This electronic thesis or dissertation has been downloaded from the King’s Research Portal at https://kclpure.kcl.ac.uk/portal/

The impact of respiratory motion correction methods on tumour detection and quantification in Positron Emission Tomography

Polycarpou, Irene

Awarding institution:
King's College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENCE AGREEMENT

Unless another licence is stated on the immediately following page this work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International licence. https://creativecommons.org/licenses/by-nc-nd/4.0/

You are free to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 07. Jan. 2025

The impact of respiratory motion correction methods on tumour detection and quantification in Positron Emission Tomography

IRENE POLYCARPOU

DIVISION OF IMAGING SCIENCES & BIOMEDICAL ENGINEERING
KING'S COLLEGE LONDON

A THESIS SUBMITTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
2014

To my beloved family...

Στην αγαπημένη μου οικογένεια...

Table of contents

List of figures

List of tables

List of equations

Acronyms

Symbols

2D	Two dimensional
3D	Three dimensional
4D	Four dimensional

C

CHO	Channelized Hotelling observer
CR	Contrast Recovery
CT	Computed Tomography

E

EM	Expectation Maximisation

F

^{18}F	Fluorine-18
FBP	Filtered Back Projection
FDG	2-deoxy-2-(18F)fluoro-D-glucose

FOV	Field Of View
FWHM	Full Width at Half Maximum

H

HD	High Definition
HO	Hotelling observer

L

LOR	Line of Response

M

MAP	Maximum A Posteriori
MCIR	Motion Compensated Image Reconstruction
ML	Maximum Likelihood
MLEM	Maximum Likelihood Expectation Maximisation
MRI	Magnetic Resonance Imaging
MSE	Mean Squared Error

N

NMR	Nuclear Magnetic Resonance
NPWMF	Non prewhitening matched filter

O

OP	Ordinary Poisson
OSEM	Ordered Subsets Expectation Maximisation
OSL	One Step Late

P

PCA	Principle Component Analysis
PET	Positron Emission Tomography
PMT	Photomultiplier Tube
PSF	Point Spread Function
PVE	Partial Volume Effect
PWMF	Prewhitening matched filter

R

RF	Radio Frequency
ROC	Receiver Operating Characteristic Curve
ROI	Region Of Interest
RPM	Real-time Position Management
RTA	Reconstruct-Transform-Average

S

SNR	Signal to Noise Ratio

SPECT	Single Photon Emission Computed Tomography
SUV	Standardized Uptake Value

T

TFE	Turbo Field Echo
TOF	Time of Flight

U

US	Ultrasound
UTE	Ultra-Short Echo Time
UV	Ultraviolet

Thesis Abstract

Respiratory motion degrades PET spatial resolution and image quality limiting the potential benefits from increased resolution. Motion correction is complicated by limitations of the poor statistical quality of PET data and there is still not a robust method available. Motion correction can be implemented at different stages of data processing either during or after reconstruction and may result in substantial improvements in image quality. The recent development of whole body PET-MRI scanners might provide a potential solution for motion correction since internal organ motion could be measured concurrently with PET using MRI. However, although there have been various proposed methods for motion correction, there is not sufficient evidence in the current literature to answer which method is better in clinical practice and investigating the impact of motion correction on lesion detectability.

The aim of this thesis is to assess respiratory motion correction in terms of its quantitative accuracy and detectability performance to determine its potential for improved and early cancer diagnosis. This thesis is based on numerical 4D simulated PET data using real MRI data. Motion correction is investigated based on MRI-derived motion fields as could be obtained from a simultaneous PET-MRI acquisition. As a first step, this thesis aims to understand the behaviour of different approaches to motion-corrected image reconstruction in terms of convergence rate and the properties of the reconstructed images obtained. This thesis then deals with the impact of respiratory motion on lesion detectability. A comprehensive assessment is performed using different amplitudes for lesion displacement due to respiration and different respiration patterns derived from actual patient respiratory traces. The impact on the detectability is compared with that achievable by a higher resolution scanner in order to investigate the importance of correcting for motion to realise the benefit from the increased resolution of future PET scanners.

Acknowledgements

First of all, I am extremely grateful to my supervisors Prof Paul Marsden and Dr Charalampos Tsoumpas for their guidance and support throughout all these years. I am truly thankful to them for sharing their expertise in the field and mentorship which were vital in pursuing this thesis.

I wish to acknowledge all the people that have provided assistance and support throughout this thesis. Special thanks to Dr Andy King for providing the motion modelling used for the simulations and Dr Christian Buerger for providing the real MR data. I am grateful to Dr Miho Shidahara for technical support on the channelized Hotelling observer, Dr Anastasia Chalkidou for performing the FDG uptake values measurements from clinical PET/CT acquired data, Dr Paul Schleyer for providing the respiratory signal data and Dr Andre Salomon for discussions on simulation data realism. I would also like to thank Dr Georgios Angelis and Dr Kris Thielemans for giving advice on many aspects of my work. Finally, I would like to thank Dr Lefteris Livieratos, Dr Stelios Christofides and Mr Georgios Soultanidis for our valuable discussions.

I dedicate this thesis to my family who have always been a tremendous source of encouragement throughout my life and without their endless love and support I would not have the goals I would have to strive.

Chapter 1

Introduction

1.1 Positron emission tomography (PET) in oncology

Across the world a million people die every year from lung and liver cancer (Ferlay *et al.* 2010). Early detection and accurate staging of disease, requiring accurate detection and localisation of small cancerous lesions, is important for more successful and more cost-efficient treatment. Studies have reported prolonged and improved survival of patients who have undergone therapy at the earliest stage of cancer in the lungs (Henschke *et al.* 2003) and liver (Adam and Vinet 2004).

Medical imaging techniques such as positron emission tomography (PET) play an essential role in the management of cancer (Hillner *et al.* 2008). PET is a non-invasive molecular imaging technique which provides images of biochemical processes in a patient by detecting the radiation emitted from radiolabelled tracer molecules injected into the patient. PET is now widely established in oncology as a powerful tool not only for diagnosis and staging (i.e. determine how far the disease has progressed) but also for radiotherapy planning and measuring the recurrence and progression of cancer to assess therapy response (Vansteenkiste and Stroobants 2001; Rohren *et al.* 2004).

PET image quality in clinical oncology is usually characterised based on two tasks:

i. How well lesions can be distinguished from normal tissues?

ii. How well the true tracer uptake can be recovered?

Both tasks depend on the quality of the images. For the first task a lesion is usually distinguished from normal tissues by a human observer. However, in cases where the use of human observers is not practical such as when various parameters are investigated numerical observers have been developed that predict human detection performance such as the channelized Hotelling observer (CHO). The CHO is based on the ability of the human eye to discriminate certain frequencies. It transforms the image from the spatial domain to the frequency domain and extracts the various frequencies using appropriate frequency bandwidths channels. Finally, using the statistical properties of the images the probability of detectability is computed. For the second task, after identifying a region of higher tracer uptake than the surrounding tissues, the region is usually evaluated by estimating the tracer uptake in a region. An index often used to characterise the accumulation of the tracer is the standardized uptake value (SUV). SUV is defined as the concentration of activity in the region of interest (ROI) divided by the average activity concentration in the patient. SUV is used widely clinically especially for staging and quantification of tumour response to therapy (Erdi *et al.* 2000; Weber *et al.* 2003). For example, accurate quantification is important to distinguish between malignant and benign pulmonary nodules.

Depending on the type of radiotracer injected, PET can provide information on different biochemical functions. In clinical practice the most common radiotracer currently used is a glucose analogue labelled with the radioisotope fluorine - 18 called fluorodeoxyglucose (^{18}F-FDG) or simply FDG. ^{18}F-FDG accumulates in regions undergoing high metabolic activity such as the brain, liver and malignant tumours, where rate of uptake is related to glucose metabolism. Consequently, accumulation (uptake) of ^{18}F-FDG by the tissue is directly related to its metabolic state (normal or hyper or hypo-metabolic), and increased uptake may indicate the presence of malignant tumour cells. Thus, PET has the ability to non-invasively detect functional changes in vivo.

1.2 Hybrid PET/MRI imaging

However, PET lacks anatomical information. Combination of the functional information offered by PET with the information offered by anatomical imaging modalities such as computed tomography (CT) or magnetic resonance imaging (MRI) allows localization of the increased radioactivity distribution. The first combination of PET with CT was developed in 2000 (Beyer *et al.* 2000) and has been proved effective in increasing the accuracy of diagnosis and cancer management compared with PET as a stand alone imaging modality (Antoch *et al.* 2003). CT data are acquired sequentially with PET data, there is an increase in the patient radiation dose and CT has relatively poor ability to differentiate between soft tissues. Compared with CT, MRI is advantageous in terms of better soft tissue contrast with no extra patient radiation dose. Simultaneous PET/MRI acquisition allows spatial and temporal correlation ensuring that the patient is being imaged in the same physiological state during both PET and MRI.

1.3 Motion blurring limitations and challenges in ^{18}FDG PET

PET acquisition usually requires 3 to 5 min per bed position (approximately 15 cm axially) and so there is patient internal organ motion such as cardiac contraction and respiration. Cardiac contraction affects mainly the heart and is a major concern of cardiac imaging applications rather than for oncologic imaging as is the focus of this project. On the other hand, respiration is more complex as it affects multiple organs such as the liver and the heart since they are being pushed and pulled by the diaphragm and other connected tissues.

For normal breathing, the motion of the diaphragm due to respiration is in the range of 15 to 20 *mm* (Schwarz and Leach 2000). Considering displacement of lesions in the lung and liver, a number of investigations reported that the displacement caused by respiratory motion is usually less than 10 *mm* for lung lesions but it may be more than 20 *mm* for lung and liver lesions close to the diaphragm (Seppenwoolde *et al.* 2002; Erdi *et al.* 2004). Motion of that range was also reported by a review of respiratory motion provided by Langen and Jones (2001). Considering that conventional PET scanners have a spatial resolution of approximately 5 to 6 *mm* motion of that range is larger than the resolution and therefore is expected to induce blurring and distortions in the images limiting the image spatial resolution (Xu *et al.* 2011; Nehmeh *et al.* 2004; Visvikis *et al.* 2006). Moreover, with the advances in resolution of future scanners (Schäfers 2008a) the effect of motion is expected to be even more substantial (Daou 2008).

Motion blurring can cause underestimation of the tracer uptake and reduce tumour detectability (Nehmeh *et al.* 2002; Liu *et al.* 2009). An example of the effects of respiratory motion is shown on simulated images from a 5 *min* PET acquisition with and without motion in figure 1.1. Considering that the biggest clinical application of PET in oncology is to detect tumour metastases for staging, and to quantify the tracer uptake it is important to investigate if the correction of motion can affect the image interpretation and clinical decision making based on these two aspects of image quality. Based on these, respiratory motion blurring contributes to the clinical challenges of detecting a small tumour or detecting a tumour with low uptake.

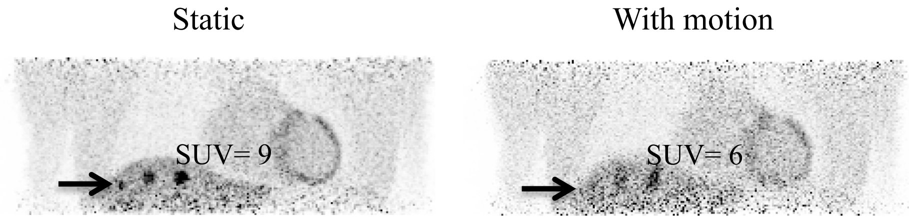


Figure 1.1: Effects of respiratory motion on simulated PET images: Motion blurring reduces lesion detectability (i.e. arrows point out a lesion clearly visble in image without motion but hardly detectable in image with motion) and causes underestimation of the tracer uptake (i.e. as measured by SUV).

1.4 Motion correction in ^{18}FDG PET

A common method to minimize respiratory motion effects is the use of respiratory synchronized acquisition (i.e., gating) which relates the motion signal based on respiratory amplitude or phase yielding multiple gates. However, each gate contains only a proportion of the measured data resulting in reconstructed images with lower signal-to-noise ratio (SNR) than using all data. Thus, a technique accounting for motion effects which uses all measured data in order to improve the SNR is expected to provide superior results. One approach to motion correction of gated data consists of two steps, (i) estimation of motion and (ii) the incorporation of motion information into the reconstruction of the gated data.

The motion fields are usually estimated by appropriate registration of the gated data with a selected reference image. To estimate the motion fields CT, MRI, or from PET images can be used. The problems encountered with CT is that the motion can be different for sequential CT and PET acquisitions. This aspect could be handled by measuring motion directly from PET, but PET data are noisy and in addition it would be difficult to generalize this technique to any radiotracer. The development of whole-body simultaneous PET-MRI might enable concurrent measurement of the internal organ motion with MRI. Accurate measurement of motion is one of the motivations for combining MRI with PET.

After estimating the motion information, there are two popular motion correction approaches - apply the estimated motion information either (i) after or (ii) during the reconstruction of the gated images known as reconstruct-transform-average (RTA) and motion compensated image reconstruction (MCIR) respectively. RTA is based on independent reconstructions of each gate using conventional algorithms which are then transformed to a reference position and averaged. In MCIR, all gated data are used, incorporating motion information within reconstruction. Current PET scanners give the reconstructed images directly and any correction (e.g. attenuation, scatter) is applied in the reconstruction. Therefore, the user does not have usually the ability to reconstruct data from the scanner and the only solution is to correct the already reconstructed images, using the RTA method. However, as MCIR is potentially a better approach, it is important to see what are the relative disadvantages and advantages of the two methods.

The first aim of this thesis addresses whether it is advantageous to perform motion correction inside the reconstruction and when (i.e. reconstruction parameters) do we need to do this by evaluating the impact of both methods on quantification and noise propagation.

The second part of this thesis deals with the important issue in clinical oncology of lesion detectability. It investigates if, when and to what extent motion correction implemented within PET reconstruction improves lung and liver lesion detectability for lesions of various sizes and signal to background ratios. Using different respiratory signals derived from actual patients, the impact of the breathing pattern on detectability was also determined. The impact of respiratory motion correction on detectability is further investigated as a function of the scanner spatial resolution.

1.5 Thesis objectives and structure

The specific objectives of this thesis are the following:

i. to understand the impact of motion correction methods on quantification by accurately characterizing the image bias and variance

ii. to understand the impact of motion blurring and motion-correction on detecting tumours as a function of scanner resolution, breathing pattern, motion amplitude and tumour size, tracer uptake

In the first part of chapter 2, a brief introduction to the basic principles of PET is given and its clinical applications in oncology are presented. Then the reader is introduced to the concepts and terminology used during data acquisition, storage, correction and reconstruction of PET data. Furthermore, chapter 2 describes the impact of image quality on the main tasks of PET in oncology and the basic principles of assessing image quality; the concept of image quality is fundamental for the understanding of this thesis. This is followed by a brief description of the main factors that degrade PET images and a general overview of the respiratory motion correction problem. Finally, chapter 2 reviews the most widely used approaches for PET motion correction.

Chapter 3 describes in detail the methodology used for simulating realistic numerical 4D PET data based on real segmented MRI dynamic data as used in this thesis. Two methodologies are presented simulating either 4D gated-PET datasets or real-time 4D PET datasets.

Chapter 4 presents an evaluation of the relative impact of implementing motion correction after or during image reconstruction. The two approaches are evaluated statistically in terms of convergence rate and the properties of the reconstructed images are obtained. The causes of any differences between the methods are also evaluated and presented. This chapter also demonstrates how overall error can be examined in terms of different components, and bias and noise trade-offs can be evaluated at voxel-level.

Chapter 5 investigates the impact of respiratory motion and spatial resolution of PET imaging systems on liver and lung lesion detectability. The results of a comprehensive assessment are presented by investigating different amplitudes for the lesion displacement due to respiration and different respiration patterns derived from actual patient respiratory traces. The impact of different lesion sizes and lesion to background ratios is also investigated.

Finally, chapter 6 summarises the findings and outcomes of this thesis and outlines potential areas for future research.

1.6 Contributions in the thesis

This thesis is the work of the candidate - specific areas where others have contributed specialist input are listed below. The open source software library for Tomographic Image Reconstruction (STIR) (Thielemans *et al.* 2012) was used for the data simulations and reconstructions (*http://stir.sourceforge.net/*).

The algorithm used for motion modelling creation was provided by Dr Andy King (King *et al.* 2011). A freely available registration algorithm (Buerger *et al.* 2011) was used (*https://www.isd.kcl.ac.uk/internal/hyperimage/*). The respiratory signals have been provided by Dr Paul Schleyer.

The MR segmentations were performed by the candidate manually using intensity thresholds to extract the region of interest. Although the implementation of RTA and MCIR methods was developed by the supervisor Dr Charalampos Tsoumpas optimization, debugging as well as thoroughly evaluation have been performed by the candidate. A methodology was also developed by the candidate to calculate values such as bias, standard deviation and mean squared error at the voxel-level using bash scripts. Finally, the Channelized Hotelling observer was developed by the candidate with the advice of Dr Miho Shidahara as described in detail in chapter 5.

1.7 List of publications

Peer-reviewed journal publications

[1] **Polycarpou I**, Tsoumpas C, King A and Marsden P K 2014, Impact of respiratory motion correction and spatial resolution on lesion detection in PET: A simulation study based on real MR dynamic data, *Physics in Medicine and Biology* **59**(3) pp. 697–713,

[2] Kotasidis F A, Tsoumpas C, **Polycarpou I** and Zaidi H, 2014 A 5D Computational Phantom for Pharmacokinetic Simulation Studies in Dynamic Emission Tomography, *Computerized Medical Imaging and Graphics* (in press)

[3] Tsoumpas C, **Polycarpou I**, Thielemans K, Green Buerger K, King A, Schaeffter T and Marsden P K 2011, The effect of regularisation in motion compensated PET image reconstruction: A realistic numerical 4D simulation study, *Physics in Medicine and Biology*, **58**(6) pp. 1759–1773

[4] **Polycarpou I**, Tsoumpas C and Marsden P K 2012, Analysis and comparison of two methods for motion correction in PET imaging, *Medical Physics*, **39**(10) pp. 6474–6483

[5] **Polycarpou I**, Thielemans K, Manjeshwar R, Aguiar P, Marsden P K and Tsoumpas

C 2011, Comparative evaluation of scatter correction in 3D PET using different scatter-level approximations, *Annals of Nuclear Medicine*, **25**(9) pp. 643–649

Conference records and abstracts

[1] **Polycarpou I**, Tsoumpas C and Marsden P K, 2013 Effect of respiratory motion correction and spatial resolution on tumor detection in PET , *in 2013 IEEE Nuclear Science Symposium*, Seoul, Korea, (oral)

[2] Kotasidis F A, **Polycarpou I**, Tsoumpas C and Zaidi H, 2013 A 5D Anthropomorphic numerical phantom for respiratory-gated parametric imaging simulation studies in dynamic emission tomography, *in 2013 IEEE Nuclear Science Symposium*, Seoul, Korea, (poster)

[3] Soultanidis G, **Polycarpou I**, Weissler B, Lerche C W, Mackewn J, Tsoumpas C, Schulz V and Marsden P. K., 2013 Demonstration of motion correction for PET-MR with PVA cryogel phantoms, *in 2013 IEEE Nuclear Science Symposium Conference Record*, Seoul, Korea, (poster)

[4] Balfour D, **Polycarpou I**, Marsden P K and King A, 2013 Respiratory motion correction of PET imaging using a combination of PET and MR data, *in 2013 IEEE Nuclear Science Symposium*, Seoul, Korea, (poster)

[5] **Polycarpou I**, Tsoumpas C and Marsden P K, 2012 Localization ROC analysis of the impact of respiratory motion blurring and correction on lesion detection in PET: A simulation study based on real MR dynamic data, *in 2012 IEEE Nuclear Science Symposium Conference Record*, Anaheim, California (poster)

[6] **Polycarpou I**, Tsoumpas C and Marsden P K, 2011 Statistical evaluation of PET

motion correction methods using MR derived motion fields, *in 2011 IEEE Nuclear Science Symposium Conference Record*, Valencia, Spain, pp. 3579–3585 (oral)

[7] Tsoumpas C, **Polycarpou I**, Buerger C, Schaeffter T and Marsden P K, 2011 The effect of regularization on image quality and quantification in motion compensated PET image reconstruction, *in 2011 IEEE Nuclear Science Symposium Conference Record*, Valencia, Spain (poster)

[8] **Polycarpou I**, Marsden P K and Tsoumpas C 2011 Evaluation of two approaches to motion-corrected PET image reconstruction, *Journal of Physics: Conference Series* 317 012001 (oral)

[9] **Polycarpou I**, Marsden P K and Tsoumpas C 2011 A comparative investigation of scatter correction in 3D PET, *Journal of Physics: Conference Series* 317 012022 (oral)

[10] **Polycarpou I**, Thielemans K, Manjeshwar R, Marsden P K and Tsoumpas C, 2010 Scatter correction in 3D PET using STIR, *in 2010 IEEE Nuclear Science Symposium Conference Records*, Knoxville, Tennessee (poster)

[11] Marsden P K, Tsoumpas C, MacKewn J, Schleyer P, **Polycarpou I**, Schaeffter T, King A, Buerger C, Vandenberghe S, Mollet P, Keereman V and Schultz V, 2010 Challenges of patient motion correction using simultaneous PET and MRI, *MR-PET Workshop*, Juelich, Germany (oral)

Chapter 2

Background and basic principles

2.1 Positron emission tomography (PET) imaging

2.1.1 Basic principles

Wrenn *et al.* (1951) introduced the concept of using positron annihilation in medicine to localise radioactive substances in brain tumours while a few years later it has been applied in the first positron scanner by Brownell and Sweet (1953). The first ring-based functional positron emission tomography (PET) for brain imaging was introduced by Ter-Pogossian *et al.* (1975), and Phelps *et al.* (1975) published the basic principles of PET as we know them today. PET is now established in medical imaging as a non-invasive modality that produces functional images of the human body (e.g. blood flow, receptor density and glucose metabolism) based on detection of photons originating from annihilations of positrons emitted by β^+ radioisotopes.

During PET imaging, a radiotracer solution is injected into the patient and then each molecule travels into the body where it decays mainly by emitting a positron. The emitted positron travels through tissue a short distance interacting with the surrounding atoms until almost complete loss of its kinetic energy when it interacts with an electron and annihilates producing a pair of high energy (i.e. 511 keV) photons (figure 2.1). The two photons travel in opposite directions and can be detected by a circular array of scintillation detectors surrounding the human body - the line that connects the two opposite detectors is known as line of response (LOR). The total number of detected events between the two opposite detectors is proportional to the total number of positron emissions along the LOR.

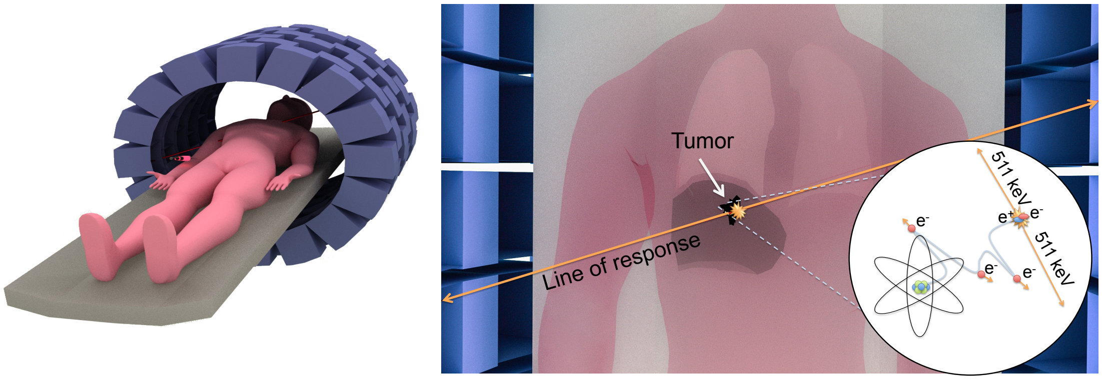


Figure 2.1: Basic principles of PET imaging: After injection of the radiotracer into the human body, each emitted positron travels through the tissue and annihilates with an electron producing two 511 keV photons which are detected by the PET scanner.

The most readily available β^+ radioisotopes used for PET imaging are ^{18}F, ^{15}O, ^{11}C and ^{13}N. The radioisotope is bound on a metabolically active biomolecule, chemical compounds, which have been designed to interact with molecules and proteins within the human body. In particular, these biomolecules are targeted at various biochemical functions in a human body (i.e. energy consumption, e.g. Alavi *et al.* 1986). Examples of such compounds are glucose, glucose analogues, water, or ammonia. Such labelled biomolecules are known as radiotracers. The radiotracer is intravenously injected into the patient in tracer quantities and a PET image is reconstructed which shows the radiotracer spatial distribution within a region of interest.

Depending on the type of radiotracer injected, PET can provide information on different biochemical functions. In clinical practice the most common radiotracer currently used is the glucose analogue labelled with the radioisotope fluorine - 18 called fluorodeoxyglucose (^{18}F-FDG) or simply FDG. ^{18}F-FDG accumulates in regions undergoing high metabolic activity such as the brain, liver and malignant tumours, where rate of uptake is related to glucose metabolism. Consequently, accumulation (uptake) of ^{18}F-FDG by the tissue is directly related to its metabolic state (normal or hyper or hypo-metabolic), and increased uptake may indicate the presence of malignant tumour cells. Thus, PET has the ability to non-invasively detect functional changes in vivo (Jones 2002).

Compared with other imaging modalities such as SPECT, PET may detect abnormalities with higher scanner sensitivity (i.e. fraction of the events generated that are detected) (Higashi *et al.* 2001). The high scanner sensitivity of PET may allow the detection and characterization of tumour as benign or malignant based on metabolic function. PET is a widely used technique in oncology for detecting tumour metastases (or staging) and quantifying the radiotracer accumulation. In addition to oncology it provides important clinical and research insights in cardiology, neurology, pharmacology and drug development.

Most currently available PET scanners are constructed from multiple rings of scintillation detector blocks surrounding the patient. The symmetrical pattern of the PET scanners allows annihilation photons emitted in opposite directions to be detected simultaneously. An example of a typical PET scanner is illustrated in figure 2.2. A detector block is comprised of an array of scintillation detectors optically coupled to photo sensors such as photomultiplier tubes (PMTs) which are the most frequently used photo sensors in PET (Casey and Nutt 1986).

Incoming 511 keV photons interact in the crystals and are converted to visible or ultraviolet (UV) photons which are subsequently converted to electrons by means of a photocathode in the PMTs and amplified to produce a signal proportional to the energy deposited in the crystal. The signal is then analysed to determine in which crystal the incoming photon was absorbed. Typically, due to size constraints, it is not possible to couple each crystal with one PMT and therefore most PET systems employ block detectors in which a block of crystal is segmented into an array of smaller crystal elements. A typical block detector is segmented into 8 × 8 detector elements and is read out by 4 PMTs as shown in figure 2.2. However, for higher resolution PET systems a larger number of detector elements has been used such as the Siemens Biograph PET scanner which has detector blocks consisted of 13 × 13 detector elements and read out by 4 PMTs (Zaidi and Alavi 2007a).

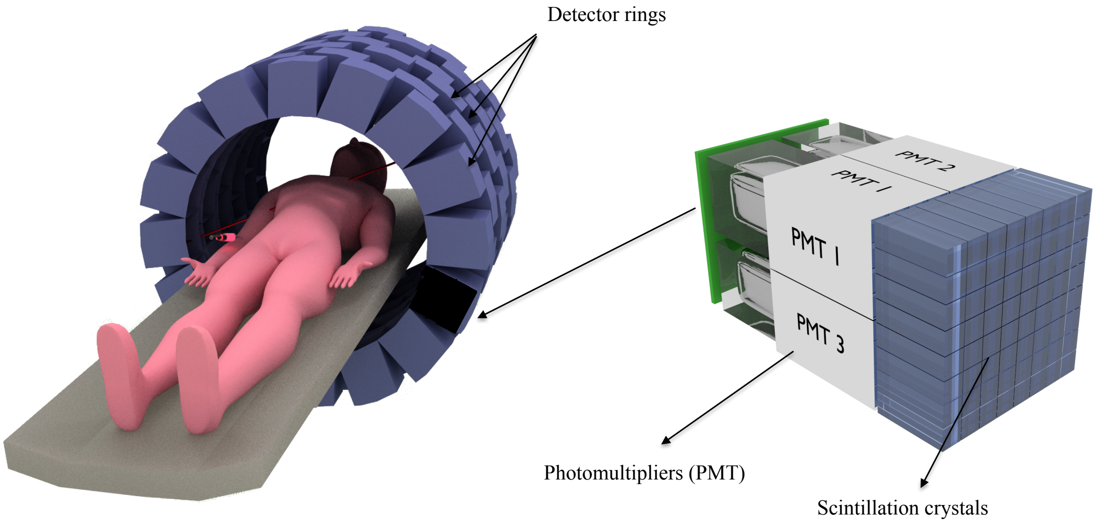


Figure 2.2: A typical PET scanner constructed from multiple rings of scintillation detector blocks. A typical PET block detector consists of scintillation crystals coupled to 4 photomultiplier tubes (PMTs). When a photon hits the crystals the light is distributed among the 4 PMTs based on the position of the crystal.

The choice of the scintillation detector is based on specific characteristics such as high stopping power, high light output and short decay time (Melcher 2000). High stopping power is required in order to minimize the distance that a photon travels until it stops after deposition of its full energy in the crystal and therefore maximize the detection of photons. For high stopping power the scintillation material must have high density and high effective atomic number. The number of visible or UV photons emitted per unit of energy deposited in the scintillation crystal, known as light output, must also be high for better localization and energy discrimination between the detected photons. Finally, short scintillation decay time is desirable to increase the ability of the detector to detect photons at high photon rates and therefore minimize dead-time count losses while a narrow coincidence timing window can be used increasing the ability to distinguish true from random coincidences. Scintillation crystals with a short decay time are often called fast scintillation crystals.

When two photons hit two opposite detectors they are analysed by a coincidence logic circuit and processor that compare the time of each event with the time of events recorded in the opposite detectors to identify if they originate from the same positron annihilation. In order to record an annihilation event, both annihilation photons need to be detected. A coincidence event is assumed to have occurred when a pair of events are recorded within a specified coincidence timing window which is essential due to the different distances from the annihilation event that the two photons have to travel and also for differences in the transformation of each signal through cables and electronics (Townsend *et al.* 1989). Also a coincidence event is valid if the energy deposited in the crystal (i.e. pulse height) by both annihilation-photons is within the selected energy window. For each accepted coincidence event, computer processing is used to assess the geometry of the active volume of the two opposite detectors to define the LOR, along which the annihilation event is assumed to have occurred. Another requirement for the acceptance of a coincidence event is that the formed LOR must be within a valid acceptance angle.

To identify two detected photons as coincidence events and record an LOR, the two photons must have energy within a specific energy window (e.g. 450-550 keV) and the time difference between their detection to be within a certain timing window (e.g. few nanoseconds wide). The total number of detected events between the two opposite detectors should be proportional to the total number of positron emissions along the LOR. Consequently, by detecting many coincidence events[¶] the activity distribution of a region can be estimated.

During PET acquisition the measured data are derived from nuclear decays, a random process that results in statistically variable data and therefore a PET image is an average of independent measurements over the duration of the acquisition. The statistical variability inherent in the limited detection of the measured data cause uncertainties (i.e. noise) given as the square root of the number of coincidence events. Lack of sufficient counts result in noisy data and a minimum number of counts are required to achieve a meaningful image. The number of measured counts depends on a number of factors such as the injected dose, acquisition duration and scanner configuration.

2.1.2 Image degradation factors

The ability of the PET to localize the point of the positron annihilation for the detected events in an image is known as spatial resolution and is empirically defined as the full width at half maximum (FWHM) of a point source distribution. Although in an ideal PET scanner the spatial resolution should be high and uniform, in reality spatial resolution is limited by several physical factors (Rahmim and Zaidi 2008). The main factors responsible for the finite spatial resolution in PET are:

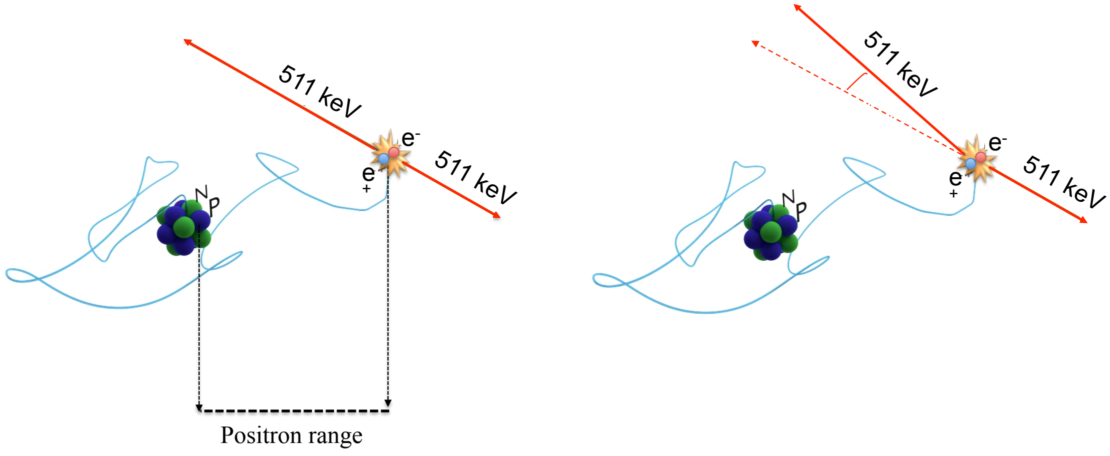
[¶] In typical clinical acquisition of the human thorax more than 10^7 events are required to obtain a statistically meaningful reconstruction of the activity distribution.

- The distance travelled by positrons before annihilation known as positron range.
- The angular deviation from 180 degrees of the emitted annihilation photons known as photon acollinearity.
- The finite crystal size in the PET detectors.
- The photon depth of interaction.

The linear distance between the nucleus from which the positron was emitted and the position of the annihilation is known as the positron range as illustrated in figure 2.3(a) (i.e. positron range is not the tortuous, continuous path of the positron). Positron range limits the accuracy of localizing the exact position where the positron was emitted and is one of the two main physical factors that deteriorate spatial resolution (Levin and Hoffman 1999).

The positron range is increased with the positron energy and decreases with increased tissue density. Consequently, different radiotracers result in different positron range distributions for different tissue types (Levin and Hoffman 1999; Bailey *et al.* 2005). In addition, from the decay of the nucleus the positrons are emitted with a continuous range of energies from zero to a maximum characteristic for each radiotracer and only a limited number have the maximum possible energy (Cherry *et al.* 2012). Due to this energy spectrum only a small fraction of positrons have enough energy to achieve the maximum positron range and a mean range is usually determined. The impact of the positron range in the PET resolution has been investigated by many authors using water as the reference medium (Levin and Hoffman 1999) or different human tissues (Sánchez-Crespo *et al.* 2004). For example, for positrons from ^{18}F, which has low maximum energy equal to 0.633 MeV and a mean energy equal to 0.202 MeV the mean positron range in water is 0.6 *mm* (Bailey *et al.* 2005).

The properties and application of the most commonly used positron emitting radionuclides as well as respective positron ranges in water are presented in table 2.1 (Bailey *et al.* 2005). Considering that the positron range depends on the atomic composition and tissue density again for positrons from ^{18}F the range as measured in human tissues is 0.54 *mm* in soft tissue and 1.52 *mm* in lung tissue (Sánchez-Crespo *et al.* 2004).



(a)

(b)

Figure 2.3: Main physical degradation factors of PET spatial resolution: (a) the distance travelled by positrons before annihilation known as positron range; and (b) the angular deviation from 180 degrees of the emitted annihilation photons known as photon acollinearity.

Although positrons lose nearly all of their momentum before annihilation, the positron and electron possess some residual momentum when they annihilate. Conservation of momentum requires that the resultant photons will not be emitted in exactly opposite directions (i.e. 180 degrees) but slightly less than 180 degrees apart, an effect known as photon acollinearity (figure 2.3(b)) (Sánchez-Crespo and Larsson

2006; DeBenedetti *et al.* 1950). The deviation from 180 degrees is approximated by a Gaussian with a FWHM of 0.5 degrees with a corresponding loss of resolution equal to 0.5R.tan(0.25°) where R is the diameter of the detector ring (Cherry *et al.* 2012). For example for a scanner of 80 cm diameter (e.g. a typical whole body scanner) the loss of resolution due to acollinearity is approximately 2 mm whereas for a smaller scanner diameter such as a 30 cm (e.g. an animal scanner) the loss is reduced to 0.7 mm. Acollinearity is the second factor affecting the spatial resolution (Levin and Hoffman 1999). Despite the physical factors, other factors may also limit the PET spatial resolution such as the geometry of the scanner.

The scanner geometry is appropriately designed not only to detect as many annihilation photons as possible but also to locate the annihilation position with certain spatial precision (i.e. spatial resolution). The finite detector size limits intrinsically the resolution. In general as shown in figure 2.4 the number of measured coincidences is reduced as the source is moved towards the edge of the FOV. For a pair of detectors the intrinsic spatial resolution is the FWHM of the measured coincidence events sensitivity when moving a point source across the detector face as shown in figure 2.4. On the scanner axis at mid-position between the two detectors the distribution is a triangle with a FWHM of half the detector width while at a position toward either detector the distribution becomes trapezoidal with FWHM equal to the detector width at the face of either detector.

Another factor of the scanner configuration that results in limited and non uniform spatial resolution is the limited accuracy with which the intersection point of the incoming photon with the detector is determined known as the parallax error (Sánchez-Crespo and Larsson 2006). For annihilation events emitted from the centre of the FOV of the scanner incoming photons enter the scintillation crystals perpendicular to the

Table 2.1: Properties of the most common positron emitting radionuclides used in PET with the positron range in water with their example compounds and applications (Bailey *et al.* 2005).

Radionuclide	Positron Energy (MeV)		Positron Range (mm)		Compound	Application
	max	mean	max	mean		
^{18}F	0.633	0.202	2.4	0.6	Fluoro-Deoxy-Glucose	glucose metabolism (oncology, cardiology, neurology)
^{15}O	1.738	0.696	7.3	2.5	Water	perfusion studies (neurology)
^{11}C	0.959	0.326	4.1	1.1	Methionine	protein synthesis (oncology)
^{13}N	1.997	0.432	5.1	1.5	Ammonia	myocardial perfusion (cardiology)

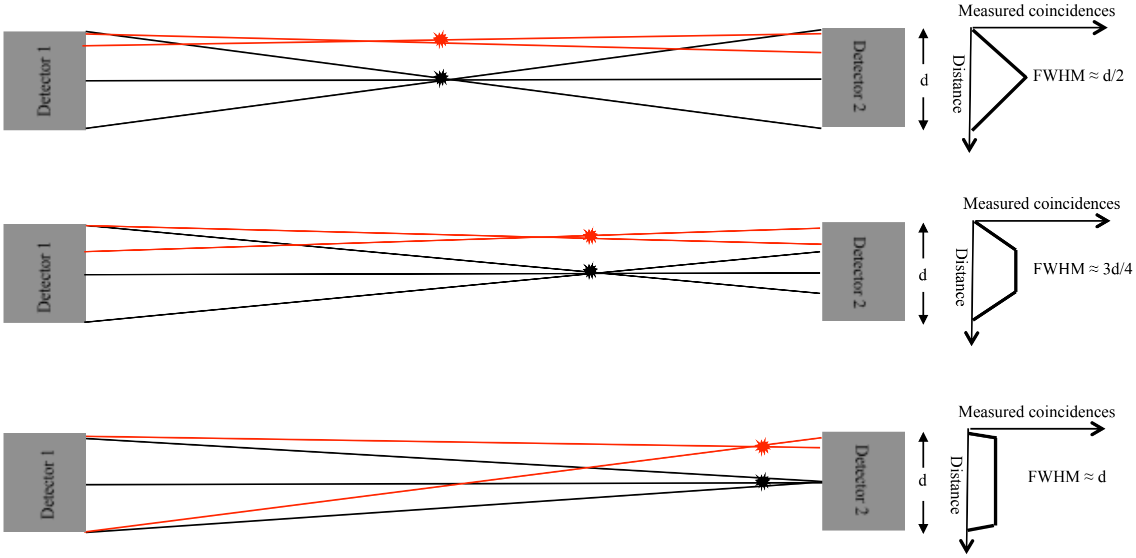


Figure 2.4: The intrinsic spatial resolution of PET detectors is determined by the finite crystal size and is defined as the FWHM of the measured coincidence events sensitivity when moving a point source across the detector face. For a point source positioned at the edge of the two detectors a limited number of coincidences is measured (red lines) whereas if the same source is positioned at the centre more coincidences are measured (black lines). As a result, the intrinsic spatial resolution is half the detector size at mid-position between the two detectors and increases with distance from the centre up to detector width at the face of either detector.

surface of the detector and although interact at different crystal depth until complete energy loss, the depth of interaction will not affect the position of the measured LOR. However, at a distance from the centre of the FOV of the scanner incoming photons enter the scintillation crystals at different angles. As a result, the photons might enter the two opposite crystals anywhere along their lengths and interact at different crystal depth or might penetrate several crystals before they interact with one of them due to the oblique angle of incidence. When there is no additional information for the depth of interaction, the interaction positions of the incoming photons to the crystals will be assigned at the most likely depth (e.g. midpoint of each crystal face) neglecting the depth of interaction and may result in mispositioning of the coincidence events as shown in figure 2.5(a).

Parallax error progressively deteriorates the spatial resolution with increasing distance from the centre of the FOV and therefore at the edge of the FOV a point source will appear as an ellipsoid (figure 2.5(b)). In addition, the error is higher when the detector is thicker and the position of the photons is away from the centre of the FOV. The parallax error can be reduced by using larger detector rings such as to reduce the angle of incidence. The correction of the parallax error can be addressed using depth of interaction information (DOI) (MacDonald and Dahlbom 1998). Despite the large interest on investigating DOI for improved image quality (Nakazawa *et al.* 2010; Thoen *et al.* 2013), for whole body PET imaging there is not currently any commercial system with DOI capability.

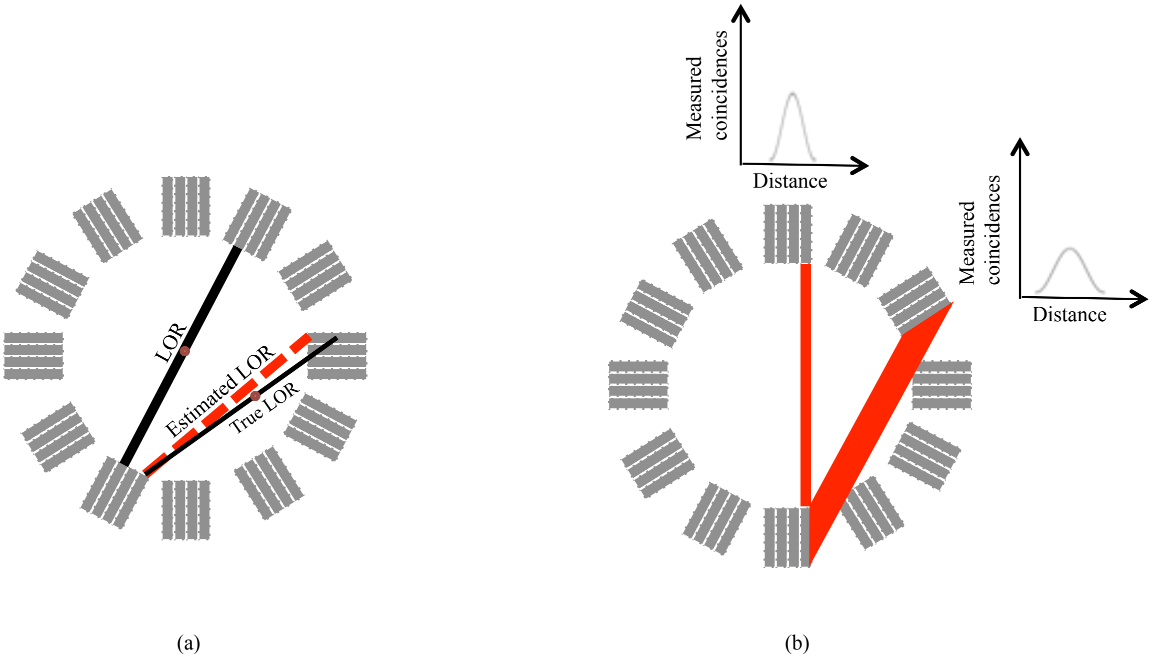


Figure 2.5: Schematic illustration of the parallax error in PET. (a) At the edge of the FOV the interaction positions of the incoming photons to the crystals will be incorrectly assigned at the midpoint of each crystal neglecting the depth of interaction. (b) Parallax error progressively deteriorates the spatial resolution with increasing distance from the centre of the FOV and causes a point source at the edge of the FOV to appear as an ellipsoid.

A combination of all the physical factors and the geometry of the scanner that limit the spatial resolution in PET determine the total spatial resolution. For example, whole body conventional human PET scanners achieve a spatial resolution approximately 5 to 6 *mm* FWHM in the centre of the scanner which decreases with the distance from the centre up to 6.5 *mm* as measured by the width of the peak produced by a scanned point-source (Teräs *et al.* 2007).

