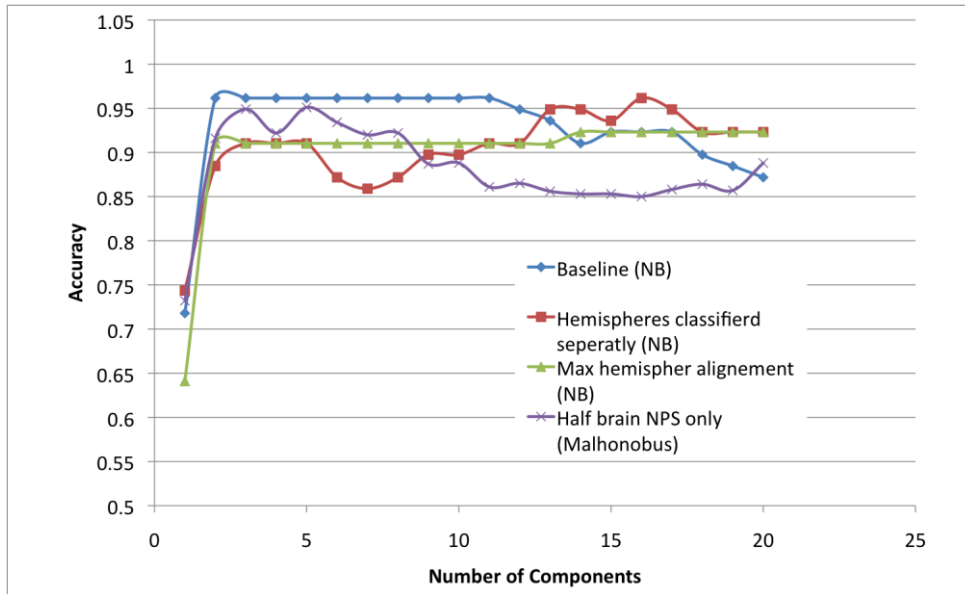


Figure 5-18 Naive Bayes analysis using Non-Symmetrical Classification

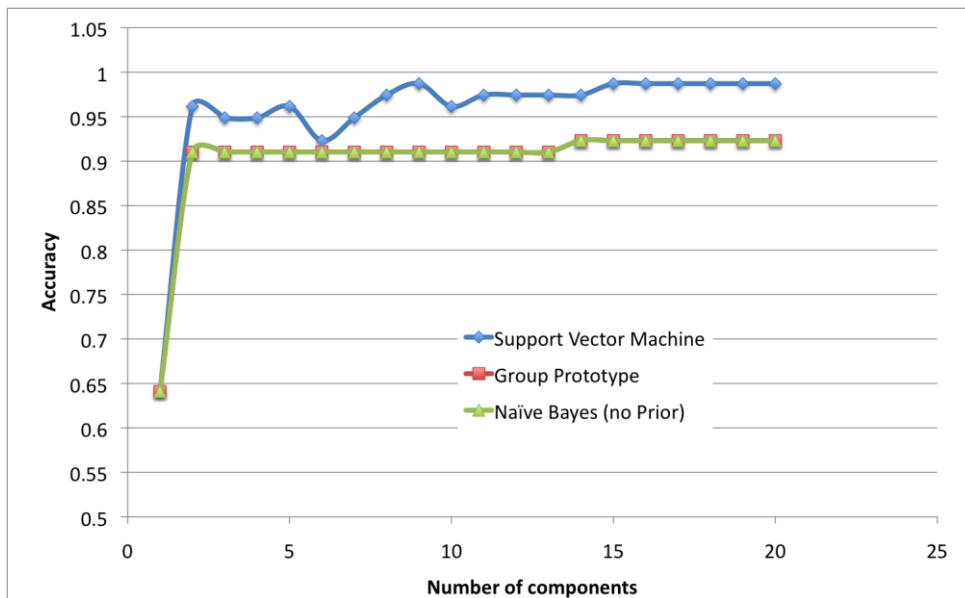


Figure 5-19 Maximum-Hemisphere aligned data classified using SVM, NB and GP systems

5.8 Overview of the SVD and machine learning systems

The best accuracy score is given by the measurements that have come from the SVM classifier. The SVM classifier gives an accuracy score of 99% using the training data in which the hemispheres with the maximum uptake are aligned

and training with between 15 and 20 components. This relates to just one patient being misdiagnosed. The broad maximum suggests a real classification result as opposed to noise in the classifier that can come about through random changes.

The 'maximum aligned hemisphere' method of allowing for unilateral disease is shown to work better than the various systems analysing half brains individually. This is perhaps surprising as this system makes certain assumptions about the disease progressions. The implied assumption is that for similar levels of asymmetry in the PS group, or that any differences in the levels of asymmetry can be accounted for within the components.

The use of confounding features was not investigated here. It was felt that the improvements that might be possible when including age dependencies would not have a great impact on what is already an almost perfect classifier. The use of these confounding features is examined further in the next chapter.

The preferred method that has been shown to give the best results consists of the following steps:

- A training dataset is acquired containing both PS and NPS subjects
- The scan data is reconstructed and minimal filtering applied
- The data is spatially registered to a template
- The full training set is count normalised using a background region as an estimate of non-specific binding
- Counts originating from outside the brain are masked out using a template
- The average counts distribution is subtracted from the dataset
- Singular Values Decomposition is used to define a feature space in 17 dimensions

- The vectors for the training data within the feature space as used to define an SVM classifier
- Data from test patients are reconstructed using the same settings
- They are then spatially normalised to the template
- The counts are normalised using the non-specific binding
- The extra-cerebral counts are masked out
- The mean training image volume is subtracted
- The test patient is mapped into the training feature space and the classification given

Although there are many steps in this procedure, only the image reconstruction and spatial alignment need any user intervention. All the other stages can be automated as they were in these investigations. Once the data is reconstructed and aligned to a template the entire process is automated and can take a matter of seconds for the entire training and classification procedure.

6 Commercially Available Methods for Analysing DAT SPECT imaging

SPECT imaging

A number of VOI based methods have been suggested for quantification in dopamine receptor and transporter imaging. These methods can be split into two subgroups depending on whether the placement of the VOIs is manual or automatic. Three of the techniques using manually based VOIs have been combined into a software package (QuantiSPECT, Mirada solutions) ¹⁵³⁻¹⁵⁵. These will be described before moving on to describe three automatic systems, BRASS (Hermes medical, Stockholm) ¹⁴⁸, Basgan (University of Genoa) and IBZMtool (University of Hamburg) ¹⁵².

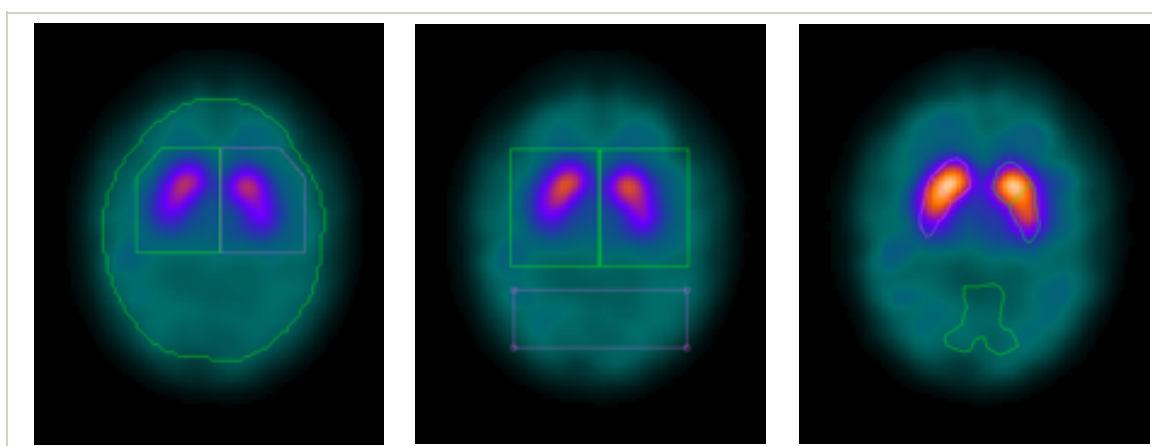


Figure 6-1 The three VOIs definition used in the QuantiSPECT software

6.1 Manually placed ROI/VOI techniques

A number of techniques have been developed that involve the manual placement of ROIs and VOIs as a measure of striatal uptake. Three of these techniques are combined in software called QuantiSPECT. Figure 6-1 shows the ROI placements from left to right, the two-box, three-box and crescent methods.

6.1.1 The 2-box Method

The 2-box method ¹⁵³ (sometimes known as the Southampton method) was introduced as a general technique for overcoming the partial volume effect by using a VOI large enough to encompass any blurring. When applied to ¹²³I-FP- β -CIT imaging, 13 transverse slices (fixed thickness 44mm) are summed together and the rest of the calculation performed in 2D. The striatal VOIs are of fixed size and are positioned manually. Their size (61*48mm) was chosen to encompass all counts from the striatum, allowing for inter-patient variability and blurring of counts. This should reduce the dependency on equipment and processing choices. The background is defined automatically using a threshold and typically encompasses most of the cortex. A threshold of 50% of non-specific binding maximum counts is used to define the boundary on heavily smoothed images. This boundary is then brought inwards by 20mm to remove any partial volume effects near the exterior of the brain. The same slices are used for both the striatum and background regions.

In retrospective studies it was found to give good correlation with clinical diagnosis ^{110,156}. The reported sensitivity and specificities varied between these studies. With processing and analysis techniques the same, the differences were put down to the referral base. Comparison with visual reporting has also been reported ⁷⁸.

Rather than the usual specific uptake ratio, they define a metric known as the Specific Uptake Size Index (SUSI). This is the specific uptake (C_s) divided by the concentration per unit volume in the background region (mean background region counts \overline{C}_{NS} divided by volume of background region V_s) (Equation 6-1).

$$SUSI = \left(\frac{C_S}{\bar{C}_{NS}} \right) \left(\frac{1}{V_S} \right)$$

Equation 6-1

SUSI has the units of volume. It is the volume of non-specific binding that contains the same activity as that in the specific binding region. SUSI can then further be converted to a specific binding index (SBI) by dividing by the volume of a standard striatum (V_{SS}) (Equation 6-2).

$$SBI = \left(\frac{C_S}{\bar{C}_{NS}} \right) \left(\frac{1}{V_S} \right) \left(\frac{1}{V_{SS}} \right)$$

Equation 6-2

The value for V_{SS} is for a standard striatal volume as the true volume of the patients' striatum is generally unknown. The figure chosen for this is based on the volumes of the striatum in their anthropomorphic phantom. It acts only as a scaling factor. Their simulation studies used a range of imaging resolutions to model a known phantom distribution. The resulting SUSI figures were shown to be independent of imaging resolution and VOI size for all but the smallest VOIs they applied. This large VOI approach has the advantage of consistency of measurements across different resolution/noise levels. However, the measure is only of gross DAT receptor numbers across the whole striatum and is unlikely to discern small changes in DAT distribution.

6.1.2 The 3-box Method

Like the 2box method, this method^{154,157,158} uses a "slab" of summed transverse slices to allow the calculation to be run in 2D. However, the thickness of this slab has to be chosen by the operator. Fixed size and shape ROIs are then positioned over the basal ganglia. The background region is of variable size and is positioned manually. Like the 2-Box method, the ROI are large compared to the

striata, to minimise the affects of image blurring, inter-subject variability and image quality parameters. The result metrics are the Total Binding Potential Index (TBPI) for each striatum, and an asymmetry index. The TBPI is calculated from the total counts in the striatum (C_S), the volume of the striatum V_S and the average counts in the background region/ non-specific binding region (\bar{C}_{NS})

Equation 6-3.

$$TBPI = 100 \frac{C_S - \bar{C}_{NS}V_S}{\bar{C}_{NS}V_S}$$

Equation 6-3

6.1.3 The Crescent Method

This model ¹⁵⁵ uses a fixed thickness slab of 12mm centred on the slice with the highest counts. Unlike the two previous methods, this uses relatively small, and anatomically shaped ROIs placed over the striatum and occipital cortex. The regions can be moved and rotated, but not resized. A further difference from the 2-and 3- box methods is the splitting of the striatum into caudate and putamen regions.

6.2 Automatically Placed ROI/VOI Techniques

6.2.1 BRASS

The Brain Registration and Analysis Software Suite (BRASS, Hermes Medical, Stockholm) is a selection of tools that can be used to automatically assess uptake in the basal ganglia. The software includes a template made of the scans of 20 healthy volunteers that have been spatially registered. A series of anatomically shaped VOIs have been defined based on this template. An implementation of

the mutual information algorithm is used to spatially normalise test patients to this template, which then allows the VOIs to be applied. The uptake is measured for left and right caudate and putamen relative to a posterior region, positioned over the occipital cortex. Figure 6-2 shows the positioning of these regions on one of the test patients. The figure shows the position of seven regions, four relating the striatum, and three non-specific binding regions. The large posterior region shown in dark blue is used to calculate the uptake ratios. The two other background regions (the light blue anterior region and green posterior-inferior region) are not used.

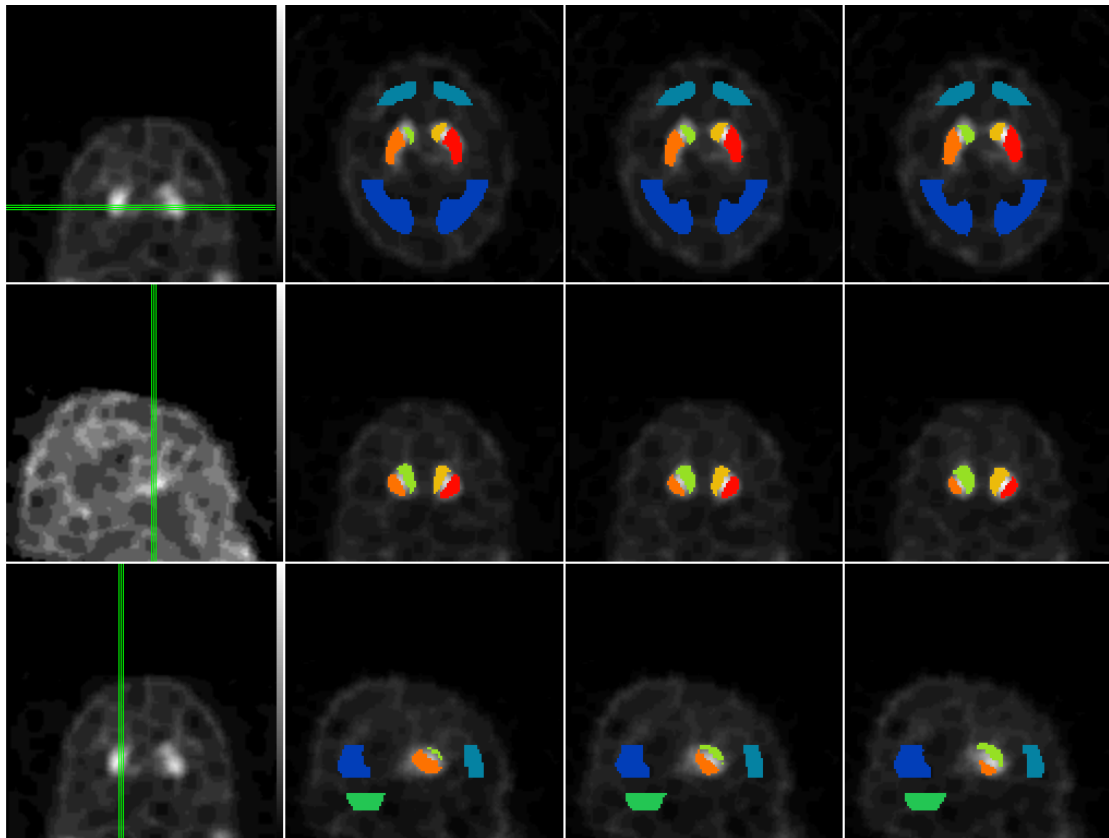


Figure 6-2 Positioning and size of the VOIs used in the BRASS analysis package

The software can be used with a choice of two templates, one produced from FBP images; the other from OSEM based images. The software has been tested in comparison with visual reporting ¹⁴⁸ and compared for images reconstructed with FBP or OSEM ¹¹⁷.