However, except from the aforementioned factors other parameters also play an important role in the final resolution such as the choice of reconstruction algorithm and parameters (e.g. image voxel size) that yield different accuracies of tracer uptake measurements and are expected to affect the final image resolution (Boellaard *et al.* 2008). For example filters that may be applied during or after the reconstruction for noise suppression may degrade the resolution of the reconstructed image with a magnitude determined by the choice of the filter parameters.

The limited resolution is one of the main reasons for the partial volume effect (PVE) (Soret *et al.* 2007). The limited system resolution introduces blurring in the image and as a result high tracer uptake objects (e.g. tumour) that are smaller than twice the PET resolution show partial loss of activity because the activity within the object "spills out" in adjacent regions with lower activity (Soret *et al.* 2007). The total counts in an area are preserved but they spread out in adjacent areas and therefore the object appears to be larger and to have a lower activity concentration than it actually has. The extent to which an object is affected by PVE is determined by its size. For large objects although will be blurred at the boundaries, the activity inside them will not change (figure 2.6).

The second aspect of PVE is the image sampling. Although the tracer distribution in a human body is continuous function, the reconstructed data are inherently discretized by a finite set of spatial basis functions (e.g. voxels). The underlying continuous distribution of the tracer uptake may not correspond to the boundaries determined from the voxel grid used to sample the image. As a result different tissue types may exist in a single voxel and the assigned tracer uptake value at the voxel will correspond to the average of all the values of the tissue types that are present within the voxel. The larger the voxel size the more this effect occurs and if the sampling is adequate (i.e. voxel size half the spatial resolution) this effect will not be significant. This effect is particularly apparent for voxels that are located on functional structure boundaries. In particular, if a voxel is located right across an edge between a high activity region and a low or no activity region, the voxel activity will be an average of the actual underlying activity distribution, thus making the edge look smoother than it actually is. Various methods have been proposed for compensating for PVE and reviewed by Rousset *et al.* (2007) and Erlandsson *et al.* (2012).

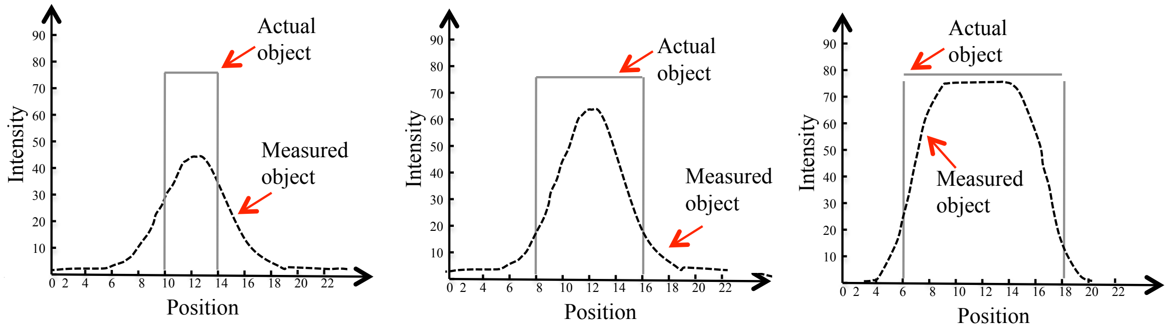


Figure 2.6: Due to the limited system resolution for objects that are smaller than twice the PET resolution the true underlying radioactivity concentration is contaminated between adjacent regions with different tracer uptake and the object appears to be larger and to have a lower activity concentration than it actually has. This effect is particularly apparent for small objects. As the size is increased the object will be blurred at the boundaries, but the activity inside will be less affected.

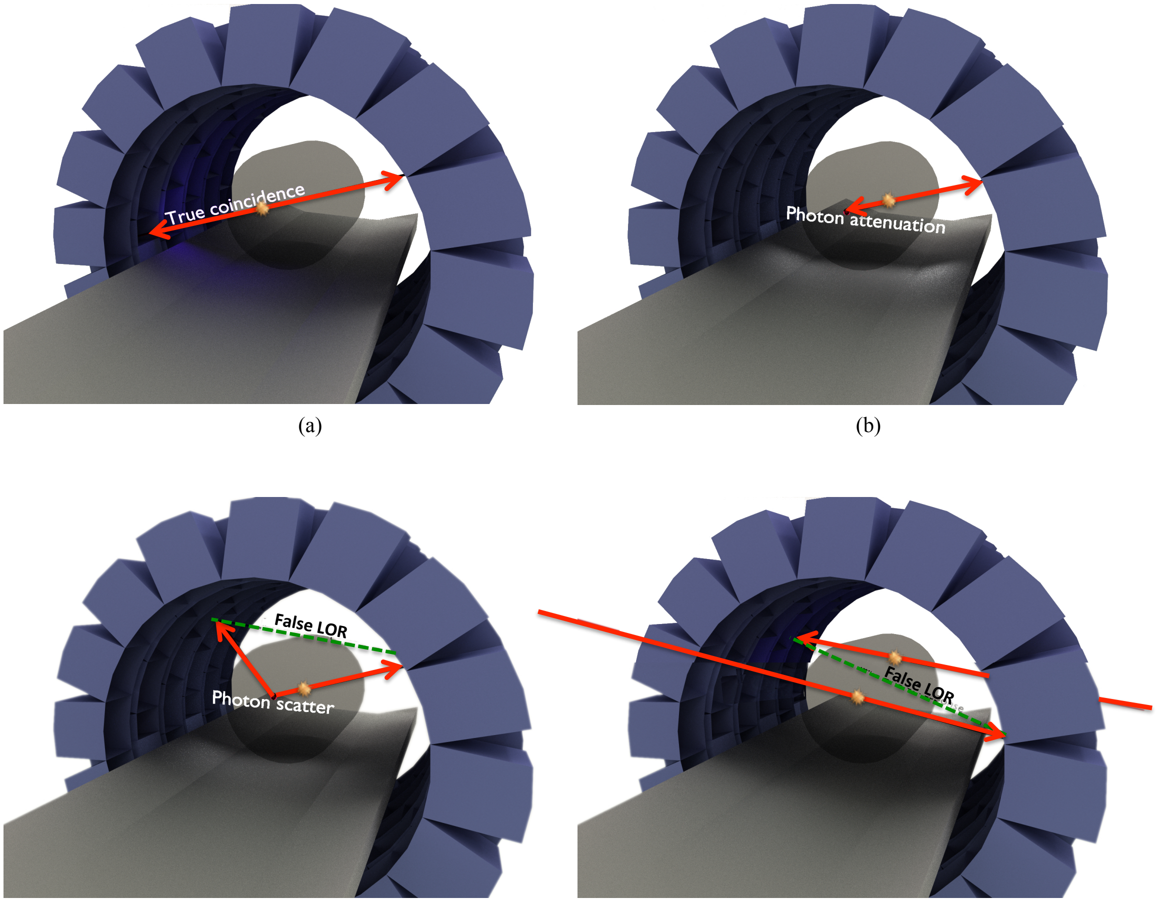
Mathematically, the blurring caused by the finite spatial resolution can be described by a 3D convolution operation using the point spread function (PSF) as representative of the system response. The PET image can be approximately modelled as the convolution of the actual tracer uptake distribution with the PSF of the scanner. The

PSF of the scanner is the distribution of the reconstruction of a point source and varies across the FOV of the scanner due to the non-uniform response of the detector. The operation is only a convolution when the PSF kernel does not vary. For the discrete case this is a linear operation, but if the kernel was spatially invariant would be a convolution operation. In addition, as the kernel varies smoothly and slowly over the FOV that it can be considered of as a local convolution.

The physical factors related to the characteristics of the radiotracers are not straightforward to change. However, more advanced image reconstruction methods have been proposed to improve image quality which will be described later. Furthermore, the detector related effects have been addressed by many studies (Budinger 1998; Adam *et al.* 2001) and are continuously improved with scanner developments. One of the recent developments is the reconstruction of images by integrating information about the time difference in the detection of the two annihilation photons a process known as time of flight (TOF) (Budinger 1983). By incorporating the TOF information into the reconstruction it allows the determination of the origin of the annihilation with higher accuracy. As a result images may be reconstructed with higher SNR properties and improved spatial resolution (El Fakhri *et al.* 2011). TOF requires faster scintillation crystals and electronics.

Except for the finite spatial resolution and the resulting PVE, PET image quality is further deteriorated by photons that may interact in the patient body and result in events contributing to background. The recorded LOR should correspond to two photons originating from the same annihilation event in order to be a true coincidence event (figure 2.7(a)). In addition to the true coincidence events, the measured events include other type of events that have been wrongly recorded within the specific energy and timing window.

One type of coincidence events are scattered events when one or both photons have undergone Compton scattering. In addition to changing the direction of the photons, scatter also reduces the energy of the photons and therefore the photon might not be detected within the energy window (i.e. photon attenuation) (figure 2.7(b)). However, some scattered photons might not lose enough energy and thus might be detected within the energy window especially if large energy windows are used. Scatter coincidence events give erroneous information about the location of the positron annihilation and therefore the coincidence event will be assigned to the wrong LOR as shown in figure 2.7(c). Scatter events may decrease the reconstructed image contrast causing inaccurate quantification (Zaidi and Montandon 2007c).



(a) (b)

(c) (d)

Figure 2.7: Types of coincidence events in PET: (a) a true event with detection of annihilation photons originating from the same annihilation; (b) a single event will be recorded if one of the two photons is scattered out of the FOV; (c) scatter event result if one or both detected photons have undergone scatter; (d) random event results from detection of two photons originating from different annihilations. Other events than true events have erroneous information about the location of the positron annihilation. The erroneous LOR is indicated with a dashed line.

Other coincidence events are random events. Random events might occur when two positrons annihilate approximately at the same time and one photon from each annihilation is not detected while the other two photons are detected. Therefore, although the photons originate from different annihilations they are accepted within the timing window, they are wrongly recorded as originating from the same annihilation, and consequently result in an incorrect LOR (figure 2.7(d)). Random events are characterised as a relatively uniform background added to the true radiotracer distribution. Random events may decrease the reconstructed image contrast and cause uncertainties in quantification values.

Events other than true events have erroneous information about the location of the positron annihilation and affect the accuracy of the estimation of the tracer uptake. However, while attenuation, mainly due to scatter, reduces the number of detected events, the random but also the scatter events increase the number of events (noise) in the data by adding a background to the true events distribution. Therefore, the number of detected true events is not the number of incident events as it would be for an ideal PET system, but is susceptible to contamination from scatter and random events decreasing the maximum possible signal to noise ratio (SNR). In particular, the noise in PET can be described by the noise equivalent count rate (NECR) given by:

$$NECR = \frac{T^2}{T + S + kR} \tag{2.1}$$

where T, S, R are the true, scatter and random events count rates respectively and k is a scalar equal to 1 or 2 for noiseless or noisy estimates of random events respectively.

The NECR is the minimum number of coincident events that would have been needed if all the incident events are considered as true events in order to obtain a statistically equivalent noise level in the presence of measured scatter and random events. The NECR is equivalent to the square of the SNR in the PET projection data and can determine the significance of a particular noise level (Badawi and Dahlbom 2005). The SNR increases as the number of detected scatter and random events is decreased.

The measured data can be corrected for the scatter and random either before the reconstruction in projection space or within reconstruction. In particular, scatter and random events after being estimated can be subtracted from the noisy measured projection data. This approach disturbs the Poisson nature of the measured data and may generate negative values in the data which will be consequently set to zero during iterative reconstruction introducing bias (Walker *et al.* 2011). Alternatively, these events can be included as additive terms in the statistical model of event detection as in the ordinary Poisson formulation of the maximum likelihood expectation maximisation (MLEM) iterative reconstruction algorithm where the model accounts for the scatter and random events (e.g. addition after forward projection operation) assuming that they contribute to the background. Including these events within the forward model of the reconstruction preserves Poisson statistics and applies the corrections separately from the forward projection and backprojection operations in the reconstruction and it does not significantly increase computational effort. Alternatively, this model can be included within the system matrix creating a less sparse system matrix which significantly increases computational effort but may improve image quality (Conti *et al.* 2012).

2.1.3 Data acquisition and storage

The circular pattern of multiple rings of PET scanners allows annihilation photons emitted in opposite directions and between adjacent rings to be detected simultaneously, three dimensional (3D) collection of data. However, a PET scanner may offer the availability to acquire data either in two dimensional (2D) or three dimensional (3D) mode. In the 2D mode a septa (usually made of lead or tungsten) is positioned in between two adjacent crystal rings. As a result, only the detection of photons between coincident detectors, either in the same ring or adjacent crystal rings, depending on the length of septa, are recorded to form a single plane. On the other hand, in 3D mode there are no septa and events which occurred between detectors by any combination of crystal rings are allowed to be detected resulting in increased scanner sensitivity (i.e. fraction of photons generated within the FOV that are detected) and SNR compared to 2D (Townsend *et al.* 1989; Cherry *et al.* 1991).

PET scanners tend to have a cylindrical geometry (figure 2.8). A coincidence event detected by two opposite detectors can be characterized by the coordinates x, y, z of detector 1, $D_1(x_1, y_1, z_1)$, and detector 2, $D_2(x_2, y_2, z_2)$, that define the LOR. The LOR can also be expressed in the spherical coordinate system by the azimuthal angle ϕ and polar angle θ as shown in figure 2.8 (Zanzonico 2004).

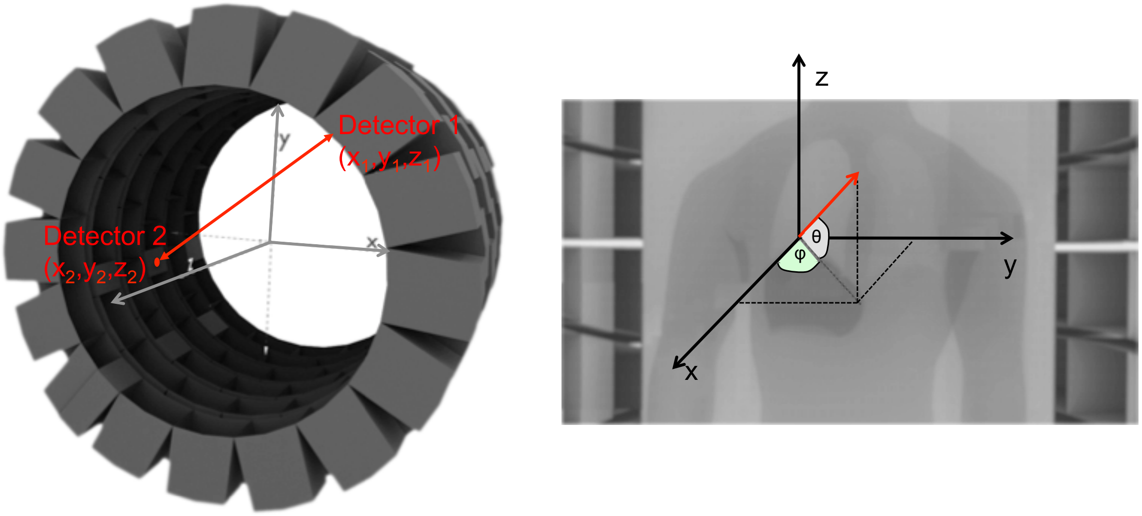


Figure 2.8: Schematic illustration of PET cylindrical geometry with axial cross section. A coincidence event detected by two opposite detectors is characterized by the coordinates of detector 1, $D_1(x_1, y_1, z_1)$, and detector 2 $D_2(x_2, y_2, z_2)$ that define the line of response (LOR). The LOR can be expressed in the spherical coordinate system by the azimuthal angle ϕ and polar angle θ.

After recording a detected coincidence event, the LOR is usually stored in list-mode or histogram format. In the list-mode format each measured coincidence event is stored in the form of a list with the coordinates of the two photons defining the LOR along with the time of detection. Each element of the list-mode format stores information for just a single coincidence event sequentially as it is detected. There are two different types of list-mode data format. In the first type the event coordinates of each detected LOR are recorded. In the second type the position of each detector is recorded in a sinogram bin and the address of the sinogram bin is recorded. With this type a single sinogram bin can be used for sub-LORs and therefore the address of the sinogram bin will not correspond to the coordinates of each recorded LOR but to the average location of all possible sub-LORs stored within the specific bin. Apart from the location of the pair of detectors and the time of the detection, additional information may be recorded such as the energy of each photon and the time difference between their detection.

In the histogram the data are stored as the sum of the measured data along the LOR within a matrix that described the location of the LOR within the FOV. The total number of detected coincidence events between any detector pair is proportional to the radioactivity distribution at each point along the LOR neglecting degradation factors (e.g. attenuation and scatter) and may be modelled as proportional to line integrals through the activity distribution, $f(x,y,z)$, along the LOR. The collection of line integrals parallel to a given direction is known as a projection.

For a 2D data acquisition the line integrals for each LOR are usually stored in the form of a sinogram, a 2D matrix with the vertical axis corresponding to the azimuthal angle ϕ and the horizontal axis corresponding to a spatial coordinate s which is related to the radial distance of each LOR from the centre of each ring. In the sinogram each bin value corresponds to a projection at a specific angle ϕ and axial offset s as illustrated in figure 2.9. Consequently, the sinogram for a single transaxial plane has dimensions equal to the number of parallel projection elements times the number of angular views. For example, storing the measured data of a point source will form a sinusoidal pattern curve and thus the name sinogram, while the sinogram of an activity distribution consists of multiple overlapped sinusoidal pattern curves.

For a 3D acquisition the polar angle θ is also used as an additional coordinate to describe the projections and the measured data are then represented as a set of sinograms with one sinogram per θ. If the events stored were measured within the same detector ring, such as in the case of a 2D acquisition, the sinograms are known as direct plane sinograms. On the other hand, if the events stored were measured between different detector rings the sinograms are known as oblique sinograms. In 3D data acquisition both direct and oblique sinograms are acquired leading to more sampling points. Consequently, the storage and computation requirements for data acquisition may be considerably higher for 3D acquisition with a large number of detector ring combinations. In order to reduce the number of acquired sinograms, different methods are used such as axial compression (Fahey 2002).

2.1.4 Image reconstruction

The measured data are used to calculate the radioactivity distribution in the patient a process known as image reconstruction. A reconstructed image spatially maps the radioactivity concentration that produced the measured data. Image reconstruction algorithms can be classified into analytic or iterative methods.

Analytic methods are based on the approximation of perfect detectors and measured data ignoring the underlying physics of the coincidence detection and the statistical noise inherent in the measured data. They provide a direct mathematical

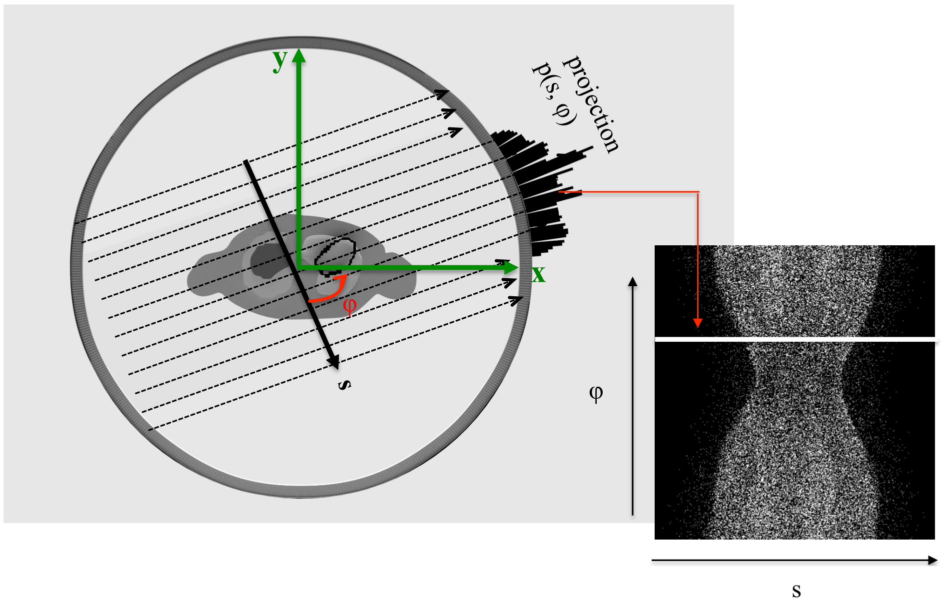


Figure 2.9: Illustration of the coordinate system used for PET data stored in a sinogram. In a sinogram the detected coincidence events are stored according to angular orientation ϕ versus its displacement s from the centre of the detector rings. The arrow points out the position in the sinogram of one projection p(s, ϕ).

solution for the recovery of the radioactivity distribution based on a line-integral model and Fourier theory. The most commonly used analytic algorithm is filtered-backprojection (FBP).

The Central Section theorem or Fourier Slice theorem states that the 1D Fourier Transform (FT) of the projection (i.e. equivalent to line integral through the radioactivity distribution) for a specific angle is equivalent to a profile at the same angle through the centre of the 2D FT of the object. Therefore, calculating the Fourier transform of the measured projections over a number of angles around the object and subsequently apply inverse Fourier transformation will give the counts distribution in the FOV.

In one approach to image reconstruction the magnitude of the values of each point in the projection can be placed back to every point in the image space along the appropriate LOR corresponding to the relevant line of integration in the radioactivity distribution a process known as backprojection. However, the exact position of the value into the image space is unknown and therefore a constant value is placed into all image voxels along the corresponding LOR. Therefore all image elements along the LOR are assigned with the value of each point in the projection. If this process is repeated for all measured angles an image of the radioactivity distribution will produced.

After simple backprojection the image is a blurred approximation of the true radioactivity distribution because the exact position of the value of each projection in the image space is unknown and is attributed to all voxels that project onto that projection ray and not only to the voxel the radioactivity is coming from. The sampling density in Fourier space is proportional to the inverse of the distance from the origin (i.e. more sampling of the Fourier transform at the origin). Therefore, each projection fills in one slice of the Fourier space resulting in oversampling in the centre and undersampling away from the centre. The blurring caused by the non uniform sampling of the Fourier space is commonly modelled as the result of a convolution of the radioactivity volume distribution and an inverse radial blurring kernel. To overcome this, the profiles are filtered prior to backprojection (e.g. ramp filter) in order to have equal contributions throughout the Fourier space a process known as filtered backprojection (FBP).

Analytic algorithms are fast reconstruction methods with low computational requirements due to the direct solution that simplifies the reconstruction. Despite this, the reconstructed image quality is limited by the approximations of the line integral model and the disregard of the statistical noise inherent in the measured data. These approximations may cause errors in quantification and noise that may appear as streak artefacts. The high noise levels can be treated by adjusting the cut-off frequency value of filters applied either during or after the reconstruction (Van Laere *et al.* 2001) but at the expense of quantification accuracy and image resolution. Due to these limitations

FBP is not usually used clinically but it often used as the appropriate method for quantification tasks.

An alternative to analytic methods is to use an iterative reconstruction algorithm. In contrast with analytic methods, iterative methods can incorporate modelling of the physical degradation factors and of the underlying physics (i.e more realistic model of the system) and the statistical variability of photon detection. Iterative algorithms may improve the image quality in terms of noise and spatial resolution compared with analytic algorithms either for the case of high count measured data where the accuracy of the model has high impact on image resolution or in the case of low count measured data where the statistical noise due to the limited detection of photons has high impact on image resolution. While, in routine clinical practice iterative algorithms have replaced analytical algorithms, in some cases where quantification is the main task of image interpretation analytic algorithms might be a preferable method.

The first step of the iterative algorithms is to model the object of the acquired data in the image (i.e. spatial) domain. Modelling of the tracer distribution is usually performed by using variables to discretize the image. Although the tracer distribution in a human body is a continuous function, the iterative algorithms approximate the continuous distribution by a finite set of spatial basis functions such as voxels or pixels, 3D or 2D image elements respectively. While, the value of each basis function is assumed to describe the activity distribution contained in the volume of the human body spanned by the voxel, the tracer distribution within the voxel is assumed to be uniform and therefore this approximation is limited by the number and dimension of the voxels used to discretize the tracer distribution in the human body. Following discretization, the system is modelled by mapping the image space to the projection space as following:

$$g_i = \vec{P}\vec{\Lambda} \tag{2.2}$$

where $\vec{\Lambda} = \{\Lambda_j\} \in \mathbb{R}^J$ is the tracer distribution for J voxels (j=1...J), $\vec{P} = \{P_{ij}\} \in \mathbb{R}^{J\times I}$ is the forward projection operator, also known as the system matrix, and $\vec{g} = \{g_i\} \in \mathbb{R}^I$ are the expected number of true events for each LOR. The element P_{ij} of the system matrix represents the probability that a photon from the j^{th} voxel is detected in the i^{th} projection element. The system matrix can include geometry, detector response, attenuation and other factors for more realistic modelling of the relationship between image space and projection space (Panin *et al.* 2006). The expected value of the measured data q_i given the measured data m_i, ($q_i = \langle m_i \rangle$), is given as:

$$q_i = \sum_{j=1}^{J} P_{ij}\Lambda_j + n_i \tag{2.3}$$

where $\vec{n} = \{n_i\} \in \mathbb{R}^I$ is the mean number of background events such as scatter and randoms (unless these are included in the system model).

The next step of iterative algorithms is to model the measured data in order to describe the statistical relationship between the measured data and the expected data. During PET acquisition the measured data are derived from nuclear decays, a random process, and therefore they are statistically variable. Photon detections are usually modelled as random variables characterised by Poisson (Shepp and Vardi 1982) or more rarely Gaussian (Fessler 1994) distribution. With the assumption of the Poisson probability, the discrete probability distribution of the expected data q_i given the measured data m_i is the following:

$$\Pr(m_i \mid q_i) = \frac{e^{-q_i} q_i^{m_i}}{m_i!} \tag{2.4}$$

To model all the projection elements i the product of the probability density function for every single element is calculated known as the Poisson likelihood, L:

$$L\left(\vec{m} \mid \vec{q}\right) = \prod_{i=1}^{I} \frac{e^{-q_i} q_i^{m_i}}{m_i!} \tag{2.5}$$

Then, after modelling the radioactivity distribution, an objective function is defined to encompass the probabilistic model that is employed to describe the noise distribution of the measured data, as well as any other model employed to describe the photon detection. The most commonly used approach in these algorithms is the maximum likelihood (ML). In ML the probability for the measured data to approach the unknown true radioactivity distribution, Λ, is to obtain a model of the mean, q, such that the probability likelihood reaches a maximum. With the assumption of a Poisson probability distribution, the optimization is usually performed in the logarithmic domain (log-likelihood). Due to the monotonic nature of the log function, this will lead to the same local maximum, which is also a global maximum, since the log-likelihood function is concave. Consequently the log-likelihood is given by:

$$\begin{aligned} lnL\left(\vec{m} \mid \vec{q}\right) &= \sum_{i=1}^{I} ln\left[\frac{e^{-q_i} q_i^{m_i}}{m_i!}\right] \\ &= \sum_{i=1}^{I} \left(-q_i + m_i ln(q_i) - ln(m_i!)\right) \\ &= \sum_{i=1}^{I} \left(-\sum_{b=1}^{J} P_{ib}\Lambda_b + m_i ln\left(\sum_{b=1}^{J} P_{ib}\Lambda_b\right) - ln(m_i!)\right) \end{aligned} \tag{2.6}$$

where q_i was substituted with $\sum_{b=1}^{J} P_{ib}\Lambda_b$. In the above equation the background term has been neglected for simplicity. To maximize the Poisson log-likelihood function its partial derivative with respect to a single voxel Λ_j is calculated and set it to zero:

$$\frac{\partial}{\partial \Lambda_j} \ln L\left(\vec{m} \mid \vec{q}\right) = \sum_{i=1}^{I} \left(-P_{ij} + \frac{m_i}{\sum\limits_{b=1}^{J} P_{ib}\Lambda_b} P_{ij} \right) = 0 \tag{2.7}$$

The above equation can not be solved directly. To estimate the Λ_j, the probabilistic model is maximized iteratively by optimizing of the objective function. The radioactivity distribution is updated until it converges to the maximiser of the objective function. Various optimization algorithms have been proposed, from gradient based to the commonly used expectation maximisation (EM) algorithm. The EM algorithm determines the maximum likelihood estimate (MLEM).

Gradient-based methods are based on adding a step size. Considering that the log-likelihood function is locally quadratic around one local maximum by adding a step size to the current estimation, Λ_j^k, the next estimate will be closer to the maximum likelihood (Qi and Leahy 2006). Using this gradient ascent the expectation maximization is given as:

$$\begin{aligned} \Lambda_j^{k+1} &= \Lambda_j^k + T_j^k \frac{\partial}{\partial \Lambda_j} \ln L\left(\vec{m} \middle| \vec{q}\right) \\ &= \Lambda_j^k + \frac{\Lambda_j^k}{\sum\limits_{i=1}^{I} P_{ij}} \left(-\sum_{i=1}^{I} P_{ij} + \sum_{i=1}^{I} P_{ij} \frac{m_i}{\sum\limits_{b=1}^{J} P_{ib}\Lambda_b^k} \right) \\ &= \frac{\Lambda_j^k}{\sum\limits_{i=1}^{I} P_{ij}} \sum_{i=1}^{I} P_{ij} \frac{m_i}{\sum\limits_{b=1}^{J} P_{ib}\Lambda_b^k} \end{aligned} \tag{2.8}$$

This is the most popular optimisation technique in PET known as the expectation maximisation (EM) algorithm. The EM algorithm is a general framework used to estimate the maximum likelihood solution through the introduction of a "complete" set of data. The "complete" data is a hypothetical, unobserved set of data, which can be thought of as the intersection between the sinogram data and the image data. The dataset $\{m_i\}$ with each element equal to the number of photons detected in sinogram bin i is incomplete dataset since there is no information from which voxel the data have been emitted. The MLEM algorithm can be derived based on the complete dataset, $\{z_{ij}\}$, where each element represents the number of detected photons in sinogram bin i that originate from the voxel j with $m_i = \sum_{j=1}^{J} z_{ij}$. The mean expected data are equal with $\langle m_i \rangle = q_i = \sum_{j=1}^{J} P_{ij}\Lambda_j$ and therefore the expected complete dataset is given by:

$$\langle z_{ij} \rangle = P_{ij}\Lambda_j \tag{2.9}$$

Both $\{z_{ij}\}$ and $\{m_i\}$ are Poisson distributions and the likelihood function of the complete dataset, $L_{CD}(\Lambda/z)$, can be written as:

$$L_{CD}(\Lambda/z) = \prod_i \prod_j \frac{(P_{ij}\Lambda_j)^{z_{i,j}} \cdot exp(-P_{ij}\Lambda_j)}{z_{ij}!} \tag{2.10}$$

To maximize the likelihood function the natural logarithm is firstly calculated known as log-likelihood function:

$$\ln L_{CD} = \sum_i \sum_b \left(z_{ib} \ln P_{ib}\Lambda_b - P_{ib}\Lambda_b\right) \tag{2.11}$$

In the above equation the factorial term has been neglected. The log-likelihood function can be maximized by taking the partial derivative with respect to a single voxel Λ_j and set it to zero:

$$\frac{\partial}{\partial \Lambda_j} \ln L_{CD} = \sum_i \left(z_{ij} \cdot \frac{1}{P_{ij}\Lambda_j} P_{ij} - P_{ij} \right) \tag{2.12}$$

Setting the partial derivative to zero, $\frac{\partial}{\partial \Lambda_j} \ln L = 0$ it follows:

$$\frac{\sum\limits_i z_{ij}}{\Lambda_j} = \sum_i P_{ij} \tag{2.13}$$

Rearranging the above equation the maximum likelihood for a single voxel j is given by:

$$\Lambda_j^{ML} = \frac{\sum\limits_i z_{ij}}{\sum\limits_i P_{ij}} \tag{2.14}$$

Although the complete dataset $\{z_{ij}\}$ is unknown, the expected values of the complete data can be used. Therefore the expectation maximization maximizes the excepted values of the complete data. The complete data can be generated by calculating the expected value of the data z_{ij} with known the measured data m_i and based on $\langle m_i \rangle = q_i = \sum\limits_{j=1}^{J} P_{ij}\Lambda_j$. The joint distribution of independent Poisson variables z_{ij} with mean values $P_{ij}\Lambda_j$ is a multinomial distribution and the probability that an event detected in the sinogram bin i originates from the voxel j is given as $(P_{ij}\Lambda_j)/(\sum\limits_{j=1}^{J} P_{ij}\Lambda_j)$. Therefore the conditional expectation of the complete data is estimated as follows:

$$\frac{z_{ij}}{m_i} = \frac{P_{ij}\Lambda_j}{\sum\limits_{j=1}^{J} P_{ij}\Lambda_j} \tag{2.15}$$

$$z_{ij}(\Lambda^k) = \frac{P_{ij}\Lambda_j^k}{\sum\limits_{j=1}^{J} P_{ij}\Lambda_j^k} m_i \tag{2.16}$$

Substituting (2.16) to (2.14) the MLEM for all the voxels is given by:

$$\Lambda_b^{k+1} = \frac{1}{\sum_i P_{ib}} \frac{\sum_i P_{ib} \Lambda_b^k m_i}{\sum_{v=1}^{J} P_{iv} \Lambda_v^k} = \Lambda_b^k \frac{1}{\sum_i P_{ib}} \sum_i P_{ib} \frac{m_i}{\sum_{v=1}^{J} P_{iv} \Lambda_v^k} \tag{2.17}$$

The EM algorithm converges monotonically to the solution that maximises the log-likelihood function of the estimates of the radioactivity distribution (2.11). Consequently, at each iteration k the updated image will increase the value of the likelihood function. Non negativity in the image domain is ensured provided that the initial image estimate is non-negative. However, this introduces bias in low count regions due to the shift of the discrete probability distribution of the Poisson distribution (Reilhac *et al.* 2008).

An accelerated version of the MLEM is the ordered subsets expectation maximization (OSEM) algorithm. The OSEM algorithm, processes the measured data in groups and instead of applying the iteration process including all projections, a small number of projections is grouped (known as subset) and the MLEM algorithm is applied to each subset. The OSEM algorithm is given by:

$$\Lambda_b^{k,q} = \Lambda_b^{k,q-1} \frac{1}{\sum_{i \in S_q} P_{ib}} \sum_{i \in S_q} P_{ib} \frac{m_i}{\sum_{v=1}^{J} P_{iv} \Lambda_v^{k,q-1}} \tag{2.18}$$

where S_q corresponds to the q^{th} subset S of the projection space.

One full iteration is completed after using all subsets. Using a subset of the measured data accelerates the convergence a factor equal to the number of subsets producing images with similar quality.

One critical characteristic of these algorithms is the convergence rate. As the EM algorithm progressively reconstructs an image, the low frequency components of the image appear within the first iterations, while as the ML estimate is approached, higher frequency components are resolved, including noise. In addition, the theoretical ML estimate may be reached only after an infinite number of iterations. Consequently, there is a bias-variance trade off.

One method to compensate for the noise deterioration in OSEM reconstruction is stopping after a small number of iterations. Early stopping might not be an efficient method since the convergence rate is noise and object dependent and therefore the convergence of different objects in the image is not achieved for all points at the same iteration number. Boellaard *et al.* (2001) investigated dynamic PET images and demonstrated that for OSEM, cold regions within a hot background converge more slowly than small hot areas within a cold background. Therefore the optimal number of iterations depends on the desired clinical benefit and the focused area as well as the count statistics data. Another way is to post-smooth the images using filters (e.g. Gaussian kernel) but the noise is controlled at the expense of resolution. An alternative approach is to generalize these algorithms, by including a prior term which favours certain characteristics in the reconstructed images (e.g. locally smooth) (Fessler 1994). Prior terms can introduce additional information about the expected image noise penalizing the loglikelihood function. This can be achieved via Bayesian estimation schemes with the most commonly of them the maximum a posteriori (MAP) algorithm (Hebert and Leahy 1989). Green (1990) proposed the one-step-late (OSL) method to apply the EM algorithm to the MAP problem where a penalty term is determined by the use of the previous estimate, hence the name one-step-late. The prior usually constrains the solution by penalizing the intensity variations in the adjacent voxel values. For example by using anatomical prior information the noise is reduced within the regions while allowing for large variations between different voxels across the boundaries. However, the image quality depends on the choice of the prior/penalty parameters.

Another feature of statistical iterative image reconstruction algorithms is that there are inherent constraints such as non-negativity. The non-negativity constraint does not allow negative values to be assigned to an image voxel. Therefore, when high variance (i.e. low count) data lead to a probability that a voxel will be assigned a negative value the OSEM bounds the intensities distributions at zero and the distributions do not remain symmetric. Whilst this non-negativity constraint is in keeping with a priori knowledge of a non-negative activity distribution, it is expected also to result in an overestimation bias for low count regions (Grezes-Besset *et al.* 2007; Ahn and Fessler 2004). Therefore linearity is not preserved with iterative reconstruction, since the reconstruction of the sum of two sinograms is not the same as the sum of the independent reconstruction of the two same sinograms. The extent of this inequality depends on the object, background activity levels, counting statistics, physical factors, number of iterations and any applied penalization factors or filtering (Barrett *et al.* 1994; Wilson and Tsui 1993; Boellaard *et al.* 2001).

2.1.5 Image quality characterization and evaluation

The most common application of PET in oncology is for detecting tumour metastases for staging and quantifying the radiotracer accumulation. An objective definition of image quality is complex because it depends on the task being performed. In clinical practice, image quality is usually characterised on how well the true tracer uptake can be recovered (i.e. estimation task) and how well lesions from normal tissues can be distinguished (i.e. classification task). PET has limitations on both tasks due to the physics of PET (e.g. spatial resolution, scatter, randoms and attenuation) and the data processing methods (e.g. reconstruction method). These two tasks can be considered to characterise and evaluate the PET image quality.

2.1.5.1 Estimation tasks

In estimation tasks, the aim is to recover quantitative values such as standardized uptake values (SUVs) from the images. The most common parameter used to measure tracer uptake is the SUV. In clinical practice there are two ways commonly used to report the SUV, the mean and the maximum SUV. The mean SUV is the average value of all the voxels in the ROI and therefore includes information from multiple voxels whereas the maximum SUV is the maximum voxel value within the ROI. Consequently, the mean SUV depends on the definition of the ROI and is less sensitive to the noise of the image than the maximum SUV. The maximum SUV is the highest voxel value within the ROI. Based on the assumption that the voxel with the highest activity value is included in the predefined region the maximum SUV is independent of the definition of the region and consequently less dependent on the observer than the mean SUV. However, the maximum SUV is more susceptible to noise than the mean SUV. As an alternative the peak SUV has been suggested (Wahl *et al.* 2009) defined as the average SUV within a group of voxels in a small size region including the voxel with the highest value within the ROI. The peak SUV is less susceptible to noise than maximum SUV and less sensitive to the definition of the ROI than mean SUV. Although peak SUV was initially defined as a spherical region of 1 cm^3 size (Wahl *et al.* 2009), various other sizes and shapes have been investigated that include the maximum voxel value but they are not necessarily centred on it (Vanderhoek *et al.* 2012).

A common approach used to assess the quantitative performance of an image reconstruction method is to determine its accuracy and precision. The ideal image is on average the true image (unbiased) whose values do not deviate from their mean (zero variance). Given that PET imaging is based on nuclear decay which is random process, if the same object is imaged multiple times, the acquired data will have statistical variations. The ideal reconstruction method will produce images that are minimally biased with respect to the true distribution and that have low variance.

Bias is a measure of the accuracy of a method and is given by the difference between the expected value and the true value. The variance is a measure of the precision of the estimate (how much the estimate varies around its average value). Variance may be evaluated by repeated estimations of the same value expressed in terms of the difference between the investigated data around its mean value. The sum of the squared deviations from the mean of all measurements divided by the number of measurements minus one defines the variance whereas the square root of the variance is known as standard deviation. Another measurement is the mean squared error (MSE) calculated as the sum of the square of standard deviation and the squared bias.

The estimation of quantitative values such as the tracer uptake is commonly performed by placing a 2D or 3D region of interest (ROI) around a target (e.g. tumour) in the image of interest. However, instead of calculating the figure of merit from the statistics of a region in a single reconstructed image, another method is to calculate the figure of merit voxel by voxel in multiple realizations of reconstructed images. In contrast with a single image, the voxel-by-voxel analysis of multiple images provides the expectation of the image estimate (average image over multiple realizations of reconstructed images).

2.1.5.2 Classification tasks

To assess detectability, (i.e. the ability of a method to distinguish lesions from normal tissues), various figures of merit can be used such as detection sensitivity (i.e. fraction of positive cases correctly identified as positive), detection specificity (i.e. fraction of negative cases correctly identified as negative) and accuracy (i.e. fraction of cases identified correctly, whether positive or negative). Another way to assess the diagnostic performance of a method is the receiver operating characteristic (ROC) curve.

A ROC curve illustrates the trade-off between the detection sensitivity and one minus the detection specificity as a function of various decision thresholds. A ROC study is performed using a set of images usually with lesion (i.e. positive image) or without lesion (i.e. negative image). The images are given to an observer who is asked to indicate the confidence level of the presence or not presence of a lesion (e.g. five levels of confidence). Using the confidence levels some parameters can be calculated such as the true positive rate (TPR) which is the fraction of positive images correctly identified as positive by the observer (i.e. detection sensitivity). Similarly, the false positive rate (FPR) is the fraction of negative images incorrectly identified as positive by the observer. Based on these two parameters the true negative rate (TNR) can be calculated which equals one minus the FPR and is known as detection specificity.

The ROC curve then is generated by plotting TPR versus FPR for various thresholds (i.e. degrees of confidence) as illustrated in figure 2.10. The ROC curve lies above or on the ascending 45-degree diagonal and as the distance of the curve increases from the 45-degree line the better the detectability performance increases from no useful information to the ideal.

One important characteristic of the ROC curve is that it not only describes the ability of a method to correctly identify the positive images but also the false positives. It must be noted, that in cases where multiple lesions are present in a single image the traditional ROC curve can not be applied as the FPR can not be calculated. For such cases an alternative method has been developed known as free-response operating characteristic (Chakraborty 2000).

However, instead of just comparing the ROC curves of various methods in order

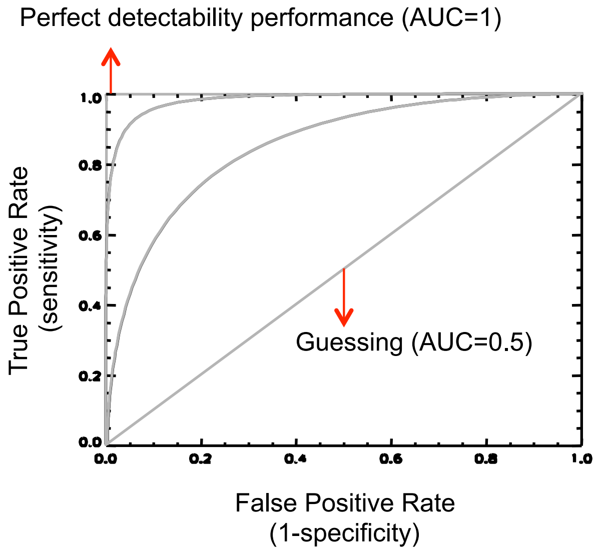


Figure 2.10: Example of receiver operating characteristic (ROC) curves. A ROC curve is a plot of the detection sensitivity versus one minus the detection specificity. The area under the ROC curve (AUC) quantifies the detectability performance of a method. The best possible prediction would yield a point here with (0,1) which represents 100% sensitivity (no false negatives) and 100% specificity (no false positives).

to evaluate their detectability performance, a single value is usually derived from the curves to characterise the detectability performance. The most common measurement to define the diagnostic performance from a ROC curve is the integral of the curve, the area under ROC curve (AUC), that describes the accuracy of correctly detecting tumours. An ideal result would have maximum detection sensitivity and also perfect separation of the distributions of the diseased and healthy cases with respect to the range of possible test results and thus makes easier the determination of a diseased from a healthy case. For a result of 100% detection sensitivity and 100% specificity (i.e. perfect detectability performance) the AUC will equal to 1 which is the maximum value while as the sensitivity and specificity are decreased the AUC is decreased. As shown in figure 2.10 for the diagonal the AUC equals 0.5 and corresponds to a random classification result (i.e. guessing) with no useful information. The diagonal line divides the ROC space; points above the diagonal represent classification results that are better than random. The properties of the AUC have been analysed in the literature (Swets and Pickett 1982; Metz 1986, 2000). In cases of clinical studies this method is limited due to the lack of information about the real presence of the tumours.

2.1.5.3 Numerical observers

Tumour detection is traditionally achieved using human observers. However, for evaluation of detectability many observers and images are required in order to obtain statistically meaningful results and especially in cases where a large number of parameters are to be evaluated this becomes an impractical process. To minimize or replace the performance of various human observers, model numerical observers have been proposed (Barrett *et al.* 1993).

Numerical observers are applied to images f and output a test statistic value λ, over a given image which can be compared to a threshold; if the value is below the threshold the given image is classified as background images or if it is above the threshold it is classified as having signal. The most common numerical observers are the linear observers which use linear discriminant functions to calculate the test statistic. In general, for a linear numerical observer the test statistic, λ, for a given image f is calculated by (Barrett and Myers 2003):

$$\lambda(f) = W^T \cdot f \tag{2.19}$$

W is the discriminant function known as the model observer template and T denotes the matrix transpose operation.

Usually there are two classes of the input images, a signal absent image (i.e. background), f_B, and a signal present image (i.e. lesion present), f_L. In the simple case of the detection of a known signal superimposed on a known background the only source of randomness in the data is the variation in the observer measurement. If the variation in the measurement is modelled as normally distributed (i.e. Gaussian) and non-correlated (i.e. white) the discriminant that provides maximum separation of the probability density functions of the data from the two classes is defined by their mean difference. Therefore, the difference in the means of the data belonging to the two classes is the optimal way to discriminate the two classes and can be seen as the ideal observer template for this case known as non-prewhitening matched filter (NPWMF) (Barrett *et al.* 1993). The test statistic λ for the NPWMF observer applied to an image f is described by (Barrett and Myers 2003):

$$\lambda_{NPW}(f) = \left[< f_2 > - < f_1 > \right]^T \cdot f \tag{2.20}$$

where $< f_k >$ is the ensemble average over all the data belonging to class k (where k=1,2). For a binary classification task where there are two classes of images, k equal to 1 corresponds to the signal absent image f_B and k equal to 2 corresponds to the signal present image f_L. The angular brackets denote an average.

NPWMF is the ideal observer when the noise is normally distributed and white (Barrett *et al.* 1993). However, as the human visual system does not have the same visual response to excitations with different spatial or temporal frequencies it has been suggested to incorporate in the NPWMF model observer information about the human visual system properties such as the eye contrast sensitivity function (i.e. ability of the eye to perceive a signal as a function of its frequency content) to increase its performance (Burgess 1994).

Furthermore, NPWMF observer uses has no information about noise correlation in the image such as any variations in the signal or the background. Due to this, it has been found to yield high correlation with human for images containing uncorrelated noise but if there is inherent randomness in the signal detection task, its performance depends on the correlations in the pattern of the background and/or the signal in the image and may yield a poor performance level. In such cases the noise inherent in the image must be decorrelated (i.e. data whitening). To do this, the image is filtered with the inverse of the covariance matrix and the result is correlated with the difference of the mean images, a process known as matched-filtering (Barrett and Myers 2003). This observer is known as prewhitening matched filter (PWMF) due to the data whitening before the matched filter operation. Under the assumption that the covariance matrix is the same for both classes of images the prewhitening operation is performed by multiplying each with the square root of the inverse of the covariance matrix and the test statistic for the PWMF, λ_{PWMF}, is given as follows:

$$\lambda_{PWMF}(f) = \left[\langle f_B \rangle - \langle f_L \rangle\right]^T \cdot K^{-1/2} \cdot K^{-1/2} \cdot f = \left[< f_B > - < f_L >\right]^T \cdot K^{-1} \cdot f \qquad (2.21)$$

where K is the covariance matrix. For white noise only the diagonal elements of the covariance matrix of the background have a non zero value. Therefore, under the assumption of equal diagonal terms the covariance is equal to the identity matrix and the PWMF becomes the NPWMF observer.

However, both NPWMF and PWMF observers are performed under the assumption of a known signal superimposed on a known background task so that the only randomness in the task is the measurement noise. In more complex situations in addition to the measurement noise, the data may also be random because of the randomness of the objects being imaged and the covariance matrices of the data from the two classes can no longer be assumed as identical. In cases where the signal or background are random or even the noise is not Gaussian the test statistic can be constrained from the outset to be linear and this is known as the Hotelling observer (Barrett *et al.* 1993). The Hotelling observer requires estimation of two scatter matrices, the interclass (S_1) and intraclass (S_2). The interclass scatter matrix S_1 measures the distance between the means of the two classes and is given by:

$$S_1 = P_1 P_2 \left(\langle f_2 \rangle - \langle f_1 \rangle \right) \left(\langle f_2 \rangle - \langle f_1 \rangle \right)^T \tag{2.22}$$

where P_1 and P_2 is the probability of occurrence of class 1 and 2 respectively.

The intraclass scatter matrix S_2 is the average covariance matrix across all the data from the two classes and is given by:

$$S_2 = \sum_{K=1}^{2} P_k K_k = \sum_{K=1}^{2} P_k \left\langle \left[(f - < f_k >) \, (f - < f_k >)^T \right] \right\rangle \tag{2.23}$$

where K_k is the ensemble covariance matrix of the data for the class k (where k=1,2).

Based on the scatter matrices a measure of the separability of the two classes, called the Hotelling trace J, is defined as (Smith and Barrett 1986; Barrett *et al.* 1993):

$$J = tr(S_2^{-1} S_1) \tag{2.24}$$

where tr denotes the trace of the matrix (i.e. sum of the diagonal elements of the matrix).

The Hotelling observer can be considered as the linear observer that use the scatter matrices and optimizes the Hotelling trace J in order to achieve maximum discrimination performance (Barrett *et al.* 1993). For a binary classification problem where there are two classes of images the optimal linear discriminant for the Hotelling observer is:

$$\lambda_{HO}(f) = \left(\langle f_2 \rangle - \langle f_1 \rangle \right)^T S_2^{-1} g \tag{2.25}$$

However, to obtain a stable statistical definition of the signal and the background a large amount of images is required. In addition, a number of studies have shown that the Hotelling observer has higher performance than the human observers as it does not take into account any information about the human visual system (Barrett and Myers 2003). To overcome these issues the Hotelling observer is constrained to a number of frequency selective channels and is known as channelized Hotelling observer (CHO) (Myers and Barrett 1987; Gifford *et al.* 2000; Gallas and Barrett 2003).

Firstly, the human eye is sensitive only to the total power in a series of frequency channels and not to individual frequencies. Therefore, the entire frequency content of an image within a given frequency channel is considered as a single value by the human eye and channels can be added to act as frequency selective filters in a similar way with the human visual system.

The second advantage of adding channels is that the number of elements of the covariance matrix as seen through the channels is reduced to a limited set of channels; if an image with N number of voxels is processed through M number of channels the image vector consisted of N elements is reduced to a feature vector consisted of M elements (i.e. each feature is the output of a frequency channel). As a result the Hotelling observer methodology is applied to the M element feature vector instead of the N element image vector and the estimation of the first and second order statistical information is used, as estimating the statistics of the whole image would not be possible. Therefore, the test statistic for the CHO, $\lambda_{CHO}(f)$, is given by:

$$\lambda_{CHO}(f) = \left(\langle Uf_2\rangle - \langle Uf_1\rangle\right)^T S_2^{-1} Uf = \left(\langle r_2\rangle - \langle r_1\rangle\right)^T K_U^{-1} r \qquad (2.26)$$

where the channel space is defined by a matrix U and $r_k = Uf_k$ is the mean for the data f of class k after transformation through the channels. Under the assumption that the two classes have the same probability of occurrence, ($P_1 = P_2 = 50\%$) the K_U is matrix averaged from the covariance matrices for the two classes derived after transformation through the channels.

The prewhitening operation using the K_U matrix aims to decorrelate the noise inherent between the channel responses and not the noise between the image elements (i.e. voxels) as in the case of Hotelling observer. CHO has proven to be good predictor of human visual performance in clinically relevant visual tasks such as the detection of lesions in images with correlated noise such as PET (Kadrmas *et al.* 2009; Lartizien *et al.* 2004; Shidahara *et al.* 2006; Nuyts *et al.* 2009).

2.1.6 Standard data corrections for quantification

The measured data are affected by a number of factors, including physical factors such as photon attenuation and scatter of the detected events as well as scanner geometric effects and detector efficiency variations. These factors may cause inaccurate quantification of tracer uptake, increased image noise and artefacts and thus they need to be accounted for before or during the image reconstruction process.

When an annihilation photon interacts with matter it changes direction and loses energy, a process known as Compton scatter and it may be absorbed within the body (i.e. photon attenuation) or detected by a detector resulting in an inaccurate definition of the annihilation event. Scatter results in diffuse background counts in reconstructed images, reducing contrast and distorting the relationship between image intensity and activity concentration.

Various methods for scatter correction have been proposed, such as: energy-based techniques; convolution or deconvolution; and simulation-based (Zaidi and Koral 2004). Energy-based methods use dual (Grootoonk *et al.* 1996; Adam *et al.* 2000), triple (Shao *et al.* 1994) or multiple (Bentourkia *et al.* 1993) energy windows to reject the detection of some scattered coincidences based on the energy difference between scattered and unscattered events. However, some unscattered photons deposit only a portion of their energy in the detectors and therefore total discrimination of the scattered and unscattered events based on their energy is not feasible. Another approach for scatter correction are the convolution (Bailey and Meikle 1994) or deconvolution (Floyd *et al.* 1985) methods. These methods assume that the scatter distribution is slowly spatially varying and use a scatter kernel to estimate the distribution of scatter from the photopeak measured projections via convolution-subtraction or deconvolution. This method is limited because the parametrisation of the scatter distribution is difficult especially for 3D acquisition.

Finally, simulation-based scatter correction methods model the scatter distribution by using emission and attenuation data with a combination of the physics of photon interactions and and detection process. This method has been applied based on Monte Carlo simulation which provides a model of the whole scatter distribution (Holdsworth *et al.* 2002; Guérin and El Fakhri 2008) or on single scatter simulation (SSS) (Watson *et al.* 1996; Watson 2000; Accorsi *et al.* 2004) which models only the single scatter distribution. For the SSS multiple scatter coincidences are modelled by appropriate scaling. The most common way is tail-scaling that fit the tails of the scatter distribution to the tails of the measured data (i.e. measured coincidences outside the patient) (Watson 2000). The tails of the measured data may consist of scattered events originating from the patient but also from outside the FOV or other factors and therefore scaling may also compensate for these factors which might not be otherwise simulated. However, tail-scaling may become unstable in cases where the tail region is very small (e.g. large patients). To overcome this disadvantage Thielemans *et al.* (2007) proposed to use all measured data to calculate the scale factors (i.e. total-scaled) providing a potentially more accurate scatter estimation procedure. Nevertheless, there are cases where multiple scatter broadens the scatter distribution and thus it is wider than the distribution of single scattered events. This broader distribution of multiple scatter may lead to a mismatch between single and total scatter distribution. Thus, scaling of the single scatter estimation might not be an accurate approximation of the total scatter distribution, and simulation of double scatter events might be necessary (Tsoumpas *et al.* 2005; Qian *et al.* 2010).

Photons are attenuated while passing through the human body mainly due to Compton scatter but also due to photoelectric effect which is very unlikely at 511 keV photon energies (Bailey *et al.* 2005). The photon attenuation depends on the distance that photons have to travel through an attenuating medium to reach the detectors, the energy of the photon and the density of the medium. The annihilation photons usually have to travel different distances inside the human body and thus they are attenuated to a different degree (Bailey *et al.* 2005). The extent of attenuation is quantified mathematically by a coefficient providing a measure of the photon attenuation probability per unit length of the attenuating medium.

Attenuation correction factors can be determined for every LOR so that the detection probability along a given LOR is not dependent on the point of the source along that line but only on the total path length that is followed by both photons (van Velden *et al.* 2008). Due to this, the spatial distribution of the attenuation coefficients in the FOV of the scanner (i.e. attenuation map) can be derived by the use of an additional external radionuclide transmission scanning (e.g. ^{153}Gd or ^{68}Ge), X-rays or CT scanning. The attenuation coefficients are energy dependent and can be directly estimated for a 511 keV energy source or can be derived using appropriate scaling. In PET/CT which is currently the gold standard for PET studies and has emerged to replace the PET as a stand alone modality, attenuation correction can be performed using CT data (Kinahan *et al.* 1998; Burger *et al.* 2002; Bai *et al.* 2003). For CT, the energy of the photons is much lower than 511 keV, usually around 50 to 90 keV. The attenuation map for PET is therefore derived by converting from the CT image the Hounsfield units which are measured for x-ray energies to suitable attenuation coefficients values for 511 keV photons by applying appropriate conversion functions for each tissue (Kinahan *et al.* 2003).

Attenuation correction can be performed either prior to reconstruction by multiplying the projection data with the appropriate correction factors, after reconstruction in the image space (Chang 1978) or in the step of reconstruction by computing weights from the attenuation correction factors map for each projection bin, and using them as multiplicative terms for each bin (Michel *et al.* 1998).

The fraction of prompt events due to random coincidences is a function of the positron emissions per second. As with scatter events, random events may decrease the reconstructed image contrast and cause uncertainties in quantification. Random events are usually estimated using a delayed or a shifted coincidence timing window (Meikle and Badawi 2005). The estimated random events can then either be subtracted from the measured events before reconstruction or can be accounted for within reconstruction as a background term along with other corrections. Another way to estimate the random coincidences is by using the rate of detected single events for a given detector pair. Random coincidences can be reduced by reducing the size of the coincidence timing window but at the expense of losing true coincidences.

The raw data can not be reconstructed assuming that all coincidence events are detected with the same scanner sensitivity. In practice there are variations mainly caused by three factors: differences in detector efficiencies, physical variations in the detector crystal and variation in the photomultiplier gain. These variations must be corrected otherwise they will lead to artefacts and inaccuracies in the quantification of the reconstructed images. Such correction, known as normalization correction, ensures that all LORs joining detectors in coincidence have the same effective sensitivity.

Correction factors are calculated for each LOR (i.e. normalisation coefficients) with two main approaches. The first and most common method, known as direct normalisation, determines the normalisation coefficients by performing a high-statistics scan where every possible LOR is illuminated by the same source and therefore all LORs are exposed to the same number of annihilation events (Badawi *et al.* 1998). This method typically uses a rotating line positron source in order to measure the sensitivity of each LOR within the FOV. Then, the normalisation coefficients are assumed to be proportional to the inverse of the measured counts for each LOR. The second way to calculate these coefficients, known as component-based normalisation (Casey *et al.* 1995) is to model the coefficients into a series of individual components and compensate for each of them separately. These components are typically calculated separately by performing one scan with a rotating line source, which provides the geometric characteristics and a second scan with a cylindrical source, which provides the crystal efficiencies. In addition, these components usually include pair-wise detector efficiencies, crystal interference, axial and radial geometric factors, as well as dead-time and count-rate dependencies (Badawi and Marsden 1999; Rodriguez *et al.* 2007).

As it was described before compared to any corrections applied prior to reconstruction, modelling the corrections within the reconstruction preserves Poisson statistics. Although there are a number of other degradation factors such as the PVE only the aforementioned are routinely dealt with in clinical practice.

2.2 Magnetic resonance imaging

Magnetic resonance imaging (MRI) is a non-invasive imaging modality that produces anatomical and functional images of the human body by measuring the magnetic properties of hydrogen nuclei inside the body. Most currently available MRI scanners are constructed from a strong magnet (1.5 to 3 Tesla magnetic field for most conventional scanners) surrounding the patient, radiofrequency coils and gradient coils. An example of a typical MRI scanner is illustrated in figure 2.11.

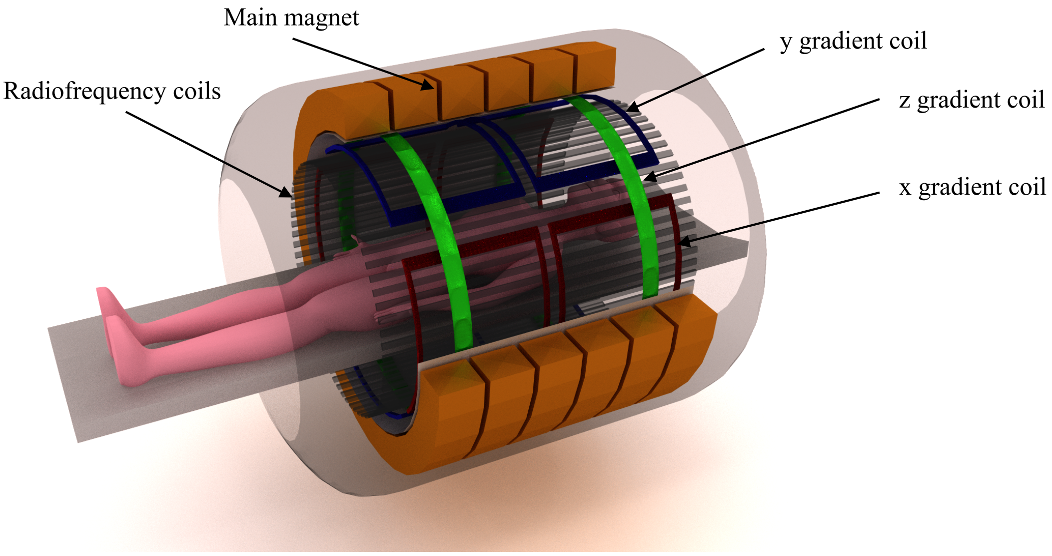


Figure 2.11: A typical MRI scanner is constructed from a strong magnet surrounding the patient, radiofrequency coils and gradient coils.

MRI is based on the detection of signal derived from the nuclear magnetic resonance (NMR) property of nuclei of atoms, in particular hydrogen, within a tissue in the body. Hydrogen is one of the main nuclei in a human body with the highest abundance occurring in water followed by fat. The hydrogen nucleus has one proton with spin equal to 1/2 resulting in a small magnetic moment[‡].

Typically, the hydrogen protons in a tissue are oriented in random directions resulting in no net magnetization at the macroscopic level. By subjecting the hydrogen nuclei to a strong external magnetic field the protons will align either parallel (i.e. spin up) or anti-parallel (i.e. spin down) to the direction of the external magnetic field and precess with a frequency proportional to the frequency of the external magnetic field known as the Larmor frequency ω_0 given by the following equation:

$$\omega_0 = \gamma B_0 \tag{2.27}$$

where γ is the gyromagnetic ratio and B_0 is the strength of the external magnetic field. This procedure at equilibrium creates a net magnetization M_0 in the direction of the external magnetic field.

To change the spin state of the protons energy must be applied at the Larmor frequency. Through the application of a tuned radio frequency (RF) field equal to the Larmor frequency of the nuclei[§] and perpendicular to the external magnetic field, the proton spins will precess in a transverse plane to the external magnetic field and will be aligned such that they precess in phase. The RF field is applied using appropriate coils and forces the protons to spin at a particular frequency and direction causing a deflection of the net magnetization. Figure 2.12 shows the effect of applying the RF pulse on the equilibrium net magnetisation. The net magnetisation M_0 is pushed towards the x-y plane (i.e. transverse plane) while it precesses around the z-axis at the Larmor frequency. As the RF pulse is turned off, due to interactions with the protons themselves and the protons with the surrounding molecules (i.e. lattice) the

---

[‡] In the nucleus of an atom, protons spin on their axis and as a result a small magnetic field is produced. For a nucleus with an even number of protons the magnetic fields tend to cancel each other. On the other hand, nucleus with an odd number of protons, such as the hydrogen which has one proton, has a net magnetic moment.

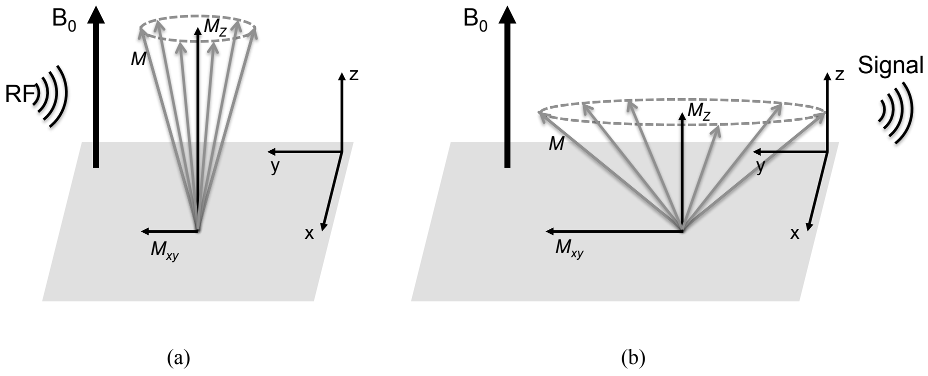
spins slowly realign back from the transverse magnetization to their original orientation (i.e. direction of the applied external magnetic field), a process known as relaxation. During relaxation the protons lose the energy absorbed from the RF pulses, emit radio frequency radiation, and generate the characteristic free induction decay (FID) curve.

Relaxation is due to two processes (McRobbie *et al.* 2006). The excited protons due to interactions with the surrounding molecules (i.e. lattice) return to the initial magnetization, before the RF field, parallel to the applied external magnetic field. The rate of this process is described with an exponential function with a characteristic time constant T1 known as longitudinal relaxation time. Furthermore, the other relaxation process is due to interactions with the protons themselves causing a loss of coherent precession of protons in the transverse plane. The rate of this process is described with an exponential function with a characteristic time constant T2 known as transverse relaxation time. The T1 and T2 relaxation time constants depend on environmental factors such as the degree of binding of the spinning protons within the neighbouring molecules and they also characterise inherent properties of tissues and are unique to every tissue at a given strength of magnetic field and therefore are responsible for the contrast between different tissue types in MRI. By changing the RF pulses the differences of the T1 and T2 relaxation times can be used to create contrast among different tissues by creating T1-weighted or T2-weighted images during which the T1 or T2 relaxation time respectively influences more the image contrast.

To create an image you need to know the origin of the signal within the patient. By exciting all the protons at the same time, the emitted signal would have no information about the position of the hydrogen nuclei in space as the Larmor frequency depends on the magnitude of the external main field. To achieve spatial localization of the

---

§ For a scanner of 1.5 to 3 Tesla external magnetic field the Larmor frequency falls within the radiofrequency band of the electromagnetic spectrum.



(a) (b)

Figure 2.12: Effect of RF pulse on the net magnetization (a) The spin vector precesses in the direction of the main static magnetic field B_0 (b) after applying RF pulse the spin vector precesses about the z-axis in the direction of the B_1 field generated from the RF pulse until returning back to its equilibrium state.

signal, field gradients are used to create a spatial dependency of the Larmor frequency a process known as spatial encoding. Three orthogonal gradient coils are used able to generate magnetic fields linearly variable in magnitude in one of the three axis in the coordinate system.

The first step of the spatial encoding is the slice selection. Following the RF pulse if a spatially varying gradient is applied, the excited protons will resonate at different frequencies which depend on their location to the gradient field. Using gradient fields that cause the Larmor frequency to vary along the direction in which the gradient is applied, spins within the slice will precess with a particular frequency and a cross-section of tissue (i.e. slice) can be selected.

The central frequency of the RF pulse determines the position of the slice along the direction of the gradient and the bandwidth of the RF pulse determines the thickness of the slice. Therefore, by adjusting the direction and frequency bandwidth of the RF pulse to correspond to the Larmor frequency range within the desired slice the RF pulse can only excite the slice with nuclei with Larmor frequency within the selected frequency range. This gradient field is known as slice-selection gradient and although they can be applied in any direction of the scanner (i.e. transverse, coronal, sagittal or oblique), they are usually applied in the transverse direction. The MRI images are usually acquired as 2D slices along the direction of the chosen axis.

After slice selection, the spatial distribution of the protons within the particular slice must be found for the other two axes. This is achieved using a frequency-encoding and phase-encoding gradient to encode the other two axes by frequency and phase respectively. The frequency-encoding and the phase-encoding gradient coils are then used to reconstruct the selected 2D slice. For a transverse slice acquisition usually the x direction is the frequency encoding and the y direction the phase encoding.

A phase encoding gradient is applied in a direction perpendicular to the slice selective gradient. Precessional frequency of the protons is increased or decreased depending on their position relative to the gradient. This introduces a spatially dependent phase shift between nuclei in different regions in the FOV that is preserved once the gradient is removed. At this moment, all spins within the slice return to their original precession frequency, caused by the static magnetic field. Nevertheless, spins which precess faster during the time that the gradient was on rotated farther than the spins which precess slower, thus they have different phases of rotation and the phase encoding gradient divides the slice into columns. The next step is to apply a frequency encoding gradient at the time that the RF signal is received by the patient. When the gradient is switched on similar to the slice selective gradient the precessional frequency of the spins along the axis of the gradient is altered linearly. The frequency encoding gradient changes the precessional frequencies for each row of the slice and when the gradient is turned off some spins have different phases, thus the slice is divided into rows. This gradient is also called readout because it is applied during the collection of the data (i.e. readout phase).

During image acquisition this sequence of pulses is repeated, increasing the amplitude of the phase encoding gradient each time to collect all data needed to produce an image. Having introduced spatially varying phase and frequency shifts we can recover this information from the signal emitted as the spins decay using Fourier techniques.

The MRI signals encoded in the three directions are reconstructed to give an image of which the intensity describes the relaxation properties in the subject tissue. MRI compared with PET provides anatomical soft tissue contrast (e.g. tissue specific contrast agents) and does not use ionizing radiation. MRI is particularly well suited for imaging soft tissue which has a high concentration of water molecules. On the other hand, for cortical bone the T2 is as short as it takes for the emitted signal to decay during acquisition and therefore it is very difficult to measure (Du *et al.* 2007). Due to this a special acquisition has been developed known as ultra-short echo time (UTE).

MRI is an anatomical imaging modality which, due to its high resolution and short acquisition times, has now been established in oncology, where it is especially useful for the detection of metastases and haematologic malignancies in the lung and liver.

2.3 Hybrid PET/MRI imaging

PET is a molecular imaging modality but with poor visualization of anatomical information. In current clinical practice it is often required to combine the functional information of PET with anatomical information such as CT or MRI to obtain localization of the increased radioactivity distribution.

Several studies have shown that the combination of functional and anatomical information from PET and CT respectively (Beyer *et al.* 2000) provides more information for clinicians for more accurate diagnosis and management of cancer compared with CT or PET alone (Antoch *et al.* 2003; Kim *et al.* 2005). PET combined with CT systems are now used routinely clinically in oncology for lung and liver cancer. However, for some applications, the performance of PET/CT is limited by the low soft-tissue contrast of CT.

On the other hand, MRI is an anatomical imaging modality with significantly less sensitivity for the measurement of functional and molecular information than PET. Compared with CT, MRI is advantageous in terms of soft tissue contrast with no extra patient radiation dose. A typical FDG PET scan using 370 MBq of injected tracer will result in a patient radiation dose of approximately 7 mSv whereas a CT acquisition will add an extra whole body dose of less than 3 mSv according to the current low-dose CT scanners (Xia *et al.* 2012). The fusion of PET and MRI has been of great interest in the literature (Pichler *et al.* 2010; Ratib and Beyer 2011). Studies have shown advantages of combining PET/MRI over PET/CT on lesion delineation and characterization for hepatic (Donati *et al.* 2010) and liver cancer (Schreiter *et al.* 2012). A more recent study also showed that PET/MRI provided better diagnosis for liver lesions than PET/CT (Beiderwellen *et al.* 2013).

Simultaneous PET/MRI acquisition allows spatial and temporal correlation ensuring that the patient is being imaged in the same physiological state during both PET and MRI. The spatial correlation of PET/MRI may be advantageous for anatomically related tasks. For example, PET recognizes areas of abnormal metabolic behaviour (e.g. cancer) while the addition of MRI anatomical information provides the site of the malignancy. Hence, in practice, MRI can identify the location and size of a tumour with great precision and PET can define the metabolic activity of the tumour. For example, the size of a tumour may not alter after therapy and MRI will identify no change. However, PET may show significant reduction in metabolic activity in the location of the tumour, which can be associated with a positive response to therapy.

Beyond the benefit on anatomical related tasks such as tumour location and size, the spatial and temporal correlation of PET with MRI allows the use of anatomical information of MRI images for other tasks such as motion correction. The accurate measurement of motion is one of the motivations for combining MRI with PET. As it is described later, motion correction methods usually require prior knowledge of motion, the accuracy of which affects the accuracy of motion correction. Motion information for PET images has been previously extracted using PET alone or from CT. However, PET images are noisy thus it is difficult to derive motion information from them, while CT gated data have better quality but result in an additional radiation dose. PET and CT data are not acquired simultaneously, so CT motion information may differ from motion occurring during PET. On the other hand, MRI can in principle offer continuous motion monitoring with much better temporal and spatial resolution than PET with no additional dose. Therefore, it is expected that simultaneous PET/MRI acquisition may be able to provide the necessary information to correct for the motion in PET (Tsoumpas *et al.* 2010; Dikaios *et al.* 2012b). Finally, PET/MRI offers the opportunity to perform other studies using MRI sequences such as diffusion and perfusion imaging, MRI spectroscopy or functional MRI.

The integration of MRI with PET started in the late 1990's (Slates *et al.* 1999) and was suggested before PET/CT became commercially available (Pichler *et al.* 2010; Ratib and Beyer 2011). PET and MRI can be simultaneous or sequential (Zaidi and Del Guerra 2011b). The first human PET insert was introduced by Siemens for brain for 3 $Tesla$ (Schlemmer *et al.* 2008) and 7 $Tesla$ (Judenhofer *et al.* 2007). Later, Philips installed the first whole-body Ingenuity TF PET/MR, with sequential PET and MRI acquisition (Zaidi *et al.* 2011a). In the same year, Siemens released the first whole-body PET/MRI system with capabilities of simultaneous PET and MRI acquisitions, the Biograph mMR (Delso *et al.* 2011). The role of PET/MRI, particularly in oncology is promising. While PET/CT is now used routinely clinically in oncology as the reference method for whole body staging and management of cancer, PET/MRI may become a standard hybrid imaging modality in the future. PET/MRI is of emerging interest for abdominal imaging (Schwenzer *et al.* 2012) and whole-body oncology (Drzezga *et al.* 2012) but also for several clinical applications such as tumour neuroimaging (Boss *et al.* 2010), prostate cancer (Bouchelouche *et al.* 2010), cardiac metabolism (Büscher *et al.* 2010) and cardiac function (Nekolla *et al.* 2011).

2.4 Respiratory motion blurring in PET

In typical clinical settings a whole body PET ^{18}F-FDG scan usually requires 3-5 minutes per bed position (5-8 bed positions for a whole body acquisition) to accumulate a sufficient amount of counts and therefore patient motion is inevitable. During PET acquisition patient motion may constitute accidental voluntary motion such as head movements and unavoidable motion such as respiratory and cardiac motion. The patient voluntary motion during the acquisition can be generally controlled by the patient. Cardiac contraction affects mainly the heart and is a major concern of cardiac imaging applications rather than for oncologic imaging as is the focus of this project. On the other hand, during respiration not only the lungs oscillate between contraction and relaxation but also multiple other organs are affected such as the liver and the heart since they are pushed and pulled by the diaphragm and other inter-connected tissues. However, the diaphragm as well as the thoracic and abdominal organs have displacements to a variable degree in different directions. In this thesis the subject of interest is the respiratory motion and other causes of motion will not be investigated.

For normal breathing, the amplitude of diaphragm motion due to respiration is about 15 to 20 *mm* in the vertical direction (Schwarz and Leach 2000) and it may be more for deep breathing. A number of investigations reported that the displacement caused by respiratory motion is usually less than 10 *mm* for lung lesions but it may be more than 20 *mm* for lung and liver lesions close to the diaphragm (Seppenwoolde *et al.* 2002; Erdi *et al.* 2004). Motion of that range was also reported by a review of respiratory motion provided by Langen and Jones (2001). Understanding the behaviour of tumour motion under breathing is important for addressing the motion effect. Based on studies of twenty patients, Seppenwoolde *et al.* (2002) measured the displacement of tumours in lungs due to motion and showed that although motion mainly affects the axial direction, motion has also been reported in lateral and anteroposterior as well.

Considering that many of the conventional PET scanners have a spatial resolution of approximately 5 to 6 *mm* FWHM, motion is larger than the resolution and therefore is expected to affect image quality. Respiratory motion induces blurring and distortions in the images limiting the image spatial resolution (Xu *et al.* 2011; Nehmeh *et al.* 2004; Visvikis *et al.* 2006). Moreover with the advances in resolution of future scanners (Schäfers 2008a) the motion effect is expected to be even more substantial (Daou 2008) and the development of accurate motion correction methods is now a major task in medical imaging research. For example, the intrinsic resolution of the scanners offered by Siemens mCT with TOF capabilities and ultra high-definition (HD) PET[†], allows a uniform spatial resolution throughout the transverse FOV of 2 *mm* FWHM (Panin *et al.* 2006).

[†] HD PET scanners incorporate the detector PSF into the reconstruction to improve the accuracy of the detector to determine the point of intersection of the incoming photon. Incoming photons might penetrate several crystals before they interact with one of them causing non-uniform and reduced spatial resolution.

2.4.1 Effect of respiratory motion blurring

Motion causes uncertain quantification and reduces tumour detectability through blurring. Several studies using simulations, phantom or patient data reported underestimation of the radioactivity distribution measured in the images affected by respiratory motion. The magnitude of the effect of respiratory motion blurring is inversely related to the size of the tumour. In particular, Boucher *et al.* (2004) used an elliptical phantom to simulate lung tissue and motors to apply a simple periodic motion only in the vertical direction. The findings showed that for a displacement of 20 *mm* a decrease of the radioactivity distribution equal 21% for a 19.4 ml sphere diameter and 41% for a 1.2 *ml* sphere diameter. Using a similar methodology, Pevsner *et al.* (2005) for a displacement of 20 *mm* showed also underestimation of the tracer uptake concentration by as much as 75% depending on the sphere diameter. Using simulations of a dynamic digital anthropomorphic (NCAT) phantom, Visvikis *et al.* (2004) investigated the effect of motion based on a realistic breathing pattern derived from real CT scans but only using one respiratory cycle with phase based gating (10 frames of 0.5 *s* each). Similarly with previous studies the maximum displacement of the diaphragm simulated was 20 *mm* and a contrast loss of up to 80% depending on lesion size and location was reported.

However, all these studies investigated periodic motion or a simple respiratory cycle. In real studies with more complex breathing patterns, the extent of the impact of motion blurring on quantification accuracy may be different. In addition, although respiratory motion mainly affects the vertical direction, motion has also been reported in lateral and anteroposterior directions as well (Seppenwoolde *et al.* 2002). Using a simple oscillatory motion to drive the phantom movement varies from realistic patient respiratory motion patterns and also does not simulate the variation in real patient respiratory motion patterns. Breathing irregularities are expected to affect the impact of motion. Consequently, the amplitude of the effect of motion blurring is expected to be influenced by the lesion position and type but also on physiological issues related with the patient specific breathing patterns.

These parameters were characterised by Liu *et al.* (2009) who applied real patient respiratory traces from 5 *min* acquisition to a digital anthropomorphic phantom to simulate more realistically the effect of motion. The findings showed larger effect of respiratory motion as the size of the tumour was decreased and as the motion amplitude was increased. Furthermore, in this study 1295 real patient respiratory traces were grouped in three breathing patterns. The results showed that patients who spent more time in the expiration position are less affected by respiratory motion than patients who have more irregular breathing pattern. In particular, the results reported for lesions of 10 *mm* diameter the mean lesion maximum SUV was 28% underestimated based on 11 *mm* diaphragm motion for breathing characterised by large quiescent period fractions.

Apart from phantom studies, the effect of motion blurring on quantification accuracy was also reported on studies with patient data. Erdi *et al.* (2004) analysed eight lung lesions from five patients and reported up to 30% reduction of the lesion SUV due to respiratory motion in the PET images. The findings showed that the extent of these effects depend on the tumour size and the pattern of breathing with the amplitude of the displacement.

Despite the large amount of studies based on simulations there is limited number of studies on human data. One limitation of the studies using patient data is that it is not possible to obtain the true tracer uptake value of a lesion which is theoretically derived from a patient who does not breathe. For this reason most researchers use images obtained during end-inspiration and end-expiration as representative of a motion free image (Erdi *et al.* 2004). Another way is to use a deep inspiration breath hold acquisition during which image data are acquired at a single breathing position and compared to the normal breathing acquisition. This may reduce the respiratory blurring and can be used as a gold standard for the motion free image but these images have high noise (Meirelles *et al.* 2007).

The inaccurate underestimation of the tumour tracer uptake measurements due to respiratory motion during acquisition results in decreased contrast and the ability to detect tumours may be affected.

In typical clinical acquisitions, several studies demonstrated that the smallest volume of lesions detectable is affected by underestimation of the true tracer uptake of small lesions that may be attributed partially to the limited resolution of the scanner. Khalaf *et al.* (2008) in a study of 173 patients evaluated pulmonary lesions based on the SUV and found underestimation of the SUV for the small tumours (less than 10 *mm*). It has been further reported that as a result of not recovering the true tracer uptake of the tumour due to the limited resolution, the contrast is reduced and consequently the ability of PET to detect small or low tracer uptake lesions is deteriorated especially for small lesions. In a study with patient data Lowe *et al.* (1998) reported the correlation of PET detectability performance with the tumour size by showing up to 80% lower detection sensitivity for the lung lesions smaller than 15 *mm* than the 92% reported for larger lesions. Other studies reported also low performance of PET to recover small lesions (i.e. less than twice the scanner resolution) in phantom (Raylman *et al.* 1999; Farquhar *et al.* 1999) and real patient studies (Higashi *et al.* 2001; Nomori *et al.* 2004).

Considering that the recovery of small lesions is already a challenging task in PET imaging deterioration of the spatial resolution and underestimation of the tumour tracer uptake measurements due to the respiratory motion blurring will impede the ability to detect abnormalities. Although there have been several studies focusing on the impact of respiratory motion on quantitative accuracy, only a limited number of them demonstrated the impact of motion blurring on the detectability of lesions. One reason for the lack of detectability studies is the lack of gold standard in patient studies because it is not possible for the patients to hold their breath during acquisition. Another reason is that it is very difficult to perform sufficient number of patient studies to ensure accurate statistical power.

Despite the blurring in the reconstructed images, respiratory motion may affect the spatial correlation of a PET image with images from other imaging modalities. For example, during PET attenuation correction using transmission data such as CT if the attenuation map does not align with the PET data due to different breathing conditions respiratory motion may lead to distortions in the image (Goerres *et al.* 2002, 2003; Nehmeh *et al.* 2002; Beyer *et al.* 2003, 2004; Osman *et al.* 2003). In PET a normal breathing protocol is applied but the lengthy acquisition results in data that are an average over multiple respiratory cycles. In clinical practice, attenuation data are usually acquired before or after PET and this may cause erroneous spatial correlation between them. In particular although a normal breathing protocol is also applied for CT, the high speed result in a snapshot of one respiratory position without containing motion blurring and may only match one arbitrary bit of PET. Also any artefacts in the transmission image caused by motion will be introduced to PET images (Sarikaya *et al.* 2003). This artefact must be corrected in order to avoid a false positive result. Motion induced artefacts during attenuation correction may result in inaccurate tumour localization and estimation of tracer uptake affecting mainly the region near the dome of the liver and the myocardium (Pan *et al.* 2005; Schäfers and Stegger 2008b; Townsend 2008).

2.5 Respiratory motion compensation in PET

Respiratory motion results in blurring and artefacts in PET images, adversely impacting the usage of PET images for cancer diagnosis and staging as well as radiation therapy planning which are dependent on the tracer uptake, size and position of the lesion. The motion artefacts are expected to be particularly severe for lesions near the boundary between the lung and the liver. In addition to the location of the lesion, the extent of motion artefacts is expected to be dependent on the tracer uptake and size. As described earlier, the PET image quality may be characterised on how well the true tracer uptake can be recovered and how well the lesions can be distinguished from normal tissues. Based on these, motion blurring contributes to the clinical challenges of detecting small tumours or detecting tumours with low uptake. Correction of motion will help for more accurate image interpretation and clinical decision making.

2.5.1 Data respiratory gating

2.5.1.1 Extraction of respiratory signal

Various methods have been proposed to record respiratory traces either during acquisition using external motion monitoring systems or after acquisition of the acquired data. A direct method to measure respiratory motion is measuring the lung volume during respiration using a spirometer. However, such equipment is not practical and not well tolerated by the patients due to the long duration of a PET acquisition. Due to this, alternative approaches to measure respiratory motion have been suggested using other parameters such as the displacement of abdominal or chest wall. This can be measured either using hardware methods or directly from the acquired data.

Hardware methods usually measure the respiratory signal with an external surrogate which is usually the displacement of the abdominal wall or chest circumference such as the use of sensors that measure pressure differences (Klein *et al.* 1998; Guivarch *et al.* 2004) or the use of a video camera to track reflective markers (Nehmeh *et al.* 2002; Dawood *et al.* 2007). Temperature sensitive devices have also been proposed that determine the respiratory motion by measuring the airflow in the lungs (Boucher *et al.* 2004). Of the large number of hardware-based methods some are commercially available. Among them are the real-time position management (RPM) and the AZ-733V. The RPM tracks the vertical motion of a box positioned in the abdominal or chest region using an infrared video camera (Pan *et al.* 2004). The RPM method has been used by several studies and proved to produce accurate and reproducible measurement of the respiratory signal (Nehmeh *et al.* 2002). The AZ-733V measures the pressure differences from the expansion of the thoracic cavity during respiration using a pressure sensor that is placed into a belt that is fastened around the abdominal or chest of the patient (Seppenwoolde *et al.* 2007). Hardware-based methods apart from the additional equipment that require to be applied are based on the approximation that the movement of the outside surface of the body represents the actual lung movement.

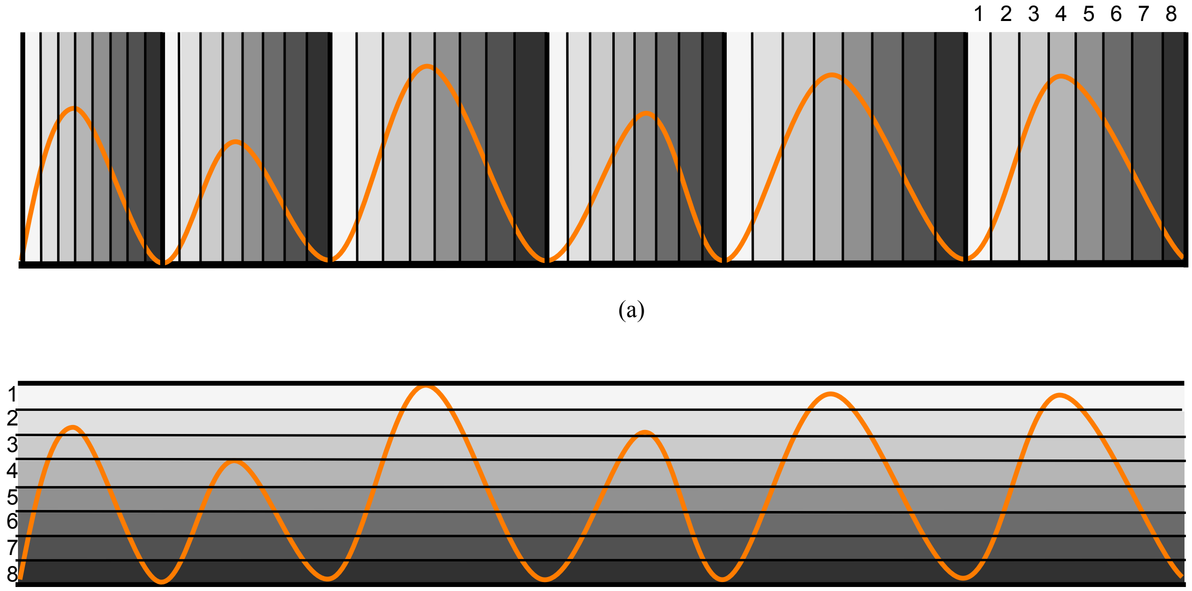
An alternative method to determine the respiratory signal is by measuring it from the actual acquired data. Data-driven methods usually measure displacement of the image elements (e.g. pixels) in the reconstructed images. Some methods are based on defining the displacement of an ROI in the dynamically reconstructed images (Bundschuh *et al.* 2007). However, these methods require many images due to the long duration of a PET acquisition making them not practical to be used clinically. Other methods have used raw data before reconstruction to measure the displacement of specific pixels by measuring their intensity variations (Schleyer *et al.* 2009, 2011). Other methods measure the counts in the raw data (Büther *et al.* 2009) or using CT together with PET raw data and apply principle component analysis (PCA) (Thielemans *et al.* 2013).

2.5.1.2 Data gating approaches

Using all the above mentioned methods a respiratory signal is created which is used to gate the acquisition. Gating divides the respiratory cycle into a number of gates (i.e. period in which the anatomy is assumed to be in a similar spatial location) according to the phase or amplitude of respiration and acquired data are stored for each gate as illustrated in figure 2.13. In a PET gated acquisition the respiratory signal is tracked over multiple cycles and the time information of each event is correlated with the motion signal. The raw list-mode data are then binned into multiple sinograms corresponding to the different gates of the tracked motion. Under the assumption that the tracked signal correlates with the motion of all organs of a patient in a periodic way for all respiratory cycles of the entire acquisition, a near motion-free image may be reconstructed for each gate.

The phase-based technique divides every cycle between two points of the maximum inspiration phase considering the time duration and neglecting the depth of breathing. Thus two inspiration phases from different cycles that may have different amplitudes, if they have the same time phase will be added together. This method results in gates of the same duration ensuring similar counts among the gates. However, studies have shown that the phase correlation of motion changes over time for the same patient and therefore phase gating may not reflect the motion amplitude for all respiratory cycles (Ozhasoglu and Murphy 2002).