Another study ¹⁴⁴ showed that for a specific classification task (PS vs. atypical PS), visual assessment could be better than semi-quantitative methods. The given explanation was that the VOI based models are unable to identify changes in the patterns of uptake.

6.2.2 BasGan

The basal ganglia matching tools package (BasGan) ¹⁰² was developed by the University of Genoa. They used a standard anatomical atlas ¹⁵⁹ on which they defined volumes of interest for the caudate and putamen manually. These defined VOIs were then convolved with a smoothing filter (a Gaussian function with 10mm FWHM) to produce a template. Test patients are then automatically aligned to this template. The main difference between this and the BRASS package is that once aligned, a partial volume effect correction is applied to the images, before caudate and putamen uptakes are given.

6.2.3 IBZM Tool

The IBZM tool ¹⁵² was originally designed for D2 imaging, but has been extended to cover DAT based investigations. The software is in the form of an add-on set of functions to the popular SPM software (see section 6.5). The overall methodology is similar to BRASS, in that it allows spatial normalisation to a template followed by VOI based measures of uptake. The entire brain (minus the striata) is used as a background region, and uptake values given for left and right striata.

6.2.4 Comparison of VOI Techniques

A comparison of some of these techniques (QuantiSPECT and BRASS) was published for a set of 22 patients (11 PS and 11 ET) ⁷⁸. Long term follow up for these patients was not available, so a visual report of the images was used to define the two groups. The report focussed on the inter- and intra- operator variability for the techniques as well as reporting the concordance with visual reporting.

6.3 Size and Shape Analysis

The property of interest may be the size or shape of an organ. This is often used for classification of CT images (e.g. RECIST criteria ¹⁶⁰) where the size of lesion is measured along its longest axis. Other applications have used the volume and/or ratios of lengths. Staff et al. ¹⁶¹ applied automatic techniques to delineate the ROI describing the striatal activity. They used the ratio of the length and width of this shape as a measure of disease progression. Using the data from 52 scans (including 20 controls and 5 non-PS patients), and long term follow up as their gold standard, ROC analysis showed similar performance to the putamen-background ratio (from the crescent ROI method in QuantiSPECT) and visual analysis.

6.4 Fractal Analysis

The heterogeneity of a ROI (e.g. the variance of the counts divided by the mean counts for a ROI) depends on the ROI size, the object being imaged, and the imaging process (e.g. resolution effects, non-uniformities in the scanner). Fractal models measure how image heterogeneity changes with ROI size through a

metric called Dispersion. Kuikka and co-workers ¹⁶² used these processes to investigate 2D transverse slices from DAT imaging studies. The phantom data matched known values. One draw back of this approach is that an estimate of the system heterogeneity must be made, either by statistical approaches, or by dual isotope scanning, and the results are sensitive to this estimate. This technique has yet to be applied as a diagnostic classification method.

6.5 Voxel-by-voxel Analysis – Statistical Parametric Mapping

Multivariate approaches have been used in some medical imaging problems. These are typically used to show similarities and differences between groups of data but have also been applied to classification.

Statistical Parametric Mapping (SPM) ¹⁶³ works on a voxel level to define the mean and standard deviation for one or more groups of images. These standard deviations are then used to identify voxels that fall outside a particular probability level. Since a brain image will typically be made of around 5×10^5 voxels, special precautions must be made in relation to the problem of multiple comparisons. If all the voxels in the image were uncorrelated, a Bonferroni or similar correction would be appropriate. The approach developed by Friston ¹⁶³ involves some smoothing of the image data thus ensuring a level of correlation between neighbouring voxels. The general linear model (GLM) can then be used to identify statistically relevant differences between groups by identifying clusters of significant voxels. The statistical power of the GLM when two groups are defined means SPM is generally used as a method for identifying differences between groups of patients, rather than classification.

Scherfler et al. ⁵⁷ investigated differences in ¹²³I-β-CIT uptake in patients with IPD and MSA-P. They found that both had reduced uptake in the striata when compared with healthy controls. The MSA-P patients were found to have statistically significant reductions in the midbrain compared to IPD. In similar work with ¹²³I-Beta-CIT, Nocker et al ⁵⁶ found significantly lower uptake in the dorsal brainstem of MSA-P. On follow up scans the MSA-P patients were found to have more rapid decline in caudate and putamen uptake compared to PD.

Cuberas-Borros et al. ¹³² used SPM to confirm the appropriateness of a VOI based technique. They found only the striatum contained significant differences between IPD and controls/drug induced PD.

In a PET based study, Wang et al. ¹⁶⁴ used ¹⁸F-FP-β-CIT and found significant decreases in striatum, and in particular posterior putamen areas. Their VOI measurements showed significant negative correlations between caudate, putamen and posterior putamen regions to UPDRS motor scores.

Colloby ¹³³ and colleagues used SPM to investigate differences between groups of patients with DLB, Alzheimer's, IPD and healthy controls imaged with ¹²³I-FP-β-CIT. They found significant differences between AD and DLB, and between IDP and controls. No significant differences were seen between DLB and IPD. They used SPM as a classification tool by quantifying differences between a test patient, and a template formed from scans of healthy volunteers.

Others have reported on regional cerebral blood flow differences in PS ⁶¹ and DLB ⁴³ and glucose metabolism in MSA ⁵⁸.

6.6 The BRASS package

6.6.1 Testing Methods

The two clinical datasets described in chapter 4 were used to test the BRASS system for DATSCAN quantification. The images for the 116 patients in dataset-A were reconstructed with the fixed settings described in chapter 4. The data for the 78 patients in dataset-B were reconstructed with a range of filter settings. The data were then imported into the BRASS software. There are two ^{123}I -FP- β -CIT templates in this software one for FBP data and one for OSEM data. The latter was used for this analysis. The software runs using predefined settings, which cannot be altered. Each image was analysed individually, using its own spatial registration, with the uptake figures being calculated after this registration process. The program does not allow uptake calculations without this registration, or using the registration parameters of another image. A visual check was performed for each image at the time of testing to confirm correct alignment.

The uptake figures (for the left and right caudate and putamen) were recorded for each filter setting. The uptake ratios were analysed to give the minimum uptake for each subject for both caudate and putamen. This data was transferred into a statistics program (SPSS statistics, v19.0, IBM, USA) for further analysis. The uptake score for the two groups of patients were analysed using an unpaired t-test with a significance level of 0.05.

The two uptake ratios were used as predictor variables in ROC analysis using the known classification as the outcome. The Area Under the Curve (AUC) was calculated as a function of post-reconstruction filter.

6.6.2 Testing the Brass Package - Results

BRASS Automatic alignment results

For the 115 patients image volumes in dataset-A, no misalignment errors were noted. Of the 78 patient volumes in dataset-B the number of patients for which the spatial alignment was deemed to be unsuccessful (based on visual assessment) was seen to be variable and possibly related to the filter settings. The whole brain volume as produced by the OSEM reconstruction package was used as the input for the BRASS program. Each image volume was analysed individually. The spatial normalisation step could therefore produce different transformations for the various image volumes for a single patient. It was noticed during this process that a number of different spatial alignments were produced.

For some of the alignments produced by the BRASS software the resulting positions of the uptake VOIs were obviously incorrect. A summary is shown in Table 6-1. There appears to be general trend in that the smoothest images tend to have more misalignment errors. One patient was badly aligned for all filter settings. The alignment for this patient is shown in Figure 6-3. A number of attempts to aid this realignment were attempted, including manual alignment prior to processing and cropping the image boundaries. These were not successful. Since the background region is outside the brain, the counts in that region are too low, and hence the striatal binding measures were very high. The uptake scores for all the miss-aligned data were included as “not-a-number” i.e. incorrectly classified, so that these results could be compared with those from

other classification models, but would not skew the means and standard deviations.

Table 6-1 Summary of alignment errors for BRASS program

	Post reconstruction filter (cycles/cm)									
	0.5	0.7	0.9	1.0	1.2	1.4	1.6	1.8	2.0	NF
Correctly aligned	75	76	77	76	77	77	77	77	77	75
Alignment error	3	2	1	2	1	1	1	1	1	2

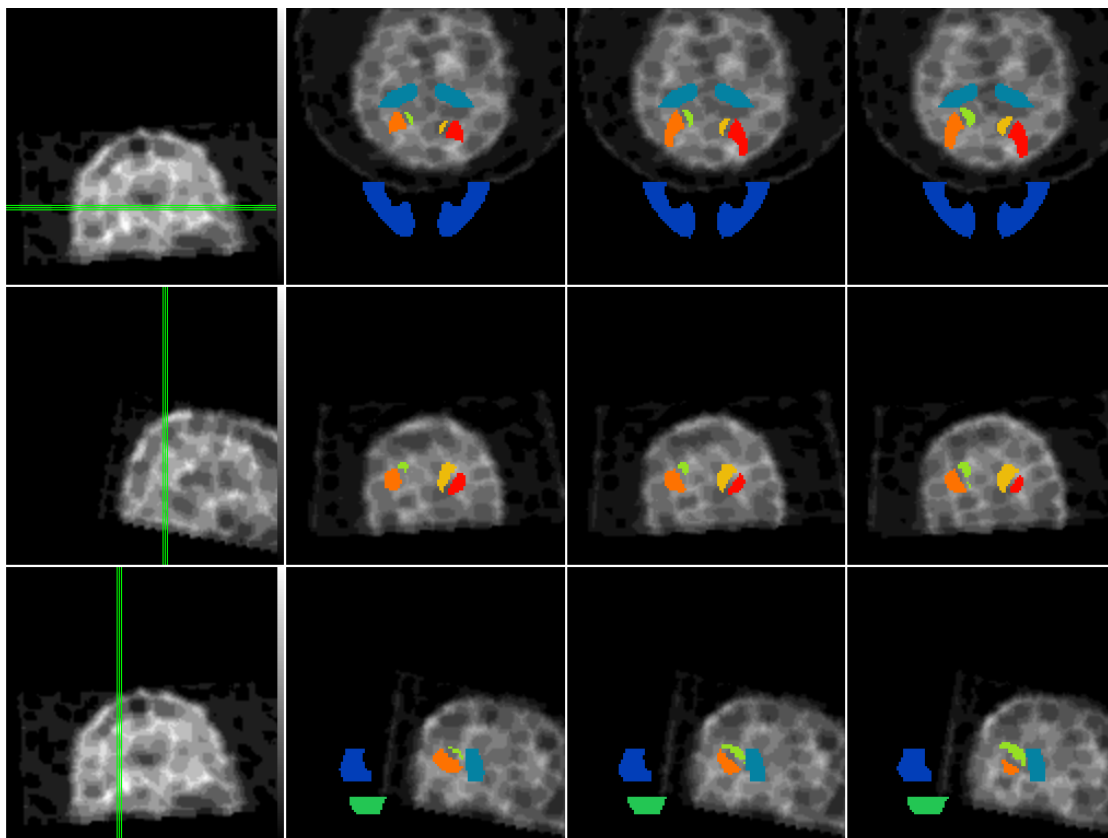


Figure 6-3 Example misregistration of scan using the BRASS software

BRASS Uptake Measurement - Results

The average caudate and putamen uptake ratios for the two patient groups are shown in Figure 6-4 and Figure 6-5 respectively. The error bars displayed show +/- one standard deviation. The non-filtered data is inserted using a dummy filter setting of 2.4 cycles/cm to allow it to be visualised.

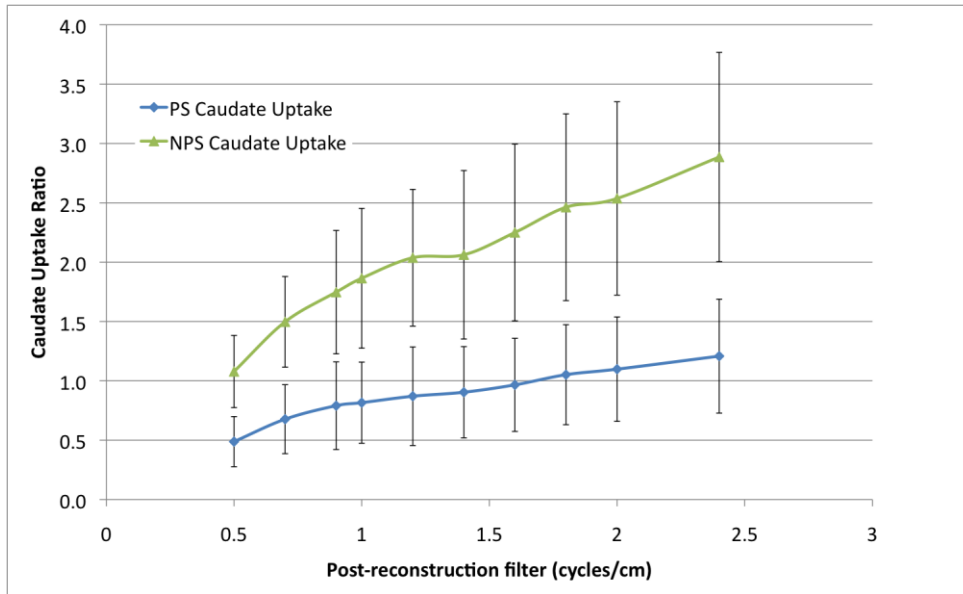


Figure 6-4 BRASS-Caudate uptake as a function of post reconstruction filter

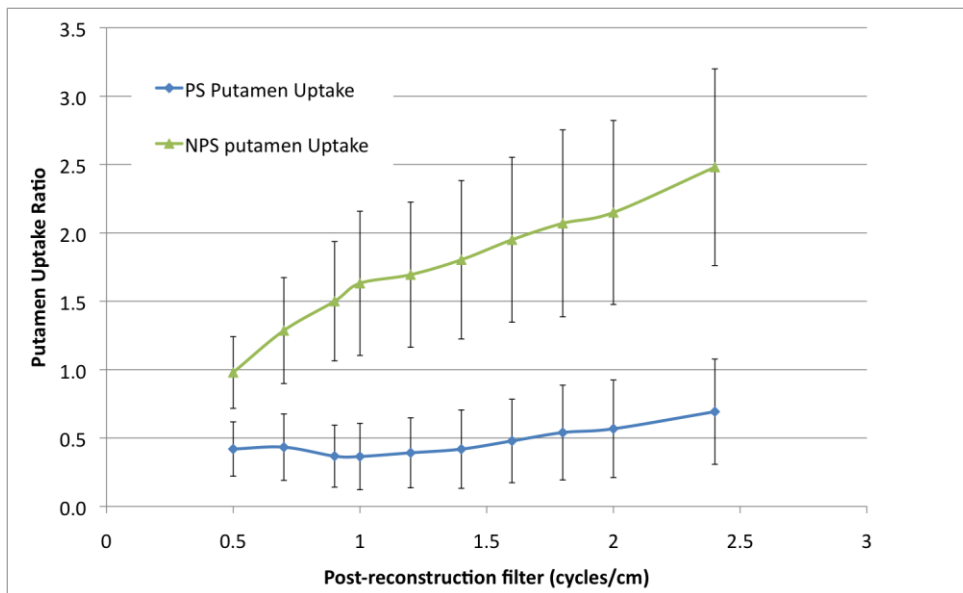


Figure 6-5 BRASS-Putamen uptake ratio as a function of post-reconstruction filter

For both caudate and putamen uptake ratios there was a statistically significant difference between the PS and Non-PS groups. The putamen results show better separation between PS and Non-PS than the caudate uptake results.