On the other hand, the amplitude-based gating scheme splits the respiratory cycle with respect to the amplitude of the signal and only phases with the same breathing magnitude will be added together. As the total time for each gate amplitude can vary



(a)

(b)

Figure 2.13: Illustration of the procedure followed for (a) phase and (b) amplitude based gating of a respiratory signal. For this example in both cases the respiratory signal is divided in eight gates which are represented by eight different grey shades numbered from one to eight.

this method may result in different number of counts. The advantage of this method has been reported by Dawood *et al.* (2007) who showed that respiratory gating methods based on amplitude result in less noisy data than methods based on phase. However, under regular breathing, phase and amplitude based gating may perform similarly. On the other hand, under variable breathing amplitude and period, the amplitude based gating may outperform phase based gating (Dawood *et al.* 2007).

Based on the approximation that images at discrete phase or time bins within the respiratory cycle are motion free, respiratory data gating is a common solution to reduce motion blurring in PET images (Dawood *et al.* 2007; Nehmeh *et al.* 2002). Optimized gated acquisition aims to achieve the best temporal resolution (i.e. accurate motion detection) and least possible noise in the images. While the temporal resolution is improved with the number of gates the SNR of the each gated image is decreased since they have a smaller proportion of the total measured data.

To optimize the temporal resolution (i.e. accurate motion detection) and noise trade-off the optimal number of gates must be determined. Dawood *et al.* (2009) proved that with respect to the accurate motion detection and the level of noise the amplitude based gating allows the use of more gates compared to phase based gating and showed that the optimal number of gates for time-based respiratory gating is six whereas for amplitude-based it is eight. Although, this might be used as a general indication of the optimal number of gates it must always be taken into account that the number of gates is related with the count statistics and consequently with the scanner sensitivity or patient.

The limitation of single gate reconstruction is that each gate contains only a proportion of the measured data resulting in reconstructed images with lower signal-to-noise ratio (SNR) than un-gated data that contain all measured data (Boucher *et al.* 2004). As a result, the advantages of minimizing motion are at the expense of reduced statistics of each individual gate (Visvikis *et al.* 2004). Thus, a technique accounting for motion effects which uses all measured data in order to improve the SNR is expected to provide superior results.

Different motion correction methods have been proposed that make use of all the gated measured data together with motion information in order to compensate for motion but also to improve SNR. Respiratory motion correction is a difficult problem due to the poor statistical quality of PET data and many approaches have previously been investigated and are reviewed in Rahmim *et al.* (2007); Nehmeh and Erdi (2008); Nehmeh (2013). Motion correction is divided into two parts. First of all motion is estimated and then is corrected.

2.5.2 Image registration

Image registration is the process of spatial alignment of two images in order to warp the one image to the other such that a feature in the first image is at the same position with a corresponding feature in the second image. Image registration is extensively used in medical imaging of the same imaging modality (intramodality registration) or from different imaging modalities (intermodality registration) (Hutton *et al.* 2002). An overview of image registration and its steps is provided by Crum *et al.* (2004). In general the image registration can be divided in the following components:

- Similarity: Calculation of how well the transformed image matches the target image.
- Transformation: Determination of the transformation that relates the positions of the corresponding features of the two images. The transformation controls the elasticity of the deformations described by the number of parameters used to describe the deformations (e.g. number of degrees of freedom) and the domain applied (e.g. locally or globally).
- Optimization: Calculation of transformation parameters that maximize or minimize (depending on the matching criterion) the similarity measure also known as cost function in order to improve the alignment of the two images.

The transformation warps the source image to the target image according to the deformation parameters. The transformation operator can be applied globally to the whole image or locally to a specific region. The choice of the transformation operator depends on the type of deformation. Several types of transformation operators exist in the literature such as rigid with the more complex affine and non-linear transformations having successively more degrees of freedom.

Considering an orthogonal coordinate system (x,y,z), a rigid transformation allows 3D rotation and 3D translation resulting in 6 degrees of freedom. During this type of motion no shape or size changes are allowed. The rigid motion due to the small number of degrees of freedom used to parameterize the movement usually models simple motion for example head motion. For brain imaging, the motion has been successfully modelled as rigid in many studies (Bloomfield *et al.* 2003).

Affine transformation is an extension of rigid motion and allows twelve degrees of freedom, the 3 rotational and 3 translational components of rigid with additional 3 scaling and 3 shears components. Although, for small deformations rigid or affine transformation can be a simple way to approximate motion, the motion of human body is complex (e.g. chest and abdomen motion due to respiration) and requires flexibility in matching local image details (Crum *et al.* 2004). Non-linear transformations have more degrees of freedom to describe the more complex movement.

Many algorithms have been suggested for non-linear image registration (Crum *et al.* 2004). One method of non-linear transformation is based on firstly defining the deformation of a given set of control points followed by interpolation of motion fields between these points. In particular, for a given image voxel only the closest control points influence its displacement, therefore the registration accuracy is highly dependent on the selection of the control points. Various transformation methods have been proposed with the ability to model local deformations such as the spline transformations (i.e. piecewise polynomials). During spline-based transformations, given a set of control points splines are applied to each control point in the image and the image is interpolated such that the location of each point at the first image matches the corresponding point at the second.

Although applying local non-linear transformations has been shown to produce accurate motion fields even for the complex motion of respiration, it is highly computationally complex due to the large number of degrees of freedom. An alternative and more efficient technique is a combination of local and global transformations such as rigid and affine (Little *et al.* 1996; Commowick *et al.* 2008). However, the particular approaches are based on predefined regions and are not very practical when registration is subject dependent and these techniques must be performed for each image separately.

An alternative approach uses hierarchical multiple local registrations in which the image is subdivided into rectangular blocks and each block has its local transformation. To combine multiple local registrations in an adaptive hierarchical structure the whole FOV is firstly selected to be registered and then is divided into multiple blocks using the centre of each block as a splitting point Hierarchical local registrations have been applied for local affine firstly in 2D (Likar and Pernus 2001) and extended to 3D (Andronache *et al.* 2008) transformations but without any consideration for the underlying content of the image blocks. As a result areas with large residual motion may be subdivided causing registration artefacts at the block edges.

Buerger *et al.* (2011) suggested multiple local affine registrations in an adaptive hierarchical structure during which instead of subdividing the image into equal size blocks, the image is divided in subblocks that correspond to areas of anatomy that have similar motion and these subblocks are registered to the reference image. To do this the splitting point is not the centre of each block but is adaptively selected based on the underlying motion. By applying adaptive splitting of the image, regions with similar motion may be considered as one block and therefore large deformations are allowed to be processed in one block. Similar to the method proposed by Likar and Pernus (2001) and Andronache *et al.* (2008) after splitting the image into blocks, interpolation is applied to determine an overall smooth deformation field that combines all local affine transformations of the registered blocks to a global non-linear free-form transformation and the procedure of blocks splitting, registration and interpolation is repeated until final transformations are calculated. This technique was compared to the local rigid and local affine approaches and was found more efficient and accurate especially for large deformations and is applied for the purposes of this thesis[||] (Buerger *et al.* 2011).

After defining the type of transformation, the similarities of the two images are measured iteratively until an optimum transformation is determined that maximizes the matching criterion of details between the two registered images. The most commonly used similarity measures are the sum of squared differences (Crum *et al.* 2004), the correlation coefficient and mutual information (Wells *et al.* 1996). Sum of squared differences calculate voxel similarities from the voxel intensity values directly assuming identical intensities. Similarly the correlation coefficient calculates the voxel similarities directly from the voxel intensity values but assumes a linear relationship between the intensities of the images. Both the sum of squared differences and the correlation coefficient cost functions assume identical or linear relationship and therefore are limited to intramodality registration. Mutual information determines the relationship among the image histograms or feature space (i.e. statistical dependency between the images) based on probabilities without assuming identical or linear relationship and therefore it is applicable to intermodality registration.

Finally, when the transformation type and the cost function are chosen iterative optimization calculates improved transformation parameters by usually maximizing or minimizing (depending on the type of cost function) the similarities between the

---

[||] The registration algorithm is available at http://www.isd.kcl.ac.uk/internal/hyperimage.

two images that are registered. Finding the optimum transformation parameters is challenging in cases when there are a large number of degrees of freedom.

Image registration has been used in medical imaging for estimating organ deformations. Images (e.g. PET, CT or MRI) acquired at different positions of the motion cycle are spatially registered to the image at a reference position to extract motion information. The problems encountered with estimating the motion information from CT is that motion can be different between sequential CT and PET acquisitions and there is extra patient radiation dose. This aspect could be handled by measuring motion directly from PET, but PET data are noisy and in addition it would be difficult to generalize this technique to any radiotracer. The development of whole-body simultaneous PET-MRI might enable concurrent measurement of the internal organ motion with MRI. Accurate measurement of motion is one of the motivations to combine MRI with PET. In our study image registration is applied on dynamic MRI data to derive motion fields in order to use them to firstly simulate concurrent PET data and then correct them for respiratory motion. However, motion estimation using registration of gated images is limited to the number of gates used. An alternative way is to use a motion model and derive transformations in real-time (McClelland *et al.* 2013).

After estimating motion the next step to correct the gated measured data for this motion. Motion correction approaches can be classified into three main groups.

1) event rebinning in which the LORs are realigned based on known motion information

2) reconstruct-transform-average (RTA) in which the motion information is applied to reconstructed gated images

3) motion compensated image reconstruction (MCIR) in which the motion is incorporated into the reconstruction algorithm

2.5.3 Event rebinning

One of the early approaches for motion correction involved incorporation of motion information prior to reconstruction into the sinogram, by repositioning the LOR of each event based on the known motion information to where they would have been detected if there were no motion.

Motion compensation directly in the projection data space has been originally suggested for correction of rigid motion such as brain imaging (Daube-Witherspoon *et al.* 1990; Bloomfield *et al.* 2003) where the motion can be measured by motion-tracking devices and list-mode data aligned based on the motion information. Later, Livieratos *et al.* (2005) proposed to correct for respiratory motion of the heart using rigid registration of respiratory gated PET images to align the LOR of each event to the reference position in list-mode data. However, the heart in this study was isolated to rigid motion. In real clinical acquisitions the heart has also non-linear motion due to respiration but the choice of whether the heart motion can be modelled as rigid or non-linear depends also on the acquisition parameters (e.g. scanner resolution). Later on, this approach was further extended to be applied for respiratory motion compensation, by modelling the motion as affine (Lamare *et al.* 2007a) and non-linear (Detorie and Dahlbom 2007). Although these studies showed that affine or non-linear transformation of the list-mode data can improve quantification accuracy for lesions in the lungs, it does not significantly correct the respiratory effects for organs below the diaphragm such as liver, and stomach and it may only be a viable approach for correcting motion if focusing on the heart or lungs with associated lesions (Lamare *et al.* 2007a).

However, motion may cause events that would have normally been detected in some LORs, to exit axially or through detector gaps the PET scanner undetected. Motion correction of individual LORs can correct detected events but does not compensate for events that have not been detected outside the FOV (Rahmim *et al.* 2004). On the other hand, the transformation of LORs along which the events are detected to LORs along which the events would have been detected, if there was no motion, may correspond to no actual detector pairs and thus these LORs also need to be properly modelled. To compensate for this, Thielemans *et al.* (2003) proposed different approaches such as scaling for the partially measured parts of the sinogram.

2.5.4 Image based motion correction

One respiratory motion correction method proposed by several authors is to transform the sequence of gated images. This image-based method, combines independent reconstructed gates which are then aligned to a reference position and averaged to compensate for the motion without loss of count statistics. This method is henceforth referred to as reconstruct-transform-average (RTA). To do this different registration algorithms have been proposed. In the context of correcting for respiratory motion, Dawood *et al.* (2006, 2008) used an optic flow method to align non-attenuation corrected PET images with the motion restricted to elastic transformations. Despite this the attenuation correction is expected to be affected by motion and it is an important correction for obtaining accurate quantification in PET images, it was not explored by these studies. Bai and Brady (2009, 2011) further explored the image based method by applying attenuation correction to the images after motion correction. Using their proposed method attenuation correction is applied voxelwise in the images. In contrast with the previous studies, in Bai and Brady (2009, 2011) estimation of non-linear motion was allowed by the use of a regularized registration algorithm with B-splines.

All these studies used motion estimation from registration of gated images and therefore only a limited number of transformations are able to be performed. Respiratory motion correction using this framework of transforming already reconstructed images was applied by King *et al.* (2011) for real time motion correction using a previously established MRI motion model derived using dynamic MRI data for three acquisitions (normal, deep, fast).

During RTA the motion corrected image is obtained post-reconstruction, without taking into account the non-linearity characteristic of iterative reconstruction. Although this method maintains the counts of measured data, it consists of a partition of counts in each gate and thus the effects of non-linearity that characterises iterative reconstruction will be exaggerated.

However, RTA is a practical approach because current PET scanners give the reconstructed images directly and include several corrections inside the reconstruction making it difficult in practice to reconstruct data from the scanner differently and therefore to correct the already reconstructed images may be the only possible way.

2.5.5 Reconstruction based motion correction

Another approach to compensate for respiratory motion is by incorporating the motion information within reconstruction known as motion compensated image reconstruction (MCIR). These algorithms are usually extensions of the standard algorithms.

Early implementation of this method proposed spatial or temporal prior/penalty terms in a 4D MAP likelihood framework to enforce consistency between the image frames in the gating dimension (Lalush *et al.* 1996; Gilland *et al.* 2002; Cao *et al.* 2003;

Gravier and Yang 2005). Therefore, these methods make use of temporal correlations between the estimated radioactivity distribution at the different respiratory positions. However, the performance of these methods depends on the choice of the prior/penalty parameters, mostly selected empirically based on a trial-and-error procedure, and these may erroneously affect the final reconstructed image. These approaches have been tested for brain and cardiac studies and found to improve image quality only for simple motion fields such as brain and motion due to the cardiac cycle whereas for respiratory motion no relevant applications have been found in the literature.

An alternative, more efficient approach consists of incorporating motion information in the probability matrix as part of the reconstruction without the need of prior/penalty terms. Incorporating the motion into reconstruction without the need of prior parameters was proposed by Qi and Huesman (2002) to compensate for rigid motion while Jacobson and Fessler (2003) proposed a method that also allows non-linear transformations to be considered. Later, in the context of rigid transformation Rahmim *et al.* (2004) incorporated motion information into the reconstruction to compensate for the brain motion. The method of Rahmim *et al.* (2004) was extended by (Qiao *et al.* 2006; Li *et al.* 2006; Lamare *et al.* 2007b) to account for non-linear motion and used time/gate-varying system matrix to integrate the gated measured data into a single comprehensive objective function. The motion information was derived by non-linear registration of CT data.

In contrast with previous studies, Jacobson and Fessler (2003) proposed a similar approach but to jointly estimate the motion information as well as the underlying image activity distribution directly from PET images during which the distribution and the deformation field are updated per iteration. The motion information was derived as compositions of interpolation functions and affine expressions of the unknown emission values. However, due to the high noise of PET images it is more desirable that the motion parameters are calculated in advance using CT or MRI. In addition a drawback of the joint motion and image estimation is the computational cost, therefore the method was evaluated only on a small 2D phantom.

However, in order to reduce the computational demands of MCIR, Qiao *et al.* (2007) proposed region based registration for respiratory motion correction. The proposed method instead of calculating and applying the motion correction into the whole images of the FOV uses only a predefined region. The proposed method has been validated using both a physical cylinder phantom with spheres and numerical simulations of the anthropomorphic NCAT phantom. The findings showed faster and improved accuracy of motion estimation by focusing just in specific regions but no difference in the reconstruction task. However, during this method the region of interest must be predefined. In addition, in practice various regions may require motion correction and as the number of regions is increased, it limits the efficiency of the method on accelerating the motion estimation task. More recent studies also compensated for respiratory motion using time/gate-varying system matrix into the reconstruction (Guérin *et al.* 2011; Chun *et al.* 2012a) but used tagged MRI to estimate the motion. Based on this framework with the difference of applying the motion information through image deformations prior to forward projection and following backprojection in the reconstruction process Dikaios and Fryer (2009, 2011a,b) also compensated for respiratory motion.

All these studies performed the motion correction using gated data. However, motion estimation using registration of gated images is limited to the number of gates used. An alternative way is to use a motion model and derive transformations in real-time (McClelland *et al.* 2013). Based on an MRI motion model Reyes *et al.* (2007)

applied respiratory motion correction. However, the motion model was derived using only two MRI image frames.

Chapter 3

Simulation of simultaneous PET-MRI data

3.1 Introduction

The increased interest on evaluation and comparison of developed algorithms for motion correction requires an accurate representation of the motion as observed in real subjects during PET acquisition. However, for extensive evaluation of different methods using various parameters a substantial clinical data base with real studies is required which takes money and time to be created. As an alternative, simulations have been developed based on computational phantoms that represent the human anatomy (Zaidi and Tsui 2009). Using simulations large datasets can be created for different disease states providing also a ground truth which is not known in real clinical studies.

Many phantoms have been developed usually from segmented high resolution anatomical images (e.g. CT or MRI) such as the widely used NCAT and its enhanced version XCAT (Segars and Tsui 2009) that model the human thorax anatomy and simulate both cardiac and respiratory motion. Computational phantoms have been extensively used in a large number of applications, such as the optimization of the detection system, the validation of reconstruction algorithms, the study of physical effects including radiation attenuation, scatter and motion (Zaidi and Xu 2007b; Tsoumpas and Gaitanis 2013a).

To incorporate motion into the numerical phantoms many studies used registration of dynamic images with sufficient temporal and spatial resolution (e.g. CT and MRI) to derive voxel deformations and use them to warp the 3D numerical phantom (Tsoumpas *et al.* 2011). For example, in the case of the XCAT phantom, respiratory and cardiac motion information has been derived from CT data resulting in more than 100 time frames over the cardiac cycle and 20 time frames over the respiratory cycle (Segars and Tsui 2009).

Numerical phantoms are integrated in software tools to simulate PET acquisition generating raw data in the form of sinograms. The most commonly used method is based on Monte Carlo simulations such as the open source software GATE (Jan *et al.* 2004). Monte Carlo simulations can model the PET process with high accuracy and precision, but their high computational demands has led to the development of fast analytic methods such as the open source software library for Tomographic Image Reconstruction (STIR) (Thielemans *et al.* 2012) or ASIM (Comtat *et al.* 1999). STIR was the research tool used in this thesis[¶].

Tsoumpas *et al.* (2011) developed a scheme for simulating realistic dynamic PET data using real MRI acquisitions and analytic simulations. Using an extension of the proposed approach temporally and spatially correlated MRI and PET datasets were generated, similar to those obtained from PET/MRI acquisitions.

This chapter provides a description of the dynamic PET simulation toolkit used in this thesis to study the performance of motion correction in PET and can be used in general to investigate the performance of different reconstruction and correction methods.

[¶] STIR consists of functions, classes and utilities for different imaging modalities such as PET. In STIR, image and projection data can be manipulated and images can be reconstructed.

3.2 Methods

3.2.1 4D numerical phantom using MRI-derived motion fields

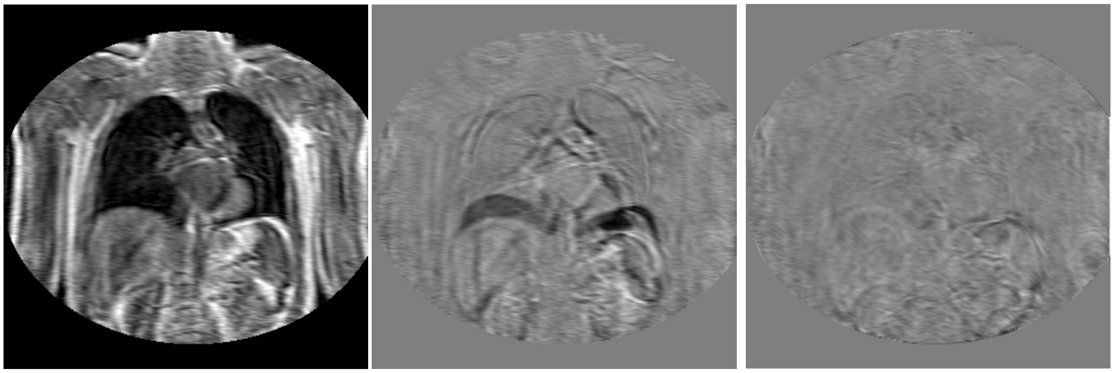
An anthropomorphic numerical phantom of the thorax was derived from the real dynamic MR dataset by simulating normal FDG activity distributions. A 4D dynamic MRI dataset of the thorax was obtained by a colleague on a 1.5 Tesla Philips Achieva^{TM} scanner with a 32-channel coil using a T1 weighted turbo field echo (TFE) sequence with repetition time equal to 3.3 *ms* and echo time equal to 0.9 *ms* and a flip angle of 10°. An ECG triggered dynamic 3D acquisition was applied with parallel imaging sensitivity encoding factor of 2 in anterior-posterior and 4 in right-left direction to speed up the acquisition so as to be possible to acquire the whole thorax volume. The whole thorax volume was acquired with a FOV of $500 \times 450 \times 245$ mm^3. The volunteer was breathing deeply to maximize the magnitude of the motion caused by respiration. The MRI dataset consisted of thirty five 3D images with a reconstructed image resolution of $1.5 \times 5 \times 1.5$ mm^3 (feet-head, right-left, anterior-posterior) and temporal resolution of 0.7 *s* per time frame. An amplitude signal was generated from a navigator placed on the diaphragm and the thirty five 3D MRI images were distributed over a range of respiratory positions covering a single respiratory cycle. One frame was selected (i.e. mid-inspiration) as the reference image.

As proposed by Tsoumpas *et al.* (2011) the 3D MRI reference image was manually segmented into different tissue types using intensity thresholds to extract the region of interest. As a next step uniform FDG standardized uptake values (SUVs) as typically measured in a clinical acquisition were assigned to each type (table 3.1). To derive these values the mean FDG uptake values for lungs, liver, myocardium, soft tissue and spinal cord were measured on real clinical PET/CT acquired data with the help of a clinician colleague (Dr Anastasia Chalkidou). The anatomical information of CT was used in order to segment the regions of interest in PET images and then calculate the mean SUV. To do this the HERMES medical image processing tool was used (http://www.hermesmedical.com) was used and the anatomical information of CT was used in order to segment the regions of interest and then calculate the mean SUV. Then spherical lesions were created and embedded at different locations in the image with various characteristics. The phantom developed by Tsoumpas *et al.* (2011) had three lesions only in the lungs whereas this phantom had six lesions in the lungs and three in the liver as described in table 3.1. Similarly, an anatomical 511 keV attenuation map was generated with attenuation values as following: for air 0 cm^{-1}, for soft tissue 0.096 cm^{-1} and for lung tissue 0.0248 cm^{-1}.

Table 3.1: Simulated radioactivity distribution in units of standardized uptake values (SUVs) and lesion diameter.

Tissue Type	SUV
Lungs	0.5
Liver	2.5
Myocardium	3.2
Soft tissue	1.0
L1 (diam.: 10 mm)	4.5
L2 (diam.: 16 mm)	4.5
L3 (diam.: 16 mm)	7.5
L4 (diam.: 10 mm)	6.5
L5 (diam.: 16 mm)	6.5
L6 (diam.: 16 mm)	9.5

After sorting the thirty five MRI images over a range of respiratory positions covering a single respiratory cycle, ten frames were selected on the basis of amplitude as the gates of the respiratory cycle including the reference image that was segmented in the previous subsection. A non-linear registration algorithm (Buerger *et al.* 2011) was applied to register each of the nine gates to the reference gate. In this way transformation fields (i.e. displacement vectors) describing the motion estimates from locations of the reference gate to the locations of each other gate were obtained. Figure 3.1 (a) displays the MRI image of the mid-inspiration gate and figure 3.1 (b) and (c) the difference between the MRI images of the end-inspiration and end-expiration gates before and after transformation respectively.



(a) (b) (c)

Figure 3.1: Visual demonstration of inspiration-expiration misalignment: (a) MRI image of the mid-inspiration phase; (b) difference between the MRI images of the end-inspiration and end-expiration phases (c) difference between the MRI images of the end-inspiration phase and the end-expiration phase after transformation to the end-inspiration using MRI derived motion field.

Respiratory motion was incorporated in the numerical phantom by applying the calculated MRI deformation fields to the phantom. Therefore the radionuclide distribution was transformed to each of the ten gate positions according to the MRI motion fields. The diaphragm motion was measured as the difference between the highest and lowest diaphragm edge locations between inspiration and expiration images. The results showed that the diaphragm motion amplitude was equal to approximately 24 mm. A similar procedure was followed to generate a gated anatomical 511 keV attenuation map.

3.2.2 Real-time 4D numerical phantom using MRI-derived motion model

An alternative way that allows real time datasets to be simulated is to incorporate motion in the numerical phantom using a motion model. The proposed approach by Tsoumpas *et al.* (2011) was extended to incorporate realistic models of respiratory motion into the numerical phantom. Datasets with continuous motion were generated for studying further the impact of motion by including the effect of intra-frame motion. To do this real respiratory signals derived from PET-CT images are combined with MRI-derived motion modelling.

3.2.2.1 3D numerical phantom

In this part the aim was to develop a more realistic version of the numerical phantom developed above that could be used for studying the effect of respiratory motion on detectability.

Firstly, the anatomy of the lungs was extended in order to include heterogeneity such as pulmonary vessels. Furthermore, more accurate attenuation maps were used by including information for bone regions extracted using an MRI UTE acquisition (Buerger *et al.* 2012). Anatomical information was obtained from a breath-hold UTE scan for a healthy volunteer on a 1.5 Tesla Philips Achieva^{TM} scanner with a 32-channel coil. An ECG triggered MRI scan was also applied to acquire information for myocardium. The procedure of the acquisition was similar as used before in section 3.2.1 with additional information of cortical bone.

The 3D MRI image was manually segmented into different tissue types, and uniform FDG SUV values were assigned to each tissue type as before in section 3.2.1 with the additional inclusion of bone with an SUV of 2.3. Lesions were embedded at different locations in the image with various characteristics based on the aim of the investigation. Similarly, an anatomical 511 keV attenuation map was generated with attenuation values as described before in section 3.2.1 but also with the inclusion of bone with an attenuation value of 0.15 cm^{-1}.

3.2.2.2 MRI-derived motion model

A subject-specific motion model was used derived from a series of real dynamic MRI data (King *et al.* 2011, 2012). For the same volunteer as before, 105 dynamic 3D MRI images of the thorax were acquired with a reconstructed image resolution of 1.5 × 5 × 1.5 mm^3 (feet-head, right-left, anterior-posterior) and temporal resolution 0.7 s per time frame. The first thirty five images were acquired during normal breathing, the next thirty five during fast breathing and the last thirty five during deep breathing.

Each of the 105 images was registered to the reference breath hold image using hierarchical registration. In particular, for each control point in the image the displacements in x, y, and z directions were calculated and plotted as functions of the head-foot diaphragm translation signal (i.e. surrogate signal). A curve based on second order polynomial functions is fitted to the x, y, and z displacements as a function of surrogate values (i.e. non-linear function allows capturing of non-linear motion). The variation of the motion fields as a function of the surrogate was modelled separately for inspiration and expiration to compensate for hysteresis effects (i.e. differences between the breathing curves followed during inspiration and expiration).

By using high temporal resolution and three types of acquisition (i.e. normal, fast and deep breathing) respiration of different amplitudes and frequencies is included in the formation of the model in order to capture respiratory cycle variation as expected in real subjects. For a given surrogate value (e.g. diaphragm position in a respiratory signal) and breathing direction (i.e. inspiration or expiration) the model calculates motion fields.

3.2.2.3 Respiratory signal database

Respiratory signals describing the vertical displacement of the diaphragm as a function of time, were obtained by a colleague from thirty four ammonia (NH_3) cardiac perfusion PET studies of anonymized freely breathing patients using a data driven method (Schleyer *et al.* 2009, 2011). The data driven method used in this study extracts the motion signal by measuring the intensity change in the pixels and it was validated against the real-time position management hardware method (RPM) which is the standard method used clinically to acquire a respiratory signal.

Each signal was acquired for 3 min of 0.1 s duration for each time point. Examples of these signals are presented in figure 3.2. The reason that all the thirty four patients studies in the database were investigated was to determine the maximum, minimum and mean motion amplitude in order to calibrate the expected motion amplitude range during a PET study. To do this, for each of the thirty four patient curves, the diaphragm motion was measured as the difference between the highest and lowest diaphragm edge locations on the gated no attenuation-corrected PET images. The results showed that the maximum and minimum diaphragm motion amplitude as measured in PET images were equal to approximately 20 and 6 mm respectively. This range is consistent with a previous study by Liu *et al.* (2009) who reported a maximum displacement 20 and 15 mm and a minimum displacement 5 and 4 mm for the left and right diaphragm respectively.

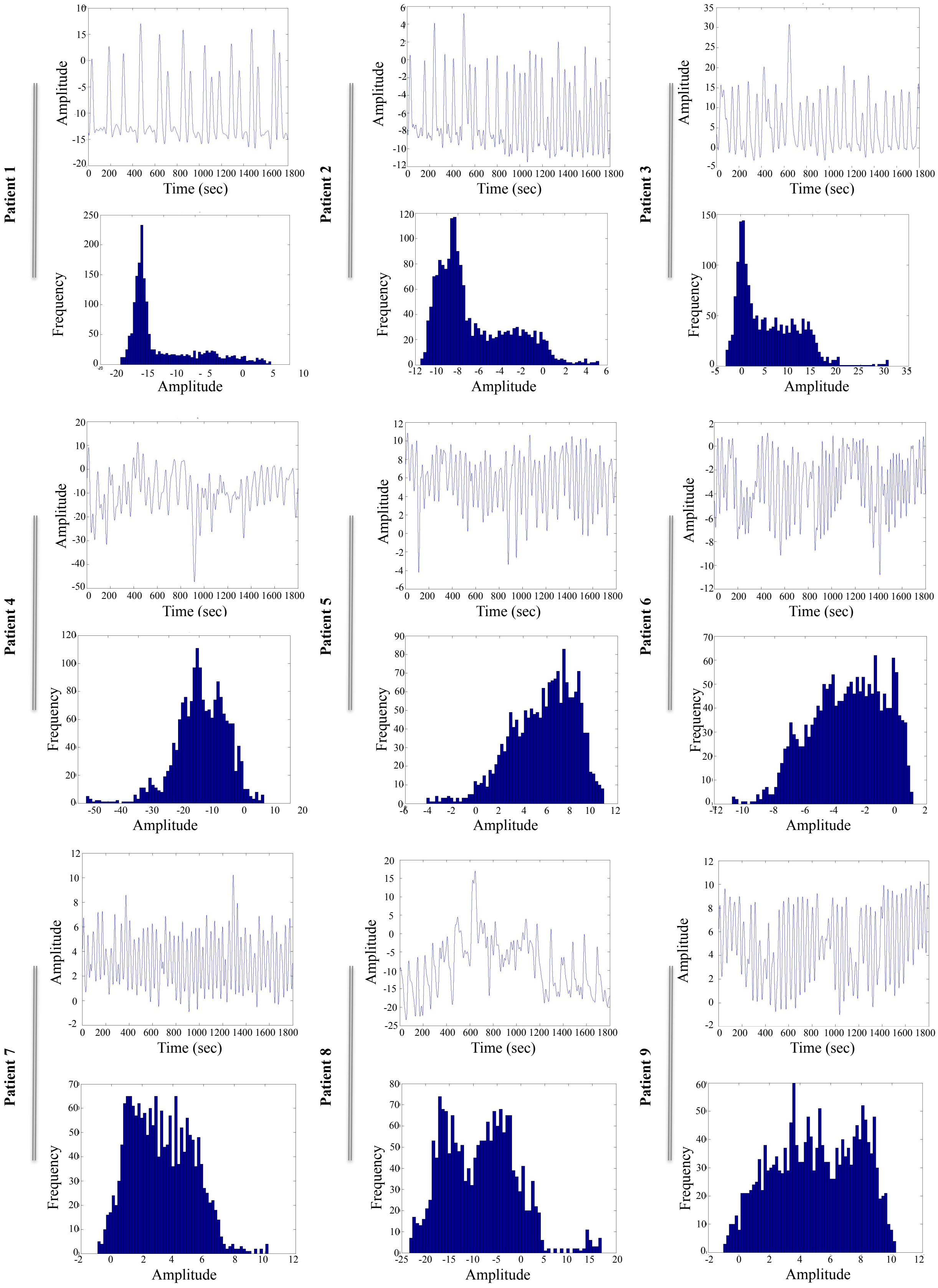


Figure 3.2: Examples of breathing curves and corresponding displacement histograms for nine patients.

As shown by Liu *et al.* (2009), the variability of the different respiratory patterns can be classified into three types based on the corresponding displacement histogram. For each of the thirty four signals in the database of thorax PET studies a displacement histogram was created in order to group the curves to these three groups. After classification of all the respiratory curves into the three types, three respiratory signals were selected according to their displacement histogram such that each signal was representative of one of the three different breathing patterns. Only one representative example was selected for each of the three breathing pattern classifications since the aim of this study was to demonstrate any possible differences using different patterns and not to investigate the individual characteristics of each study. The selected respiratory signals and their corresponding displacement histograms are illustrated in figure 3.3. The first signal is described by long quiescent motion periods (type-A) accounting for approximately 60% of patients, the second signal by regular quiescent motion periods (type-B) accounting for approximately 20% of patients and the third signal is characterised by random baseline shifts (type-C) accounting for approximately 20% of patients.

With the use of the selected breathing patterns a database with variable diaphragm motion amplitude was created based on the above measurements of expected motion amplitude range in a 3 *min* acquisition. To do this, each of the three signals was firstly shifted to have zero median displacement and then scaled to nine diaphragm amplitudes ranged from 4 to 20 *mm* with a step of 2 *mm* creating a database of breathing type-A, type-B and type-C each one consisting of nine signals.

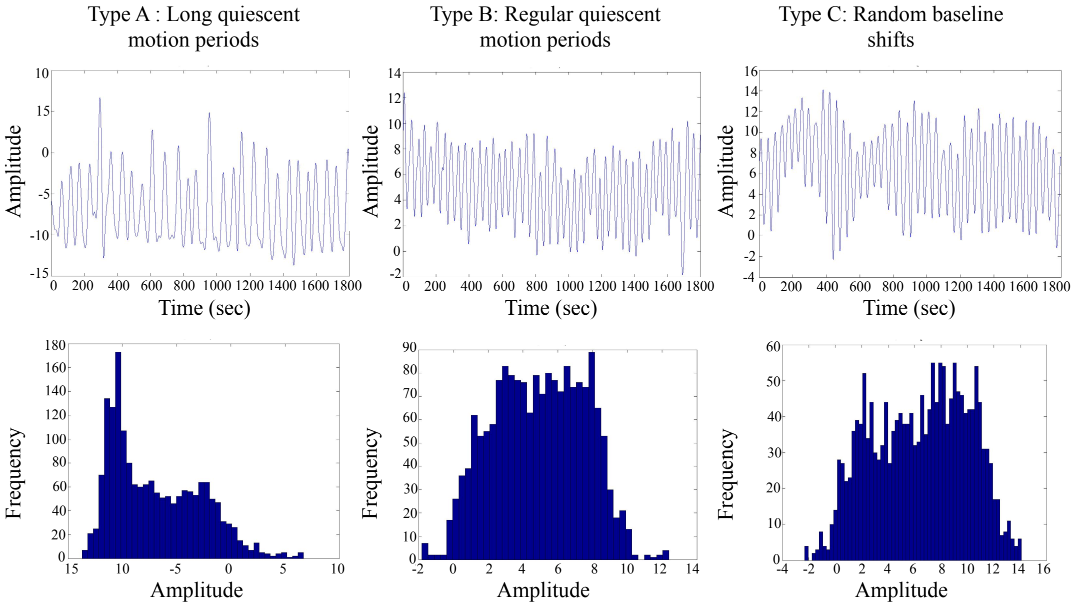


Figure 3.3: Breathing curves and corresponding displacement histograms for the three breathing types: (a) long quiescent motion periods (type-A) accounting for approximately 60% of patients, (b) regular quiescent motion periods (type-B) accounting for approximately 20% of patients and (c) random baseline shifts (type-C) accounting for approximately 20% of patients.

3.2.2.4 4D computational phantom

Given a respiratory signal, the motion model is used to calculate motion fields necessary to transform the reference distribution (i.e. the 3D tracer uptake distribution) to the relevant respiratory positions and generate a 4D phantom. Note that no redistribution of radioactivity within the body was simulated. The FDG distribution was warped using transformations calculated from the motion model for each sampling point every 0.1 s during a 3 min acquisition (i.e. 1800 sampling points) as illustrated in figure 3.4. Similarly, the motion fields were applied to the anatomical attenuation maps to create dynamic datasets.

Furthermore, for each dynamic PET dataset, the data were gated based on amplitude as described in section 2.5.1 to generate eight partitions (gates) over the full motion cycle. To do this, for a given breathing type and motion amplitude, the corresponding amplitude histogram was firstly divided into eight equal amplitude bins (i.e. gates). Then each of the 1800 images in a dynamic phantom based on its respiratory position it was assigned to the corresponding gate. An averaged image was then formed for each gate creating a 4D dynamic FDG distribution comprising of eight gated images. Since, the PET acquisition system model is linear, averaging the FDG distributions prior to simulating the projection data will give the same result as simulating continuous projection data and gating afterwards.

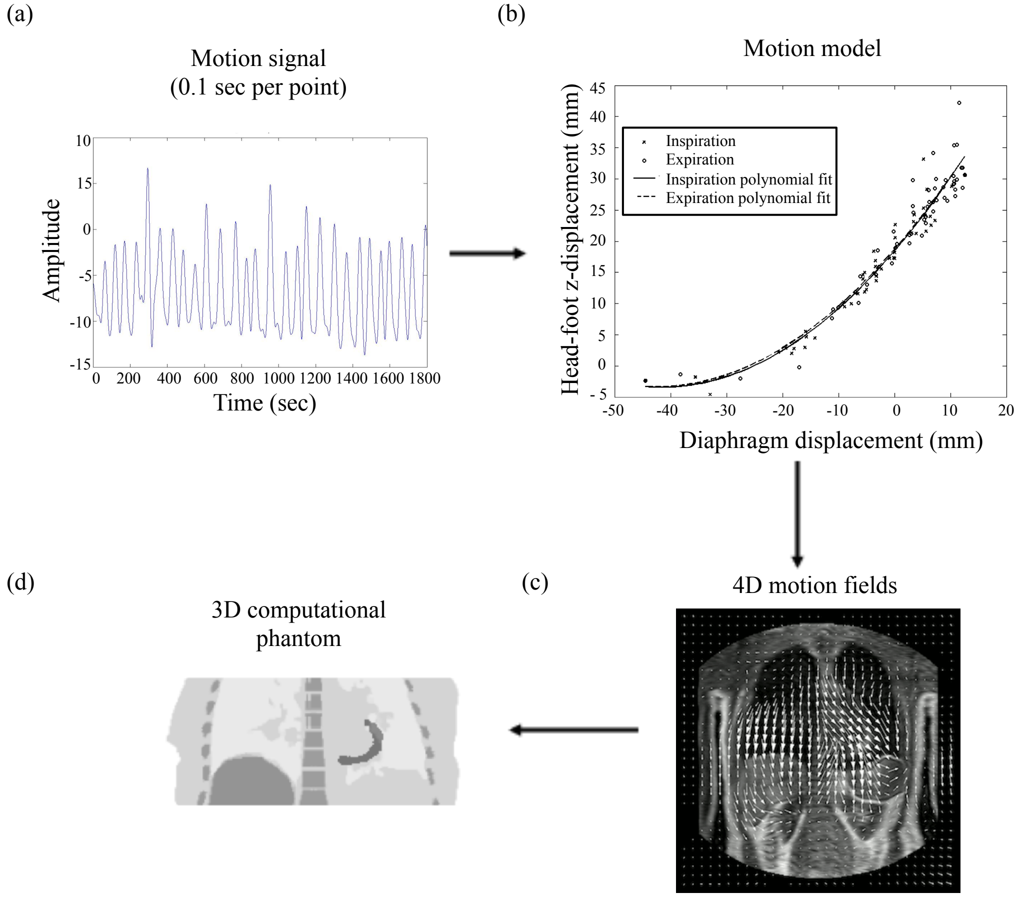


Figure 3.4: The 4D simulation is based around a motion model that describes the 3D trajectory (b) of each point in a grid covering the torso (c) as a function of the displacement of the diaphragm (a). The 4D description of the motion in the torso is then used to create the 4D radionuclide distribution model (d).

3.2.3 PET simulation

The output from both techniques is a dynamic FDG distribution. Each gated image was used as the input to an analytic simulation of a PET acquisition to create projection data as described by Tsoumpas *et al.* (2011). The gated emission maps were forward projected and raw data were generated in the form of sinograms for the Philips Gemini scanner (Surti *et al.* 2007). The simulated scanner is made of 44 detector rings of 90.034 cm inner ring diameter with an axial extent of 0.4 cm. 644 detectors exist in each ring (see also Appendix A1). 3D projection data (maximum ring difference: 43, span: 3) were simulated based on Siddon's ray tracing algorithm with ten rays per tube of response (see also Appendix A2).

Following forward projection, photon attenuation, scatter and Poisson noise can be included in each respiratory position. Attenuation coefficient projection data was estimated from the attenuation images using the same analytic forward projector settings and applied by multiplication to each respiratory gate. Scatter was simulated analytically using the single scatter simulation formula (Watson *et al.* 1996) as implemented in STIR (Tsoumpas *et al.* 2004; Polycarpou *et al.* 2011) (Appendix B). A global scaling was applied such that the total amount of the simulated scatter events corresponded to 33% of total measured events. Noise was added to each respiratory gate accordingly to the purposes of the investigation.

The resulting projection data are then reconstructed using standard or novel reconstruction algorithms. In this thesis STIR was also used for the reconstructions.

3.3 Results

3.3.1 4D Gated-PET datasets using MRI-derived motion fields

One gated simulated numerical phantom was created simulating motion from a single respiratory cycle consisting of ten respiratory positions without intraframe motion.

Figure 3.5 and figure 3.6 show examples of slices through the images of the simulated radioactivity distribution images without and with lesions respectively. Images are presented for a single respiratory position and for the average of the simulated radioactivity distribution for ten respiratory positions created with the use of dynamic MRI images consisting of 10 gates of 0.7 *s* per time frame. For the emission datasets figure 3.7 shows example of the result of forward projecting the radioactivity distribution images presented in figure 3.5. The image blurring introduced by motion is evident in the data with motion compared to the data without motion.

Figure 3.5: Examples of transverse, coronal and sagittal planes of the simulated radioactivity distribution for (a) one respiratory position and (b) average of ten respiratory positions as created by dynamic MRI images consisting of 10 gates of 0.7 *s* per time frame.

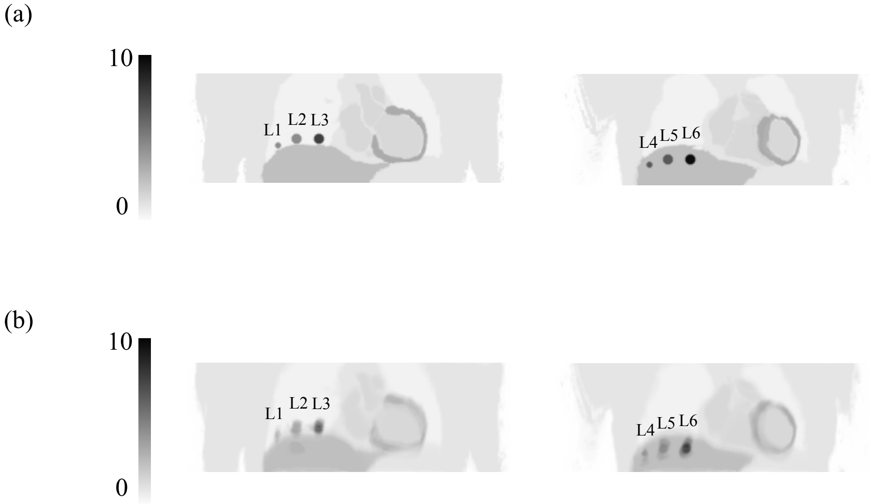


Figure 3.6: Examples of coronal planes of the simulated radioactivity distribution with lesions for (a) one respiratory position and (b) average of ten respiratory positions as created by dynamic MRI images consisting of 10 gates of 0.7 *s* per time frame.

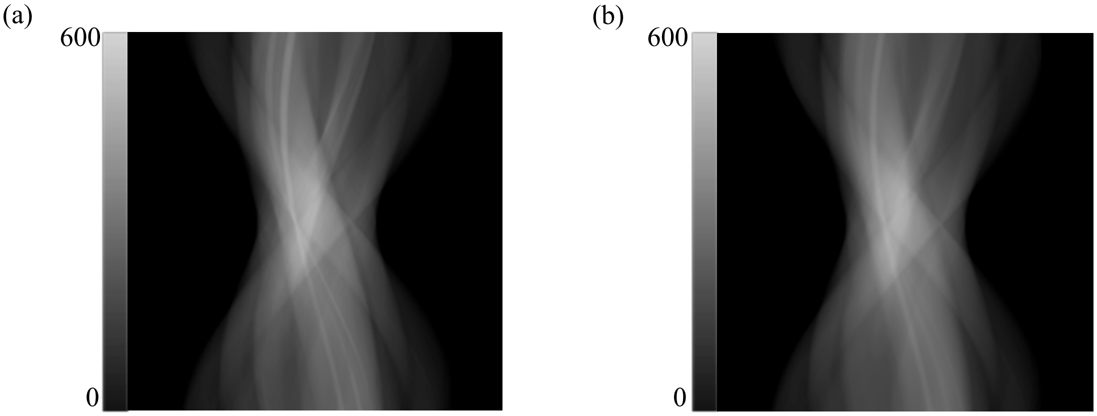


Figure 3.7: Sinograms corresponding to the simulated radioactivity distribution for (a) one respiratory position and for (b) ten averaged respiratory positions as created by dynamic MRI images consisting of 10 gates of 0.7 *s* per time frame.

3.3.2 Real-time 4D PET datasets using MRI-derived motion model

A database of a real-time simulated numerical phantom with variable diaphragm motion amplitude and breathing patterns was created. Figure 3.8 shows an example of slices through the images of the simulated radioactivity distribution for a single respiratory position and for the average of the simulated radioactivity distribution for

1800 respiratory positions as created with the use of the MR motion model combined with a respiratory signal of 3 min of 0.1 s duration for each time point. The selected respiratory signal is for type-C breathing with 20 mm maximum motion amplitude. For the emission datasets figure 3.9 shows example of the result of forward projecting the radioactivity distribution images presented in figure 3.8.

Figure 3.8: Examples of transverse, coronal and sagittal planes of the simulated radioactivity distribution for (a) one respiratory position and (b) average of 1800 respiratory positions as created with the use of an MRI motion model combined with a respiratory signal of 3 min of 0.1 s duration for each time point.

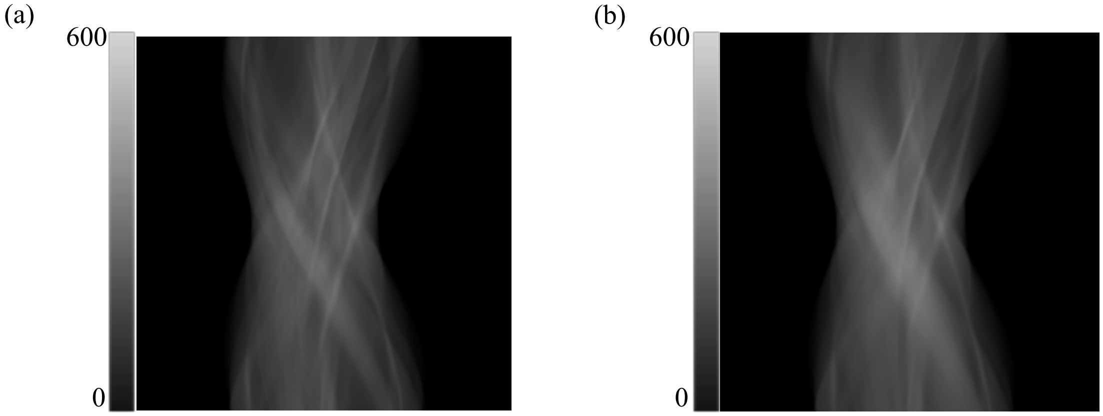


Figure 3.9: Sinograms corresponding to the simulated radioactivity distribution for (a) one respiratory position and (b) 1800 averaged respiratory positions as created with the use of an MRI motion model combined with a respiratory signal of 3 min of 0.1 s duration for each time point.

3.4 Discussion

This chapter presents the methodology that can be followed in order to create dynamic PET datasets based on a sequence of real MRI acquisitions followed by a combination of segmentations and image registrations. Two alternative methodologies for producing dynamic PET radioactivity distribution are described.

The first approach is based on registration of dynamic MRI data over a single respiratory cycle to calculate motion information. Motion information is then applied to warp the 3D tracer uptake distribution to the relevant respiratory positions and generate a 4D phantom. This approach models a single respiratory cycle.

However, a realistic breathing cycle is not perfectly periodic and may have cycle-to-cycle variations in magnitude, as well as longer-term variations in both magnitude and period (Sonke *et al.* 2008; Siebenthal *et al.* 2007). Consequently, real respiratory signals representative of three breathing patterns expected to be observed during a PET acquisition (Liu *et al.* 2009) can be used with variable motion amplitude in order to provide a better understanding of the extent of motion blurring in both normal and extreme breathing cases and define when motion is important or not.

The second approach, is based on a motion model that is combined with real respiratory signals to estimate the motion information for any position in the signal. Similarly, a 4D phantom is created using the estimated motion information to transform the 3D tracer uptake distribution to any position in the signal. Using a motion model any respiratory position can be generated to simulate real time data as observed in real subjects. Considering that in real studies respiratory motion is only in part periodic this method allows the simulation of motion trajectories according to the amplitude and frequency of motion observed in real subjects. In addition, this method allows simulations of temporal resolution similar to the PET imaging system being simulated.

In this thesis a subject-specific motion model was used. On the other hand, the proposed simulation scheme could be further extended by the use of a global motion model to provide motion information without the need for patient specific MRI acquisitions (Fayad *et al.* 2012).

Furthermore, in the example of the motion model in figure 3.4 the plot represents the displacements for a specific control point in the image as a function of the head-foot diaphragm translation signal. Although the plot for the specific control point shows almost identical relationship between inspiration and expiration this is only an example and it is not the case for all control points in the image. In general, in the specific motion model used in this thesis the variation of the motion fields as a function of the surrogate are modelled separately for inspiration and expiration. Previous studies have shown differences between the trajectories followed during inspiration and the trajectories followed during expiration (i.e. hysteresis effects) especially for anterior-posterior displacements but also for superior-inferior displacements (King *et al.* 2009). Lack of modelling the different paths followed by a tumour during inspiration and expiration may lead to underestimation of errors in describing the motion especially for patients with an irregular breathing pattern. Therefore, the complexity of having inspiration and expiration separately is necessary in order to capture respiratory cycle variation to represent the actual real motion observed in patients.

The degree of blurring is evident in the data with motion in both gated-simulation and real-time simulation. The simulations can be a useful tool to evaluate the accuracy of motion correction methods. However, the gated-simulation with no intra-gate motion can be a simple approach to investigate the properties of the motion correction algorithms without being affected by the motion correction accuracy and the consequent transformation errors that may occur when intra-gate motion is present.

On the other hand, real-time simulation provides a more realistic and comprehensive view of motion effect as observed in real acquisitions and might be a preferable method to study the resolution recovery rather than the properties of the algorithms.

In addition, in real clinical applications a lung lesion that is visible in a PET scan may appear in a CT scan with attenuation equivalent to soft tissue density and be therefore classified as soft tissue. On the other hand, there are cases where lung lesions that are revealed in a PET scan may not be visible in the attenuation images and may be classified as lung rather than soft tissue during attenuation correction. In the simulation procedure used in this thesis we considered lung lesions not visible in attenuation images. Under the assumption that attenuation of the lung lesions is equivalent to lungs, during both the creation of simulations and reconstruction, the absence of lung lesions in the attenuation images is not expected to affect the conclusions derived from this thesis. However, in the case that lung lesions were assumed to behave as soft tissue density masses in the attenuation images there are two scenarios. Firstly under the assumption of no motion artefacts, in the presence of lung lesions in attenuation images during the creation of the simulations, their absence during attenuation correction may result in undercorrection (no compensation for loss of counts due to attenuation) which is expected to result in underestimation of the tumor tracer uptake. In a previous study Martinez-Möller *et al.* (2009) investigated attenuation correction using attenuation maps with and without compensating for the lung lesions and showed that the absence of the lung tumor in the attenuation map may result in a variation in the tumor SUV of 35%. Furthermore, in the presence of motion any mismatches between the emission and attenuation image is expected to introduce additional uncertainties in the quantification accuracy (Killoran *et al.* 2011).

Finally, the proposed methodology can be a fast and practical approach to simulate a large number of dynamic PET datasets allowing also the opportunity to generate multiple statistical realizations of any given dataset. The two methodologies presented in this chapter are used in the next chapters to study the impact of respiratory motion correction on lesion detection and quantification in PET. Depending on the level of approximation that is required for a study, the appropriate methodology can be followed.

3.5 Conclusion

A scheme was described for simulating realistic dynamic PET data using real MRI acquisitions. This approach allows incorporation of realistic models of respiratory motion to generate temporally and spatially correlated MRI and PET datasets, as those expected to be obtained from PET/MRI acquisitions. Simulations with realistic anatomy and motion trajectories as those observed in real subjects can help to investigate the performance of different reconstruction and correction methods (e.g. motion).

Chapter 4

Analysis of two methods for respiratory motion compensated reconstruction

4.1 Introduction

Respiratory motion correction is a difficult problem due to the poor statistical quality of PET data and many approaches have previously been investigated. The two most popular motion correction approaches in PET are RTA and MCIR as described in chapter 2. Depending on the stage that the correction is applied the final motion corrected image is expected to have different image properties. Theoretically, an average of independently reconstructed gates is expected to be affected by the low counts of each gate and therefore to provide inferior quality images, in terms of both quantitative accuracy and noise characteristics, than the use of all gates in a single reconstruction (Lamare *et al.* 2007b). This is in agreement with the theoretical comparison of RTA and MCIR by Asma *et al.* (2006) who predicted better image quality, in terms of both quantification accuracy and noise properties, for MCIR. Another recent theoretical investigation showed that the MCIR has lower or comparable variance to RTA (Chun and Fessler 2013). However, the results were not assessed for low counts where the statistical noise introduces bias and affects the accuracy of the reconstructed images (Walker *et al.* 2011).

Simulation studies showed that although MCIR recovers lesion contrast better and is quantitatively more accurate than RTA (Dikaios and Fryer 2011b), at late iterations RTA may be less noisy and has better lesion detectability than MCIR (Thielemans *et al.* 2006). Although the low noise observed in RTA is contrary to theoretical expectations, this has not been further communicated in the literature. Another more recent study by Dikaios and Fryer (2011a) showed that MCIR is more robust quantitatively than RTA, attributing the differences to interpolation errors introduced during transformation. However, there is still not sufficient evidence in the current literature to answer the question of which is the best practical method for motion correction.

This chapter compares RTA versus MCIR and investigates the differences between them in terms of bias and noise. Real MRI dynamic data are used to obtain motion information. Evaluation is performed using voxel-based statistical analysis with an adequate number of iterations to study quantification accuracy. In contrast with other studies, we perform an analysis with respect to the number of iterations for reconstruction without or with post-filtering using a high number of independent realisations to ensure sufficient statistics. This chapter also demonstrates a paradigm of a robust methodological approach on how to separate the overall error into different components and evaluate both bias and noise trade-offs at the voxel-level.

4.2 Methods

4.2.1 Statistical comparison of RTA and MCIR

4.2.1.1 Simulated 4D PET data

A realistic spatially and temporally coregistered dynamic PET dataset comprising of ten gated PET images was simulated as follows, using the procedure described in section 3.2.1. Ten frames distributed over a range of respiratory positions on the basis of amplitude were used as the gates of the respiratory cycle. Transformation fields describing the motion estimates from locations of the reference gate to the locations of each other gate were obtained using a non-linear registration algorithm (Buerger *et al.* 2011). In addition, the corresponding inverse MRI motion fields were numerically estimated from the same motion vectors (Crum *et al.* 2007).

The MRI reference image was manually segmented into different tissue types, and uniform FDG standardized uptake values (SUVs) as typically measured in a clinical acquisition were assigned to each type (table 4.1). Six spherical lesions of different contrasts and sizes were manually embedded at different locations in the lungs and in the liver as shown in figure 4.1 and as described in table 4.1. Then the radionuclide distribution was transformed to each of the ten gate positions according to the MRI motion fields.

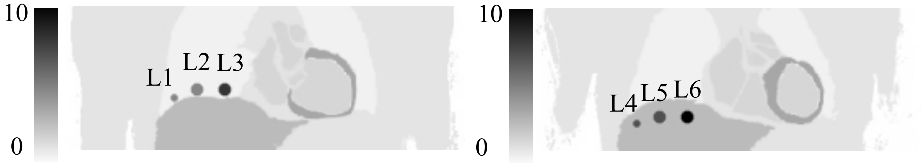


Figure 4.1: Coronal planes of the simulated radioactivity distribution with six lesions of different contrasts and sizes added at different locations in the lungs and in the liver as described in table 4.1.

Table 4.1: Simulated radioactivity distribution in units of standardized uptake values (SUVs) and lesion diameter.

Tissue Type	SUV
Lungs	0.5
Liver	2.5
Myocardium	3.2
Soft tissue	1.0
L1 (diam.: 10 mm)	4.5
L2 (diam.: 16 mm)	4.5
L3 (diam.: 16 mm)	7.5
L4 (diam.: 10 mm)	6.5
L5 (diam.: 16 mm)	6.5
L6 (diam.: 16 mm)	9.5

Each gate of the 4D generated activity distribution was used as the input to an analytic PET-acquisition simulation to generate raw data in the form of sinograms for the Philips Gemini scanner. Following forward projection, attenuation and Poisson noise were included in each respiratory position. No scatter or random coincident events were simulated. Noise was added to each respiratory gate such that the overall counts of the whole dataset correspond to a 5 min 18F-FDG 3D PET clinical thorax acquisition (i.e. approximately 50 million coincidences). One hundred realisations of the acquisition process were simulated[¶]. To generate multiple realisations of the acquisition process statistical noise was included taking different seeds from the random number generator for each gate and realisation.

4.2.1.2 Motion compensated image reconstruction

In this study, both motion correction methods, RTA and MCIR, were developed within STIR. During RTA, motion correction is performed by first reconstructing independently the raw data of each respiratory position using the OSL algorithm as illustrated in figure 4.2. Then the reconstructed image of each gate is transformed to the reference position using known motion fields. The same motion fields were used both to create the simulations and to correct for motion. The transformed gates are then averaged to produce a motion-corrected image:

$$\Lambda_{vg}^{(s+1)} = \Lambda_{vg}^{(s)} \frac{1}{\sum\limits_{b\in S_l} P_{bv} A_{bg} + \beta_g \frac{\partial}{\partial \Lambda_{vg}} E_v^{(s)}} \sum_{b\in S_l} P_{bv} \frac{Y_{bg}}{\sum\limits_{\tilde{v}} P_{b\tilde{v}} \Lambda_{\tilde{v}g}^{(s)} + \frac{B_{bg}}{A_{bg}}} \tag{4.1}$$

$$\Lambda_v = \frac{1}{G} \sum_g W_{g\rightarrow 1} \Lambda_{vg} \tag{4.2}$$

[¶] Part of the datasets are available at: http://www.isd.kcl.ac.uk/pet-mri/simulated-data/.

where:

- $\Lambda_{vg}^{(s)}$ is the estimated radioactivity concentration at voxel v, subiteration number s and gate g.

- Y_{bg} is the number of coincident photons measured events of each detector pair (bin) b and gate g.

- S_l corresponds to the l^{th} subset of the projection space, which is divided into L total subsets.

- s is the subiteration number and $l = s \bmod L$. A set of L subiterations comprises a full iteration.

- P_{bv} is the geometric system projection matrix.

- $W_{1\rightarrow g}$, $W_{g\rightarrow 1}$, represent the backward and forward warping operations respectively that move the activity from voxel in gate g to the voxel in the reference gate 1.

- G is the total number of gates.

- A_{bg} and B_{bg} are the attenuation coefficient and the background term respectively for each bin and gate.

- β_g is the penalization factor.

- E is the energy function.

The penalty function used in this study was based on the median root prior (Alenius and Ruotsalainen 1997, 2002; Bettinardi *et al.* 2002). The penalty function is calculated based on the difference of the voxel value with the median among neighbouring voxels around voxel v for the previous iteration denoted as M_v as described by:

$$\beta_g \frac{\partial}{\partial \Lambda_{vg}} E_v^{(s)} = \beta_g \frac{\Lambda_{vg}^{(s)} - M_v^{(s)}}{M_v^{(s)}} \tag{4.3}$$

In addition, a gated attenuation map was used in the reconstruction in order to model the motion in the attenuation map and for each gate the attenuation correction factors were calculated using the corresponding attenuation map.

For MCIR the motion transformations are incorporated directly into the reconstruction as illustrated in figure 4.3. The motion corrected image is reconstructed using the following iterative formula which is based on conventional OSL, including motion via the forward/backward transformation operators (Dey and King 2009):

$$\begin{aligned}\Lambda_v^{(s+1)} = \Lambda_v^{(s)} & \frac{1}{\sum\limits_{b \in S_l, g} \sum\limits_{g} W_{g\rightarrow 1} \left(P_{bv'} A_{bg} \right) + G\beta_g \frac{\partial}{\partial \Lambda_{vg}} E_v^{(s)}} \\ & \times \sum_{b \in S_l, g} \sum_{g} W_{g\rightarrow 1} \left(\frac{Y_{bg}}{\sum\limits_{\tilde{v}} P_{b\tilde{v}} W_{1\rightarrow g} \Lambda_{\tilde{v}}^{(s)} + \frac{B_{bg}}{A_{bg}}} P_{bv} \right)\end{aligned} \tag{4.4}$$

Based on the implementation of the expectation maximization update in the conventional OSL algorithm, the forward warping operation, $W_{1\rightarrow g}$, should be the transpose of the motion matrix, $W_{g\rightarrow 1}$. However, similarly with previous studies (Dey and King 2009; Dikaios and Fryer 2011a,b; Ouyang *et al.* 2013), in this thesis the inverse of the motion matrix was used rather than its transpose. If the motion matrix is orthogonal, then the transpose and the inverse are the same, and therefore the expectation maximization is still obtained even if the inverse of the motion matrix is used. On the other hand, if the motion matrix is simply not orthogonal, using the inverse rather than the transpose of motion matrix results in an approximate implementation of the expectation maximization update, with the level of approximation corresponding to the level to which the transpose and inverse of motion matrix give different results.

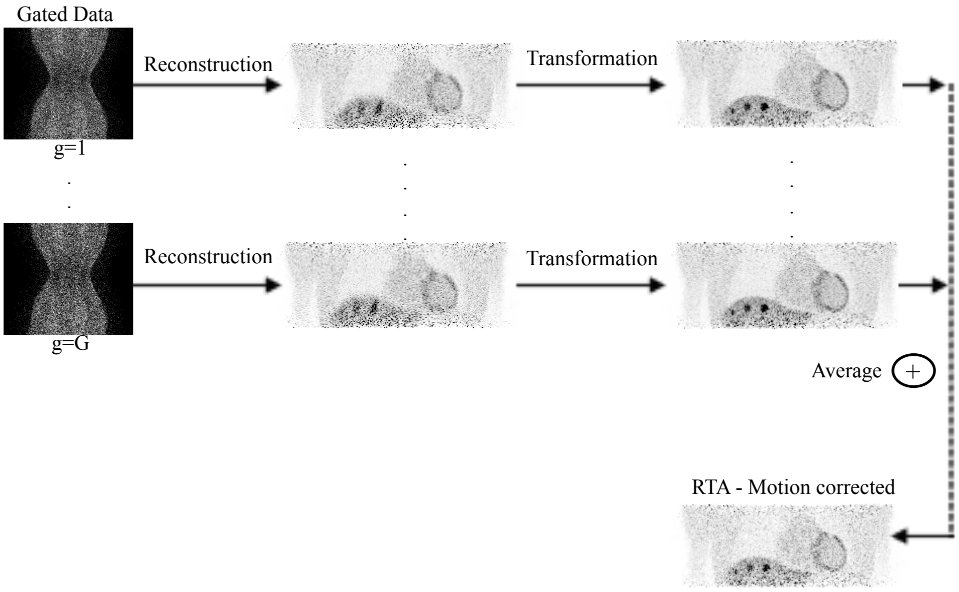


Figure 4.2: Reconstruct-transform-average motion correction method: Independent reconstructions of each gate are transformed to a reference gate and averaged.

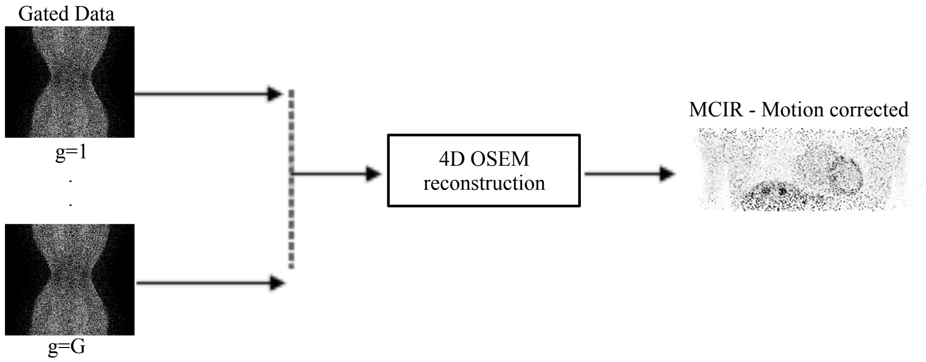


Figure 4.3: Motion compensated image reconstruction motion correction method: Motion information is incorporated within the reconstruction.

4.2.1.3 Evaluation datasets

To understand the errors introduced during the different motion correction methods the investigation was divided into two parts: First, RTA and MCIR methods were evaluated with the motion fields set to zero for both simulation and correction. Then the same reconstructions were evaluated using motion fields as described above. Applying RTA and MCIR methodology with zero motion fields enables evaluation of the methods without being affected by transformations and the consequent inaccuracies due to interpolation that may occur when motion is present. For these cases, reconstructions were performed for β_g equal to zero (i.e. standard OSEM) and with 23 subsets and for different iterations from 1 to 20. Each volume consisted of 250 × 250 × 87 voxels each with an isotropic size of 2 mm. Furthermore, for the next part of the chapter which studies the impact of post-reconstruction filtering, each reconstructed realisation image was smoothed with a 4 mm isotropic FWHM Gaussian kernel. The following datasets were reconstructed and evaluated:

- Reference: The reference gate (i.e. gate 1) was reconstructed using the conventional OSEM algorithm with the same standard count statistics as for the entire gated dataset.

- Non-motion corrected (Non-MC): The ten respiratory positions were averaged and reconstructed using the conventional OSEM algorithm.

- RTA with zero motion (Gate1-RTA): The reference gate (i.e. gate 1) was reconstructed by applying RTA with the use of zero motion fields. The ten gates were simulated as different noise realisations of the reference gate.

- MCIR with zero motion (Gate1-MCIR): The reference gate was reconstructed by applying MCIR with the use of zero motion fields. The ten gates were simulated as different noise realisations of the reference gate.

- Motion corrected with RTA (MC-RTA): The ten respiratory positions were reconstructed using the RTA motion correction method with MRI-derived motion fields.
- Motion corrected with MCIR (MC-MCIR): The ten respiratory positions were reconstructed using the MCIR motion correction method with MRI-derived motion fields.

The choice of the number of iterations was selected after a preliminary study which showed that for the reference reconstructed gate the mean values over the 100 realizations for single voxels selected in the centre of the lung and liver lesion approached a stabilized value after 20 iterations (figure 4.4). In addition, the difference between the standard deviation values of the single voxels, after 15 iterations with the value after 20 iterations is small and therefore 20 iterations can be considered a satisfactory number of iterations to study quantification accuracy for the specific type of simulation setup used in this study.

Secondly, the advantages gained by the two different methods by including median-root-prior regularization within the reconstruction have also been studied. For the purposes of this part of the chapter the degree of regularization was determined by changing the value of the weighting factor β_g in (4.1) and (4.4). The kernel size for the regularization was 3 voxels isotropic. Different weights have been studied for the regularization term (i.e. β_g: 0, 0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 50, 100). These weights were constant and spatially invariant throughout each reconstruction. For practical reasons, as a first step only one reconstruction realisation was obtained for all the different values of the regularization term in order to determine the optimal value of the regularization term. Twenty independent noise realisations have been reconstructed using the optimal values of β_g for both RTA and MCIR.

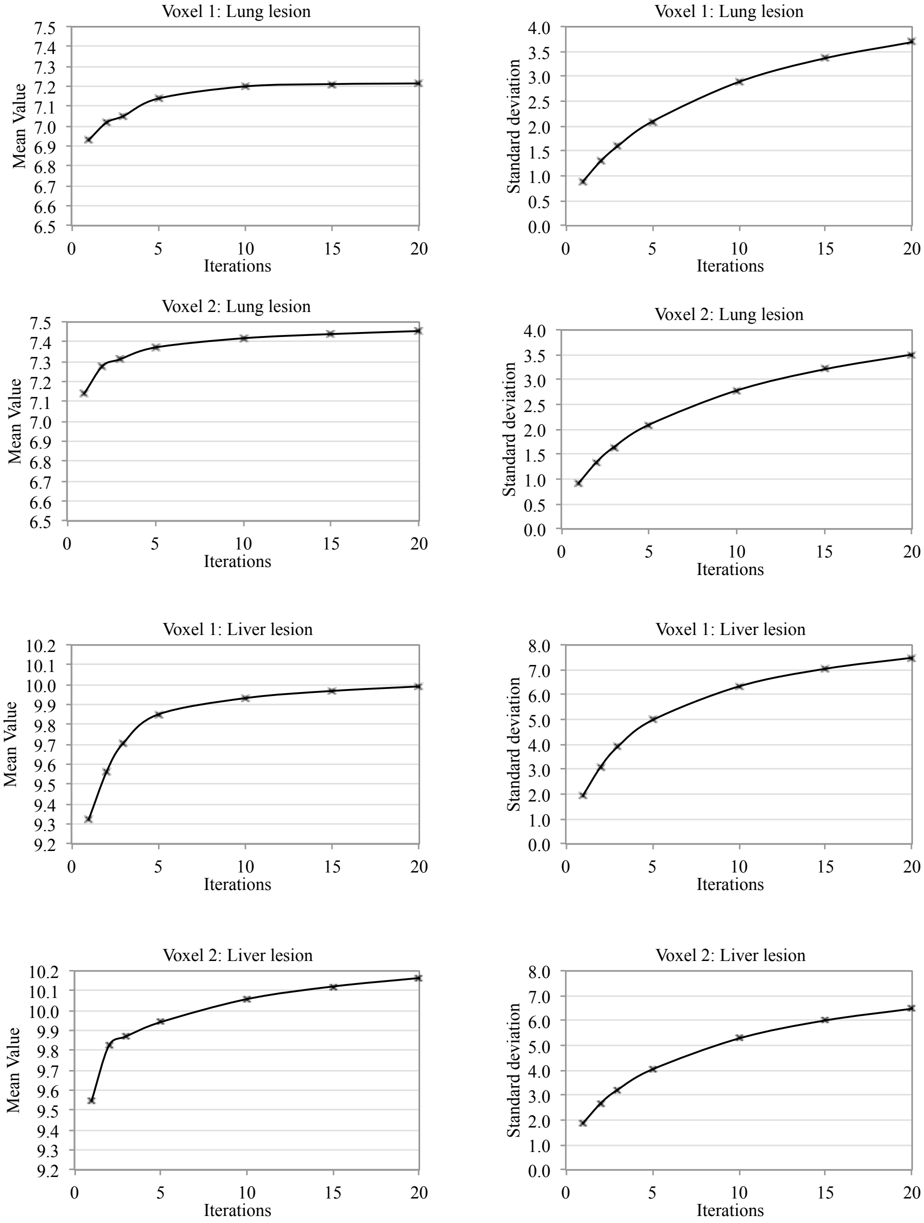


Figure 4.4: Mean and standard deviation values over the 100 realizations with respect to the number of iterations (with increment of 23 subsets) for the reference reconstructed image. Results are displayed for four single voxels, two adjacent voxels in the centre of the lung lesion and two adjacent voxels in the centre of the liver lesion.

4.2.1.4 Image assessment

Figures of merit were computed from all the reconstructed images produced from each numerical experiment. As it was mentioned in chapter 2 the advantage of simulated data for assessing or comparing algorithms is that the gold standard (i.e. true images) are known and therefore the precision (i.e. noise) and accuracy (i.e. bias) of the reconstructed images can be assessed for each algorithm. The following figures of merit were calculated:

Bias: An image of the absolute percentage bias with respect to the true simulated radioactivity distribution was determined as follows:

$$\begin{aligned} Bias(\%) &= \left| \frac{\frac{1}{N}\sum_{i=1}^{N}\Lambda_i - \Lambda_{ideal}}{\Lambda_{ideal}} \right| \times 100 \\ &= \left| \frac{\bar{\Lambda} - \Lambda_{ideal}}{\Lambda_{ideal}} \right| \times 100 \end{aligned} \tag{4.5}$$

where Λ_{ideal} is the reference image, Λ_i is the reconstructed image of the i^{th} noise realization and $\bar{\Lambda}$ is the mean image calculated as the average of the sequence of N realisations of the reconstructed image.

Noise: The noise was evaluated through the use of the standard deviation (std) calculated over the sequence of N reconstructed images:

$$std = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(\Lambda_i - \bar{\Lambda})^2} \tag{4.6}$$

Mean squared error (MSE): MSE was calculated as the sum of the square of the standard deviation and the squared bias:

$$MSE = (std)^2 + (bias)^2 \tag{4.7}$$

Contrast recovery (CR): The percentage CR was calculated according to:

$$\begin{aligned} CR(\%) &= \frac{C_{rec}}{C_{ref}} \times 100 \\ &= \frac{(I_{lesion}/I_{backg})_{rec} - 1}{(I_{lesion}/I_{backg})_{ref} - 1} \times 100 \end{aligned} \tag{4.8}$$

where C_{rec} and C_{ref} are the contrast values in the reconstructed image of interest and the reference image, respectively. I_{lesion} and I_{backg} are the SUV values in the lesion and background, respectively. For the background region a spherical region of interest (ROI) of 10 mm diameter was placed close to each lesion. For lesions 1-3 the ROI was placed inside the lungs and for lesions 4-6 inside the liver. The lesion ROIs were created in the simulated images in order to contain the actual lesion without excluding edge voxels and therefore including any partial volume effect. Any underestimation or overestimation of the volume caused by the two motion correction methods was not assessed in this investigation.

The results were also assessed using a statistical significance test. The Wilcoxon signed-rank test was chosen as the different measurements may not follow a normal distribution. For the bias, standard deviation, and MSE all voxel pairs between MCIR and RTA were compared separately for each lesion, whereas for the CR of each lesion the values from each reconstructed realisation pair were compared. The statistical significance threshold was set to 0.001 and p-values below this threshold were considered statistically significant.

4.2.2 Investigation of uncertainties introduced in RTA and MCIR

An investigation has been performed to study the hypothesis that the differences observed between RTA and MCIR in section 4.2.1 depend on the the non-linearity effects of iterative reconstruction which are exaggerated due to the relatively low counts in each gate. To examine this hypothesis the two methods were compared by applying zero motion fields for simulation plus correction using an iterative non-linear algorithm, such as the OSEM, and an analytic linear algorithm, such as the filtered backprojection (FBP). Therefore without any complications related to the properties of the motion information, RTA and MCIR were compared in terms of bias introduced due to the reconstruction properties. Then as a next step using only the OSEM algorithm the datasets were reconstructed by including the motion effect for simulation plus correction.

4.2.2.1 Simulated 4D PET data

A PET dataset derived from a real dynamic MRI dataset comprised of ten gated images was used (figure 4.5, table 4.2) as developed by Tsoumpas *et al.* (2011). Each of the PET gates was forward projected simulating the Philips Gemini scanner. Afterwards, attenuation and Poisson noise were added to each gate such that the overall counts correspond to a 5 *min* PET clinical thorax acquisition. 100 realizations were simulated. Transformations describing the motion between the ten respiratory positions and a reference position were obtained by applying a hierarchical affine image registration procedure (Buerger *et al.* 2011) to each of the 10 gates in the MRI dataset.

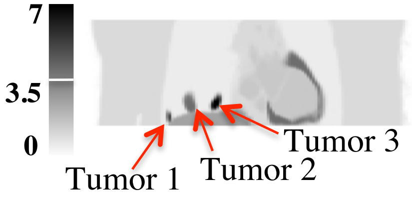


Figure 4.5: Coronal plane of the simulated radioactivity distribution with three lesions of different contrasts and sizes added at different locations in the lungs as described in table 4.2.

Table 4.2: Simulated radioactivity distribution in units of standardized uptake values (SUVs).

Tissue Type	SUV
Lungs	0.5
Liver	2.5
Myocardium	3.5
Soft tissue	1.0
Tumour 1	3.48
Tumour 2	3.48
Tumour 3	6.95

4.2.2.2 Evaluation datasets

Reconstructions were performed using OSEM (23 subsets with 1-40 iterations) and FBP (with ramp filter). Each slice consisted of 167 × 167 pixels with size 4 *mm* each, and the entire volume consisted of 87 slices with 2 *mm* thickness. The comparisons were assessed by computing the voxel-wise figure of merit bias using the images of the 100 realizations. Bias was determined as the difference with respect to the true simulated radioactivity as described in (4.5).

4.3 Results

4.3.1 Statistical comparison of RTA and MCIR

Figure 4.6 shows coronal planes of the mean images for the reference reconstructed, the non-motion corrected and following motion correction images using RTA and MCIR. Results are displayed at 20 iterations. Without motion correction, images are blurred and the small lesions are hardly visualized. However, after motion correction using either RTA or MCIR the image blurring is improved and all lesions are visible.

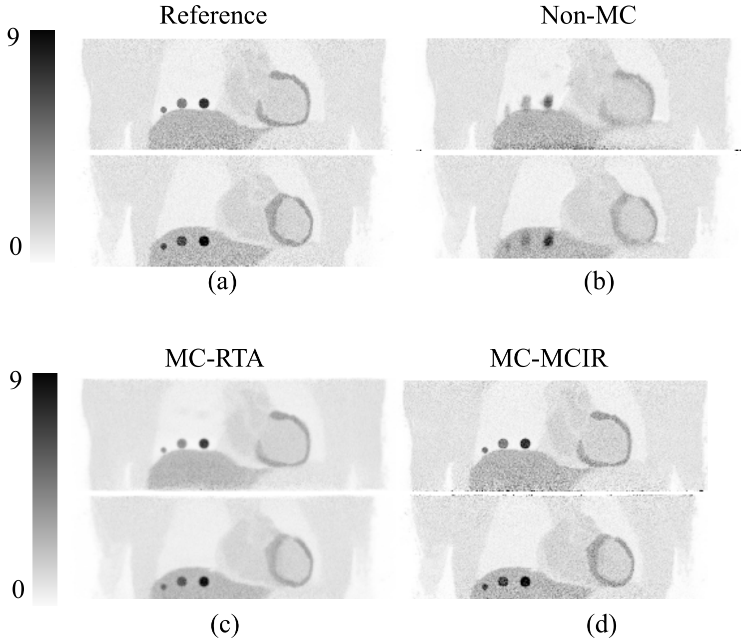


Figure 4.6: Coronal planes of mean images (i.e. average image across all 100 noisy realisations) for the reference reconstructed gate (a); non-motion corrected (Non-MC) (b); and following motion correction using RTA (MC-RTA) (c); and MCIR (MC-MCIR) (d). Results are displayed at 20 iterations. Motion correction improves motion blurring.

Table 4.3 and table 4.4 present mean and percentage CR values with standard errors for the reference reconstructed image and images reconstructed before and after motion correction. The results are presented for the six different lesions as shown in figure 4.1 and described in table 4.1. Without motion correction there is erroneous estimation of the radioactivity distribution especially for the small lesions. After motion correction using RTA there is only partial recovery of the PET signal. On the other hand, motion correction using MCIR is successful in the recovery of the true values of all lesions with CR more than 90% for all lesions.

Table 4.3: True and mean values with standard errors for the different lesions. Results are displayed at 20 iterations for the: reconstructed reference image and before motion correction (Non-MC) and after motion correction using RTA (MC-RTA) and MCIR (MC-MCIR).

SUV	True	Reference	Non-MC	MC-RTA	MC-MCIR
Lesion 1	4.5	4.26 ± 0.25	1.68 ± 0.16	2.87 ± 0.09	4.37 ± 0.31
Lesion 2	4.5	4.26 ± 0.25	2.35 ± 0.18	3.22 ± 0.08	4.30 ± 0.30
Lesion 3	7.5	7.33 ± 0.34	4.37 ± 0.26	6.14 ± 0.13	7.40 ± 0.39
Lesion 4	6.5	6.27 ± 0.56	3.72 ± 0.43	4.94 ± 0.19	6.42 ± 0.66
Lesion 5	6.5	6.36 ± 0.54	4.13 ± 0.43	5.35 ± 0.19	6.40 ± 0.70
Lesion 6	9.5	9.30 ± 0.69	6.30 ± 0.57	7.82 ± 0.23	9.23 ± 0.78

Table 4.4: Mean percentage CR values with standard errors for the different lesions over the 100 realisations. Results are displayed at 20 iterations for the: reconstructed reference image and before motion correction (Non-MC) and after motion correction using RTA (MC-RTA) and MCIR (MC-MCIR).

CR (%)	Reference	Non-MC	MC-RTA	MC-MCIR
Lesion 1	93 ± 7	30 ± 3	61 ± 8	94 ± 14
Lesion 2	96 ± 6	48 ± 3	70 ± 9	91 ± 14
Lesion 3	95 ± 5	57 ± 4	83 ± 7	96 ± 13
Lesion 4	97 ± 3	31 ± 3	67 ± 5	99 ± 12
Lesion 5	104 ± 5	42 ± 3	78 ± 6	98 ± 13
Lesion 6	98 ± 5	55 ± 3	82 ± 6	98 ± 12

Figure 4.7 shows coronal planes of the bias and standard deviation images for the RTA and MCIR reconstructions with the motion fields set to zero for both simulation and correction as described in the methodology subsection (section 4.2.1.3). Figure 4.8 shows coronal planes of the bias and standard deviation images for the motion corrected data. The results are displayed at a low number of iterations as typically used in the clinic (i.e. 2 iterations), for a case expected to have low bias (i.e. 10 iterations), and for a case expected that the reconstruction has not necessary reach convergence but has reached an approximately stabilized value over iterations satisfactory to study the quantification accuracy as has been shown in section 4.2.1.3 (i.e. 20 iterations). The results illustrate substantial differences between the two methods. For a low number

of iterations the MCIR bias is localized around edges and approaches a uniform small value as the number of iterations increase. On the other hand, for RTA high bias remains throughout. Standard deviation images show that noise is stabilized to a constant value at a small iteration number for RTA, while for MCIR it progressively increases with the number of iterations.

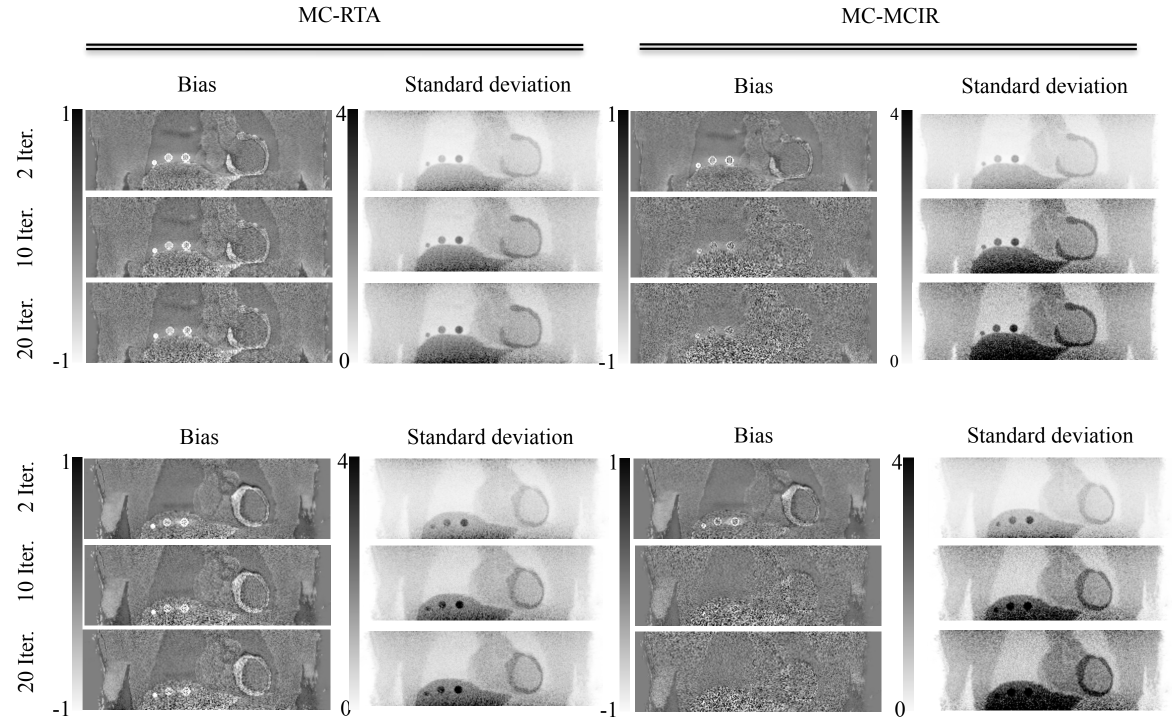


Figure 4.7: Coronal planes of the bias and standard deviation images for reconstructions with zero motion fields for simulation plus correction for RTA (Gate1-RTA) and MCIR (Gate1-MCIR). For 2 iterations the MCIR bias is localized around edges and approaches a uniform small value at 20 iterations. On the other hand, RTA has high bias for all iterations especially at the edges of the lesions independent of the iteration number. Standard deviation is stabilized to a constant value for RTA, while for MCIR it progressively increases with the number of iterations.

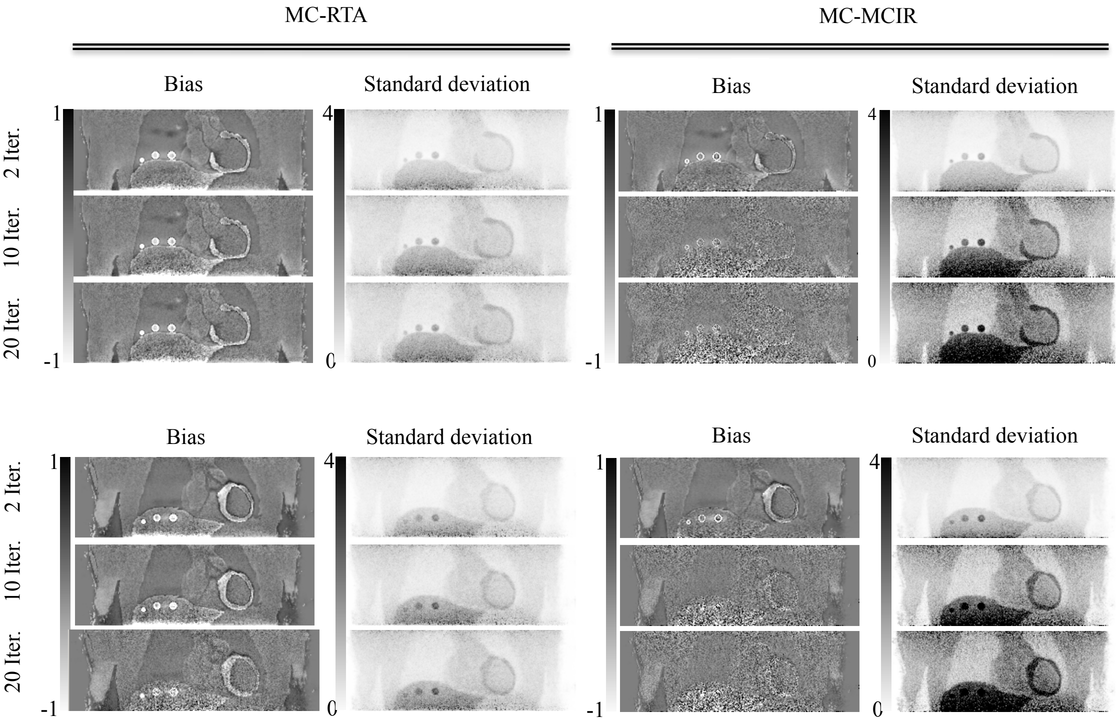


Figure 4.8: Coronal planes of the bias and standard deviation images after motion correction using both RTA (MC-RTA) and MCIR (MC-MCIR). For 2 iterations the MCIR bias is localized around edges and approaches a uniform small value at 20 iterations. On the other hand, RTA has high bias for all iterations especially at the edges of the lesions independent of the iteration number. Standard deviation is stabilized to a constant value for RTA, while for MCIR it progressively increases with the number of iterations.

Figure 4.9 shows the absolute percentage bias versus the number of iterations before and after motion correction using RTA and MCIR as well as the RTA and MCIR reconstructions with the motion fields set to zero for both simulation and correction as described in the methodology subsection. Results are displayed for the average of the lesions in the lungs and liver correspondingly. Reconstructions with RTA and MCIR with zero motion fields have different performance: at a low number of iterations (i.e. 1-2), bias is similar in both methods, i.e. approximately 30% and 18% for lung and liver lesions, respectively, whereas at late iterations (i.e. 20) MCIR bias in lungs and liver is 5% and 2%, respectively, while RTA bias is 18% and 14% for lungs and liver, respectively. Introducing motion into the simulations and correcting for it, MCIR performance has slightly reduced bias, while RTA has clearly increased bias.