The Area Under the Curve (AUC) results for both BRASS caudate and putamen uptake ratios are shown in Figure 6-6. The putamen uptake ratio is shown to be a better classifier than the caudate uptake ratio for all filter settings. There is a drop in AUC at the two extremes of filter setting. There is a broad maximum in

the AUC for the putamen uptake ratio corresponding to filter settings in the 1-1.4cycles/cm range.

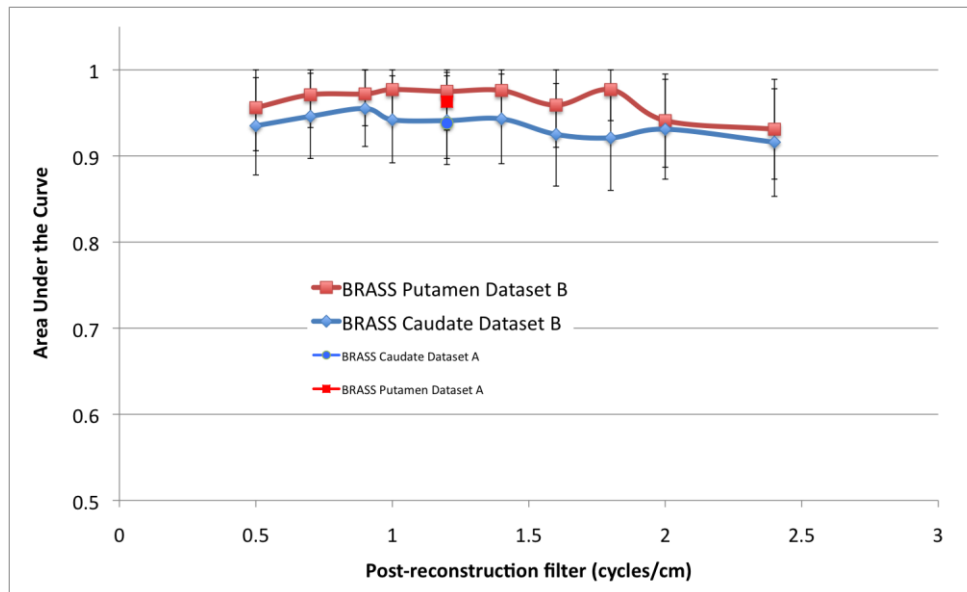


Figure 6-6 BRASS ROC analysis results as function of post reconstruction filter

6.6.3 Testing the BRASS system – discussion

The putamen uptake ratio is shown to give very good accuracy in classifying between the two disease states. This is likely to be due to the normal progression of disease. This will generally start in the posterior section of the putamen and extend anteriorly with disease progression. With such a progression the last part of the striatum to be affected is the caudate, which can show good uptake even at relatively advanced disease.

The increase in uptake ratios follows the expected trend with increasing filter cut-off frequency. When wider smoothing filters are applied they have the effect of blurring some of the counts within the striatum into the surrounding areas. The effect is greater for higher contrast studies, as the blurring of counts occurs in both directions – from the striatum in to the background and from the background into the striatum.

The drop in accuracy with the three smoothest filters can be explained by the realignment problems for some patients. The drop in accuracy with sharper filters/no post-reconstruction filter may be in part due to the misregistered patients, but this does not explain this entire drop in AUC. The increased noise levels in these images may have lead to changes in accuracy at these points, or undetected misregistrations that have altered the uptake.

6.7 The QuantiSPECT packages

6.7.1 Testing methods

The two clinical datasets described in chapter 4 were analysed using the QuantiSPECT software. Although this software contains a module to perform FBP reconstruction, it was not used. Instead the various reconstructed data from the Hermes system (as described in Chapter 3) was used. All three analysis packages were used to assess the uptake in the clinical datasets.

As the uptake ROIs/VOIs were assessed visually there were no perceived errors in placement. However, a number of the studies were considerably harder to align. These were generally the studies where there was very low striatal uptake. An example of one of these problematic image volumes is shown in Figure 6-7.

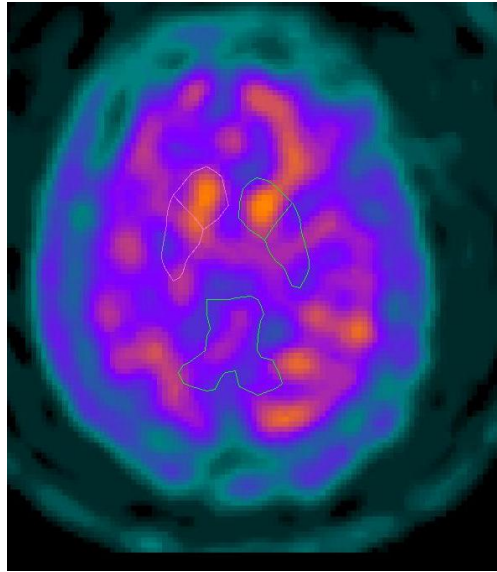


Figure 6-7 QuantiSPECT positioning uncertainty

The three methods produce eleven uptake measurements. In all but one of these measurements lower figures are correlated with PS and higher figures with Non-PS diseases. So for these readings the minimum score for each patient was used in the analysis. The symmetry index is the one figure that does not follow this pattern as it scales from minus to plus infinity, with zero being normal. To allow for analysis along with the other metrics the absolute of the asymmetry score was subtracted from 1000.

6.7.2 The 2-Box Method - Results

The mean SBI for the two groups of patients is shown in Figure 6-8 as a function of reconstruction post filter. The error bars show one standard deviation for the two groups. There is good separation for both groups at all filter frequencies. There is an obvious increase in the SBI when very smooth filters are used.

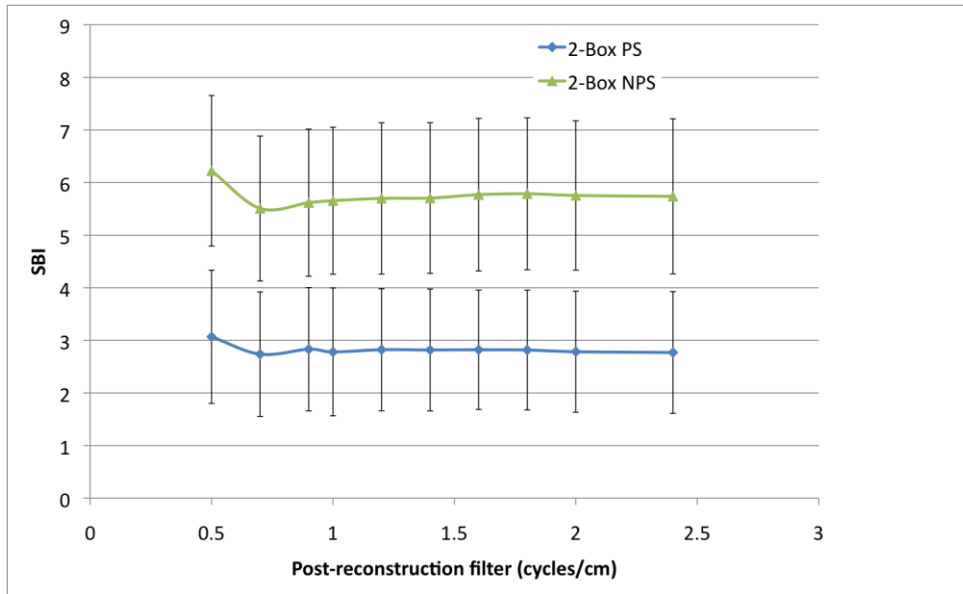


Figure 6-8 2-Box Striatal Binding Index as a function of the post reconstruction filter

The results of the ROC analysis in the form of the AUC are shown in Figure 6-9 as a function of filter cut-off frequency. The AUC from the original dataset is also shown corresponding to a post-reconstruction filter of 1.2cycles/cm for comparison.

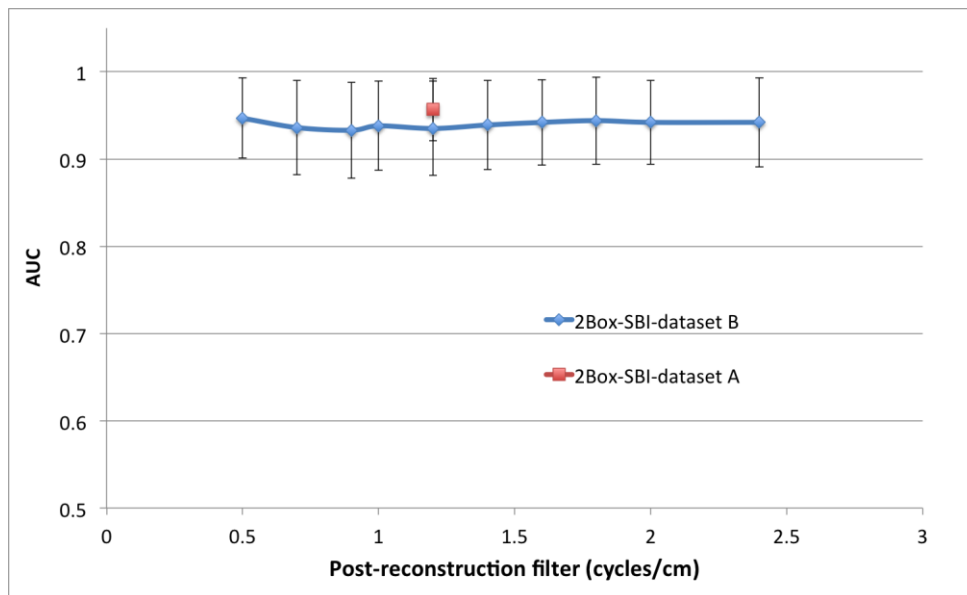


Figure 6-9 2-Box ROC analysis as a function of post-reconstruction filter

6.7.3 The 2-Box Method - Discussion

The SBI curves for the two groups of patients show very uniform response and are nearly independent of image filtering level. Only at very smooth filters does

the SBI for the two groups change. This uniformity of response is shown in the constant AUC results. One of the goals of the authors originally proposing this technique was to create a resolution insensitive system and these results confirm this property.

6.7.4 The 3-Box Method - Results

The mean minimum TBPI figure for the two groups of patients as a function of the post reconstruction filter is shown in Figure 6-10. The error bars reflect one standard deviation in the results. The mean of the absolute asymmetry indexes for the two groups is shown in Figure 6-11 as a function of reconstruction filter.

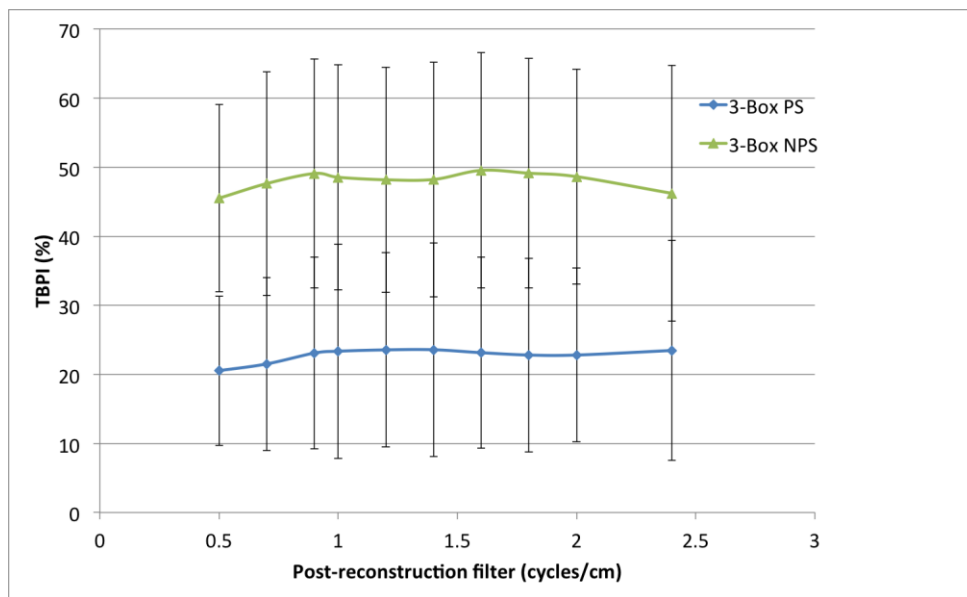


Figure 6-10 3-Box TBPI as a function of post reconstruction filter setting

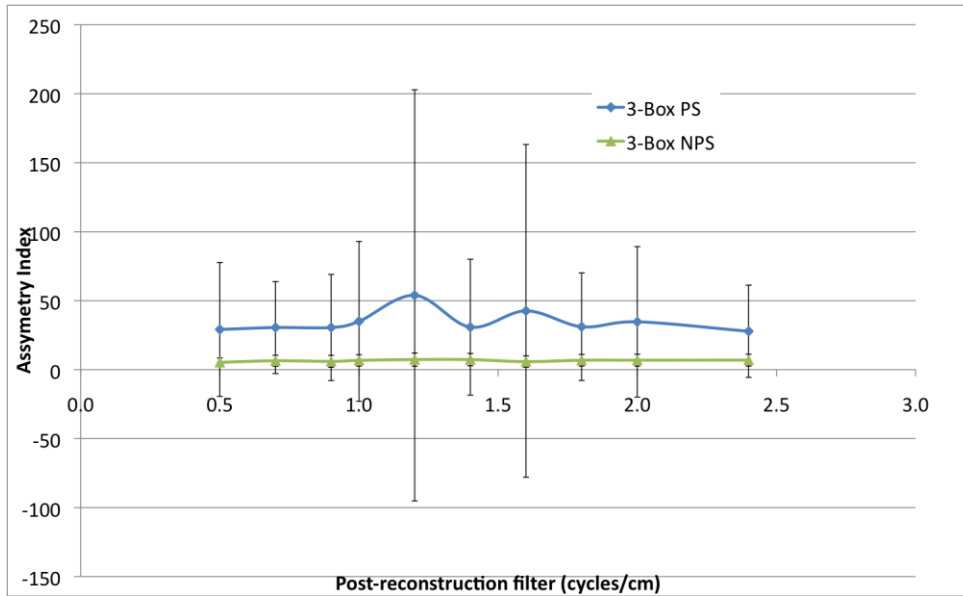


Figure 6-11 3-Box Asymmetry index as a function of post reconstruction filter

The results of the ROC analysis in terms of AUC for both the TBPI and the asymmetry index are shown in Figure 6-12.

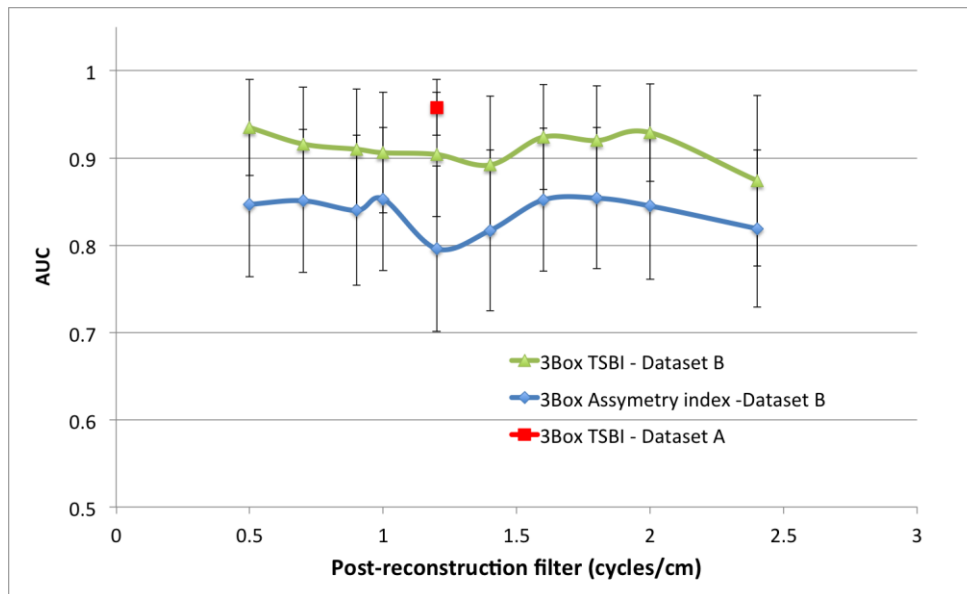


Figure 6-12 3-Box ROC analysis as a function of filter cut-off

6.7.5 The 3-Box Method - Discussion

As with the 2-box analysis, the large VOI size of the 3-box analysis gives a very uniform response with resolution. However, the TBPI results show very large variation. There is considerable overlap between the two groups when the standard deviation is taken into account. The main difference between this

technique and the 2-box system is the placement of the background region. For the 2-box method this is automatically defined and encompasses the vast majority of the cerebellum. In the 3-box system the background region is moveable and re-sizable, although in this work the default size was used. With a fixed size and shape, the variations in this region can come from two factors.

The first factor is blurring of counts from the striatum. The recommended placement of this region is posterior to the striatal regions. In published work others have positioned it abutting the striatal regions ⁷⁸. Here the placement was centrally positioned, posterior to the striatal regions, with a gap of 4 voxels to try to limit the level of spill-over from the striatum.

A second variation comes from the size, shape and position of the ventricles relative to the position of this VOI. The ventricles hold cerebrospinal fluid in a central cavity in the middle for the brain and have zero DAT uptake. The size and shape of these vary between patients and can grow in size with age.

The asymmetry index shows very large variations in magnitude – especially when the uptake is low. It gives poor results as a classifier on its own. It may be possible to include this as a second step of a decision tree classifier.

6.7.6 The Crescent method - results

The mean of the minimum striatal uptakes is shown in Figure 6-13 as a function of reconstruction filter. A sharp drop off in measured uptake is seen at the smoother filter settings. The same over all patterns are seen in the putamen uptake ratios (Figure 6-14). Above $f_{co}=1.0$ cycle/cm caudate and putamen uptake ratios show no dependency on filter. Both curves show good separation between

the two disease states. As with the BRASS system, the putamen uptake curves give better differentiation.

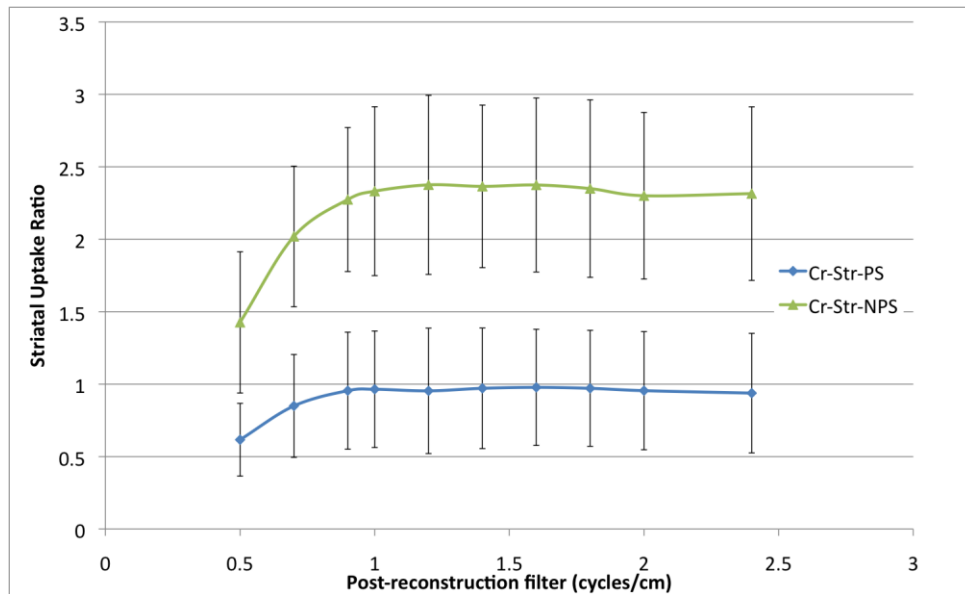


Figure 6-13 Crescent method striatal uptake as a function of reconstruction filter

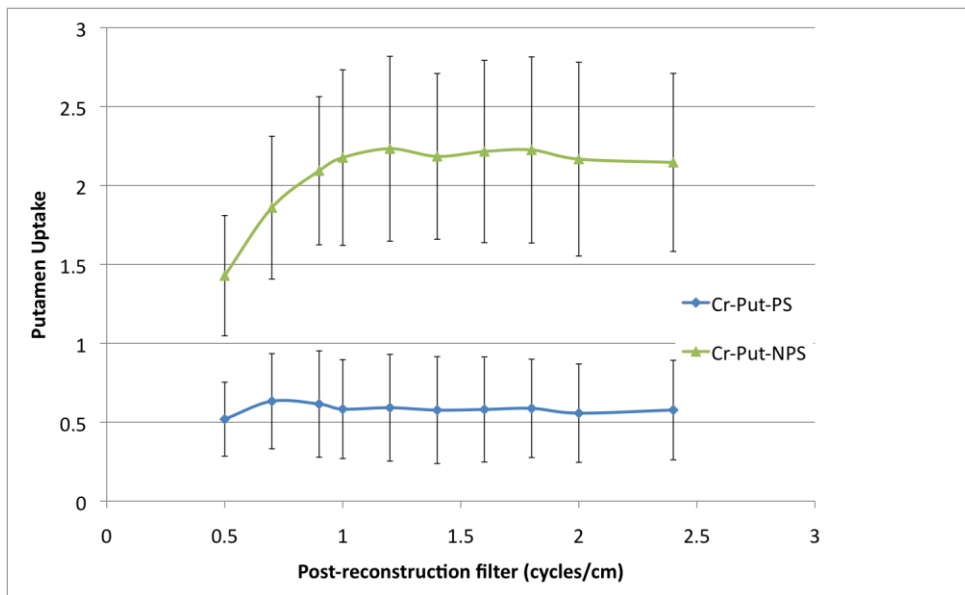


Figure 6-14 Crescent method putamen uptake as function of reconstruction filter

The ratio of caudate to putamen counts is shown in Figure 6-15 as a function of smoothing filter setting. The ratio is constant for filters above 1.0cycles/cm. At lower frequencies the ratio approaches 1 as the smoothing filter averages the signals from the two regions.

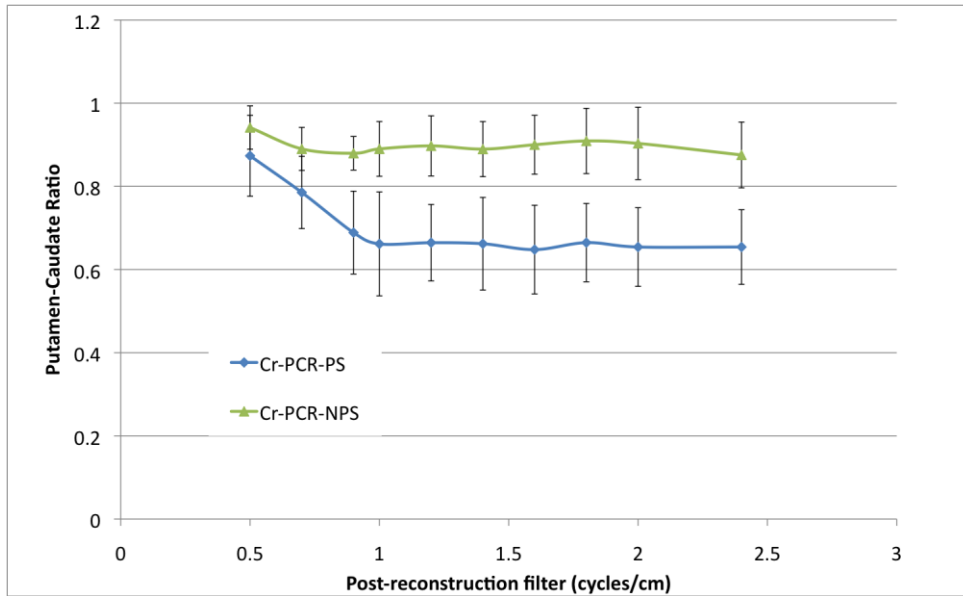


Figure 6-15 Putamen to Caudate ratio as a function of reconstruction filter

The usefulness of these metrics as a classifier was tested using ROC analysis. Figure 6-16 shows the AUC for these results as a function of filter cut-off frequency.

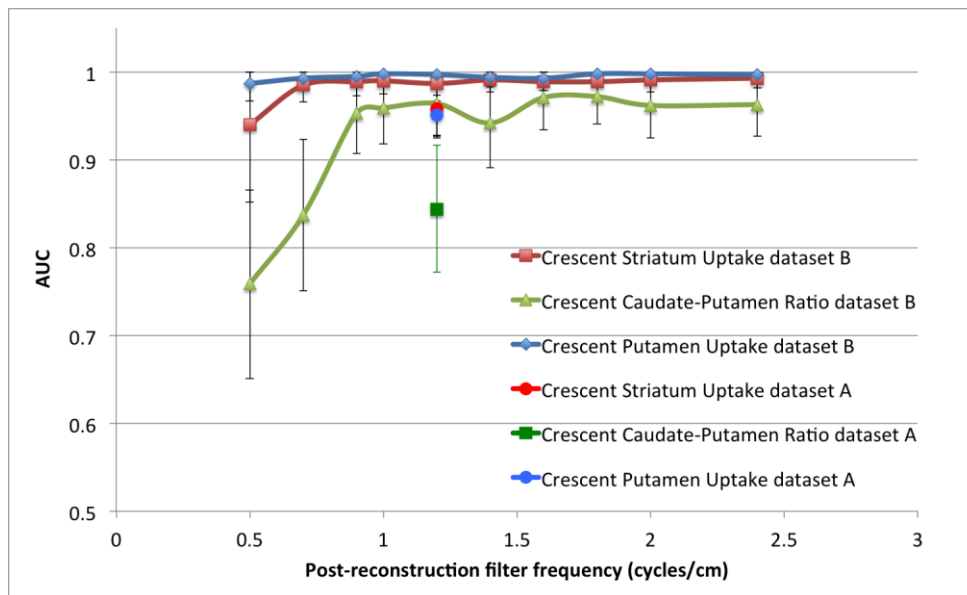


Figure 6-16 ROC results for the crescent method

6.7.7 The Crescent method - discussion

The results from dataset-B show almost perfect AUC scores for the striatum and caudate results for all but the smoothest of filter settings. The putamen uptake ratio is particularly good as a classifier.

The results from dataset-A show lower AUC scores than dataset B. However only the difference in PCR AUC score is statistically significant.

6.8 Comparison of datasets

The two datasets are described in chapter 4. They were acquired with different matrix sizes and scan timing. The composition of known diagnosis is slightly different but not statistically significant. The AUC results were compared for the data from dataset-A and the data from dataset-B that had been smoothed with a 1.2cycles/cm filter. The comparison was made using the method outlined by Hanley and NcNeil ¹⁶⁵ using a significance level of $p=0.05$.

The two large VOI analysis modes (2box and 3box) both have higher AUC scores for the older dataset. The two methods using the smaller VOIs achieved similar results for both datasets, with slightly lower values for the older data set. However these differences were not statistically significant. The one pair of ROC curves that were found to be statistically significant were those from the PCR method, which had a particularly poor AUC result for dataset A.

6.9 Reconstruction setting comparisons

The variation in accuracy with filter setting was investigated using the better performing analysis techniques – the crescent and BRASS based uptake ratios and the 2-box model. The AUC is plotted as a function of the smoothing filter

setting in Figure 6-17. The VOI techniques described here show stable performance with image filtering. These 5 methods are very stable for the range of filter 1.0-1.4 cycles/cm.