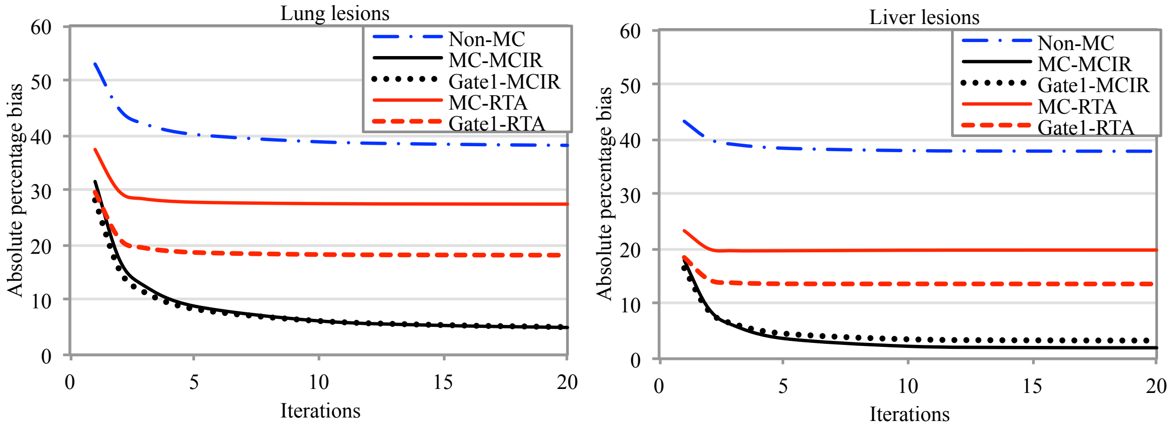


Figure 4.9: Absolute percentage bias with respect to the number of iterations (with increment of 23 subsets) for images with no-motion correction (Non-MC), following motion correction using RTA (MC-RTA) and MCIR (MC-MCIR), and reconstructions with zero motion fields for simulation plus correction for RTA (Gate1-RTA) and MCIR (Gate1-MCIR). Although correction with both RTA and MCIR improves quantitative accuracy, MCIR outperforms RTA. In contrast to RTA, MCIR reconstruction with zero motion fields has similar performance to introducing motion into the simulations and correcting for it.

For similar datasets as shown in figure 4.9, figure 4.10 shows the absolute percentage bias versus standard deviation at different iteration numbers, while figure 4.11 illustrates the MSE versus the number of iterations. At 1-2 iterations MCIR has smaller bias than RTA but similar noise. However, MCIR advances along the bias-standard deviation curve faster than RTA as the number of iterations increase. In terms of MSE, at 1-2 iterations MCIR provides superior performance to RTA for most occasions, while for a high number of iterations RTA is extensively better than MCIR.

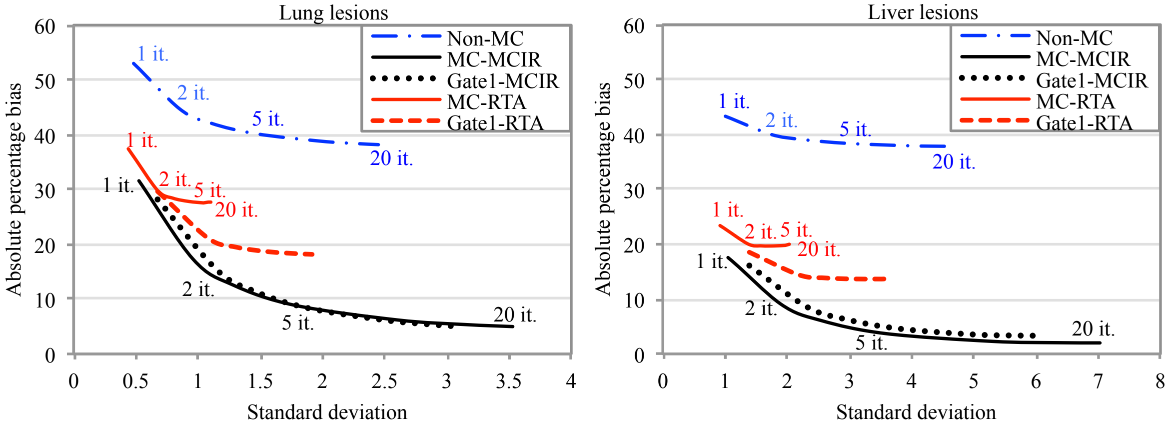


Figure 4.10: Absolute percentage bias versus the standard deviation at different number of iterations for the images with no motion correction (Non-MC), following motion correction using RTA (MC-RTA) and MCIR (MC-MCIR), and reconstructions with zero motion fields for simulation plus correction for RTA (Gate1-RTA) and MCIR (Gate1-MCIR). At low iterations MCIR has smaller bias than RTA but similar noise while MCIR advances along the bias-standard deviation curve faster than RTA as the number of iterations increase.

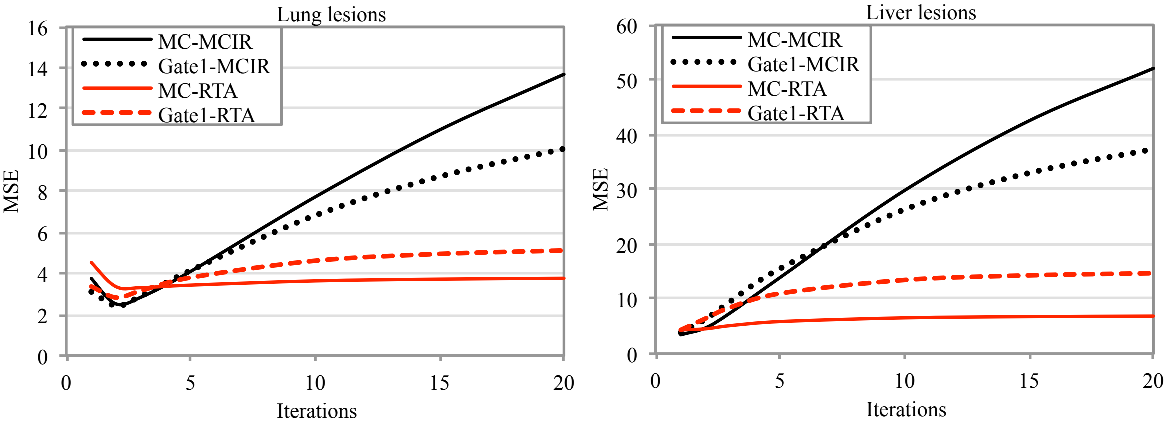


Figure 4.11: Mean squared error with respect to the number of iterations (with increment of 23 subsets) after motion correction using RTA (MC-RTA) and MCIR (MC-MCIR), and reconstructions with zero motion fields for simulation plus correction for RTA (Gate1-RTA) and MCIR (Gate1-MCIR).

Figure 4.12 shows coronal planes of the images produced with one realisation after motion correction using RTA and MCIR, before and after smoothing with a 4 *mm* isotropic FWHM Gaussian kernel. The images are presented at 20 iterations. Prior to smoothing MCIR images are noisier than RTA, while after smoothing the two methods have more similar visual quality. Figure 4.13 and figure 4.14 illustrate the mean percentage CR and the MSE, respectively, versus the number of iterations for the different lesions after motion correction. Results are displayed before and after smoothing the 100 reconstructed realisations.

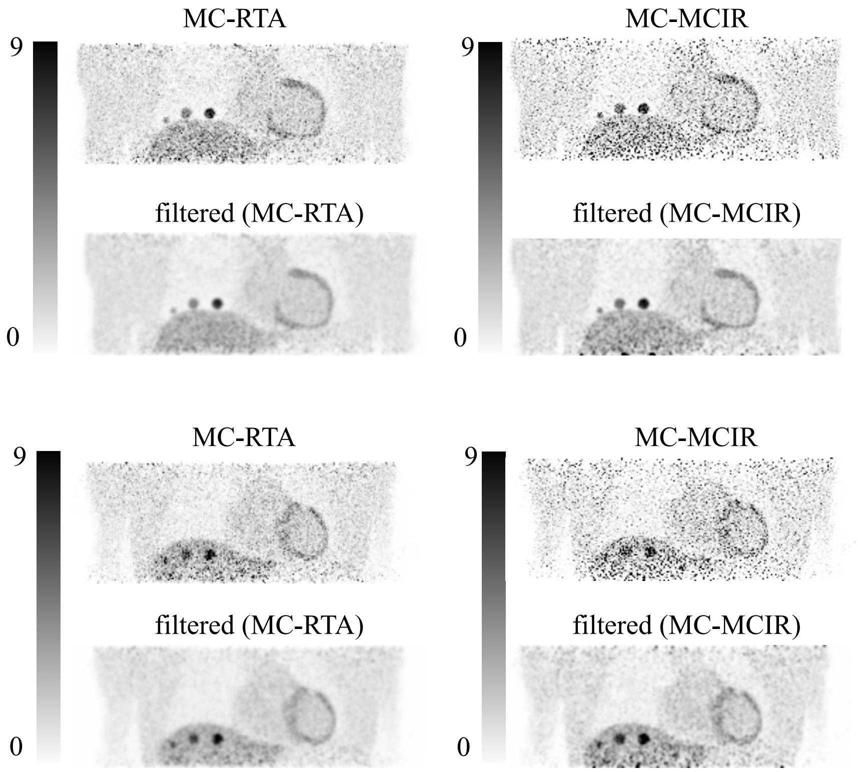


Figure 4.12: Coronal planes of the image produced with one realisation without and with post-reconstruction filtering after motion correction using RTA (MC-RTA) and MCIR (MC-MCIR). Results are displayed at 20 iterations. Smoothing reduces noise and improves visual quality especially for MCIR.

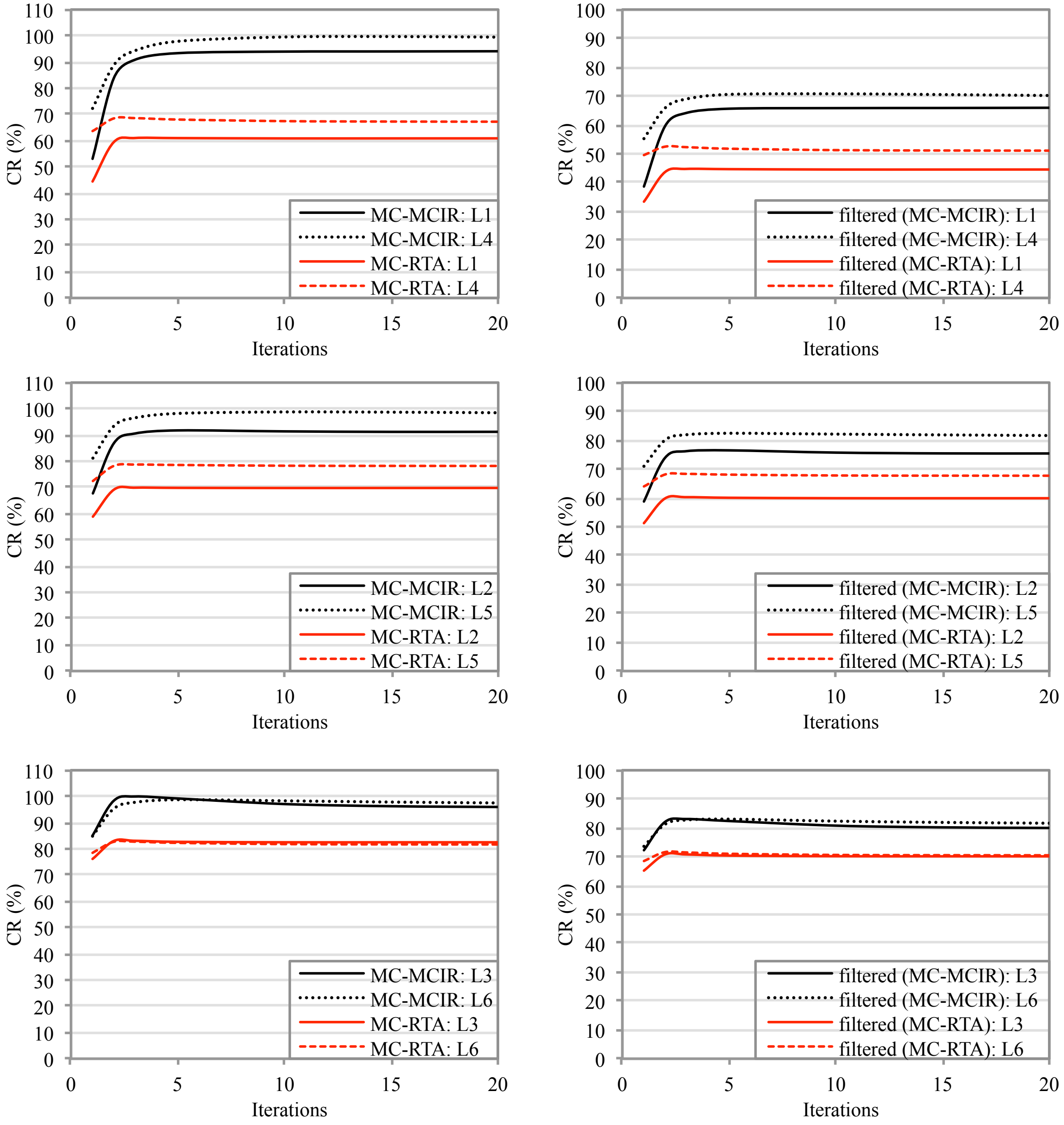


Figure 4.13: Mean percentage contrast recovery with respect to the number of iterations (with increment of 23 subsets) for the different lesions over the 100 realisations without and with post-reconstruction filtering after motion correction using RTA (MC-RTA) and MCIR (MC-MCIR).

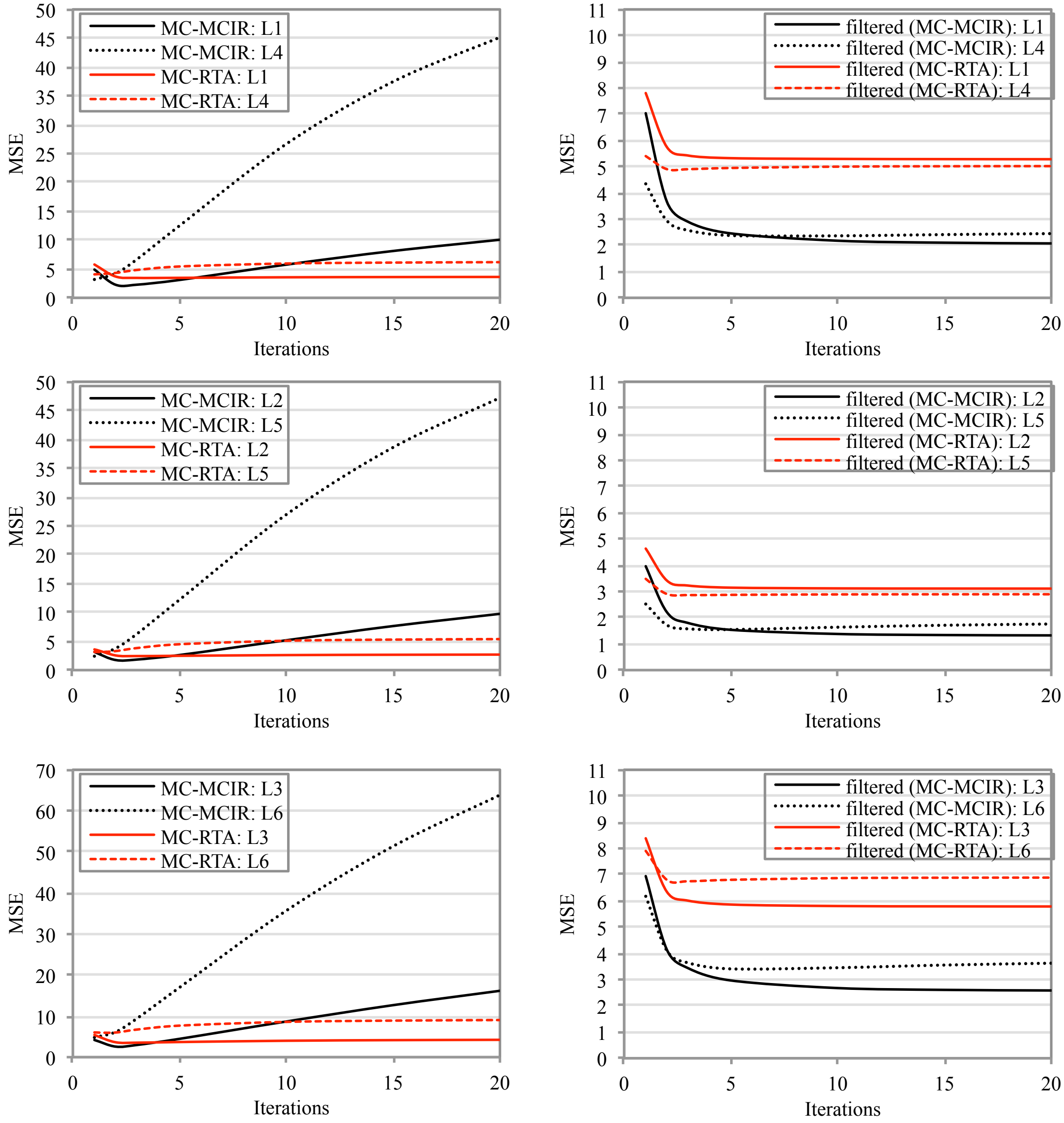


Figure 4.14: Mean squared error with respect to the number of iterations (with increment of 23 subsets) for the different lesions over the 100 realisations without and with post-reconstruction filtering after motion correction using RTA (MC-RTA) and MCIR (MC-MCIR).

Table 4.5 shows the aggregate results (i.e. average over the six lesions) with standard errors for the mean percentage CR and the MSE at different iteration numbers after motion correction using RTA and MCIR without and with post-reconstruction filtering. The standard errors correspond to the precision of the average values over the six lesions and therefore the variation in the average values with post-reconstruction filtering is not necessarily lower than without post-reconstruction filtering.

Table 4.5: Average percentage CR and MSE values with standard errors over the six lesions at different iteration numbers. Results are displayed without and with post-reconstruction filtering after motion correction using RTA (MC-RTA) and MCIR (MC-MCIR).

	Iteration	MC-RTA	Filtered (MC-RTA)	MC-MCIR	Filtered (MC-MCIR)
CR (%)	1	66 ± 6	55 ± 6	74 ± 6	61 ± 6
	2	74 ± 4	61 ± 5	91 ± 2	74 ± 4
	3	74 ± 4	61 ± 5	95 ± 2	76 ± 4
	5	74 ± 4	61 ± 5	96 ± 1	77 ± 3
	10	73 ± 4	61 ± 5	96 ± 1	76 ± 3
	15	73 ± 4	61 ± 5	96 ± 1	76 ± 3
	20	73 ± 4	61 ± 5	96 ± 1	76 ± 3
MSE (SUV^2)	1	5 ± 0.5	6 ± 1	4 ± 0.4	5 ± 1
	2	4 ± 0.5	5 ± 1	3 ± 1	3 ± 0.4
	3	4 ± 1	5 ± 1	5 ± 1	3 ± 0.4
	5	4 ± 1	5 ± 1	9 ± 3	2 ± 0.3
	10	5 ± 1	5 ± 1	18 ± 6	2 ± 0.3
	15	5 ± 1	5 ± 1	26 ± 8	2 ± 0.3
	20	5 ± 1	5 ± 1	32 ± 10	2 ± 0.3

For the figures of merit bias, standard deviation, CR and MSE, the lesion values between MCIR and RTA at different numbers of iterations (e.g. 1, 2, 3, 5, 10, 15, and 20) are compared using the Wilcoxon signed-rank test. Without post-reconstruction filtering, there are statistically significant differences between RTA and MCIR for all lesions in terms of bias, standard deviation, and CR. In the case of MSE, there are only a few instances without statistically significant difference between RTA and MCIR. In particular, these differences are not significant at: two iterations for lesion 4 ($p = 0.15$)

and lesion 6 ($p = 0.45$), at three iterations for lesion 3 ($p = 0.014$), and at five iterations for lesion 1 ($p = 0.17$) and lesion 2 ($p = 0.15$). With post-reconstruction filtering for all lesions there are statistically significant differences between post-reconstruction filtered RTA and post-reconstruction filtered MCIR for bias, standard deviation, CR and MSE (p less than 0.001) except for standard deviation at first iteration in lesion 4 ($p = 0.25$).

Figure 4.15 shows coronal planes of the images produced with one realisation after motion correction using RTA and MCIR for different values of the penalization factor (i.e. β_g: 0, 0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 50, 100). Results are displayed at 10 iterations. For β_g equal to 50 the RTA and MCIR result in similar performance based on visual appearance.

Figure 4.16 shows coronal planes of the mean, bias and standard deviation images for the motion corrected images using both RTA (MC-RTA) and MCIR (MC-MCIR) using penalization factor β_g equal to 2 for the RTA and 10 for the MCIR. The results have been obtained using twenty noise realisations and are presented at 10 iterations.

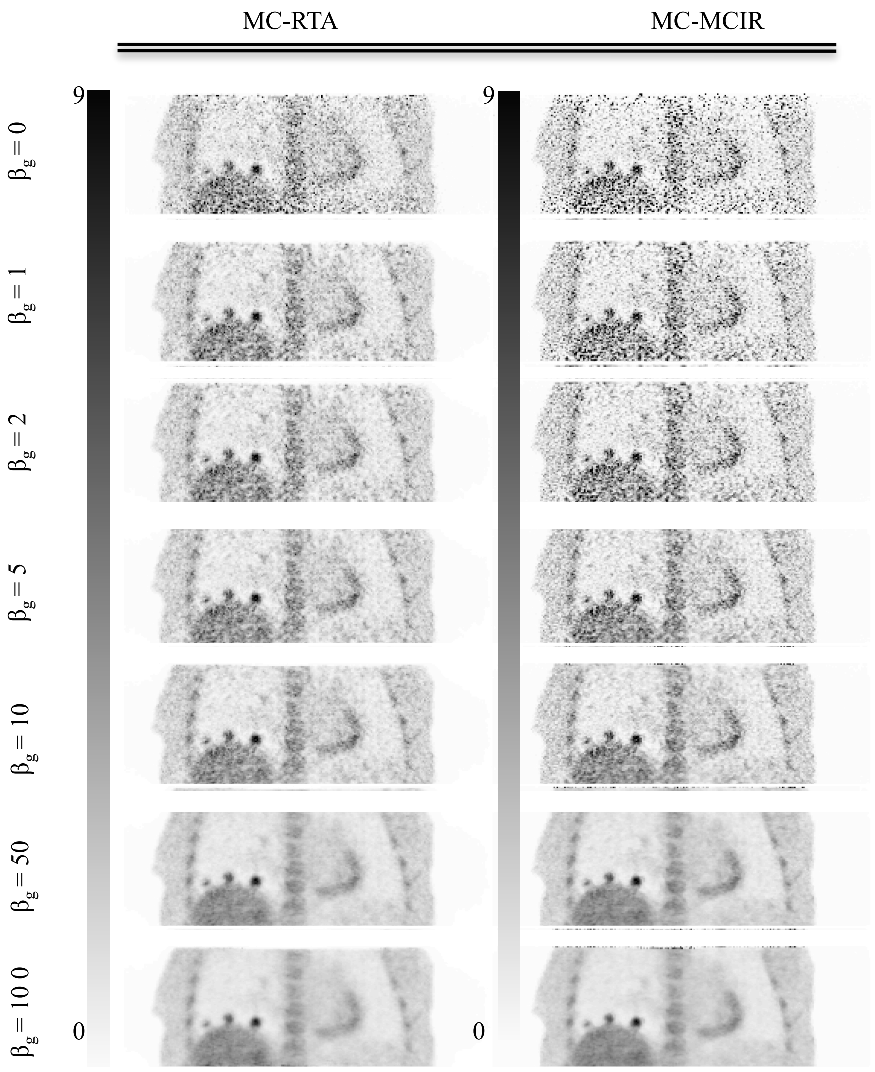


Figure 4.15: Coronal planes of the image produced with one realisation after motion correction using (MC-RTA) and MCIR (MC-MCIR) for different values of the penalization factor (i.e. β_g: 0, 0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 50, 100). Results are displayed at 10 iterations.

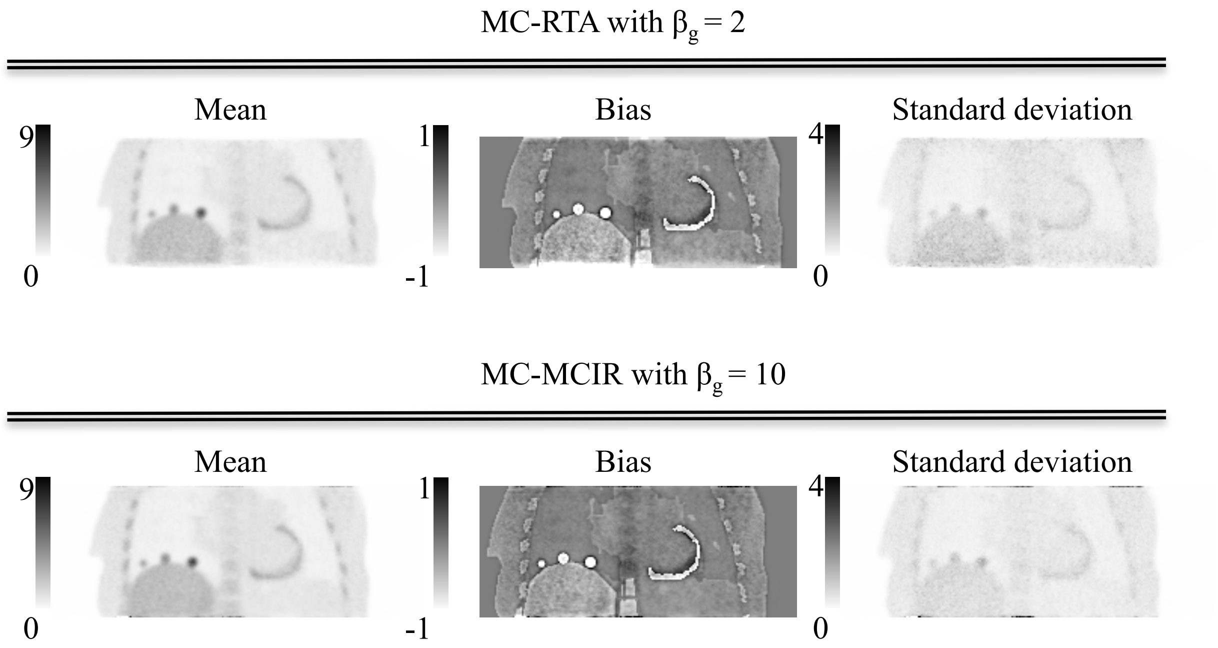


Figure 4.16: Coronal planes of the mean, bias and standard deviation images after motion correction using both RTA (MC-RTA) and MCIR (MC-MCIR) using penalization factor β_g equal to 2 for the RTA and 10 for the MCIR. The results have been obtained using twenty noise realisations and are presented at 10 iterations.

4.3.2 Investigation of uncertainties introduced in RTA and MCIR

Figure 4.17 shows the bias images for reconstructions with zero motion fields for simulation plus correction for RTA and MCIR using the reconstruction algorithms FBP and OSEM at 1, 5 and 40 iterations with 23 subsets. FBP reconstructed images demonstrate equal bias for RTA and MCIR. On the other hand, substantial differences are observed between RTA and MCIR using OSEM; overestimations and underestimations in low and high counts areas respectively.

Figure 4.18 presents the plots of the absolute percentage bias with respect to the number of iterations for reconstructions with zero motion fields for simulation plus correction for RTA (i.e. Gate1-RTA) and MCIR (i.e. Gate1-MCIR) for the FBP and OSEM algorithms. Results of including the motion effect for simulation plus motion correction using RTA and MCIR are also presented for the OSEM algorithm. Without including the motion effect, FBP reconstructed images result in identical bias for RTA and MCIR while OSEM reconstructed images show higher bias for RTA compared to MCIR. After introducing motion into both simulation and correction the bias in MCIR remains low while RTA shows increased bias demonstrating that interpolation errors are introduced during RTA motion correction.

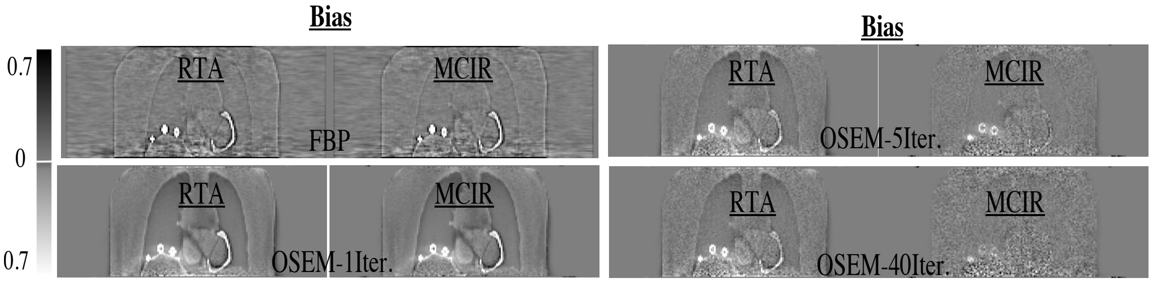


Figure 4.17: Coronal planes of the bias images for reconstructions with zero motion fields for simulation plus correction for RTA and MCIR. Results are displayed for the reconstruction algorithms FBP and OSEM (1, 5 and 40 iterations).

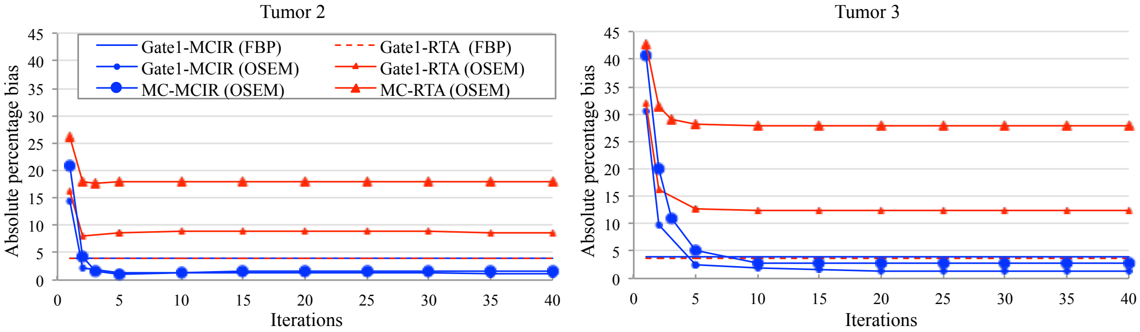


Figure 4.18: Plots of the absolute percentage bias with respect to the number of iterations (with increment of 23 subsets) for reconstructions with zero motion fields for simulation plus correction for RTA (Gate1-RTA) and MCIR (Gate1-MCIR) for the FBP and OSEM reconstruction algorithms. Results are also plotted for images from simulations which include the motion and plus motion correction using RTA (MC-RTA) and MCIR (MC-MCIR) for the OSEM reconstruction algorithm.

4.4 Discussion

The main aim of this chapter was to understand the behaviour of the two most common motion correction approaches, RTA and MCIR, in terms of convergence and image properties. As can be seen in figure 4.6, table 4.3 and table 4.4 motion correction using either RTA or MCIR improves motion blurring and quantification accuracy. The two methods behave differently: MCIR has better CR than RTA (table 4.4) and negligible bias, whereas RTA has high bias but minimum noise (figure 4.8). The high bias in RTA was also reported in other studies, (Thielemans *et al.* 2006; Dikaios and Fryer 2011a,b) while the different noise characteristics between RTA and MCIR were reported only by Thielemans *et al.* (2006).

Applying RTA and MCIR methodology with zero motion fields, and therefore avoiding any consequent inaccuracies due to interpolations that may occur during transformations, RTA and MCIR have different performances. As shown in figure 4.7 for reconstructions with zero motion fields for simulation plus correction for 2 iterations the MCIR bias is localized around edges and approaches a uniform small value at 20 iterations while RTA has high bias for all iterations especially at the edges of the lesions and the myocardium indicating resolution degradation. Considering the absence of motion effects the higher bias observed in RTA compared to MCIR can be attributed to resolution degradation due to low count statistics. This is also confirmed quantitatively in figure 4.9. For low numbers (i.e. 1 or 2) of iterations, bias is similar in both methods. In particular at 1 iteration bias for both RTA and MCIR is approximately 30% and 18% for lung and liver lesions, respectively, whereas at 20 iterations MCIR bias in lungs and liver is 5% and 2%, respectively, while RTA bias in lungs and liver is 18% and 14%, respectively.

After the introduction of motion into the simulations and correction for it, MCIR performance is unchanged, whereas the RTA shows even higher bias (figure 4.8). A possible explanation for the additional bias in RTA is the interpolation errors introduced during the transformation procedure where RTA appears to be more sensitive than MCIR. During RTA equation (4.2) each reconstructed gate is transformed once and thus errors will have a linear effect on the corrected image. On the other hand, during MCIR equation (4.4) the transformation is not applied on the image directly but on the update image at each iteration: the same transformation is used prior to forward projection and following back-projection and the errors are reduced (Dikaios and Fryer 2011a).

One source of bias in RTA is the relatively low count statistics (Walker *et al.* 2011) in each gate (i.e. 1/10 of a 5 min acquisition), while MCIR considers all counts simultaneously. The low count statistics may cause very slow convergence or not allow convergence to the correct value. In particular, within the framework of OSEM algorithm, non-negativity constraint does not allow negative values to be assigned to an image voxel and when low count data lead to a probability that a voxel will be assigned a negative value the OSEM bounds the intensities distributions at zero. As a result for high variance data the mean of the intensity distributions changes and the distributions do not remain symmetric (Walker *et al.* 2011) and is expected to result in an overestimation bias for low count regions (Grezes-Besset *et al.* 2007; Ahn and Fessler 2004). Therefore linearity is not preserved in the OSEM; the reconstruction of the sum of two sinograms, similar to the methodology followed in MCIR, is not the same as the sum of the independent reconstruction of the two same sinograms, similar to the methodology followed in RTA.

The effect of the non-linearity of the OSEM at low counts statistics on the performance of RTA and MCIR was also confirmed by a study using a linear (i.e. FBP) and an iterative (i.e. OSEM) reconstruction algorithm. The methodology followed by the two methods in the absence of motion was compared; reconstructing the sum of a number of sinograms (MCIR) or adding independent reconstructions of sinograms (RTA). As illustrated in figure 4.17, FBP reconstructed images demonstrate equal bias for RTA and MCIR in areas and edges which can be attributed to reconstruction artefacts. On the other hand, substantial differences are observed between RTA and MCIR using OSEM; overestimations and underestimations in low and high counts areas respectively. For all regions, higher bias is observed for RTA than MCIR while for both methods the bias is progressively decreased as the number of iterations is increased (figure 4.18). Considering that images are motion free, the high bias in RTA can be attributed to the reconstruction properties (i.e. OSEM non-linearity effects) which are exaggerated due to the relatively low counts in each gate while MCIR uses all data simultaneously and thus it is less susceptible to non-negativity constraints.

The investigation comparing OSEM with FBP provides a better understanding of the differences observed between RTA and MCIR. With the use of a linear reconstruction algorithm the sum of independently reconstructed projection data (i.e. RTA) is the same with summing all the projection data and then reconstructing them (i.e. MCIR). On the other hand, with the use of a non-linear algorithm the sum of independently reconstructed projection data (i.e. RTA) is prone to the non-linearity effects which are exaggerated due to the low counts present in each gate. Furthermore, figure 4.18 indicates that after introducing motion into both simulation and correction, bias in MCIR is low. On the other hand, RTA shows increased bias after introducing motion, suggesting that interpolation errors are introduced during the RTA motion correction.

Even though other studies (Dikaios and Fryer 2011a,b; Thielemans *et al.* 2006) reported high bias in RTA, our results show that RTA errors occur from both the transformation procedure and the non-linearity characteristics of OSEM (i.e. first step of RTA). OSEM non-linearity characteristics result in bias of approximately 18% and 14% for lungs and liver lesions correspondingly, while interpolation errors cause additional bias of approximately 9% and 6% for lungs and liver lesions, respectively. Furthermore, it is interesting to note that after introducing motion, MCIR has marginally even less bias than MCIR without motion. This might be explained by the use of all data in one reconstruction taking advantage of regions that appear in different voxels of multiple gates providing denser discretization sampling Dikaios and Fryer (2012a). Based on this Wallach *et al.* (2012) enhanced the image resolution by using motion information in a sequence of respiratory gated datasets representing different views of the same object.

Although resolution degradation may be caused during RTA due to both low count statistics and interpolation errors, noise is very small and stabilizes to a constant value, whereas in MCIR noise becomes progressively greater with the number of iterations. At a low number of iterations (e.g. 1 or 2), currently the typical setting for clinical PET, MCIR outperforms since it has a smaller bias than RTA but similar noise, while MCIR advances along the bias versus standard deviation curve faster than RTA as the number of iterations increase (figure 4.10).

The different noise characteristics of RTA and MCIR have been reported only by Thielemans *et al.* (2006). It has been assumed that the summation of independently reconstructed gates will result in inferior quality images compared to the use of all gates in a single reconstruction (Lamare *et al.* 2007b) and that therefore MCIR will provide better signal to noise properties than RTA. Similar conclusions were also drawn by a theoretical study by Asma *et al.* (2006) using penalized likelihood reconstruction where it was shown that MCIR would provide lower noise than RTA with or without penalization in both methods. Another recent theoretical study based on penalized likelihood reconstruction by Chun and Fessler (2013) showed that noise in MCIR is expected to be lower or comparable to noise of RTA. However, the finding of our study shows that at high number of iterations (e.g. 20) RTA provides superior performance to MCIR in terms of MSE due to reduced and stabilized noise over iterations (figure 4.11) indicating that RTA may outperform MCIR in lesion detectability (Thielemans *et al.* 2006). The single realisation images in figure 4.12 provide a visual demonstration that at a high number of iterations MCIR may have worse detectability properties than RTA due to high noise. For a small number of iterations (e.g. 2) MCIR provides superior performance to RTA for nearly all occasions, while at a high number of iterations (e.g. 20) RTA provides superior performance to MCIR in terms of MSE ($p < 0.001$) (figure 4.14).

Furthermore, in the current simulations no blurring was used to mimic scanner resolution, something that may result in aliasing and cause variance magnification use effects that are lessened with RTA via image interpolation. However, the differences in bias and noise characteristics of RTA and MCIR observed in this chapter are in agreement with a preliminary study presented in Appendix C that uses more realistic simulations confirming that the noted effects are indeed a characteristic of the RTA and MCIR methods and not a detail of the simulations or an aliasing effect.

The choice of the optimal method depends on the figure of merit referred to (i.e. bias, standard deviation, MSE). RTA causes high levels of overestimation and underestimation in cold and hot regions, respectively due to bias induced by the non-negativity constraints and interpolation errors resulting in resolution degradation. It is found that for a small number of iterations, MCIR outperforms in the aspect of bias and MSE and has similar noise distribution to RTA. However, for a large number of iterations RTA has improved MSE compared to MCIR due to the bias induced from non-negativity constraints and interpolation smoothing. Consequently, overall this study indicates that MCIR may provide superior performance overall to RTA providing noise can be minimized.

A common way to treat noise in clinical practice is early stopping or post-reconstruction filtering or a combination of early stopping and post-reconstruction filtering. In this way, MSE performance is improved at the expense of CR and resolution. In the particular simulations of this study without filtering, MCIR successfully recovers the true values of all lesions with a CR of more than 90%, whereas after post-reconstruction filtering CR is more than 65% (figure 4.13). Following smoothing MCIR outperforms RTA (with the same level of smoothing) in terms of both CR and MSE ($p < 0.001$). Without post-reconstruction filtering the MSE values are not affected by the lesion size: lesions of different size but of the same SUV and background (e.g. lesion pairs 1/2 and 4/5) show similar performance. After post-reconstruction filtering, the MSE values are more affected in the small lesions (e.g. lesions 1 and 4).

The aggregate results for the CR and the MSE presented in table 4.5, even if they are lesion dependent, provide an indication of the trend for the optimal number of iterations for the specific figures of merit. To minimize the MCIR MSE without post-reconstruction filtering requires 2 iterations while with post-reconstruction filtering needs 3 to 5 iterations. In order to maximize CR, MCIR needs 3 to 5 iterations, with or without post-reconstruction filtering. On the other hand, the best case for optimal CR and MSE for RTA without or with post-reconstruction filtering is at 2 iterations. After 2 iterations post-reconstruction filtered MCIR outperforms RTA with or without post-reconstruction filtering in both MSE and CR. In particular, regarding the optimal iteration number for each figure of merit: After post-reconstruction filtering MCIR results in 42% lower MSE and 4% higher CR than without post-reconstruction filtering RTA. It must be noted that while post-reconstruction filtering in RTA deteriorates both MSE and CR on average 25% and 18%, respectively, post-reconstruction filtering in MCIR deteriorates CR on average 20% but improves the MSE 33%.

The performance of post-reconstruction filtered MCIR is likely to depend on the smoothing level and the reconstruction algorithm parameters chosen (e.g. number of iterations). Different smoothing levels and number of iterations may be required based on the desired contrast recovery or noise level and diagnostic task. For example, the optimum number of iterations with respect to MSE is at about 5 iterations and CR at about 5 iterations with the use of an isotropic 4 *mm* Gaussian filter (table 4.5). However, MCIR that is stopped early at 2 iterations and post-reconstruction filtered produces similar results to non filtered RTA in terms of CR. This demonstrates that in the current clinical protocols for FDG PET acquisitions, motion correction with RTA is expected to be a valid method.

However, post-reconstruction filtering is expected to be a good approach to regularize MCIR when the quantification is not the main issue since the noise is controlled at the expense of quantification accuracy. In order to fully exploit the high quantification accuracy of MCIR and RTA, methods to minimize noise while retaining the quantification accuracy may be superior to post-reconstruction filtering such as introducing appropriate regularization into MCIR reconstruction.

The findings of this study show that introducing regularization into the reconstruction results in similar performance between the RTA and MCIR in terms of both bias and noise properties as shown in figure 4.16. However, the performance of regularization depends on the value of the penalization factor β_g as it has been shown to affect the RTA and MCIR in a different way (figure 4.15). In order to obtain similar noise levels between RTA and MCIR the latter requires a higher value of β_g. Another advantage gained by including median-root-prior regularization within the reconstruction is that RTA bias is decreased for small β_g (e.g. 0.1) Tsoumpas *et al.* (2013b). The advantages gained by including median-root-prior regularization within the reconstruction of both motion correction methods are presented in detail in Tsoumpas *et al.* (2013b). Further research is needed in better regularisation approaches (Wang and Qi 2012).

Furthermore, a recent study showed improvements in the performance of RTA by reducing the bias using registration-weighted motion correction (Dikaios and Fryer 2012a). Another possible way of dealing with the high bias in RTA is using a better warping function in order to reduce the interpolation errors during the transformation. In addition, as the voxel size is decreased interpolation errors are expected to be decreased but this may not be an efficient method to deal with the bias in RTA: as the voxel size is decreased the low count statistics create an additional limitation.

In this study no scatter or random coincident events were simulated. In the simulations used for the purposes of this thesis the available method for scatter correction is the single scatter simulation (SSS) based method with appropriate scaling to compensate for the multiple scatter coincident events (see Appendix B). In the case that scatter was included in the simulations, the scatter distribution would be estimated for each respiratory position and used for the simulations of the projection data for each respiratory position. However, for the scatter correction, during reconstruction of the non motion corrected image, the scatter distribution is estimated with non-spatially correlated emission and attenuation data due to the effect of motion. As a result of this misalignment the measured data are not aligned to the scatter points in the attenuation image introducing uncertainties in the scatter distribution. In addition, the effect of motion is evident during scaling of the SSS which is based on the tails of the measured data. If the measured data are not aligned with the attenuation map, the coincidences might be wrongly measured outside the patient resulting in wrong tails and therefore wrong scaling factor introducing inaccuracies to the estimated scatter distribution and consequently uncertainties in the quantification accuracy of the reconstructed image. On the other hand, during reconstruction of the motion corrected image, with both RTA and MCIR methods, the scatter distribution corresponding to each respiratory position is used and therefore the results are not expected to change. In the case of random events which tend to be uniformly distributed in the FOV the effect of motion is expected to be small. Increasing the realism of the modelling of the photon detection by including scatter and random events is expected to affect the results of the effect of motion on PET but not the differences observed between the two algorithms which is the point under investigation in this study.

Another limitation of the simulation datasets used for the purposes of this investigation is that the same projector was used for simulation and reconstruction. In the case of having a different projector as in the real cases edge artefacts might be introduced due to the mismatches such as overestimation of the blurring kernel used for the reconstruction compared to the true system PSF when using a broader blurring resolution kernel than the actual system PSF. This would result in inaccurate modelling of the system matrix and therefore it would be more complex to define the origin of the observed errors when comparing two motion correction methods. The purpose of this choice was to increase as much as possible the matching of forward projector and backprojector in order to simplify the problem and isolate the effect of motion to clearly characterize the discrepancies between the two motion correction methods.

Finally, it is important to mention that in this investigation the same motion fields were used both to create the simulations and to correct for motion, and intra-gate motion was not simulated. However, in reality there will be discrepancies between actual and estimated motion due to the dynamic acquisition and the inaccuracies of the motion estimation. Although this is a limitation of this study, using the same motion fields to both generate and correct the data favours motion correction accuracy and allows the investigation of reconstruction both methods without being affected from any errors introduced from motion estimation inaccuracies.