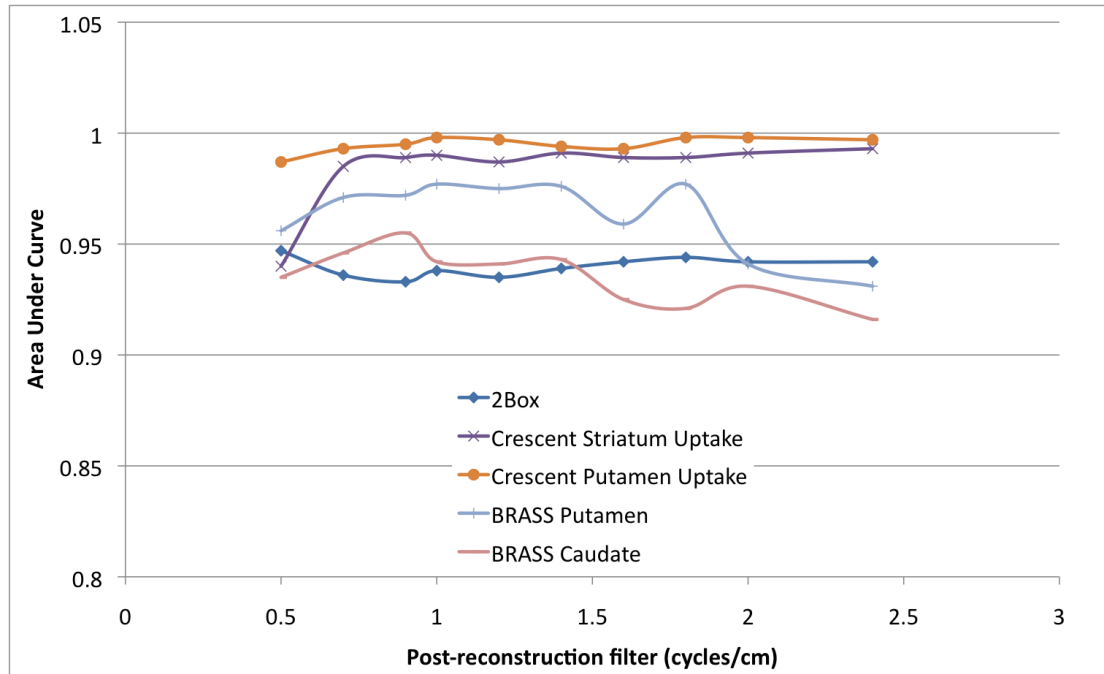


Figure 6-17 AUC as a function of filter setting for a range of analysis techniques

6.10 Method comparisons - sensitivity and specificity

For all of the results described above the cut-off value for the various uptake metrics has been used as a variable in ROC analyses. To use these results as a classifier, a specific cut-off must be chosen to give the diagnosis. As shown in the ROC curves, a compromise must be made between the sensitivity and specificity of the tests. Where this compromise falls depends on the use of the test, the severity of consequence for false negative or false positive results, and on the prevalence of the disease in the test population. In this work we compare the maximum accuracy for the test population investigated here. The maximum accuracies achievable with the different VOI classifiers are shown in

for dataset B. The maximum accuracy for each method is shown as a function of filter frequency in Figure 6-18.

The results for the 2 BRASS methods include between 1 and 3 missed diagnosis that were a direct result of misalignment. The uptake figures for those patients were not included, and a null results entered. Each incorrectly classified subject means a drop in accuracy of 1.3%. If these misaligned patients are not included in the analysis, the BRASS- putamen system reaches a peak accuracy of 0.97.

From Figure 6-18 it can be seen that some of the methods are strongly influenced by image filtering, in particular the putamen-caudate ratio and the asymmetry index. The Crescent-Putamen method gives the highest accuracy of all the techniques. It also gives consistent performance across the levels of image filtering.

Table 6-2 Comparison of optimised VOI methods using dataset B

	Cut-off	Sensitivity	Specificity	Accuracy
Brass –Caudate	1.77	0.96	0.77	0.91
Brass –Putamen	1.00	0.95	1.00	0.96
2-Box	5.14	0.98	0.73	0.91
3-Box	40	0.96	0.73	0.90
3-Box AI	11*	0.79	0.86	0.81
Cr Striatum	1.70	0.96	0.95	0.96
Cr-PCR	0.84	0.95	0.95	0.95
Cr-Putamen	1.52	0.98	1.00	0.99

***All readings below cut-off values give positive diagnosis except for 3-Box-AI which is give a negative diagnosis if below the cut-off.**

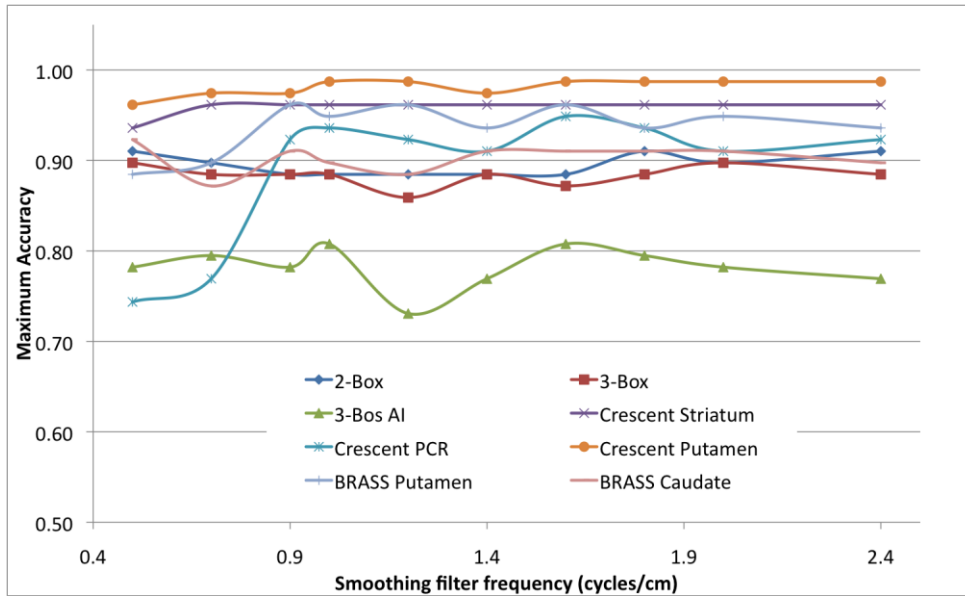


Figure 6-18 Maximum achievable accuracy for each classification model as a function of filter frequency

The maximum accuracy achievable for the datasets A and B (using the optimum filter setting for dataset B) is shown in Table 6-3. The cut-off values in this table are different for some of the measurements as the original dataset used attenuation correction.

Table 6-3 Comparison of optimised VOI methods using both datasets

	Cut-off Dataset A	Accuracy Dataset A	Cut-off Dataset B	Accuracy Dataset B
Brass –Caudate	1.68	0.87	1.77	0.91
Brass –Putamen	1.19	0.93	1.00	0.96
2-Box	4.85	0.91	5.14	0.91
3-Box	44	0.91	40	0.90
3-Box AI*	-	-	11	0.81
Cr Striatum	1.5	0.91	1.70	0.96
Cr-PCR	0.86	0.81	0.84	0.95
Cr-Putamen	1.2	0.91	1.52	0.99

*All readings below cut-off values give positive diagnosis except for 3-Box-AI which is give a negative diagnosis if below the cut-off. Asymmetry index not recorded for dataset A

6.11 Variation with Age

The uptake scores for the non-PS subjects data used above was used to test the use of patient age as a confounding feature. The data used to give the best

classifications (used in Table6-2) was used. The uptake scores were plotted as a function of patients' age (in years, at time of scan). A linear regression was performed on this data, to show the age dependence of these uptake figures. A summary is shown in Table 6-4. The average change in uptake values across the measurements was 6.6% per decade.

Table 6-4 Age dependence of uptake figures

	Slope	Intercept	r ²	% change per decade
Brass -Caudate	-0.025	3.52	0.36	-7.1%
Brass -Putamen	-0.027	3.03	0.39	-8.9%
2-Box	-0.07	10.2	0.48	-6.9%
3-Box	-0.52	78.3	0.28	-6.6%
Cr Striatum	-0.017	3.31	0.22	-5.1%
Cr-Putamen	-0.017	3.22	0.24	-5.3%

The data was used to calculate the 95% confidence intervals for the prediction of uptake measure, given age. Due to the low number of subjects in the Non-PS grouping, the confidence intervals for this data are correspondingly wide.

Figure 6-19 shows the PS and NPS data for the 2-Box analysis as a function of age. A regression line has been fitted to the Non-PS data. The classification cut-off (defined above, Table6-2) is shown giving a misclassification of 7 patients. The 95% confidence intervals from the regression are also shown. Although there is clearly age dependence in the uptake score, examination of this graph shows that a better classification cannot be given using this extra information.

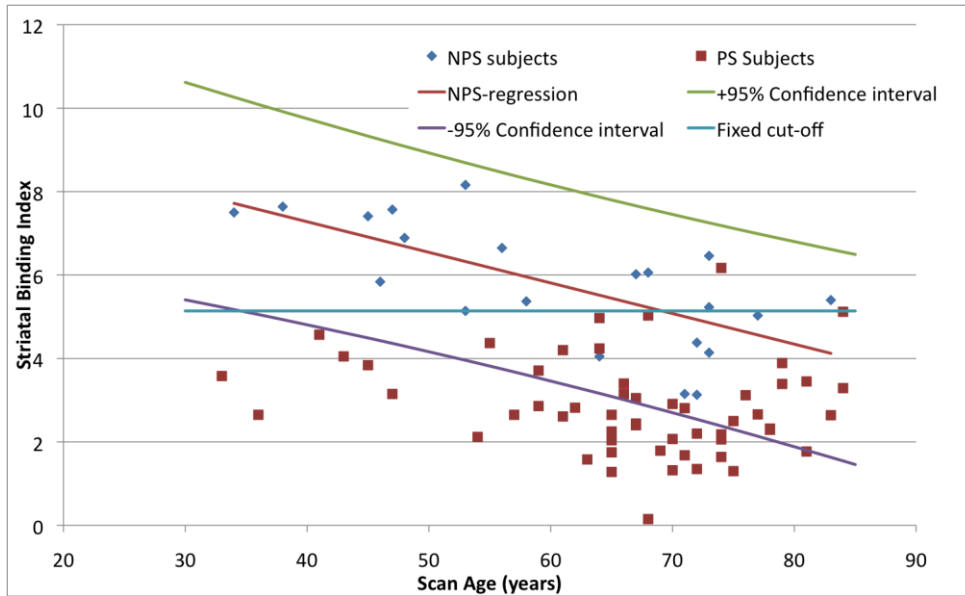


Figure 6-19 Age dependence of 2Box striatal binding index

Figure 6-18 shows the same analysis performed on the crescent method putamen uptake ratio. There is a single misclassified patient when using the non-age dependent cut-off (1.52). Again, the age dependence is seen in the NPS subjects, however this extra information does not improve the classification.

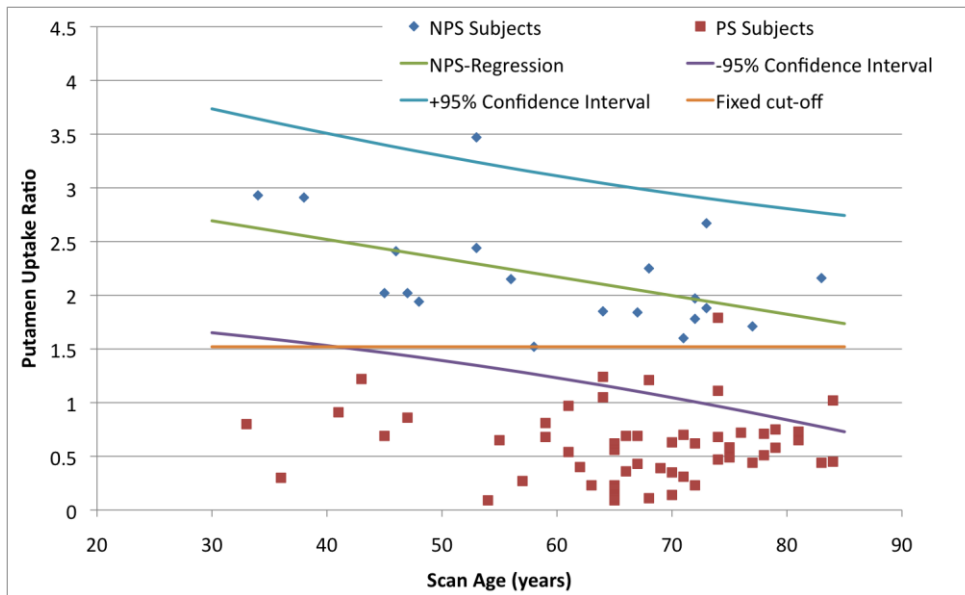


Figure 6-20 Age Dependence of Crescent-Putamen uptake ratio

6.12 Variation with age – discussion

The regression analysis on non-PS patients shows an age dependency for all uptake measures. The average reduction in DAT uptake measurements of 6.6%

compares well with published figures of 4.1%,¹⁵⁰ 4-6.7%¹⁶⁶ and 6.5%¹⁶⁷. The caudate-to-putamen ratio does not exhibit an age related change that suggests that the age related changes in caudate and putamen are similar. This age dependence should allow an age related cut-off to be used to give better classification however this has not been the case for this data. Part of the reason for this is the number of subjects in the training set. Fewer subjects in the NPS group mean quite widely spaced 95% confidence intervals. The change in uptake with age is one of many factors that affect the uptake measurements. There is still a wide variation in uptake in the normal population that is not accounted for by age. In the test-retest variability measurements published by Booij et al.⁴⁸ over 7% variability was seen in the healthy control group. The non-PS groups used here is, because of its origin, not a healthy control group. They are patients with movement disorders that are thought to not have any of the Parkinsonian syndromes. It is likely that such a group will have a normal variance as great if not greater than the healthy control group used in that work. Other factors that may contribute to the variation in uptake in the non-PS group could include subject medication history, and possibly scatter and attenuation effects.

6.13 Commercially available methods – discussions

In this chapter, four methodologies for assessment of DAT images were described and assessed under a range of conditions.

The dependence on smoothing filter in for the crescent method uptake measurements is much lower than that of the BRASS methodology. This difference may be down to the size of the VOIs used. Unfortunately, it was not

possible to assess the sizes of the VOIs used in these two techniques. A secondary factor may be in the performance of the image registration program used in BRASS. Some misregistrations were found when analysing the second patient group, but it is possible that some other more subtle misregistration were not noticed during visual checking. The large VOI methods showed very stable output with filter smoothing.

The two methods that measure putamen uptake separately, the BRASS-Put and Crescent-Put measurements, performed the best across the two patient datasets.

7 Comparison of Classification systems

In the previous two chapters a wide range of classification techniques have been tested using a large dataset of clinical images from a carefully controlled clinical audit database.

7.1 Overall accuracy

The best performing commercial system was the crescent method using the putamen uptake index. This gave 99% accuracy for training set B. The novel SVD based system developed in this work gave the same accuracy for the same data set.

7.2 Variation with image smoothness

Over all the various classification models here showed good stability in performance with image filtering levels. This was despite the wide range of image qualities tested.

There were some patterns in the performance when compared to filtering levels. The large VOI system showed the best performance with very smooth images. The 2-box method showed good uniformity of accuracy with image filtering. The 3-box method did not have this same uniformity most likely because of the variation in background VOI placement with respect the ventricles.