4.5 Conclusion

This investigation has revealed the statistical and computational differences between RTA and MCIR. The main findings are the substantially different convergence and statistical performance between them. The choice of the optimal method depends on the figure of merit referred to (i.e. bias, standard deviation, MSE, CR). RTA causes high levels of overestimation and underestimation in cold and hot small regions, respectively, due to the non-linearity of OSEM at low count statistics and interpolation errors resulting in resolution degradation. It is found that for a small number of iterations (e.g. 2), MCIR outperforms in the aspect of bias and has similar noise properties to RTA. On the other hand, for large number of iterations RTA has better MSE than MCIR. This study indicates that MCIR may provide superior performance overall to RTA, but only if noise is minimized. However, in current clinical applications where quantification is not the main objective, RTA can be a practical and simple method to correct for motion.

Chapter 5

Effect of respiratory motion correction and spatial resolution on lesion detection in PET.

5.1 Introduction

One of the biggest clinical applications of PET in oncology is staging of disease for which it is necessary to detect small and low uptake lesions with as high a sensitivity as possible. However, lesion detection in PET is inherently limited by the spatial resolution of the scanner which underlines the importance of improved resolution in future scanners (Schäfers 2008a). Despite the improvements in PET scanner technology, respiratory motion is an unavoidable factor that severely degrades the effective spatial resolution (Xu *et al.* 2011) limiting lesion detection performance for chest and abdominal imaging. Respiratory motion typically causes displacements of less than 10 *mm* for lung lesions but this may increase to more than 20 *mm* for lung and liver lesions close to the diaphragm (Seppenwoolde *et al.* 2002; Langen and Jones 2001). Considering that conventional PET scanners have a spatial resolution of approximately 6 *mm*, motion of that range effectively reduces the spatial resolution. Resolution is the limiting factor for detecting small lesions. With the potential improved resolution of future scanners the effects of motion are likely to become relatively more important (Daou 2008).

Despite the interest in the development of respiratory motion correction methodologies and the benefits of them on quantitative imaging tasks, the improvement in sensitivity for lesion detection has not yet been rigorously evaluated. A study by Marache-Francisco *et al.* (2010) has shown that motion correction can improve detectability in the liver, but only a small effect is seen in the lungs, and this may be attributed to the extent of the displacement caused by motion that varies with the location of the lesions or the accuracy of the motion fields used for correction. Another recent study by Park *et al.* (2012) reported significantly better detection performance for lung lesions with than without respiratory motion correction. However, in both these studies the extent of the improvement as a function of lesion characteristics was not investigated while the study was restricted to one respiratory signal (i.e. motion amplitude and breathing pattern) and assumed conventional PET scanner resolution.

The aim of this chapter is to understand the impact of respiratory motion correction in PET on detecting lesions in the lungs and liver as a function of lesion size, tracer uptake, breathing pattern and extent of displacement. Simulated dynamic PET data were gated based on amplitude and reconstructed with a conventional image reconstruction method and with respiratory motion compensation using motion information obtained from gated MRI data. Based on the results of chapter 4 the motion information was incorporated within the reconstruction (Li *et al.* 2006; Qiao *et al.* 2006; Lamare *et al.* 2007b) to successfully recover accurate quantitative values (Polycarpou *et al.* 2012). The improvement in detectability performance resulting from motion correction was quantified using ROC methodology with CHO. All evaluations were performed for both a simulated 6 *mm* FWHM scanner spatial resolution and for 3 *mm* FWHM spatial resolution in order to determine the relative importance of motion correction as new higher resolution scanners are developed.

5.2 Materials and Methods

5.2.1 Simulated dynamic PET data

For this chapter as a first step, a population of 3D phantoms was generated by adding spherical lesions to the phantom created in section 3.2.2 with size ranging from 6-12 *mm* diameter with a step of 2 *mm* and lesion to background ratio ranging from 3:1 to 6:1 with a step of 1. For each combination of size and lesion to background ratio, one phantom was created with lesions manually embedded in the lungs and liver such that for a given phantom lesions were added to identical positions. Considering four different lesion diameters and four different lesion to background ratios, sixteen phantoms were created in total. Finally, each phantom (i.e. 3D FDG distribution) was convolved with a Gaussian kernel of 6 or 3 *mm* FWHM in x, y, z to simulate a current and a better resolution scanner respectively[¶].

As a next step, real respiratory signals were combined with an MRI-derived motion model to simulate multiple dynamic thorax PET datasets acquired from a continuously moving subject as described in section 3.2.2. Real signals representative of three breathing patterns expected to be observed during a PET acquisition (Liu *et al.* 2009) were used with variable motion amplitude in order to provide a better understanding of the extent of motion in both normal and extreme breathing cases and define when motion is important or not. Furthermore, the dynamic PET data were gated based on amplitude. Dynamic datasets were created for each of the 16 phantoms, 2 simulated resolutions, 3 breathing types and 9 motion amplitudes (figure 5.1).

[¶] Based on that the scanner PSF is commonly approximately modelled by a 3D Gaussian distribution, the total effect of spatial resolution may be approximated by convolving the FDG distribution with a Gaussian kernel featuring the same FWHM as the scanner PSF. This approximation may also account for physical phenomena affecting spatial resolution and not included in the simulation (e.g. positron range, non-collinearity effect, detector geometry blurring).

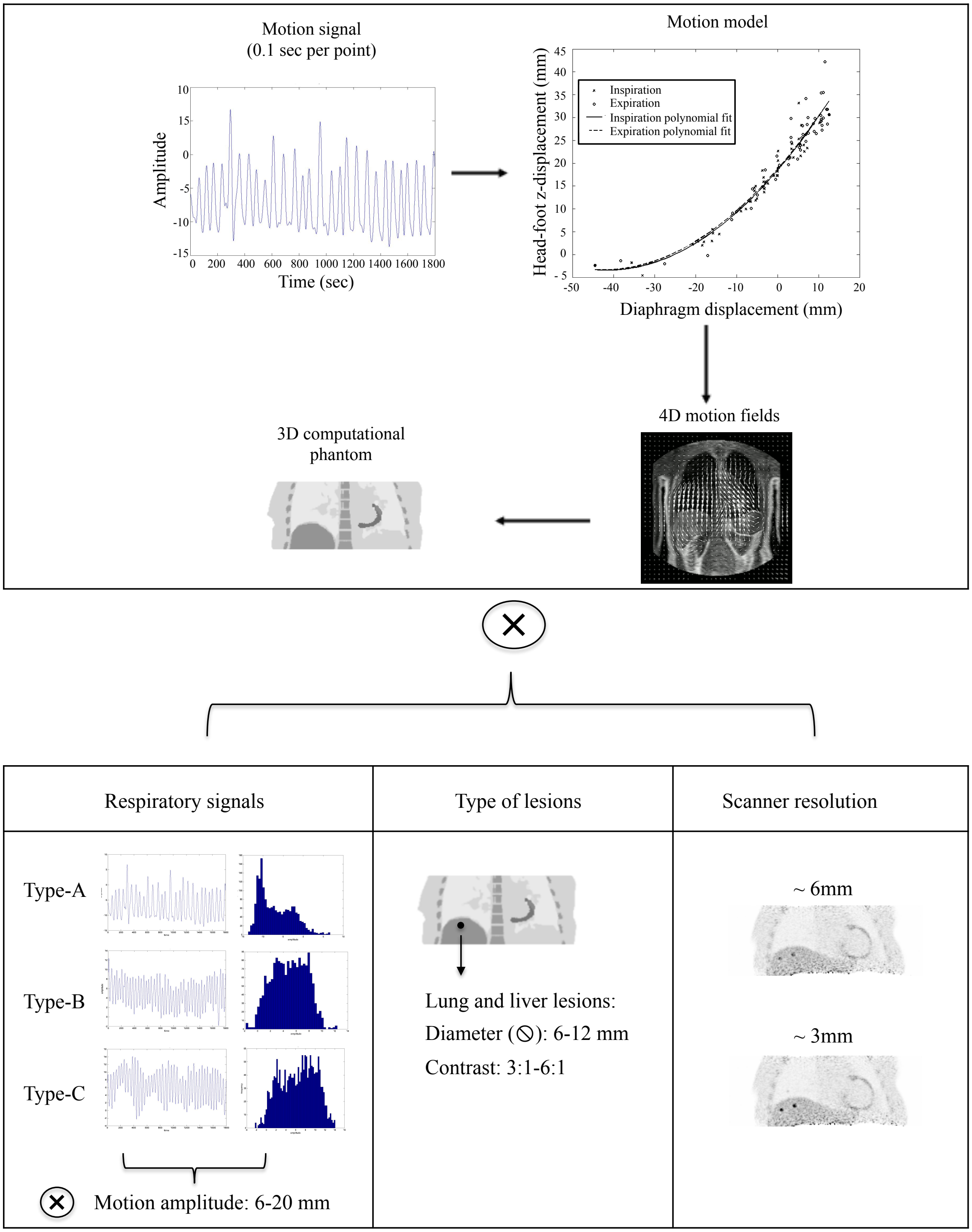


Figure 5.1: Dynamic PET datasets simulation using motion modelling.

Finally, each gated image was used as the input to an analytic simulation of a PET acquisition to create projection data as described in section 3.2.3. The projection data accounted for photon attenuation, scatter (approximately 33% of the total counts) but not for random coincidences. Statistical (Poisson) noise was added to the projection data such that the overall counts of the whole dataset correspond to 50 million unscattered coincidences (i.e. 3 to 5 min 3D PET clinical thorax acquisition). To add noise in the gated projection data, based on the corresponding amplitude based histogram (i.e. distribution of the motion amplitudes), each of the eight sinograms was weighted according to the dwell time at each displacement (i.e. histogram value). Each histogram value shows the duration that patient spends in the specific breathing position (i.e. how many times the specific amplitude was repeated in a respiratory signal). Sixty noise realisations were obtained for each investigated case.

5.2.2 Reconstruction algorithms

Each dataset was reconstructed without and with motion correction. Based on the results of chapter 4 we showed that incorporating the motion information within the reconstruction algorithm can successfully recover accurate quantitative values and therefore the motion corrected images were obtained via the forward / backward transformation operators in the reconstruction as described in equation (4.4). In all cases, reconstructions were performed for β_g equal to zero (i.e. OSEM) and with twenty three subsets and for two iterations.

The transformation operators describing the motion between the gated data (i.e. eight positions of the respiratory cycle) were estimated using voxel-wise deformation vectors, which were produced by co-registration of MRI gated datasets with a non-linear registration algorithm based on a composition of hierarchical local affine registrations recently developed by Buerger *et al.* (2011). The transpose motion matrix operation of MCIR was approximated by warping the image with the inverted motion fields which were numerically estimated from the same motion vectors (Crum *et al.* 2007). Finally, reference datasets (zero motion amplitude) were created by reconstructing only the reference gate with the same standard count statistics as for the entire gated dataset.

The selection of the reconstruction parameters have been optimized in chapter 4 which showed that for 23 subsets and for a low number of iterations (i.e. 2) as typically used in the clinic, MCIR outperforms RTA since it has the lowest possible MSE over all iterations and therefore it is expected to provide an overall good performance. Consequently, all iterative reconstructions were performed with 23 subsets and for 2 iterations. Each slice consisted of 250×250 pixels with size 2×2 mm each, and the entire volume consisted of 87 slices with 2 mm thickness.

5.2.3 Detectability performance

Model numerical observers, such as the CHO, can be used to predict human performance in clinically relevant visual tasks. As described in chapter 2 a CHO computes a test statistic λ_{CHO} over a given image f using the input image and statistics. In the CHO implementation of this study two classes of images were considered, lesion present and lesion absent.

Fourier Transformation was applied to transform a given image from the spatial domain to the frequency domain. The output of the Fourier Transformation was then multiplied with frequency selective channels with variable frequency bandwidth in order to extract certain frequencies from the image. The CHO used in this study consisted of five radially symmetric octave wide frequency channels with 3D processing (Frey *et al.* 2002). The frequency and width of the first channel was $1/64$ cycles per voxel as has been used in previous studies and which has shown good correlation with human observer performance. The CHO was applied in various adjacent planes with the central plane containing the centre of the lesion in order to approach the volumetric viewing scheme implemented by human observers in real clinical studies as discussed by Chan *et al.* (1997).

The CHO was applied by estimating first order (mean) and second order (covariance) statistics using a set of different noise realizations for lesion-present and lesion-absent reconstructed images (i.e. training images). 60 realisations for each combination of lesion size and lesion to background ratio were used for lesion present images, and the same number of images without lesions was created. The images were processed through a "masking" procedure in order to extract a region of interest (ROI) based on the known lesion location (i.e. everything outside the extracted ROI was set to zero). The ROI was spherical and was selected to be 36 *mm* diameter, three times larger than the size of the largest size tumour (i.e. 12 *mm* diameter) in order to account for sufficient statistical information regarding both the tumour and the background. The same mask was applied for both lesion present and lesion absent image and the CHO was applied to this 3D volume and not to the entire image volume.

Both the mean and covariance were calculated using the Microsoft Excel program. The mean for the lesion present images after transformation through the frequency selective channels over N noise realizations of images is obtained by:

$$\langle r_L \rangle = \langle U f_L \rangle = \frac{1}{N} \sum_{i=1}^{N} U f_{L,i} \tag{5.1}$$

where $f_{L,i}$ is a signal present image of the i^{th} noise realization and U is the channels matrix. The angular brackets denote an average over all images.

The mean for the lesion absent images after transformation through the frequency selective channels over N noise realizations of images is obtained by:

$$\langle r_B \rangle = \langle U f_B \rangle = \frac{1}{N} \sum_{i=1}^{N} U f_{B,i} \tag{5.2}$$

where $f_{B,i}$ is a signal absent image of the i^{th} noise realization.

In this implementation there is the same number of images with lesions and images without lesions. Under the assumption that the two classes have the same probability of occurrence the composite covariance matrix, K_U, is estimated by averaging the two covariance matrices for the lesion present and lesion absent data after transformation through the frequency selective channels as follows:

$$K_u = \frac{0.5}{N-1} \sum_{i=1}^{N} \left(r_{L,i} - \langle r_L \rangle \right) \left(r_{L,i} - \langle r_L \rangle \right)^T + \frac{0.5}{N-1} \sum_{i=1}^{N} \left(r_{B,i} - \langle r_B \rangle \right) \left(r_{B,i} - \langle r_B \rangle \right)^T \tag{5.3}$$

For a given image after transformation from the image domain to the frequency domain the Fourier transformed image passes through five frequency channels. The output of the channels is processed through the prewhitening operation (K_U^{-1}) and then multiplied by the ensemble difference of the means of the data, $\langle r_B \rangle = \langle U f_B \rangle$ and $\langle r_L \rangle = \langle U f_L \rangle$, a process that can be considered as matched filtering of the data to produce finally the test statistic (λ_{CHO}) (figure 5.2). The λ_{CHO} as described in chapter 2 (2.26) is given by:

$$\lambda_{CHO}(f) = \left(\langle r_L \rangle - \langle r_B \rangle \right)^T K_U^{-1} r \tag{5.4}$$

Each reconstructed image was processed through the frequency channels to estimate the test statistic λ of lesion present or lesion absent. The test statistic values were used as continuously distributed to create multiple receiver operating characteristic (ROC) curves (Swensson 1996; Metz 1986). Using the software Web-based calculator for ROC curves (http://www.rad.jhmi.edu/jeng/javarad/roc/JROCFITi.html) an ROC curve was fitted to the given data using a maximum-likelihood technique of a binormal model. This software is a freely available web version of LABROC package developed by Metz *et al.* (1988). It creates ROC curves and provides estimation of the area under each ROC curve (AUC) and its standard error using a non-parametric approximation. The significance of the differences between two different ROC curves was statistically tested using the the method of Hanley and McNeil (1982) implemented in the VassarStats software (http://vassarstats.net/).

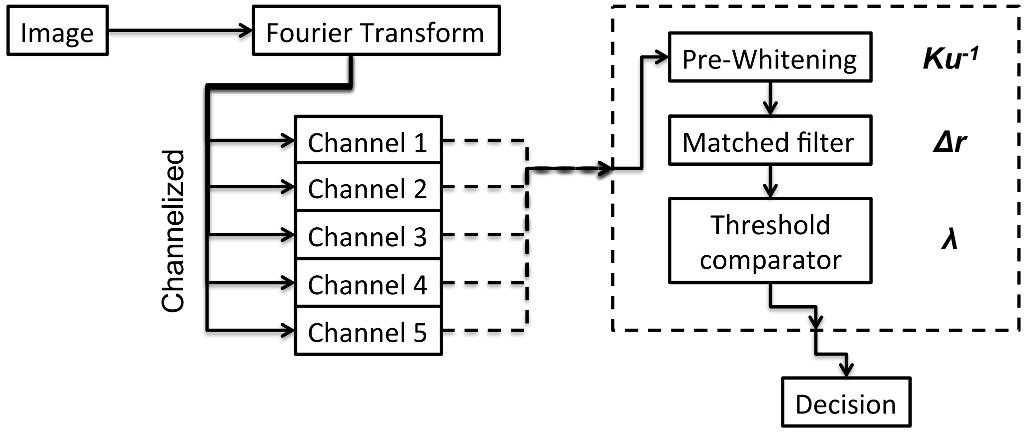


Figure 5.2: Channelized Hotelling observer procedure: The image is transformed from the image domain to the frequency domain. The Fourier transformed image passes through five frequency channels. The output of the channels is processed through prewhitening and matched filtering to produce the test statistic (λ). The value of the test statistic over a given image can be compared to a threshold to derive the decision of the CHO; if the value is below the threshold, the given image is classified as background images without lesion or if it is above the threshold, it is classified as having a lesion.

The AUC was calculated as the figure of merit of the detectability performance. An AUC value of 0.5 represents random detectability while 1 represents perfect detectability. An ROC curve was created for each case investigated; 2 regions of interest (i.e. lungs and liver) ×2 reconstruction methods (without or with motion correction) ×2 scanner resolution (i.e. 3 *mm* and 6 *mm* FWHM) ×4 size values ×4 tracer uptake values (i.e. 3:1-6:1) ×9 lesion motion amplitudes ×3 breathing types. In addition, a ROC curve was created for the reference image for 2 regions of interest (i.e. lungs and liver) ×4 size values ×4 tracer uptake values (i.e. 3:1-6:1). To derive the ratings and generate the ROC curves, the number of noise realisations used was again 60 realisations for each case investigated.

The detection performance using the CHO implemented in this study was compared with that obtained using human observers. Two observers participated in this study. The observers had to report the position of any suspected lesion and their confidence in its being present at that position on a scale from 1 to 6 (i.e. 1: definitely no lesion, 2: probably no lesion, 3: possibly no lesion, 4: possibly lesion, 5: probably lesion, 6: definitely lesion). The volumetric visualization software tool Amide (Loening and Gambhir 2001) was used. The observer evaluated reconstructed images with or without motion correction for the datasets of 20 *mm* motion under breathing of type-C. The results were analysed with the AUC calculated in a similar way as in the CHO study.

5.3 Results

5.3.1 Impact of respiratory motion on lesion detectability for three breathing types

For scanner resolutions of 6 *mm* and 3 *mm* FWHM figure 5.3 shows examples of coronal planes of lesion-present cases after reconstruction of the reference image and without correction of 20 *mm* motion for three breathing types. Each row in the figure shows different slices to emphasize a different lesion. As indicated by arrows the first and third rows show liver lesions whereas the second and fourth rows show lung lesions. The displayed lesions have a diameter of 8 *mm* with a lesion to background ratio of 3:1 and 6:1. Lesions are easily detectable in the reference images but hardly or not detectable in images with motion. In the reference images, for both 3 *mm* and 6 *mm* FWHM PET resolution scanners, the lesions are easily detectable. On the other hand, in the images with motion, for the 6 *mm* FWHM scanner resolution it is difficult or virtually impossible to detect the tumours whereas by increasing the resolution to 3 *mm* FWHM some lesions are recovered. However, lesions in images representing breathing type-A are more easily detectable compared to images representing breathing type-C especially for scanner resolution of 3 *mm* FWHM.

Figure 5.4 presents the effect of motion on lesion detection performance as measured by the AUC plotted as a function of motion amplitude. The relationship was investigated for three breathing patterns and for approximate scanner resolutions of 6 *mm* and 3 *mm* FWHM. The results are displayed for the liver lesions of 8 *mm* diameter and for low (i.e. 3:1) and high (i.e. 6:1) lesion to background ratio values. For all the aforementioned graphs a zero motion amplitude is also presented as the result of the reference reconstructed image. Considering a current (i.e. 6 *mm* FWHM) PET resolution, motion of 20 *mm* for small lesions (e.g. 8 *mm* diameter) and high (i.e. 6:1) lesion to background ratio may result in a degradation of detectability performance of up to 16% reduction in AUC under breathing of type-A, up to 28% for type-B and up to 30% for type-C. Similarly, for high PET resolution (i.e. 3 *mm* FWHM) motion of 20 *mm* for small lesions (e.g. 8 *mm* diameter) and low (i.e. 3:1) lesion to background ratio may result in a degradation of detectability performance of up to 21% under breathing of type-A, up to 29% for type-B and up to 37% for type-C.

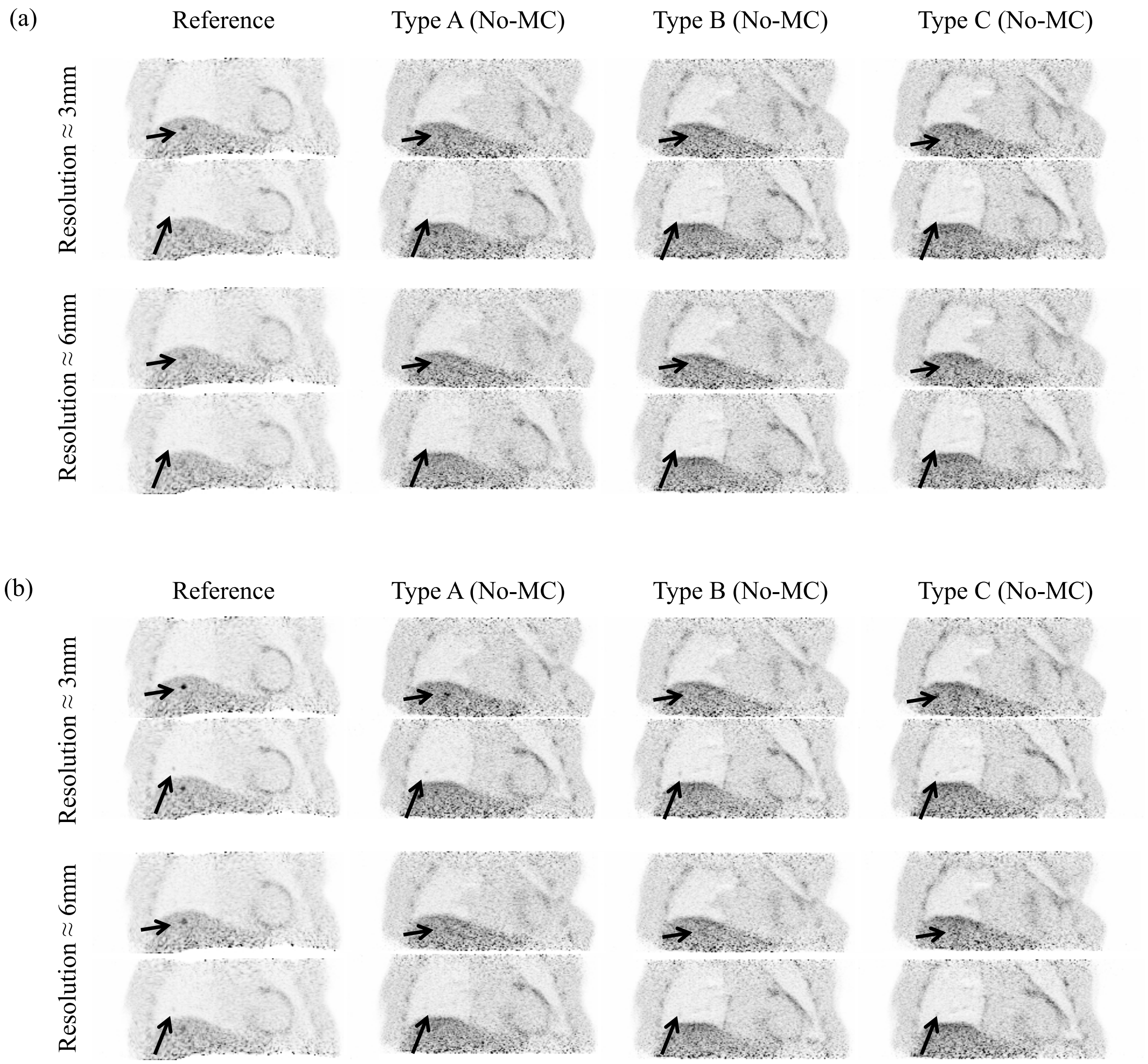


Figure 5.3: Visual demonstration of the effect of respiratory motion blurring on lesion detectability: Coronal planes after reconstruction of the reference image and prior to correction (No-MC) of 20 *mm* motion under breathing of type-A, type-B and type-C. The lesions, indicated with arrows, are characterised with a diameter of 8 *mm* with a lesion to background ratio of: (a) 3:1 and (b) 6:1. The results are displayed for PET scanner resolution of 6 *mm* and 3 *mm* FWHM. The lesions are easily detectable in the reference image but hardly or not detectable in images with motion.

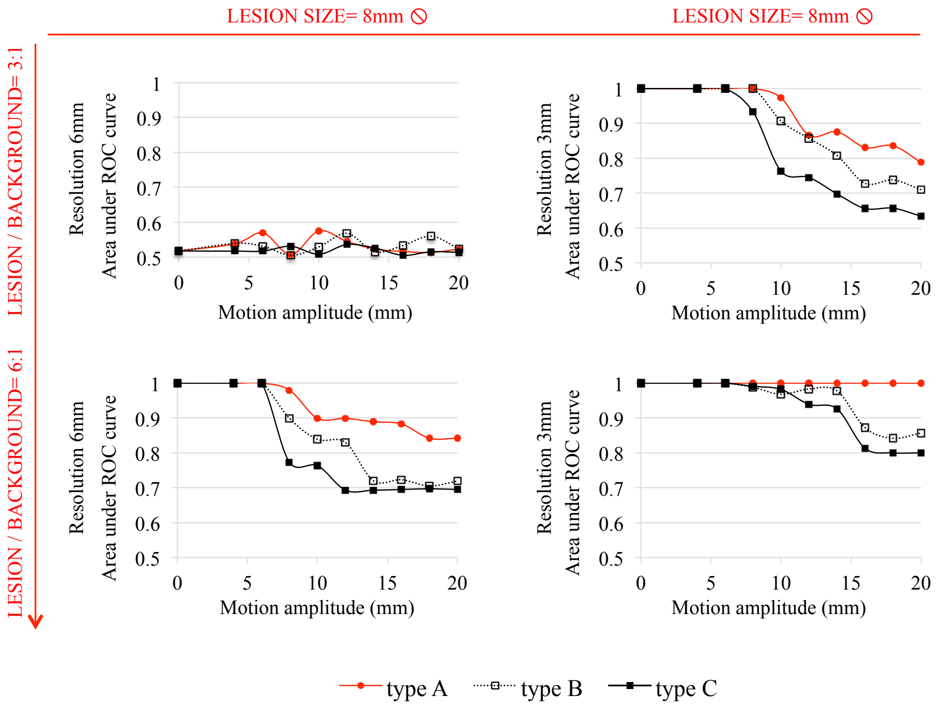


Figure 5.4: Effect of motion on area under the ROC curve (AUC) estimated using Channelized Hotelling observer as a function of the motion amplitude under breathing of type-A, type-B and type-C. The results are displayed for the liver lesions of 8 *mm* diameter and for 3:1 and 6:1 lesion to background ratio values. The impact of motion in lesion detection is higher for respiration with higher amplitudes and long-term variability (i.e. type-B and type-C) as compared with longer quiescent motion periods (i.e. type-A).

5.3.2 Impact of respiratory motion correction on lesion detectability as a function of lesion characteristics

The impact of motion correction on lesion detectability is presented for type-C which is the case expected to benefit more from motion correction as shown in section 5.3.1. For scanner resolutions of 6 *mm* and 3 *mm* FWHM figure 5.5 shows examples of coronal planes of lesion-present cases after reconstruction prior to motion correction and after motion correction of 20 *mm* motion under breathing of type-C. As indicated by arrows the first and third rows show liver lesions whereas the second and fourth rows show lung lesions. The displayed lesions have a diameter of 8 *mm* with a lesion to background ratio of 3:1 and 6:1. Lesions are easily detectable in the images after motion correction but hardly or not detectable in images with motion.

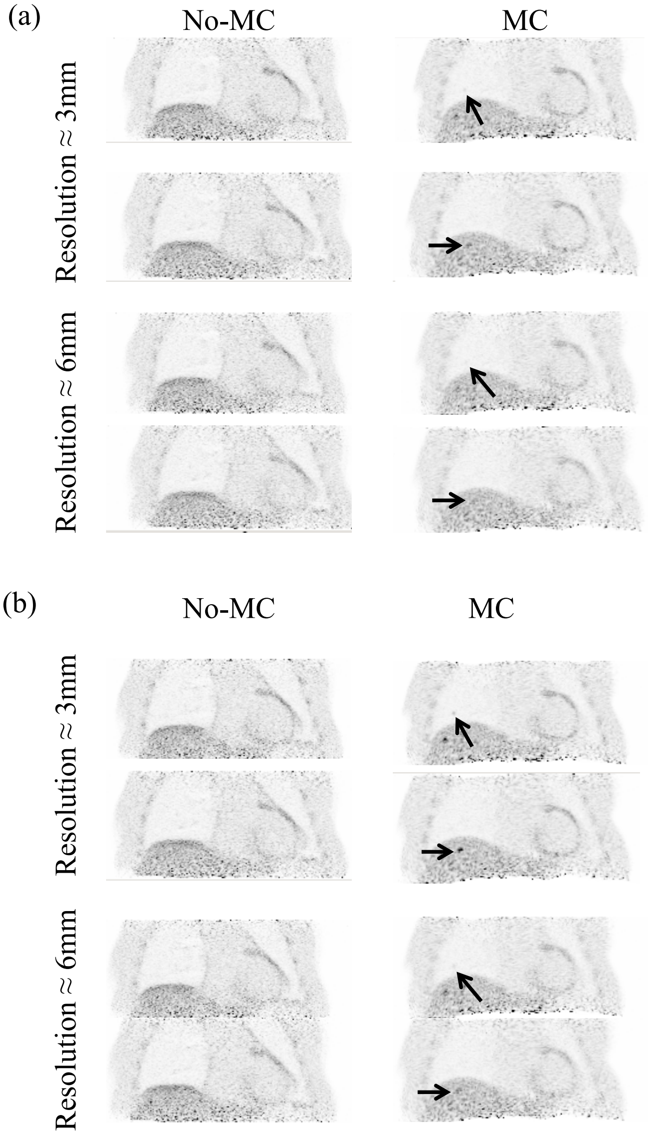


Figure 5.5: Visual demonstration of the effect of respiratory motion correction on lesion detectability: Coronal planes after reconstruction prior to motion correction (No-MC) and after motion correction (MC) of 20 *mm* motion under breathing of type-C. The lesions, indicated with arrows, are characterised with a diameter of 8 *mm* with a lesion to background ratio of: (a) 3:1 and (b) 6:1. The results are displayed for PET scanner resolution of 6 *mm* and 3 *mm* FWHM. The lesions are easily detectable after motion correction but hardly or not detectable in images with motion.

Figure 5.6 presents the performance in lesion detection before and after motion correction as a function of lesion size and tracer uptake. For each lesion size (i.e. 6 to 12 *mm* diameter) and lesion to background ratio (i.e. 3:1 to 6:1) the detection performance as measured by the AUC is plotted with respect to the motion amplitude without motion correction (No-MC) and with motion correction (MC) under breathing pattern of type-C. The results are displayed for scanner resolutions of 6 *mm* and 3 *mm* FWHM. Similar results have been observed for the lung lesions and only the results for the liver lesions are shown.

As the motion amplitude becomes larger than the scanner resolution motion correction significantly increases the detectability performance for small (less than 10 *mm* diameter) lesions (p-value less than 0.01). The highest gain in detection performance after motion correction is achieved for small (less than 10 *mm* diameter) or low tracer uptake (3:1) lesions. In particular, motion correction showed an up to 42% and 51% improvement in AUC for 6 *mm* and 3 *mm* FWHM PET resolution respectively. For example lesions of 8 *mm* and 3:1 lesion to background ratio for 3 *mm* FWHM scanner PET resolution and 20 *mm* motion are just detectable prior to motion correction (AUC= 0.63) whereas after correction they may be easily detected (AUC=0.95). On the other hand, the detection performance was only moderately changed with motion correction using a large lesion size since as the lesion size is increasing the susceptibility to motion blurring is decreased. In particular, lesions with diameter larger than 12 *mm* have negligible (i.e equal to 9% for 3 *mm* FWHM resolution and 15% for 6 *mm* FWHM resolution) reduction in AUC with motion even for low tracer uptake and high motion amplitudes (i.e. 20 *mm*) and were perfectly detected even with motion, indicating that in this case there is no apparent advantage on detectability with motion correction even for respiration characterised by randomly varying amplitude (i.e. type-C).

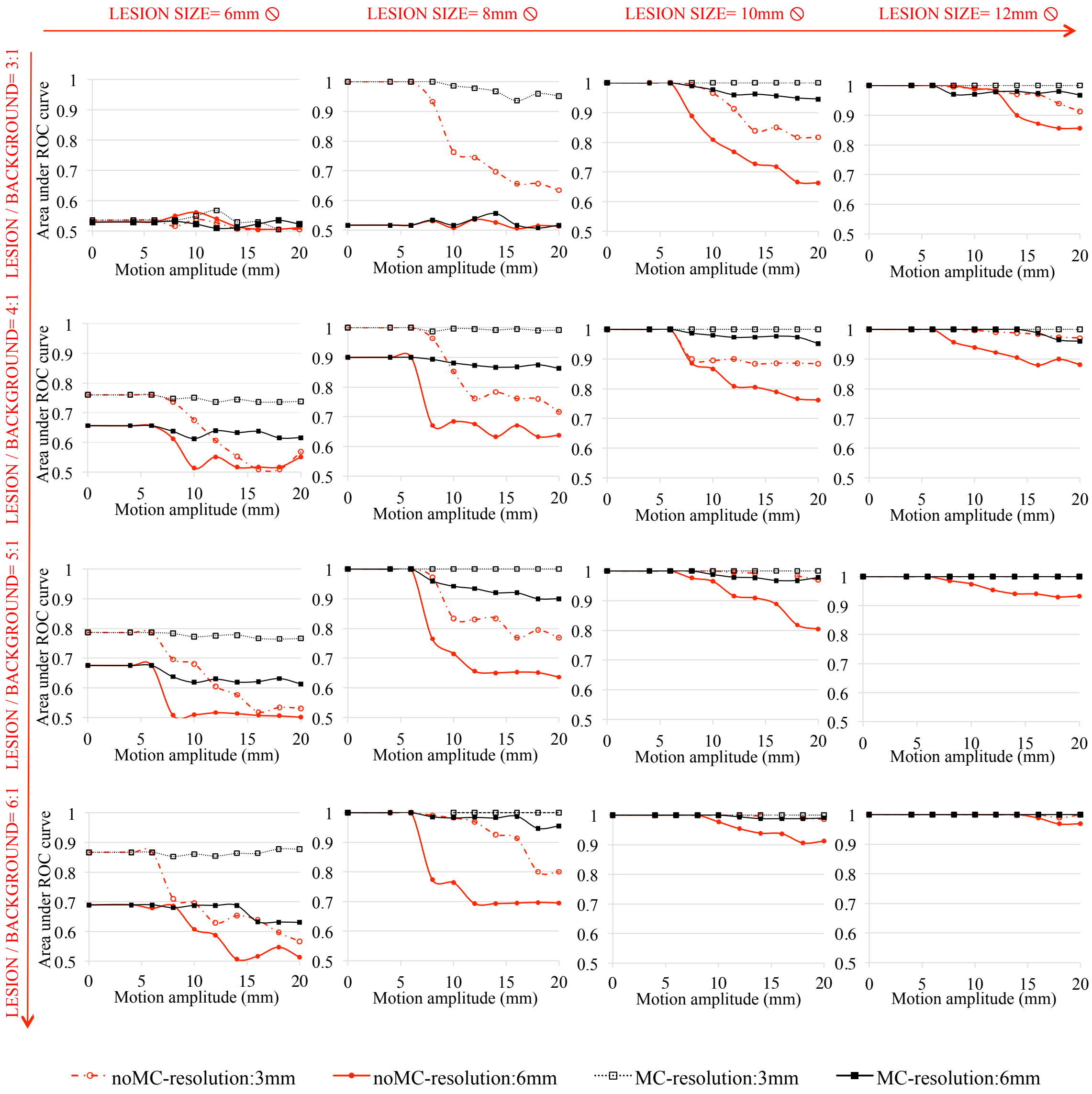


Figure 5.6: Area under the ROC curve (AUC) estimated using Channelized Hotelling observer as a function of the motion amplitude after reconstruction prior to motion correction (No-MC) and after motion correction (MC) under breathing of type-C. The results are displayed for PET scanner resolution of 6 *mm* and 3 *mm* FWHM and for liver lesions. Motion correction significantly increases the detectability performance for small lesions (less than 10 *mm* diameter).

5.3.3 Improvement in lesion detectability with respiratory motion correction and scanner resolution

Figure 5.7 shows the lesion detection performance as measured by the AUC with the standard errors for a lesion motion amplitude of 20 *mm* and breathing pattern of type-C. The results are given for scanner resolutions of 6 *mm* and 3 *mm* FWHM without motion correction (No-MC) and with motion correction (MC). The p-values of the statistical significance test of the differences between prior and after motion correction are also given. For a p-value less than 0.01 the result is considered statistically significant. After motion correction the 3 *mm* FWHM resolution outperforms the 6 *mm* FWHM resolution in terms of lesion detection performance. For example, liver lesions of 8 *mm* diameter with low (i.e. 3:1) lesion to background ratio are not detectable with a 6 *mm* FWHM resolution scanner without or with motion correction (AUC=0.51-0.52) whereas with a 3 *mm* FWHM resolution scanner they are hardly detectable (AUC=0.63) without correction but almost perfectly detectable if motion is corrected (AUC=0.95). The combination of increased scanner resolution and motion correction results in substantial improvement in detectability. For the same cases as in figure 5.7 the plotted ROC curves are shown in figure 5.8.

Figure 5.9 and figure 5.10 show the lesion detection performance as measured by the AUC for two human observers. The evaluated datasets as in figure 5.7 include a lesion motion amplitude of 20 *mm* and breathing pattern of type-C. The results are given for scanner resolutions of 6 *mm* and 3 *mm* FWHM without motion correction (No-MC) and with motion correction (MC).

Channelized Hotelling observer
Type C – motion amplitude 20 mm

LESION / BACKGROUND= 3:1

LESION SIZE= 6mm ⌀

	No-MC	MC
Resolution ≈ 6mm	0.5 ± 0.07	0.5 ± 0.07
	p = 0.48	
Resolution ≈ 3mm	0.5 ± 0.07	0.5 ± 0.07
	p = 0.48	

LESION SIZE= 8mm ⌀

	No-MC	MC
Resolution ≈ 6mm	0.51 ± 0.07	0.52 ± 0.07
	p = 0.43	
Resolution ≈ 3mm	0.63 ± 0.07	0.95 ± 0.03
	p < 0.005	

LESION SIZE= 10mm ⌀

	No-MC	MC
Resolution ≈ 6mm	0.66 ± 0.06	0.94 ± 0.03
	p < 0.005	
Resolution ≈ 3mm	0.82 ± 0.05	1
	p < 0.005	

LESION SIZE= 12mm ⌀

	No-MC	MC
Resolution ≈ 6mm	0.85 ± 0.05	0.96 ± 0.03
	p = 0.03	
Resolution ≈ 3mm	0.91 ± 0.03	1
	p = 0.01	

LESION / BACKGROUND= 4:1

LESION SIZE= 6mm ⌀

	No-MC	MC
Resolution ≈ 6mm	0.55 ± 0.07	0.62 ± 0.07
	p = 0.22	
Resolution ≈ 3mm	0.55 ± 0.07	0.74 ± 0.06
	p = 0.03	

LESION SIZE= 8mm ⌀

	No-MC	MC
Resolution ≈ 6mm	0.64 ± 0.07	0.86 ± 0.05
	p < 0.005	
Resolution ≈ 3mm	0.72 ± 0.06	0.99 ± 0.04
	p < 0.005	

LESION SIZE= 10mm ⌀

	No-MC	MC
Resolution ≈ 6mm	0.77 ± 0.06	0.95 ± 0.03
	p < 0.005	
Resolution ≈ 3mm	0.88 ± 0.04	1
	p < 0.005	

LESION SIZE= 12mm ⌀

	No-MC	MC
Resolution ≈ 6mm	0.88 ± 0.04	1
	p < 0.005	
Resolution ≈ 3mm	0.97 ± 0.02	1
	p = 0.09	

LESION / BACKGROUND= 5:1

LESION SIZE= 6mm ⌀

	No-MC	MC
Resolution ≈ 6mm	0.5 ± 0.07	0.61 ± 0.07
	p = 0.14	
Resolution ≈ 3mm	0.53 ± 0.07	0.77 ± 0.06
	p < 0.005	

LESION SIZE= 8mm ⌀

	No-MC	MC
Resolution ≈ 6mm	0.64 ± 0.07	0.90 ± 0.04
	p < 0.005	
Resolution ≈ 3mm	0.77 ± 0.05	1
	p < 0.005	

LESION SIZE= 10mm ⌀

	No-MC	MC
Resolution ≈ 6mm	0.80 ± 0.05	0.98 ± 0.01
	p < 0.005	
Resolution ≈ 3mm	0.97 ± 0.01	1
	p = 0.09	

LESION SIZE= 12mm ⌀

	No-MC	MC
Resolution ≈ 6mm	0.93 ± 0.03	1
	p = 0.02	
Resolution ≈ 3mm	1	1
	p = NaN	

LESION / BACKGROUND= 6:1

LESION SIZE= 6mm ⌀

	No-MC	MC
Resolution ≈ 6mm	0.51 ± 0.07	0.63 ± 0.07
	p = 0.11	
Resolution ≈ 3mm	0.57 ± 0.06	0.88 ± 0.05
	p < 0.005	

LESION SIZE= 8mm ⌀

	No-MC	MC
Resolution ≈ 6mm	0.69 ± 0.07	0.96 ± 0.02
	p < 0.005	
Resolution ≈ 3mm	0.8 ± 0.06	1
	p < 0.005	

LESION SIZE= 10mm ⌀

	No-MC	MC
Resolution ≈ 6mm	0.91 ± 0.03	1
	p = 0.01	
Resolution ≈ 3mm	1	1
	p = NaN	

LESION SIZE= 12mm ⌀

	No-MC	MC
Resolution ≈ 6mm	0.93 ± 0.02	1
	p = 0.02	
Resolution ≈ 3mm	1	1
	p = NaN	

Figure 5.7: Area under the ROC curve (AUC) with standard errors estimated using Channelized Hotelling observer for liver lesions with motion amplitude of 20 *mm* under breathing of type-C after reconstruction prior to motion correction (No-MC) and after motion correction (MC) for PET scanner resolutions of 6 *mm* and 3 *mm* FWHM. The p-values of the statistical significance test of the differences between prior and after motion correction are also given. After motion correction the 3 *mm* FWHM resolution has superior detection performance than the 6 *mm* FWHM resolution. The benefit of increasing the scanner resolution is small unless motion correction is applied.