A number of papers have been written discussing the measurement and use of cross camera calibration factors. In them various striatal uptake measurements have been made on anthropomorphic phantoms ^{86,168} and some have been applied to ranges of healthy volunteers ⁸⁰. The uptake measurements used in those works have been included in this project, and found to be insensitive to

image smoothing levels. Further more, these works have used single reconstruction software and parameters. It must be surmised that differences in uptake seen across scanners may be linked to the scattering and collimator effects that are prevalent in 123-Iodine imaging.

This could not be tested within this work as all the patients were being scanned on a single site and as part of clinical routine. A further study could be performed in which volunteer patients are scanned on multiple scanners from single injections.

7.3 Statistical and clinical significance of these results

Most of these semi-quantitative methods are capable of providing excellent differentiation between the disease states. The reason these techniques are so powerful is to do with the progression of the disease. In Parkinson's disease clinical symptoms only start to appear once there has been a significant loss (>50% change) in dopaminergic cells. Such significant changes lead to high diagnostic performance.

However, there remain some patients that do not get the correct diagnosis. One group of patients who have been identified are those classed as Symptoms Without Evidence of Dopaminergic Deficit. This group appear to have some form of Parkinsonian syndrome, but it does not follow the normal pattern of IPD or the other main Parkinson's-plus syndromes.

One reason that has been suggested for possible false negative scans when using ^{123}I -FP- β -CIT is the non-specificity of the pharmaceutical. The ligand is known to have an affinity for serotonin transporters (SERT) and well as DAT ³⁹ ¹⁶⁹. It has been suggested that the early stages of some Parkinsonian syndromes may

involve a compensatory up regulation of SERT in response to dopaminergic loss. At the time of writing no clear evidence of this effect has been published, but it may be a useful area for investigation in the future.

In this study the gold standard diagnosis was given by a neurologist who was personally involved in the subjects care. Whilst undoubtedly an expert, the complex nature of these diseases may allow a misdiagnosis to be given. Having said that, the near perfect classification results given by the different classification methods suggests that so such errors have been made. There are of course a number of patients in the original audit database for whom a correct diagnosis was not recorded. By leaving out these patients the overall sensitivity and specificity of these tests are greatly exaggerated. The only true gold standard for the measurement of accuracy of these tests is with post-mortem evidence and sufficient data for this is unlikely to ever be recorded.

7.4 Overview and Conclusions

The aims of this work as set out in chapter 1 and having reviewed the available results and data the following points can be made:

The gamma camera performance characteristics for ¹²³iodine use were measured and used to optimise the clinical acquisition of SPECT imaging of the brain.

The image quality for a wide range of reconstruction techniques were compared using a range of image quality metrics. In the end the exact reconstruction settings were shown to have little effect on the ultimate image quality metric- their usefulness in giving clinical diagnosis.

Four commercially available DAT SPECT analysis programs were tested and compared under a range of settings.

A collection of automatic classification techniques were developed using feature extraction and machine learning algorithms. The a performance of these algorithms was compared directly to the commercial systems to test the original hypothesis for this thesis:

“An automatic classification tool based on machine learning techniques will give better performance than currently available techniques.”

The answer to this hypothesis is negative, but not because of failing in the novel technique developed here. Rather, it is not proven because of the near perfect performance of both the novel technique, and one of the commercially available methods.

7.5 Future work:

The excellent results for the combined feature extraction and machine-learning system may be useful in the analysis of other imaging tasks, especially if the task involves two distinct image distributions. One such group of imaging techniques is the measurement of amyloid plaque load in Alzheimer’s disease.

These techniques could be applied to the database of images that are available as part of the PPMI project ¹⁷⁰. This project will include a large number of patient scans from a range of scanner. The scan subjects are a mix of healthy volunteers and treatment naïve patients with Parkinson’s disease. With improved statistics

of a much larger dataset, better distinction may be drawn between the analysis techniques described here.

Although near perfect classification was reported for these datasets, a recent publication by Segovia et al. ¹⁷¹ used the results described here as a benchmark with which to compare a new classification procedure. They utilised a PLS based transformation followed by a SVM classifier and showed an improvement over the techniques reported here.

Bibliography

- 1 Towey, D., Bain, P. & Nijran, K. Automatic classification of 123I-FP-CIT (DaTSCAN) SPECT images. *Nucl Med Commun* **32**, 699-707, doi:10.1097/MNM.0b013e328347cd09 (2011).
- 2 Towey, D. in *Creative Counting in Nuclear Medicine, IPEM meeting*.
- 3 Towey, D., Bain, P., Al-Nahhas, A. & KS., N. in *Annual Congress of the European Association of Nuclear Medicine*. OP189.
- 4 Towey, D., Bain, P., Al-Nahhas, A. & Nijran, K. in *Annual Congress of the European Association of Nuclear Medicine*. S336.
- 5 Szyszko, T. *et al*. Incidence of bilateral disease on DATscan in assessment of patients with unilateral Parkinsonian symptoms. *Eur J Nucl Med Mol Imaging* **35** S360 (2008).
- 6 Towey, D., Moysiadou, F. & Nijran, K. in *BNMS*. 471-472.
- 7 Parkinson, J. An essay on the shaking palsy. 1817. *J Neuropsychiatry Clin Neurosci* **14**, 223-236; discussion 222 (2002).
- 8 Samii, A., Nutt, J. & Ransom, B. Parkinson's disease. *Lancet* **363**, 1783-1793, doi:10.1016/S0140-6736(04)16305-8 (2004).
- 9 Jankovic, J. Parkinson's disease: clinical features and diagnosis. *J Neurol Neurosurg Psychiatry* **79**, 368-376, doi:10.1136/jnnp.2007.131045 (2008).
- 10 Ziemssen, T. & Reichmann, H. Non-motor dysfunction in Parkinson's disease. *Parkinsonism Relat Disord* **13**, 323-332, doi:10.1016/j.parkreldis.2006.12.014 (2007).
- 11 Cummings, J. Depression and Parkinson's disease: a review. *Am J Psychiatry* **149**, 443-454 (1992).
- 12 Chaudhuri, K. & Schapira, A. Non-motor symptoms of Parkinson's disease: dopaminergic pathophysiology and treatment. *Lancet Neurol* **8**, 464-474, doi:10.1016/S1474-4422(09)70068-7 (2009).
- 13 Goldstein, D., Sewell, L. & Holmes, C. Association of anosmia with autonomic failure in Parkinson disease. *Neurology* **74**, 245-251, doi:10.1212/WNL.0b013e3181ca014c (2010).
- 14 Nicpon-Marieb, E. *Human Anatomy and Physiology (The Benjamin/Cummings series in the life sciences)*. (Benjamin-Cummings Publishing Company, 1994).
- 15 Hughes, A., Daniel, S., Kilford, L. & Lees, A. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry* **55**, 181-184 (1992).
- 16 Hobson, D. Clinical manifestations of Parkinson's disease and parkinsonism. *Can J Neurol Sci* **30 Suppl 1**, S2-9 (2003).
- 17 Rehman, H. Progressive supranuclear palsy. *Postgrad Med J* **76**, 333-336 (2000).
- 18 Papp, M., Kahn, J. & Lantos, P. Glial cytoplasmic inclusions in the CNS of patients with multiple system atrophy (striatonigral degeneration, olivopontocerebellar atrophy and Shy-Drager syndrome). *J Neurol Sci* **94**, 79-100 (1989).

- 19 Graham, J. & Oppenheimer, D. Orthostatic hypotension and nicotine sensitivity in a case of multiple system atrophy. *J Neurol Neurosurg Psychiatry* **32**, 28-34 (1969).
- 20 Stefanova, N., Bucke, P., Duerr, S. & Wenning, G. Multiple system atrophy: an update. *Lancet Neurol* **8**, 1172-1178, doi:10.1016/S1474-4422(09)70288-1 (2009).
- 21 Kosaka, K. Lewy bodies in cerebral cortex, report of three cases. *Acta Neuropathol* **42**, 127-134 (1978).
- 22 McKeith, I. *et al.* Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): report of the consortium on DLB international workshop. *Neurology* **47**, 1113-1124 (1996).
- 23 Klein, J. *et al.* Neurotransmitter changes in dementia with Lewy bodies and Parkinson disease dementia in vivo. *Neurology* **74**, 885-892, doi:10.1212/WNL.0b013e3181d55f61 (2010).
- 24 Marras, C. & Lang, A. Invited article: changing concepts in Parkinson disease: moving beyond the decade of the brain. *Neurology* **70**, 1996-2003, doi:10.1212/01.wnl.0000312515.52545.51 (2008).
- 25 Beach, T. The history of Alzheimer's disease: three debates. *J Hist Med Allied Sci* **42**, 327-349 (1987).
- 26 Deuschl, G., Bain, P. & Brin, M. Consensus statement of the Movement Disorder Society on Tremor. Ad Hoc Scientific Committee. *Mov Disord* **13 Suppl 3**, 2-23 (1998).
- 27 Holmes, G. On certain tremors in organic cerebral lesions. *Brain* **27**, 327 (1904).
- 28 Lee, A. Vascular dementia. *Chonnam Med J* **47**, 66-71, doi:10.4068/cmj.2011.47.2.66 (2011).
- 29 Zijlmans, J., Daniel, S., Hughes, A., Revesz, T. & Lees, A. Clinicopathological investigation of vascular parkinsonism, including clinical criteria for diagnosis. *Mov Disord* **19**, 630-640, doi:10.1002/mds.20083 (2004).
- 30 Thanvi, B. & Treadwell, S. Drug induced parkinsonism: a common cause of parkinsonism in older people. *Postgrad Med J* **85**, 322-326, doi:10.1136/pgmj.2008.073312 (2009).
- 31 *New Tools in the Diagnosis of Parkinsonism. Imaging of the Pre- and Postsynaptic Dopaminergic System Using DaTSCAN and IBZM SPECT*. Vol. 13 (Academic Pharmaceutical Productions, 2001).
- 32 Brucke, T. *et al.* SPECT imaging of dopamine and serotonin transporters with [123I]beta-CIT. Binding kinetics in the human brain. *J Neural Transm Gen Sect* **94**, 137-146 (1993).
- 33 Innis, R. *et al.* Single photon emission computed tomographic imaging demonstrates loss of striatal dopamine transporters in Parkinson disease. *Proc Natl Acad Sci U S A* **90**, 11965-11969 (1993).
- 34 Kuikka, J. *et al.* Iodine-123 labelled N-(2-fluoroethyl)-2 beta-carbomethoxy-3 beta-(4-iodophenyl)nortropane for dopamine transporter imaging in the living human brain. *Eur J Nucl Med* **22**, 682-686 (1995).
- 35 Vermeulen, R. *et al.* Evaluation of [123I] beta-CIT binding with SPECT in controls, early and late Parkinson's disease. *Nucl Med Biol* **22**, 985-991 (1995).

- 36 Tsuchida, T. *et al.* Reproducibility of dopamine transporter density measured with 123I-FPCIT SPECT in normal control and Parkinson's disease patients. *Ann Nucl Med* **18**, 609-616 (2004).
- 37 Winogrodzka, A. *et al.* [123I]FP-CIT SPECT is a useful method to monitor the rate of dopaminergic degeneration in early-stage Parkinson's disease. *J Neural Transm* **108**, 1011-1019 (2001).
- 38 Kung, H. *et al.* Imaging of dopamine transporters in humans with technetium-99m TRODAT-1. *Eur J Nucl Med* **23**, 1527-1530 (1996).
- 39 Ziebell, M. Evaluation of the superselective radioligand [123I]PE2I for imaging of the dopamine transporter in SPECT. *Dan Med Bull* **58**, B4279 (2011).
- 40 Catafau, A. *et al.* Test-retest variability and reliability of 123I-IBZM SPECT measurement of striatal dopamine D2 receptor availability in healthy volunteers and influence of iterative reconstruction algorithms. *Synapse* **62**, 62-69, doi:10.1002/syn.20465 (2008).
- 41 Bullich, S. *et al.* Impact of scatter correction on D2 receptor occupancy measurements using 123I-IBZM SPECT: comparison to 11C-Raclopride PET. *Neuroimage* **50**, 1511-1518, doi:10.1016/j.neuroimage.2010.01.013 (2010).
- 42 Fujita, M. *et al.* Effect of scatter correction on the compartmental measurement of striatal and extrastriatal dopamine D2 receptors using [123I]epidepride SPET. *Eur J Nucl Med Mol Imaging* **31**, 644-654, doi:10.1007/s00259-003-1431-7 (2004).
- 43 Kemp, P. Imaging the dopaminergic system in suspected parkinsonism, drug induced movement disorders, and Lewy body dementia. *Nucl Med Commun* **26**, 87-96 (2005).
- 44 Okada, T. *et al.* Assessment of affinities of beta-CIT, beta-CIT-FE, and beta-CIT-FP for monoamine transporters permanently expressed in cell lines. *Nucl Med Biol* **25**, 53-58 (1998).
- 45 Booij, J. *et al.* [123I]FP-CIT SPECT shows a pronounced decline of striatal dopamine transporter labelling in early and advanced Parkinson's disease. *J Neurol Neurosurg Psychiatry* **62**, 133-140 (1997).
- 46 Booij, J. *et al.* Practical benefit of [123I]FP-CIT SPET in the demonstration of the dopaminergic deficit in Parkinson's disease. *Eur J Nucl Med* **24**, 68-71 (1997).
- 47 Booij, J. *et al.* Human biodistribution and dosimetry of [123I]FP-CIT: a potent radioligand for imaging of dopamine transporters. *Eur J Nucl Med* **25**, 24-30 (1998).
- 48 Booij, J. *et al.* Imaging of dopamine transporters with iodine-123-FP-CIT SPECT in healthy controls and patients with Parkinson's disease. *J Nucl Med* **39**, 1879-1884 (1998).
- 49 Catafau, A. & Tolosa, E. Impact of dopamine transporter SPECT using 123I-Ioflupane on diagnosis and management of patients with clinically uncertain Parkinsonian syndromes. *Mov Disord* **19**, 1175-1182, doi:10.1002/mds.20112 (2004).
- 50 Kemp, P., Clyde, K. & Holmes, C. Impact of 123I-FP-CIT (DaTSCAN) SPECT on the diagnosis and management of patients with dementia with Lewy bodies: a retrospective study. *Nucl Med Commun* **32**, 298-302, doi:10.1097/MNM.0b013e328343d4ec (2011).

- 51 Marshall, V., Patterson, J., Hadley, D., Grosset, K. & Grosset, D. Two-year follow-up in 150 consecutive cases with normal dopamine transporter imaging. *Nucl Med Commun* **27**, 933-937, doi:10.1097/01.mnm.0000243374.11260.5b (2006).
- 52 Ishikawa, T. *et al.* Comparative nigrostriatal dopaminergic imaging with iodine-123-beta CIT-FP/SPECT and fluorine-18-FDOPA/PET. *J Nucl Med* **37**, 1760-1765 (1996).
- 53 Lundkvist, C., Halldin, C., Ginovart, N., Swahn, C. & Farde, L. [18F] beta-CIT-FP is superior to [11C] beta-CIT-FP for quantitation of the dopamine transporter. *Nucl Med Biol* **24**, 621-627 (1997).
- 54 Nikkinen, P., Liewendahl, K., Savolainen, S. & Launes, J. Validation of quantitative brain dopamine D2 receptor imaging with a conventional single-head SPET camera. *Eur J Nucl Med* **20**, 680-683 (1993).
- 55 Koch, W., Hamann, C., Radau, P. & Tatsch, K. Does combined imaging of the pre- and postsynaptic dopaminergic system increase the diagnostic accuracy in the differential diagnosis of parkinsonism? *Eur J Nucl Med Mol Imaging* **34**, 1265-1273, doi:10.1007/s00259-007-0375-8 (2007).
- 56 Nocker, M. *et al.* Progression of dopamine transporter decline in patients with the Parkinson variant of multiple system atrophy: a voxel-based analysis of [(123)I]beta-CIT SPECT. *Eur J Nucl Med Mol Imaging*, doi:10.1007/s00259-012-2100-5 (2012).
- 57 Scherfler, C. *et al.* Voxel-wise analysis of [123I]beta-CIT SPECT differentiates the Parkinson variant of multiple system atrophy from idiopathic Parkinson's disease. *Brain* **128**, 1605-1612, doi:10.1093/brain/awh485 (2005).
- 58 Juh, R. *et al.* Voxel based comparison of glucose metabolism in the differential diagnosis of the multiple system atrophy using statistical parametric mapping. *Neurosci Res* **52**, 211-219, doi:10.1016/j.neures.2005.03.010 (2005).
- 59 Filippi, L. *et al.* 123I-FP-CIT in progressive supranuclear palsy and in Parkinson's disease: a SPECT semiquantitative study. *Nucl Med Commun* **27**, 381-386, doi:10.1097/01.mnm.0000202858.45522.df (2006).
- 60 Rossi, C. *et al.* "Parkinson-dementia" diseases: a comparison by double tracer SPECT studies. *Parkinsonism Relat Disord* **15**, 762-766, doi:10.1016/j.parkreldis.2009.05.012 (2009).
- 61 Imon, Y., Matsuda, H., Ogawa, M., Kogure, D. & Sunohara, N. SPECT image analysis using statistical parametric mapping in patients with Parkinson's disease. *J Nucl Med* **40**, 1583-1589 (1999).
- 62 Kemp, P. *et al.* The contribution of statistical parametric mapping in the assessment of precuneal and medial temporal lobe perfusion by 99mTc-HMPAO SPECT in mild Alzheimer's and Lewy body dementia. *Nucl Med Commun* **26**, 1099-1106 (2005).
- 63 Guedj, E. *et al.* Neuroimaging findings in a case of Holmes tremor. *Clin Nucl Med* **32**, 139-140, doi:10.1097/01.rlu.0000251948.30673.4f (2007).
- 64 Gajos, A. *et al.* The clinical and neuroimaging studies in Holmes tremor. *Acta Neurol Scand* **122**, 360-366, doi:10.1111/j.1600-0404.2009.01319.x (2010).

- 65 Barthel, H. *et al.* Concordant pre- and postsynaptic deficits of dopaminergic neurotransmission in neurologic Wilson disease. *AJNR Am J Neuroradiol* **24**, 234-238 (2003).
- 66 Doepp, F. *et al.* Brain parenchyma sonography and 123I-FP-CIT SPECT in Parkinson's disease and essential tremor. *Mov Disord* **23**, 405-410, doi:10.1002/mds.21861 (2008).
- 67 Berendse, H. & Ponsen, M. Diagnosing premotor Parkinson's disease using a two-step approach combining olfactory testing and DAT SPECT imaging. *Parkinsonism Relat Disord* **15 Suppl 3**, S26-30, doi:10.1016/S1353-8020(09)70774-6 (2009).
- 68 ANGER, H. SCINTILLATION CAMERA WITH MULTICHANNEL COLLIMATORS. *J Nucl Med* **5**, 515-531 (1964).
- 69 Mather, R. L. Gamma-Ray Collimator Penetration and Scattering Effects. *Journal of Applied Physics* **28**, 1200-1207 (1957).
- 70 BÉ, M. *et al.* Table of radionuclides (Vol. 1-A= 1 to 150). *Bureau International des Poids et Mesures, France: Sèvres* (2004).
- 71 Dobbeleir, A., Hambye, A. & Franken, P. Influence of high-energy photons on the spectrum of iodine-123 with low- and medium-energy collimators: consequences for imaging with 123I-labelled compounds in clinical practice. *Eur J Nucl Med* **26**, 655-658 (1999).
- 72 Zaidi, H. *Quantitative Analysis in Nuclear Medicine Imaging*. (Springer, 2005).
- 73 R., C. S., A., S. J. & E., P. M. *Physics in Nuclear Medicine*. (Saunders, 2003).
- 74 Shepp, L. & Vardi, Y. Maximum likelihood reconstruction for emission tomography. *IEEE Trans Med Imaging* **1**, 113-122 (1982).
- 75 Benamer, T. *et al.* Accurate differentiation of parkinsonism and essential tremor using visual assessment of [123I]-FP-CIT SPECT imaging: the [123I]-FP-CIT study group. *Mov Disord* **15**, 503-510 (2000).
- 76 Ottaviani, S. *et al.* Comparative analysis of visual and semi-quantitative assessment of striatal [123I]FP-CIT-SPET binding in Parkinson's disease. *Neurol Sci* **27**, 397-401, doi:10.1007/s10072-006-0718-4 (2006).
- 77 Badiavas, K. *et al.* SPECT imaging evaluation in movement disorders: far beyond visual assessment. *Eur J Nucl Med Mol Imaging* **38**, 764-773, doi:10.1007/s00259-010-1664-1 (2011).
- 78 Morton, R. *et al.* Comparison of different methods of DatSCAN quantification. *Nucl Med Commun* **26**, 1139-1146 (2005).
- 79 Skanjeti, A., Angusti, T., Margheron, M., Iudicello, M. & Podio, V. FP-CIT SPECT evaluation: time to go beyond visual assessment! *Eur J Nucl Med Mol Imaging* **39**, 727-728, doi:10.1007/s00259-011-2042-3 (2012).
- 80 Darcourt, J. *et al.* EANM procedure guidelines for brain neurotransmission SPECT using (123)I-labelled dopamine transporter ligands, version 2. *Eur J Nucl Med Mol Imaging* **37**, 443-450, doi:10.1007/s00259-009-1267-x (2010).
- 81 Tatsch, K. *et al.* European Association of Nuclear Medicine procedure guidelines for brain neurotransmission SPET using (123)I-labelled dopamine D(2) transporter ligands. *Eur J Nucl Med Mol Imaging* **29**, BP30-35 (2002).
- 82 Larsson, A., Ljungberg, M., Mo, S., Riklund, K. & Johansson, L. Correction for scatter and septal penetration using convolution subtraction methods

- and model-based compensation in 123I brain SPECT imaging—a Monte Carlo study. *Phys Med Biol* **51**, 5753-5767, doi:10.1088/0031-9155/51/22/003 (2006).
- 83 Shiga, T. *et al.* The effect of scatter correction on 123I-IMP brain perfusion SPET with the triple energy window method in normal subjects using SPM analysis. *Eur J Nucl Med Mol Imaging* **29**, 342-345 (2002).
- 84 Soret, M., Koulibaly, P., Darcourt, J., Hapdey, S. & Buvat, I. Quantitative accuracy of dopaminergic neurotransmission imaging with (123)I SPECT. *J Nucl Med* **44**, 1184-1193 (2003).
- 85 Iida, H. *et al.* Effects of scatter and attenuation correction on quantitative assessment of regional cerebral blood flow with SPECT. *J Nucl Med* **39**, 181-189 (1998).
- 86 Zaknun, J., Schucktanz, H. & Aichner, F. Impact of instrumentation on DaTSCAN imaging: how feasible is the concept of cross-systems correction factor? *Q J Nucl Med Mol Imaging* **51**, 194-203 (2007).
- 87 Gilland, D., Jaszczak, R., Greer, K. & Coleman, R. Quantitative SPECT reconstruction of iodine-123 data. *J Nucl Med* **32**, 527-533 (1991).
- 88 Small, A. *et al.* Downscatter correction and choice of collimator in 123I imaging. *Phys Med Biol* **51**, N307-311, doi:10.1088/0031-9155/51/17/N03 (2006).
- 89 Macey, D., DeNardo, G., DeNardo, S. & Hines, H. Comparison of low- and medium-energy collimators for SPECT imaging with iodine-123-labeled antibodies. *J Nucl Med* **27**, 1467-1474 (1986).
- 90 Gilland, D., Jaszczak, R., Turkington, T., Greer, K. & Coleman, R. Volume and activity quantitation with iodine-123 SPECT. *J Nucl Med* **35**, 1707-1713 (1994).
- 91 Lagerburg, V., de Nijs, R., Holm, S. & Svarer, C. A comparison of different energy window subtraction methods to correct for scatter and downscatter in I-123 SPECT imaging. *Nucl Med Commun* **33**, 708-718, doi:10.1097/MNM.0b013e3283537cd0 (2012).
- 92 NEMA. *NEMA Standards Publication NU 1-2007, Performance Measurements of Gamma Cameras.*
- 93 Cao, Z., Maunoury, C., Chen, C. & Holder, L. Comparison of continuous step-and-shoot versus step-and-shoot acquisition SPECT. *J Nucl Med* **37**, 2037-2040 (1996).
- 94 Huesman, R. The effects of a finite number of projection angles and finite lateral sampling of projections on the propagation of statistical errors in transverse section reconstruction. *Phys Med Biol* **22**, 511-521 (1977).
- 95 Notes for guidance on the clinical administration of radiopharmaceuticals and use of sealed radioactive sources. Administration of Radioactive Substances Advisory Committee. *Nucl Med Commun* **21 Suppl**, S1-93 (2000).
- 96 Goatman, K. A. *et al. Mathematical Techniques in Nuclear Medicine (IPEM Reports Series)*. 1st edn, (Institute of Physics and Engineering in Medicine, 2011).
- 97 Hudson, H. & Larkin, R. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans Med Imaging* **13**, 601-609 (1994).
- 98 Chang, L. A Method for Attenuation Correction In Radionuclide Computed Tomography. *IEEE Trans Nuclear Science* **ns -25**, 638-643 (1978).

- 99 Koral, K. *et al.* SPECT dual-energy-window Compton correction: scatter multiplier required for quantification. *J Nucl Med* **31**, 90-98 (1990).
- 100 Ogawa, K., Harata, Y., Ichihara, T., Kubo, A. & Hashimoto, S. A practical method for position-dependent Compton-scatter correction in single photon emission CT. *IEEE Trans Med Imaging* **10**, 408-412 (1991).
- 101 Soret, M., Koulibaly, P., Darcourt, J. & Buvat, I. Partial volume effect correction in SPECT for striatal uptake measurements in patients with neurodegenerative diseases: impact upon patient classification. *Eur J Nucl Med Mol Imaging* **33**, 1062-1072, doi:10.1007/s00259-005-0003-4 (2006).
- 102 Calvini, P. *et al.* The basal ganglia matching tools package for striatal uptake semi-quantification: description and validation. *Eur J Nucl Med Mol Imaging* **34**, 1240-1253, doi:10.1007/s00259-006-0357-2 (2007).
- 103 Penney, B. *et al.* Constrained least-squares restoration of nuclear medicine images: selecting the coarseness function. *Med Phys* **14**, 849-858 (1987).
- 104 Beekman, F., Slijpen, E. & Niessen, W. Selection of task-dependent diffusion filters for the post-processing of SPECT images. *Phys Med Biol* **43**, 1713-1730 (1998).
- 105 Association, A. P. & DSM-IV., A. P. A. T. F. o. *Diagnostic and statistical manual of mental disorders: DSM-IV-TR*, (American Psychiatric Publishing, Inc., 2000).
- 106 Rorden, C. & Brett, M. Stereotaxic display of brain lesions. *Behavioural Neurology* **12**, 191-200 (2000).
- 107 Rasband, W. S. ImageJ. *U. S. National Institutes of Health, Bethesda, Maryland, USA* (1997). <<http://rsb.info.nih.gov/ij/>>.
- 108 Jarritt, P. *et al.* UK audit of single photon emission computed tomography reconstruction software using software generated phantoms. *Nucl Med Commun* **23**, 483-491 (2002).
- 109 Seret, A. & Forthomme, J. Comparison of different types of commercial filtered backprojection and ordered-subset expectation maximization SPECT reconstruction software. *J Nucl Med Technol* **37**, 179-187, doi:10.2967/jnmt.108.061275 (2009).
- 110 Tossici-Bolt, L., Hoffmann, S., Kemp, P., Mehta, R. & Fleming, J. Quantification of [(123)I]FP-CIT SPECT brain images: an accurate technique for measurement of the specific binding ratio. *Eur J Nucl Med Mol Imaging* **33**, 1491-1499 (2006).
- 111 Popperl, G., Radau, P., Linke, R., Hahn, K. & Tatsch, K. Diagnostic performance of a 3-D automated quantification method of dopamine D2 receptor SPECT studies in the differential diagnosis of parkinsonism. *Nucl Med Commun* **26**, 39-43 (2005).
- 112 Pirker, W. *et al.* Measuring the rate of progression of Parkinson's disease over a 5-year period with beta-CIT SPECT. *Mov Disord* **18**, 1266-1272, doi:10.1002/mds.10531 (2003).
- 113 Berti, V. *et al.* Clinical correlation of the binding potential with 123I-FP-CIT in de novo idiopathic Parkinson's disease patients. *Eur J Nucl Med Mol Imaging* **35**, 2220-2226, doi:10.1007/s00259-008-0872-4 (2008).

- 114 Dobbeleir, A., Hambye, A., Vervaet, A. & Ham, H. Quantification of iodine-123-FP-CIT SPECT with a resolution-independent method. *World Journal of Nuclear Medicine* **4**, 252-261 (2005).
- 115 Dickson, J. *et al.* The impact of reconstruction method on the quantification of DaTSCAN images. *Eur J Nucl Med Mol Imaging* **37**, 23-35, doi:10.1007/s00259-009-1212-z (2010).
- 116 Hutton, B. & Lau, Y. Application of distance-dependent resolution compensation and post-reconstruction filtering for myocardial SPECT. *Phys Med Biol* **43**, 1679-1693 (1998).
- 117 Koch, W. *et al.* Is iterative reconstruction an alternative to filtered backprojection in routine processing of dopamine transporter SPECT studies? *J Nucl Med* **46**, 1804-1811 (2005).
- 118 Warwick, J. *et al.* The Role of CT-Based Attenuation Correction and Collimator Blurring Correction in Striatal Spect Quantification. *Int J Mol Imaging* **2011**, 195037, doi:10.1155/2011/195037 (2011).
- 119 Crespo, C. *et al.* Quantification of dopaminergic neurotransmission SPECT studies with 123I-labelled radioligands. A comparison between different imaging systems and data acquisition protocols using Monte Carlo simulation. *Eur J Nucl Med Mol Imaging* **35**, 1334-1342, doi:10.1007/s00259-007-0711-z (2008).
- 120 Houston, A. S., Kemp, P. M. & Macleod, M. A. A method for assessing the significance of abnormalities in HMPO brain SPECT images. *J Nucl Med* **35**, 239-244 (1994).
- 121 Lozza, C. *et al.* Executive processes in Parkinson's disease: FDG-PET and network analysis. *Hum Brain Mapp* **22**, 236-245, doi:10.1002/hbm.20033 (2004).
- 122 Nobili, F. *et al.* Cognitive-nigrostriatal relationships in de novo, drug-naive Parkinson's disease patients: a [I-123]FP-CIT SPECT study. *Mov Disord* **25**, 35-43, doi:10.1002/mds.22899 (2010).
- 123 Nijran, K. & Barber, D. Towards automatic analysis of dynamic radionuclide studies using principal-components factor analysis. *Phys Med Biol* **30**, 1315-1325 (1985).
- 124 Nijran, K. S. *Analysis of Dynamic Radionuclide Studies Using Principal Components Factor Analysis* PhD thesis, The University of Sheffield, (1984).
- 125 Barber, D. Digital computer processing of brain scans using principal components. *Phys Med Biol* **21**, 792-803 (1976).
- 126 Barber, D. The use of principal components in the quantitative analysis of gamma camera dynamic studies. *Phys Med Biol* **25**, 283-292 (1980).
- 127 Padilla, P. *et al.* Analysis of SPECT brain images for the diagnosis of Alzheimer's disease based on NMF for feature extraction. *Neurosci Lett* **479**, 192-196 (2010).
- 128 Acton, P., Mozley, P. & Kung, H. Logistic discriminant parametric mapping: a novel method for the pixel-based differential diagnosis of Parkinson's disease. *Eur J Nucl Med* **26**, 1413-1423 (1999).
- 129 Horn, J. *et al.* Differential automatic diagnosis between Alzheimer's disease and frontotemporal dementia based on perfusion SPECT images. *Artif Intell Med* **47**, 147-158, doi:10.1016/j.artmed.2009.05.001 (2009).

- 130 Lopez, M. *et al.* Automatic tool for Alzheimer's disease diagnosis using PCA and Bayesian classification rules. *Electronics Letters* **45**, 389-391, doi:10.1049/el.2009.0176 (2009).
- 131 Cortes, C. & Vapnik, V. Support-vector networks. *Machine Learning* (1995).
- 132 Cuberas-Borros, G. *et al.* Quantitative evaluation of striatal I-123-FP-CIT uptake in essential tremor and parkinsonism. *Clin Nucl Med* **36**, 991-996, doi:10.1097/RLU.0b013e3182291a7b (2011).
- 133 Colloby, S. J. *et al.* The application of statistical parametric mapping to 123I-FP-CIT SPECT in dementia with Lewy bodies, Alzheimer's disease and Parkinson's disease. *Neuroimage* **23**, 956-966 (2004).
- 134 Eberl, S. *et al.* Automated interstudy image registration technique for SPECT and PET. *J Nucl Med* **37**, 137-145 (1996).
- 135 Meyer, C. *et al.* Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations. *Med Image Anal* **1**, 195-206 (1997).
- 136 Kas, A. *et al.* Validation of a standardized normalization template for statistical parametric mapping analysis of 123I-FP-CIT images. *J Nucl Med* **48**, 1459-1467, doi:10.2967/jnumed.106.038646 (2007).
- 137 Van Laere, K., Koole, M., Versijpt, J. & Dierckx, R. Non-uniform versus uniform attenuation correction in brain perfusion SPET of healthy volunteers. *Eur J Nucl Med* **28**, 90-98 (2001).
- 138 Abi-Dargham, A. *et al.* SPECT imaging of dopamine transporters in human brain with iodine-123-fluoroalkyl analogs of beta-CIT. *J Nucl Med* **37**, 1129-1133 (1996).
- 139 De Keyser, J., De Backer, J., Ebinger, G. & Vauquelin, G. [3H]GBR 12935 binding to dopamine uptake sites in the human brain. *J Neurochem* **53**, 1400-1404 (1989).
- 140 Van Laere, K. *et al.* Experimental performance assessment of SPM for SPECT neuroactivation studies using a subresolution sandwich phantom design. *Neuroimage* **16**, 200-216, doi:10.1006/nimg.2001.1047 (2002).
- 141 Hotelling, H. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology* **24**, 417-441 and 498-520 (1933).
- 142 Pearson, K. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine* **2**, 559-572 (1901).
- 143 Hamilton, D., List, A., Butler, T., Hogg, S. & Cawley, M. Discrimination between parkinsonian syndrome and essential tremor using artificial neural network classification of quantified DaTSCAN data. *Nucl Med Commun* **27**, 939-944, doi:10.1097/01.mnm.0000243369.80765.24 (2006).
- 144 Kahraman, D., Eggers, C., Schicha, H., Timmermann, L. & Schmidt, M. Visual assessment of dopaminergic degeneration pattern in 123I-FP-CIT SPECT differentiates patients with atypical parkinsonian syndromes and idiopathic Parkinson's disease. *J Neurol* **259**, 251-260, doi:10.1007/s00415-011-6163-1 (2012).

- 145 Kim, Y. *et al.* Combination of dopamine transporter and D2 receptor SPECT in the diagnostic evaluation of PD, MSA, and PSP. *Mov Disord* **17**, 303-312 (2002).
- 146 Plotkin, M. *et al.* Imaging of dopamine transporters and D2 receptors in vascular parkinsonism: a report of four cases. *J Neural Transm* **112**, 1355-1361, doi:10.1007/s00702-005-0278-4 (2005).
- 147 Koch, W., Mustafa, M., Zach, C. & Tatsch, K. Influence of movement on FP-CIT SPECT quantification: a Monte Carlo based simulation. *Nucl Med Commun* **28**, 603-614, doi:10.1097/MNM.0b013e328273bc6f (2007).
- 148 Koch, W., Radau, P., Hamann, C. & Tatsch, K. Clinical testing of an optimized software solution for an automated, observer-independent evaluation of dopamine transporter SPECT studies. *J Nucl Med* **46**, 1109-1118 (2005).
- 149 van Dyck, C. *et al.* Human biodistribution and dosimetry of the SPECT D2 dopamine receptor radioligand [123I]IBF. *Nucl Med Biol* **23**, 9-16 (1996).
- 150 Lavalaye, J., Booi, J., Reneman, L., Habraken, J. & van Royen, E. Effect of age and gender on dopamine transporter imaging with [123I]FP-CIT SPET in healthy volunteers. *Eur J Nucl Med* **27**, 867-869 (2000).
- 151 Best, S. *et al.* Striatal dopamine transporter availability with [123I]beta-CIT SPECT is unrelated to gender or menstrual cycle. *Psychopharmacology (Berl)* **183**, 181-189, doi:10.1007/s00213-005-0158-5 (2005).
- 152 Buchert, R. *et al.* IBZM tool: a fully automated expert system for the evaluation of IBZM SPECT studies. *Eur J Nucl Med Mol Imaging* **33**, 1073-1083, doi:10.1007/s00259-006-0067-9 (2006).
- 153 Fleming, J., Bolt, L., Stratford, J. & Kemp, P. The specific uptake size index for quantifying radiopharmaceutical uptake. *Phys Med Biol* **49**, N227-234 (2004).
- 154 Costa, D. *et al.* in *Annual Congress of the European Association of Nuclear Medicine*. PS-342.
- 155 Lokkegaard, A., Werdelin, L. & Friberg, L. Clinical impact of diagnostic SPET investigations with a dopamine re-uptake ligand. *Eur J Nucl Med Mol Imaging* **29**, 1623-1629, doi:10.1007/s00259-002-0938-7 (2002).
- 156 Goethals, I., Dobbeleir, A., Ham, H., Santens, P. & D'Asseler, Y. Validation of a resolution-independent method for the quantification of 123I-FP-CIT SPECT scans. *Nucl Med Commun* **28**, 771-774, doi:10.1097/MNM.0b013e32829152c9 (2007).
- 157 Costa, D., Aston, P. & Gacinovic, S. in *Annual Congress of the European Association of Nuclear Medicine*. S103.
- 158 Pigden, I., Viljoen, M., Costa, D. & Gacinovic, S. in *Annual Congress of the European Association of Nuclear Medicine*. S112.
- 159 Talairach, J. & Tournoux, P. *Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: an approach to cerebral imaging*. (Thieme, 1988).
- 160 Eisenhauer, E. *et al.* New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* **45**, 228-247, doi:10.1016/j.ejca.2008.10.026 (2009).
- 161 Staff, R. *et al.* Shape analysis of 123I-N-omega-fluoropropyl-2-beta-carbomethoxy-3beta-(4-iodophenyl) nortropine single-photon emission computed tomography images in the assessment of patients with

- parkinsonian syndromes. *Nucl Med Commun* **30**, 194-201, doi:10.1097/MNM.0b013e328314b863 (2009).
- 162 Kuikka, J. *et al.* Imaging the structure of the striatum: a fractal approach to SPECT image interpretation. *Physiol Meas* **19**, 367-374 (1998).
- 163 Friston, K. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. (Academic Press, 2006).
- 164 Wang, J. *et al.* 18F-FP-CIT PET imaging and SPM analysis of dopamine transporters in Parkinson's disease in various Hoehn & Yahr stages. *J Neurol* **254**, 185-190, doi:10.1007/s00415-006-0322-9 (2007).
- 165 Hanley, J. & McNeil, B. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **143**, 29-36 (1982).
- 166 Varrone, A. *et al.* European multicentre database of healthy controls for [(123)I]FP-CIT SPECT (ENC-DAT): age-related effects, gender differences and evaluation of different methods of analysis. *Eur J Nucl Med Mol Imaging*, doi:10.1007/s00259-012-2276-8 (2012).
- 167 van Dyck, C. *et al.* Age-related decline in dopamine transporters: analysis of striatal subregions, nonlinear effects, and hemispheric asymmetries. *Am J Geriatr Psychiatry* **10**, 36-43 (2002).
- 168 Tossici-Bolt, L. *et al.* Calibration of gamma camera systems for a multicentre European (1)(2)(3)I-FP-CIT SPECT normal database. *Eur J Nucl Med Mol Imaging* **38**, 1529-1540, doi:10.1007/s00259-011-1801-5 (2011).
- 169 Koopman, K., la Fleur, S., Fliers, E., Serlie, M. & Booij, J. Assessing the optimal time point for the measurement of extrastriatal serotonin transporter binding with 123I-FP-CIT SPECT in healthy, male subjects. *J Nucl Med* **53**, 1087-1090, doi:10.2967/jnumed.111.102277 (2012).
- 170 The Parkinson Progression Marker Initiative (PPMI). *Prog Neurobiol* **95**, 629-635, doi:10.1016/j.pneurobio.2011.09.005 (2011).
- 171 Segovia, F., Górriz, J., Ramírez, J. & Álvarez..., I. Improved Parkinsonism diagnosis using a partial least squares based approach. *Medical Physics* **39**, 4395 (2012).

Appendix

The attached document shows the copyright application/certificate for use of copyrighted material within this thesis as detailed on page 11.

WOLTERS KLUWER HEALTH LICENSE TERMS AND CONDITIONS

Nov 20, 2012

This is a License Agreement between David J Towey ("You") and Wolters Kluwer Health ("Wolters Kluwer Health") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Wolters Kluwer Health, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3032980330352
License date	Nov 20, 2012
Licensed content publisher	Wolters Kluwer Health
Licensed content publication	Nuclear Medicine Communications
Licensed content title	Automatic classification of 123I-FP-CIT (DaTSCAN) SPECT images
Licensed content author	David Towey, Peter Bain, and Kuldip Nijran
Licensed content date	Jan 1, 2011
Volume Number	32
Issue Number	8
Type of Use	Dissertation/Thesis
Requestor type	Individual
Title of your thesis / dissertation	SPECT imaging and automatic classification methods in movement disorders
Expected completion date	Nov 2012
Estimated size(pages)	300
Billing Type	Invoice
Billing address	28 the nave Laindon Basildon, Essex SS15 5NW United Kingdom
Customer reference info	
Total	0.00 GBP
Terms and Conditions	

Terms and Conditions

1. A credit line will be prominently placed and include: for books - the author(s), title of book, editor, copyright holder, year of publication; For journals - the author(s), title of article, title of journal, volume number, issue number and inclusive pages.

2. The requestor warrants that the material shall not be used in any manner which may be considered derogatory to the title, content, or authors of the material, or to Wolters Kluwer.
3. Permission is granted for a one time use only within 12 months from the date of this invoice. Rights herein do not apply to future reproductions, editions, revisions, or other derivative works. Once the 12-month term has expired, permission to renew must be submitted in writing.
4. Permission granted is non-exclusive, and is valid throughout the world in the English language and the languages specified in your original request.
5. Wolters Kluwer cannot supply the requestor with the original artwork or a "clean copy."
6. The requestor agrees to secure written permission from the author (for book material only).
7. Permission is valid if the borrowed material is original to a Wolters Kluwer imprint (Lippincott-Raven Publishers, Williams & Wilkins, Lea & Febiger, Harwal, Igaku-Shoin, Rapid Science, Little Brown & Company, Harper & Row Medical, American Journal of Nursing Co, and Urban & Schwarzenberg - English Language).
8. If you opt not to use the material requested above, please notify Rightslink within 90 days of the original invoice date.
9. Please note that articles in the ahead-of-print stage of publication can be cited and the content may be re-used by including the date of access and the unique DOI number. Any final changes in manuscripts will be made at the time of print publication and will be reflected in the final electronic version of the issue.
 Disclaimer: Articles appearing in the Published Ahead-of-Print section have been peer-reviewed and accepted for publication in the relevant journal and posted online before print publication. Articles appearing as publish ahead-of-print may contain statements, opinions, and information that have errors in facts, figures, or interpretation. Accordingly, Lippincott Williams & Wilkins, the editors and authors and their respective employees are not responsible or liable for the use of any such inaccurate or misleading data, opinion or information contained in the articles in this section.
10. This permission does not apply to images that are credited to publications other than Wolters Kluwer journals. For images credited to non-Wolters Kluwer journal publications, you will need to obtain permission from the journal referenced in the figure or table legend or credit line before making any use of the image(s) or table(s).
11. Other Terms and Conditions:

v1.4

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK500900916. Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

**Make Payment To:
 Copyright Clearance Center
 Dept 001
 P.O. Box 843006
 Boston, MA 02284-3006**

For suggestions or comments regarding this order, contact RightsLink Customer Support: customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

Gratis licenses (referencing \$0 in the Total field) are free. Please retain this printable

license for your reference. No payment is required.