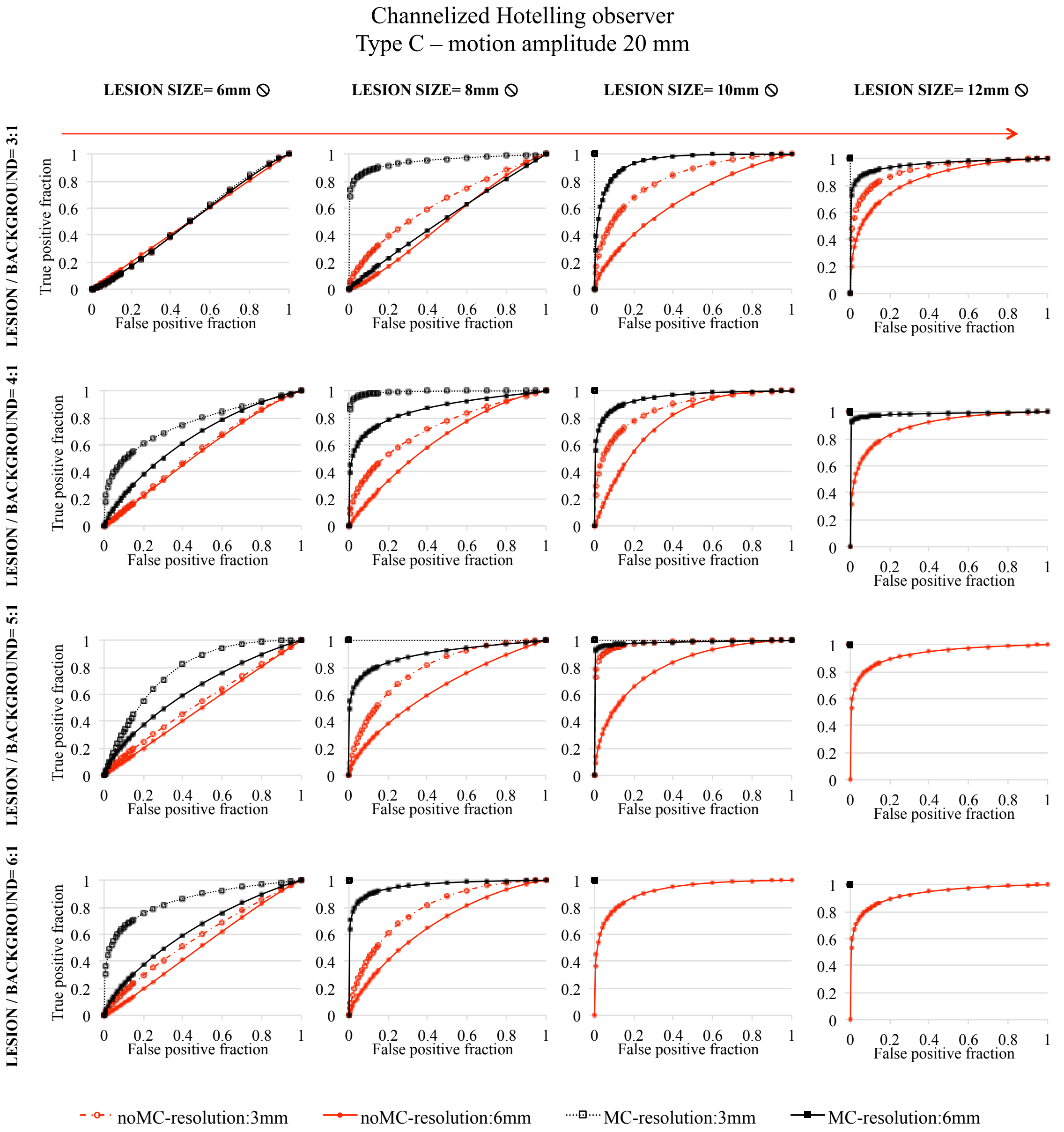


Figure 5.8: ROC curves estimated using Channelized Hotelling observer for liver lesions with motion amplitude of 20 *mm* under breathing of type-C after reconstruction prior to motion correction (No-MC) and after motion correction (MC) for PET scanner resolutions of 6 *mm* and 3 *mm* FWHM. After motion correction the 3 *mm* FWHM resolution has superior detection performance than the 6 *mm* FWHM resolution. The benefit of increasing the scanner resolution is small unless motion correction is applied.

Human observer 1
Type C – motion amplitude 20 mm

		LESION SIZE= 6mm ⦸		LESION SIZE= 8mm ⦸		LESION SIZE= 10mm ⦸		LESION SIZE= 12mm ⦸	
		No-MC	**MC**	**No-MC**	**MC**	**No-MC**	**MC**	**No-MC**	**MC**
LESION / BACKGROUND= 3:1	**Resolution ≈ 6mm**	0.5	0.54	0.50	0.58	0.63	0.75	0.67	0.96
	Resolution ≈ 3mm	0.54	0.66	0.50	0.96	0.75	1	0.79	1
LESION / BACKGROUND= 4:1	**Resolution ≈ 6mm**	0.50	0.58	0.58	0.83	0.87	1	0.91	1
	Resolution ≈ 3mm	0.50	0.75	0.58	1	0.91	1	1	1
LESION / BACKGROUND= 5:1	**Resolution ≈ 6mm**	0.54	0.75	0.50	1	0.87	1	1	1
	Resolution ≈ 3mm	0.54	1	0.75	1	0.83	1	0.96	1
LESION / BACKGROUND= 6:1	**Resolution ≈ 6mm**	0.50	0.79	0.67	1	0.95	1	1	1
	Resolution ≈ 3mm	0.54	1	0.83	1	0.96	1	0.96	1

Figure 5.9: Area under the ROC curve (AUC) obtained for human observer A. The results are presented for liver lesions with motion amplitude of 20 *mm* under breathing of type-C after reconstruction prior to motion correction (No-MC) and after motion correction (MC) for PET scanner resolutions of 6 *mm* and 3 *mm* FWHM.

Human observer 2
Type C – motion amplitude 20 mm

		LESION SIZE= 6mm ⌀		LESION SIZE= 8mm ⌀		LESION SIZE= 10mm ⌀		LESION SIZE= 12mm ⌀	
		No-MC	MC	No-MC	MC	No-MC	MC	No-MC	MC
LESION / BACKGROUND= 3:1	Resolution ≈ 6mm	0.50	0.54	0.50	0.50	0.67	0.92	0.63	0.96
	Resolution ≈ 3mm	0.54	0.50	0.50	0.95	0.75	1	0.83	1
LESION / BACKGROUND= 4:1	Resolution ≈ 6mm	0.50	0.54	0.58	0.83	0.79	1	0.83	1
	Resolution ≈ 3mm	0.50	0.67	0.58	1	0.83	1	0.92	1
LESION / BACKGROUND= 5:1	Resolution ≈ 6mm	0.50	0.62	0.50	0.90	0.79	1	0.92	1
	Resolution ≈ 3mm	0.50	0.92	0.67	1	0.87	1	0.96	1
LESION / BACKGROUND= 6:1	Resolution ≈ 6mm	0.50	0.75	0.67	0.96	0.92	1	0.92	1
	Resolution ≈ 3mm	0.58	1	0.75	1	0.96	1	1	1

Figure 5.10: Area under the ROC curve (AUC) obtained for human observer B. The results are presented for liver lesions with motion amplitude of 20 *mm* under breathing of type-C after reconstruction prior to motion correction (No-MC) and after motion correction (MC) for PET scanner resolutions of 6 *mm* and 3 *mm* FWHM.

5.4 Discussion

5.4.1 Impact of respiratory motion on lesion detectability for three breathing types

Visual interpretation of figure 5.3 shows blurring and inferior lesion contrast in PET reconstructed images with motion compared to the motion-free reference images. Lesions in the lungs and liver may be hardly detectable in the images with motion whereas they are clearly detectable in the reference image. However, the detectability performance differs among the three breathing pattern types; motion blurring is more important for acquisitions under respiration with long-term amplitude variability (i.e. type-B and type-C) as compared with longer quiescent motion periods (i.e. type-A). The edge artefacts observed in the images with motion compared to the motion-free images are a consequence of emission and attenuation map mismatches (figure 5.3) that are particularly severe at the boundaries between low and high attenuating media such as in the cases of liver and lung. These edge artefacts are not visible after applying motion correction as illustrated in figure 5.5.

Quantitative analysis of lesion detection performance as measured by the AUC presented in figure 5.4 confirms the visual observation and shows the extent of the impact of motion blurring in detectability among the different breathing types. For any breathing type, respiratory motion degrades the detectability performance, with the greatest impact of motion blurring observed for type-C whereas the least is observed for type-A; for 20 mm motion amplitude the degradation of AUC compared to the zero motion amplitude equals 16% for the type-A, 28% for the type-B and 30% for the type-C. The differences among the three breathing pattern types are expected since respiration becomes progressively more irregular from type-A to type-C. The patients in type-A tend to breathe to a similar end-expiration location and spend a higher amount of time near that location while patients with type B respiration have similar inspiration and expiration positions but spend similar amounts of time in inspiration and expiration locations. On the other hand, patients with type-C respiration cycle breathe irregularly with long term variation of the breathing cycles length and amplitudes and therefore type-C is expected to have the highest motion content in the acquired data with type A the lowest. However, it must be noted that approximately 60% of patients are expected to breathe with a type-A pattern whereas only 20% of patients are expected to breathe with type-C.

5.4.2 Impact of respiratory motion correction on lesion detectability as a function of lesion characteristics

Comparing the images in figure 5.5 and the various plots in figure 5.6 we are able to see the difference resulting from respiratory motion correction as lesion size and tracer uptake increase. Motion correction significantly increases the detectability performance with the greatest improvement achieved for the smaller lesions and lesions with low metabolic activity as long as the detectability was not limited by the scanner resolution. Large diameter or high tracer uptake lesions are less subject to motion and more easily detectable yielding little improvement by using motion correction. This is expected due to the fact that motion blurring spreads out and decreases the apparent tracer uptake (i.e. lesion contrast) (Liu *et al.* 2009) and therefore smaller lesions which are more susceptible to the PVE. Consequently the required tracer uptake in order to detect the lesions varied with the motion amplitude and lesion size.

Another feature of the results in figure 5.6 is that motion correction does not fully recover the best possible outcome of a scan without motion and the detection performance often strongly depends upon the motion amplitude. This might be attributed to the limited number of gates used as each gate is representative of a specific moment of the motion cycle, the different gates have different motion content especially under breathing with high variability between the different breathing cycles. Consequently, a higher number of gates, or real time motion correction, may be required to limit the content of motion in each gate and capture extreme phases such as a deep inspiration or expiration. In the current study 8 gates have been used as is common in clinical practice. However, the optimal number of gates is expected to be related with the scanner resolution and motion amplitude (Dawood *et al.* 2009) but this is beyond the scope of this study. Another explanation is the possible inaccuracies of the estimated motion fields since different motion fields were used to create the simulations and to correct for motion. Simulations were derived using motion fields for each point in real-time respiratory curves (i.e. 1800 points) whereas motion fields for correction were derived from registration of gated data (i.e. 8 gates). Therefore the results are prone to any errors in the motion estimation (e.g. interpolation) introduced by the registration algorithm.

5.4.3 Improvement in lesion detectability with respiratory motion correction and scanner resolution

The results indicate that in order to take advantage of the potential increased spatial resolution of future PET scanners, it becomes important to account for respiratory motion assuming that acquisition protocols remain largely similar to those currently used and investigated in this study. In particular, considering a small liver tumour for the current standard of 6 *mm* FWHM PET resolution (i.e 6 *mm* lesion diameter) even if it has high tracer uptake (i.e. 6:1) it can not be detected due to the inherently limited resolution (AUC=0.5). However, although it is likely to be recovered with a next generation scanner of 3 *mm* FWHM (i.e. AUC=0.88 for zero motion amplitude) the motion effect limits this benefit (i.e. AUC=0.57 in the non motion corrected images). Similarly, considering a very low tracer uptake (3:1) of an 8 *mm* diameter liver tumour, it is not detectable with a 6 *mm* FWHM resolution (AUC=0.51) whereas although it is likely to be recovered with a next generation scanner of 3 *mm* FWHM resolution (i.e. AUC=1 for zero motion amplitude) the motion effect limits this benefit (i.e. AUC=0.63 in the non motion corrected images). Consequently, improving spatial resolution allows not only the detection of smaller tumours but also the detection of tumours with low metabolic activity and contrast (i.e. metastases) - the detection of small and lower uptake lesions may prove important for early detection and staging.

To the best of our knowledge within the current literature there are only limited studies showing the impact of motion correction for different parameters. Marache-Francisco *et al.* (2010); Park *et al.* (2012) showed that for conventional PET scanner resolution (i.e. 6 *mm* FWHM) motion correction can improve detectability but both studies did not study the effect of the lesion characteristics or extent of motion or resolution.

Although clinically the respiratory cycle can be divided into a number of gates according to the phase or amplitude of respiration, in this investigation the gating was based only on amplitude. As it was described in chapter 2 previous studies showed that phase and amplitude based gating may result in similar performance under regular breathing while for breathing when the phase and amplitude of the motion change over time the amplitude based gating may outperform phase based gating. This is due to that with the phase-based technique the acquired data are stored to gates based on their time phase regardless of the depth of breathing and therefore data with the same phase range but with different amplitudes may be combined together in the same gate.

Consequently, the phase gating may not reflect the motion amplitude for all respiratory cycles while the amplitude-based gating scheme splits the respiratory cycle with respect to the amplitude of the signal and only the inspiration phases with the same breathing magnitude will be added together.

Similar trends were observed on the detectability performance of the two human observers (figure 5.9 and figure 5.10). In most of the cases the AUC estimated with the CHO (figure 5.7) has been shown to be in good agreement with human observers (figure 5.9 and figure 5.10) except from some cases where the CHO detectability performance was higher than the human observer performance. In particular, without motion correction for a 3 *mm* FWHM resolution for a 10 *mm* diameter liver tumour of 3:1 lesion to background ratio both human observers achieve an AUC equal to 0.75 and the CHO achieves an AUC equal to 0.82 while for a tumour of 5:1 lesion to background ratio the human observers achieve an AUC equal to 0.83 and 0.87 respectively and the CHO achieves an AUC equal to 0.97. The increased detectability performance achieved using CHO was also observed in previous studies (Gifford *et al.* 2000) and has been attributed to the fact that in the CHO a known signal is superimposed on a known background and therefore it is likely that the contrast required for human observer is higher than CHO. Although the numerical observer was generally in good agreement with human observers, the observers were not trained and therefore the results may be artificial. As a future step a broader population of expert human observers could be used to evaluate the results.

There are a number of limitations to the current study that should be considered. Firstly, all results are shown at a low number of iterations (i.e. 2) of image reconstruction with a standard OSEM algorithm as used in clinical practice. Using other reconstruction parameters (i.e. number of iterations and subsets) (Morey and

Kadrmas 2013) and advanced techniques such as time of flight and point spread function modelling in image reconstruction (Alessio *et al.* 2010; Tong *et al.* 2010) or smoothing filters and incorporating regularisation (Wang and Qi 2012; Tsoumpas *et al.* 2013b) may provide better image quality and could affect detectability performance therefore limiting any possible positive false findings that might occur due to the noise observed in OSEM reconstructed images. However, this is beyond the scope of the current study. This study was based on a simulation for a 3 min acquisition which is the current protocol in clinical practise. However, for a longer study (e.g. 60 min) motion variability is more likely to occur. This is probably the case when you do dynamic imaging of one bed position. However, in cases where you have multiple bed positions multiple times (Karakatsanis *et al.* 2013) for acquisitions with continuous bed motion that will last longer than 5 min motion and variability of breathing will be an issue to be solved.

The motion model used in this thesis has certain limitations and is in general a simplification of the actual real motion observed in patients. Real motion is characterized with intracycle and intercycle variability. Intracycle variability characterizes the differences observed between the trajectories during inspiration and the trajectories during expiration which is addressed in the motion model used in this investigation. Intercycle variability is divided in two parts. The first part is the variation in the amplitudes that are observed during breathing which is also addressed in the motion model while the second part is the variation in relationship between the surrogate value and the actual motion (i.e. different motion may correspond to the same surrogate value), something that can not be addressed in the specific implementation which has a single motion field for a single surrogate value.

Another limitation is the spherical shape of the lesions and the homogeneity of the background uptake of the simulations simplifying the detection task especially in liver. In addition, bulk motion during acquisition (i.e. unpredictable random patient motion) was not taken into account (Kolbitsch *et al.* 2012). Finally, the numerical observer was applied in the centre of the lesion eliminating the possibility of false positives in surrounding regions.

5.5 Conclusions

This chapter indicates that respiratory motion correction has a high impact on lesion detectability. Patients with large quiescent motion periods are less subject to the effects of motion than patients with regular breathing cycles or with long-term variability in respiratory cycles and higher motion amplitudes. Furthermore, lesion detection performance depends on the lesion size, position and tracer uptake. Motion blurring particularly influences small size (less than 10 *mm*) or low metabolic activity lesions highlighting that respiratory motion correction could play a key role in detecting marginally detectable lesions. Finally, this chapter underlines the importance of correcting for respiratory motion in order to benefit from the potential improved spatial resolution of future PET scanners.

Chapter 6

General summary and conclusions

Despite the advancement in the PET reconstruction methods and scanner technology, patient respiratory motion is an inevitable factor during acquisition that degrades image quality and resolution. In addition, following the continuous improvements in PET spatial resolution, subject motion is now an important source of resolution degradation, thus motion correction is a major challenge in PET imaging research. The aim of this thesis was to assess respiratory motion correction in terms of its quantitative accuracy and detectability performance in order to determine its potential for improved and early cancer diagnosis.

Two of the most common motion correction approaches that make use of all gated measured data, are RTA and MCIR. The two approaches differ in how motion correction is applied to the gated data. With RTA it is applied after reconstruction of the gated data and with MCIR it is applied during reconstruction. Both RTA and MCIR methods require knowledge of the subject motion and a key aspect of motion correction is the accurate measurement of internal organ motion. The recent development of whole body PET-MRI scanners might provide a potential solution for motion correction since internal organ motion could be measured concurrently with PET using MRI. In this thesis based on the use of 4D simulated PET data using real MRI data, respiratory motion correction is investigated with MRI-derived motion fields as could be acquired using simultaneous PET-MRI.

One of the aims of this thesis was to understand the behaviour of applying motion correction during or after reconstruction of the gated data. A statistical analysis was performed to characterise the bias and variance of the two methods. The results in chapter 4 have shown that although both approaches improve the image quality, they have different convergence and statistical performance. The extent of the different behaviour depends on the figure of merit used.

In terms of quantification accuracy, incorporating the motion correction of the gated PET data within the reconstruction process has been found to be successful with recovery of more than 90% of the true values of all regions even for lesions of small diameter (i.e. 10 mm). On the other hand, applying motion correction after reconstruction of the gated data using the RTA method causes resolution degradation as a result of high levels of overestimation and underestimation in cold and hot small regions respectively. A comprehensive evaluation of datasets with or without the occurrence of motion has shown the non-linearity of the OSEM at low counts statistics and interpolation errors as the main sources of the bias observed in RTA.

In terms of noise characteristics, the errors introduced during RTA motion correction have found to act as a low pass filter yielding a smoothed final motion corrected image with low and stabilized noise to a constant value after a large number of iterations. In contrast, MCIR noise becomes progressively greater with the number of iterations.

A common way to treat noise in clinical practice is early stopping (i.e. small number of iterations) or post-reconstruction filtering or a combination of both. In this thesis post-reconstruction filtered RTA and MCIR have been compared. Post-reconstruction filtered MCIR outperforms RTA (with the same level of smoothing) in terms of both quantification accuracy and noise properties for any iteration number even for the small lesions (10 *mm* diameter) which are affected more by smoothing. For a combination of post-reconstruction filtering and early- stopping (e.g. 2 iterations) MCIR produces similar results to unfiltered RTA. This demonstrates that in the current clinical protocols for ^{18}F-FDG PET acquisitions, motion correction with RTA is expected to be a valid method.

Although smoothing improves the noise properties, it limits the spatial resolution and quantification accuracy. The performance of post-reconstruction filtered MCIR depends on the smoothing level and the reconstruction algorithm parameters chosen (e.g. number of iterations). Different smoothing levels and number of iterations may be required based on the desired quantification accuracy or noise level and diagnostic task. This indicates that to fully exploit the high quantification accuracy of MCIR and RTA, methods to minimize noise while retaining the quantification accuracy may be superior to post-reconstruction filtering. In parallel with this work a study showed that introducing appropriate regularization into MCIR reconstruction may be the optimal choice for reconstruction with respect to quantification accuracy (Chun and Fessler 2013). As a further step chapter 4 provides a demonstration that introducing appropriate regularization to MCIR may be the optimal choice with respect to quantification accuracy. The results demonstrate that including median-root prior regularization within the reconstruction of RTA and MCIR methods can improve noise properties and also maintain the quantification accuracy. The performance of regularization depends on the value of the penalization factor and it has been shown to affect the RTA and MCIR in a different way (i.e. a stronger penalization is required for MCIR than RTA to achieve similar noise level).

Consequently, MCIR provides superior performance overall to RTA when noise is minimized. However, in current clinical applications where quantification is not the main objective, RTA can be a practical and simple method to correct for motion especially when tumour quantification is not the main task.

As a next step in this thesis, the impact of motion correction on lesion detectability in the lungs and in the liver was investigated using the MCIR method with early stopping, which was shown to provide optimal image quality. Respiratory motion correction leads to visually enhanced reconstructed PET images depicting better detectability performance (chapter 5). The results demonstrate that lesion displacement due to respiration larger than twice the scanner resolution has a high impact on lesion detectability. Lesions in the lungs and liver may be hardly detectable without motion correction. However, patients with large quiescent motion periods are less subject to the effects of motion than patients with regular breathing cycles or with long-term variability in respiratory cycles and higher motion amplitudes. The differences in detectability performance observed among the different breathing patterns indicates that the effect of motion depends on the pattern and extent of breathing and therefore if a motion model is used to estimate the motion information for motion correction, it must be either patient specific or a generic population based model able to model the breathing variability among the patients (McClelland *et al.* 2013).

Furthermore, chapter 5 shows that lesion detection performance depends on the lesion size, position and tracer uptake. Motion blurring particularly influences small size (less than 10 *mm*) or low metabolic activity lesions (3:1 lesion to background ratio) highlighting that respiratory motion correction could play a key role in detecting currently undetectable lesions. Motion correction significantly increases the detectability performance especially for small lesions and lesions with low metabolic activity on which even for a small amount of motion (i.e. 6 *mm*) motion correction seems essential for their detection. Large diameter or high tracer uptake lesions are less subject to motion and more easily detectable yielding little improvement by using motion correction and therefore the required tracer uptake in order to detect the lesions varied with the motion amplitude and lesion size.

Finally, in order to take advantage of the potential increased spatial resolution of future PET scanners, it becomes important to account for respiratory motion assuming that acquisition protocols remain largely similar or longer to those currently used and investigated in this study. Improving spatial resolution allows not only the detection of smaller tumours but also the detection of tumours with low metabolic activity and contrast (i.e. metastases) - the detection of small and lower uptake lesions may prove important for early detection and staging.

This thesis is focused on one radiotracer (FDG) because it is currently the most used. However, the methodology is applicable to different types of body motion as well as other imaging modalities such as single photon emission computed tomography (SPECT) and other radiotracers.

In this thesis concurrent PET data were simulated from dynamic MRI data as expected from a simultaneous PET-MR scanner. In the future the findings of this thesis could be evaluated using clinical data from a whole body PET-MRI scanner. The realism of the simulations was limited by not simulating the positron range in the MRI. In addition, in real studies limitations in the MRI acquisition process such as magnetic field inhomogeneities may affect the correlation of PET and MRI data. Another aspect of using simultaneous PET/MRI acquisition is that during spatial and temporal registration the PET images might not be perfectly aligned with the MRI images. To eliminate any spatial or temporal inconsistencies between PET and MRI data Balfour *et al.* (2013) proposed to correct PET data using motion fields derived from a combination of the information of an MRI derived motion model with the PET.

In addition, for the comparison of the two motion correction algorithms in chapter 4 the same motion fields were used both to create the simulations and to correct for motion, and intra-gate motion was not simulated in order to investigate the two methods without being affected from any errors introduced from motion estimation inaccuracies. Extension of this thesis can evaluate the accuracy of the motion estimation and how it affects the performance of the two motion correction methods. Motion estimation inaccuracies might propagate differently in the two motion correction methods and this aspect needs further investigation.

Results on introducing regularization into the reconstruction for motion correction are promising (chapter 4). However, this thesis has been focused on the use of one type of regularization (i.e. median-root prior). Further research is needed in better regularisation approaches (Wang and Qi 2012). In particular, a recent study by Chun and Fessler (2012b) showed that for conventional quadratic regularization motion can lead to non-uniform and anisotropic spatial resolution in areas with different displacement due to motion affecting mainly the quantification of small tumours. As a next step motion correction can be further investigated by selecting regularization parameters that can achieve isotropic and uniform spatial resolution in areas with different displacement due to motion. Finally, extension of this thesis can evaluate the impact of respiratory motion correction on image quality using advanced techniques such as time of flight and point spread function modelling in image reconstruction that are expected to provide better image quality (Alessio *et al.* 2010; Tong *et al.* 2010).

Appendix A. PET simulated data

Simulations and reconstructions have been performed using the open source software library for Tomographic Image Reconstruction (STIR) (Thielemans *et al.* 2012). STIR consists of functions, classes and utilities for different imaging modalities such as PET. In STIR, image and projection data can be manipulated and images can be reconstructed. STIR provides the flexibility to simulate cylindrical scanners of different size and implement new reconstruction algorithms (Jacobson *et al.* 2000).

STIR is based on the use of utilities with command-line functionality. The user can manipulate the function of utilities by selecting different components in text files with Interfile-like syntax known as parameter files.

The implementation details of the analytic simulations as used in this thesis are described in this appendix.

A1. Scanner settings

The parameter file used to simulate the scanner:

```
INTERFILE :=
name of data file := PhilipsGemini.s
originating system := GeminiTF
GENERAL DATA :=
GENERAL IMAGE DATA :=
type of data := PET
imagedata byte order := LITTLEENDIAN
PET STUDY (General) :=
PET data type := Emission
applied corrections := None
number format := float
number of bytes per pixel := 4
number of dimensions := 4
matrix axis label [4] := segment
matrix size [4] := 29
matrix axis label [3] := view
matrix size [3] := 322
matrix axis label [2] := axial coordinate
matrix size [2] := 5,11,17,23,29,35,41,47,53,59,65,71,77,83,
87,83,77,71,65,59,53,47,41,35,29,23,17,11,5
matrix axis label [1] := tangential coordinate
matrix size [1] := 322
```

```
minimum ring difference per segment := -43,-40,-37,-34,-31,-28,-25,-22,-19,
-16,-13,-10,-7,-4,-1,2,5,8,11,14,17,20,23,26,29,32,35,38,41
maximum ring difference per segment := -41,-38,-35,-32,-29,-26,-23,-20,-17,
-14,-11,-8,-5,-2,1,4,7,10,13,16,19,22,25,28,31,34,37,40,43
Scanner parameters:=
Scanner type := GeminiTF
Number of rings := 44
Number of detectors per ring := 644
Inner ring diameter (cm) := 90.034
Average depth of interaction (cm) := 1.5
Distance between rings (cm) := 0.4
Default bin size (cm) := 0.4
View offset (degrees) := 0
Maximum number of non-arc-corrected bins := 322
Default number of arc-corrected bins := 287
Number of blocks per bucket in transaxial direction := 6
Number of blocks per bucket in axial direction := 1
Number of crystals per block in axial direction := 44
Number of crystals per block in transaxial direction := 23
Number of detector layers := 1
Number of crystals per singles unit in axial direction := 44
Number of crystals per singles unit in transaxial direction := 23
end scanner parameters:=
effective central bin size (cm) := 0.226921
image scaling factor[1] := 1
data offset in bytes[1] := 0
number of time frames := 1
END OF INTERFILE :=
```

A2. Forward projector settings

The forward projector settings were the following:

```
Maximum absolute segment number to process: 14
Shift z-origin (in pixels): 0
projector pair type := Matrix
Projector Pair Using Matrix Parameters :=
Matrix type := Ray Tracing
Ray Tracing Matrix Parameters:=
use a slightly better approximation than simple ray tracing
disable caching := 0
store only basic bins in cache := 1
restrict to cylindrical fov := 1
number of rays in tangential direction to trace for each bin := 10
use actual detector boundaries := 1
do symmetry 90degrees min phi := 0
do symmetry 180degrees min phi := 1
do symmetry swap segment := 1
do symmetry swap s := 1
do symmetry shift z := 1
```

A3. Reconstruction settings

The following reconstruction parameter file is compatible to the local STIR version (based on 2.1 version). A variation is now available with STIR 3.0 and more information can be found in STIR user's guide documentation.

```
OSMAPOSLParameters :=
objective function type:=PoissonLogLikelihoodWith
LinearModelForMeanAndGatedProjectionDataWithMotion
PoissonLogLikelihoodWithLinearModelFor
MeanAndGatedProjectionDataWithMotion Parameters:=

input filename prefix := INPUTFILE
maximum absolute segment number to process := -1
zero end planes of segment 0:= 1
projector pair type := Matrix
Projector Pair Using Matrix Parameters :=
Matrix type := Ray Tracing
Ray Tracing Matrix Parameters:=
use a slightly better approximation than simple ray tracing
disable caching := 0
store only basic bins in cache := 1
restrict to cylindrical fov := 1
number of rays in tangential direction to trace for each bin := 10
use actual detector boundaries := 1
do symmetry 90degrees min phi := 0
do symmetry 180degrees min phi := 1
do symmetry swap segment := 1
do symmetry swap s := 1
do symmetry shift z := 1
End Ray Tracing Matrix Parameters:=
End Projector Pair Using Matrix Parameters :=

; if the next parameter is disabled,
; the sensitivity will be computed using the normalisation object
sensitivity filename:= SENSOFILE
; if next is set to 1, sensitivity will be recomputed
; and written to file (if "sensitivity filename" is set)
recompute sensitivity := 1
use subset sensitivities := 1

normalisation sinograms prefix:= NORMALIZATION
```

```
additive sinograms prefix := ADDITIVE

zoom:= 2
xy output image size (in pixels) := 250
end PoissonLogLikelihoodWithLinearModelForMeanAnd
GatedProjectionDataWithMotion Parameters:=
number of subsets:= 23
start at subiteration number:= 1
start at subset:= 0
number of subiterations:= 46
save estimates at subiteration intervals:= 46
output filename prefix := OUTPUTFILE

END :=
```

Appendix B. Scatter simulation

B1. Introduction

The purpose of this appendix was to validate the realism of simulations used in this thesis by including scatter effect. A study for scatter modelling and correction is presented based on the MSc thesis of the candidate which focused on scatter correction using analytic reconstructions and extended here for an iterative reconstruction algorithm.

As it was described in section 2.1.6 one of the most commonly used methods for PET scatter correction is the use of the analytic algorithm single scatter simulation (SSS) (Watson *et al.* 1996). In SSS emission and attenuation images are used with the scanner characteristics and the probability of the photon to be scattered and detected by a detector pair estimated using the Klein-Nishina formula for each scatter point in the attenuation image. For each detector pair the single scatter distribution is calculated as the sum of the single scatter probabilities over all scatter points. The scattering distribution is then calculated by summation of the probabilities of scatter that have been estimated for each possible scatter point in the attenuation image and every detector pair.

To accommodate for multiple scatter, the SSS estimation is usually scaled to the tails of the measured data but tail-scaling may become unstable in cases where the tail region is very small and total-scaling in which the scaling factors are determined using all measured data might be a more accurate scatter estimation procedure.

The first aim of this study was to clarify the improvements in terms of quantification if scatter correction requires double scatter simulation and to investigate how accurate it is to approximate multiple scattered events using tail or total scaled single scatter. The second aim was the evaluation of the SSS scatter correction as implemented in STIR library.

B2. Anthropomorphic numerical phantom

A whole body anthropomorphic numerical phantom (Zubal *et al.* 1994) was used. In the original phantom each slice consisted of 128 × 128 pixels with size 4 *mm* each, and the entire volume consisted of 243 slices with 4 *mm* thickness but for this investigation to approximate a larger human torso than the average the original pixel size for each slice was scaled to 5.27 mm with 3.27 mm slice thickness. Figure B1 shows example of slices through the simulated radioactivity distribution and attenuation images as described in table B1.

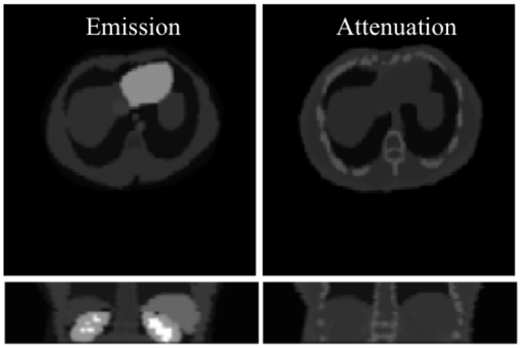


Figure B1: Examples of transverse, coronal and sagittal planes of the simulated radioactivity distribution and attenuation.

Table B1: Simulated radioactivity distribution in units of standardized uptake values (SUVs) and attenuation values in cm^{-1}.

Tissue Type	SUV	Attenuation (cm^{-1})
Lungs	0.39	0.03
Liver	2.14	0.10
Heart	3.15	0.10
Spinal cord	1.07	0.11
Ribcage sternum	1.04	0.15
Kidneys	3.51	0.10
Stomach	1.08	0.10

B3. Comparative evaluation of different scatter-level approximations

Simulations for different scatter levels and different scaling approaches simulated using the Monte Carlo simulation package known as simulation system for emission tomography (SimSET) (Lewellen *et al.* 1998) have been provided by the colleagues Dr Ravindra Manjeshwar and Dr Pablo Aguiar. 3D acquisition was simulated based on the GE Discovery^{TM} STE scanner. This scanner is made of 24 detector rings of 88.62 cm inner ring diameter with axial extent of 0.24 cm. 560 detectors exist in each ring. Standard energy settings were used for the simulation: 425 keV for the lower and 650 *keV* for the upper energy threshold. Simulated data included separate projection data for true, single scatter, double scatter and multiple scatter events resulting in 103.1 total million coincidences.

Tail-scaled single scatter data were also calculated by scaling the single scatter sinogram to the total events sinogram (i.e. sum of true, single scatter, double scatter and multiple scatter events) using the tail-fit method (Watson *et al.* 1997). Scale factors were estimated for each slice so as to be equal with the number of events in the "tails" of the total events sinogram[¶]. Using a similar approach the total-scaled single scatter were also calculated by setting the tail region to cover the entire projection data.

Reconstructions were performed with STIR using the ordered subsets expectation maximization (OSEM) algorithm (i.e. 2 iterations and 28 subsets) and the filtered backprojection (FBP) algorithm (i.e. ramp filter of 0.5 cycles cut-off frequency). Each slice consisted of 128 × 128 × pixels with size 5.27 *mm* each, and the entire volume consisted of 47 slices with 3.27 *mm* thickness. During the reconstruction the data were corrected for detection efficiencies, attenuation and interleaving. The following SimSET datasets were reconstructed and evaluated:

i. Reference: Only the true events were reconstructed.

ii. Non-scatter corrected: The total simulated events (i.e. sum of true, single scatter, double scatter and multiple scatter events) was reconstructed.

iii. Single scatter corrected: The total simulated events was reconstructed by applying single scatter correction.

iv. Single plus double scatter corrected: The total simulated events was reconstructed by applying single and double scatter correction.

v. Tail-scaled single scatter corrected: The total simulated events was reconstructed by applying tail-scaled single scatter correction.

vi. Total-scaled single scatter corrected: The total simulated events was reconstructed by applying total-scaled single scatter correction.

[¶] The "tails" is a binary sinogram obtained by thresholding the attenuation correction factors sinogram. The minimum and maximum thresholds used in this study were 0.4 and 1.4, respectively.

In the aforementioned cases for the scatter correction the corresponding scatter data is incorporated into the OSEM algorithm as an additive term, while subtracted from the total emission sinogram for the FBP algorithm.

In order to quantitatively evaluate the results, mean SUV and the corresponding standard deviation values were calculated from the ROI analysis of the FBP reconstructed images of all investigated cases for the following: kidneys, heart, liver, spinal cord, ribcage sternum and lungs. The ROIs excluded edge/boundary voxels to avoid partial volume effects.

B4. Evaluation of scatter correction using SSS implemented in STIR

The scatter distribution was computed with the SSS analytic algorithm as implemented in STIR (Tsoumpas *et al.* 2004). To estimate the SSS the GE Discovery^{TM} STE PET scanner with coarser sampling was used such that to create a scanner consisted of 6 detector rings with axial extent of 0.336 *cm* and 40 detectors for each ring. However, the data were then distributed on a rectangular sampling grid to correct for interleaving and upsampled with linear interpolation prior to reconstruction. Furthermore, for the scatter points a low resolution attenuation image was used by resampling the attenuation image such that each slice consisted of 39 × 39 × pixels with size 17.58 *mm* each, and the entire volume consisted of 14 slices with 10.90 *mm* thickness. Reconstructions were performed using OSEM and FBP with the same parameters as in the previous section. The scatter estimation was applied to the emission projection data iteratively in order to perform scatter correction (Watson *et al.* 1997) as described in 5 steps (figure B2):

i. Step 1: The normalized SimSET total emission data are corrected for attenuation and then reconstructed.

ii. Step 2: Use SSS as described on the previous section with the image obtained by Step 1. Scale the SSS estimation using tail-fit or total-fit with the same methodology as described in the previous subsection.

iii. Step 3: Subtract the scaled scatter estimation from the total SimSET emission data to scatter-correct the sinogram.

iv. Step 4: Repeat Steps 1 and 2, but in place of the SimSET total emission data of Step 1 use the scatter-corrected sinogram of Step 3

v. Step 5: Average scatter estimates from Steps 2 and 4. Use the result as final scatter estimate.

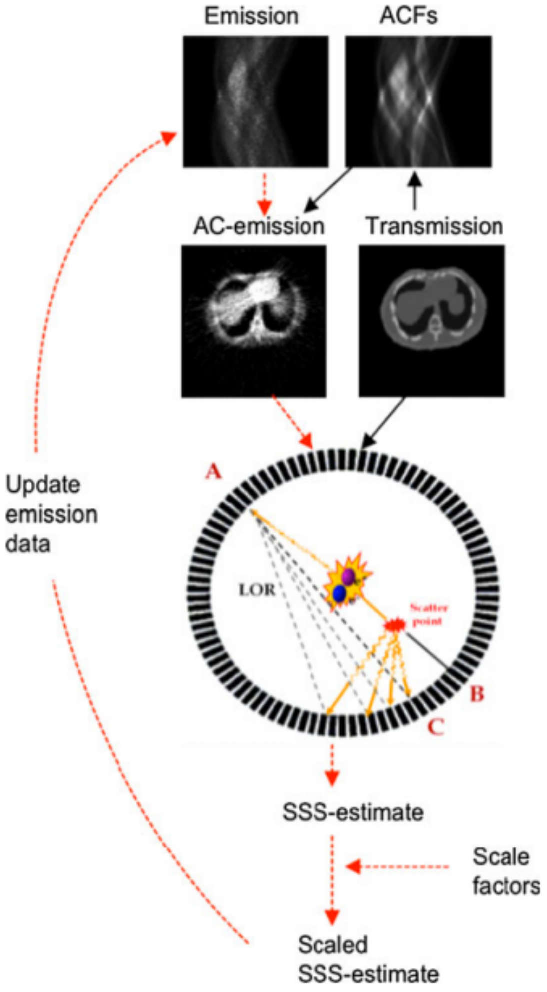
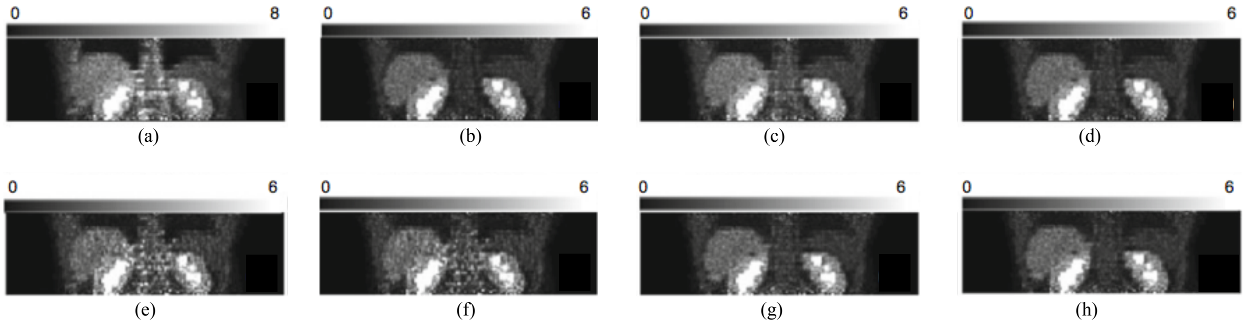


Figure B2: Schematic representation of scatter correction of the SSS algorithm with scaling.

B5. Results

Figure B3 shows coronal planes of OSEM reconstructed images for the different scatter order approximations (i.e. total non-corrected; ideal total scatter corrected; single scatter corrected; single plus double scatter corrected; tail-scaled single scatter corrected; and total-scaled single scatter corrected) as well as scatter correction using SSS implemented in STIR (i.e. tail-scaled SSS corrected; and total-scaled SSS corrected).



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure B3: Coronal planes of images for the: (a) total non-corrected; (b) ideal total scatter corrected; (c) single scatter corrected; (d) single plus double scatter corrected; (e) tail-scaled single scatter corrected; and (f) total-scaled single scatter corrected; (g) tail-scaled SSS corrected; and (h) total-scaled SSS corrected.

Figure B4 illustrates the profiles through a coronal plane of the OSEM reconstructed images of the SimSET total non-corrected, ideal total scatter corrected, single scatter corrected, and the SSS corrected using tail-fit or total-fit.

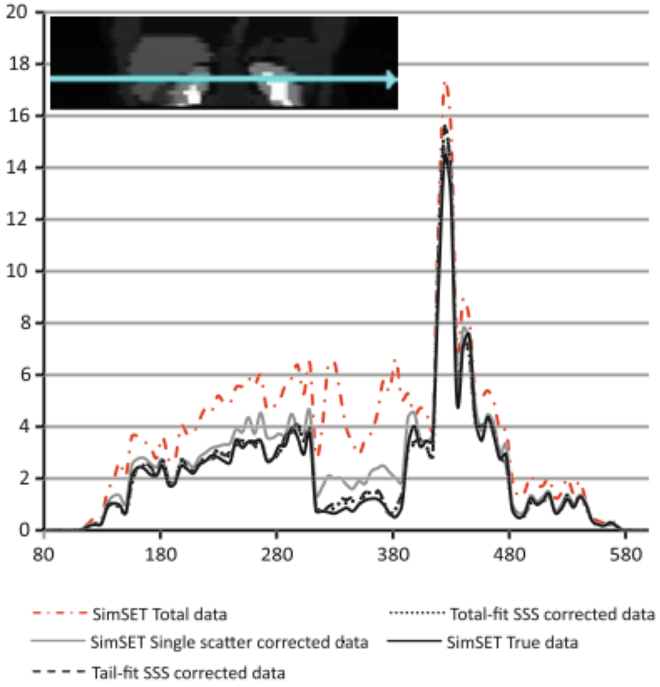


Figure B4: Intensity profiles through a coronal plane of the OSEM reconstructed images of the SimSET total non-corrected; ideal total scatter corrected; single scatter corrected; and of the tail-scaled SSS corrected; and total-scaled SSS corrected.

The SUV with the standard deviations for the FBP reconstructed images for the ideal radioactivity distribution and the different scatter order approximations are presented in figure B5.

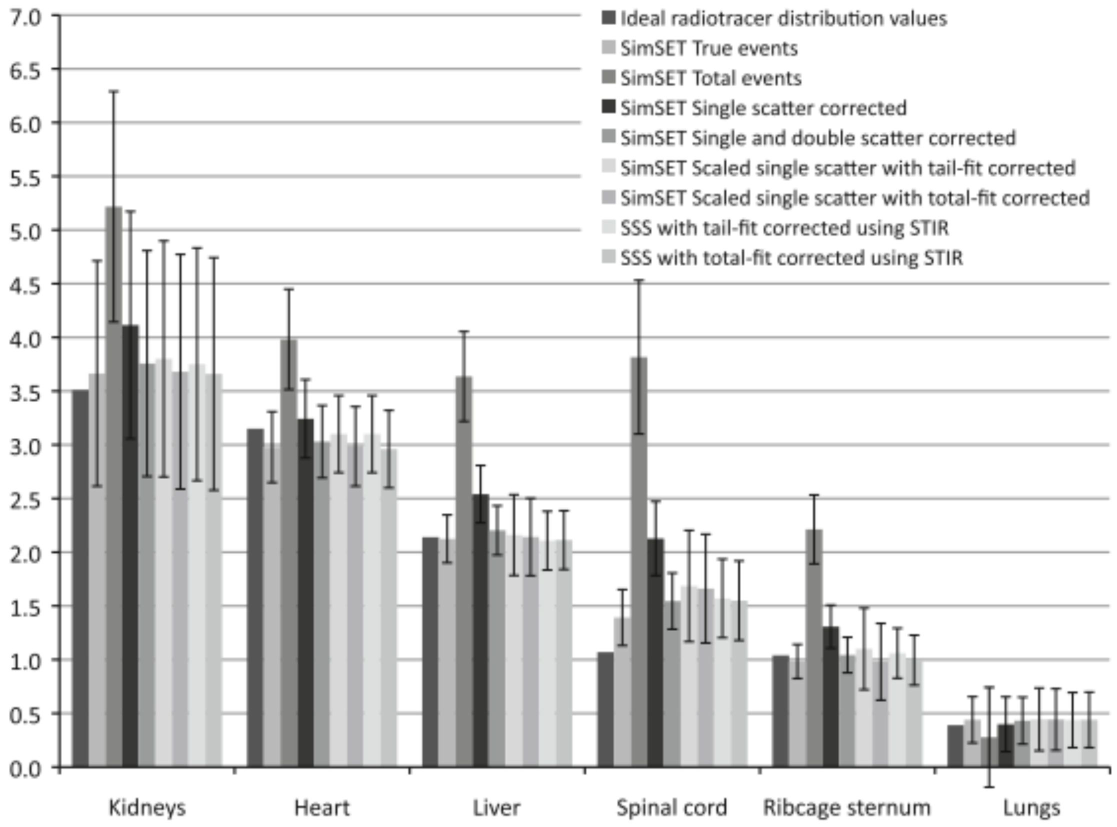


Figure B5: Ideal radioactivity distribution and SUV with standard deviation for different ROIs of the FBP reconstructed images. The results correspond to SimSET reconstructed output unless indicated as SSS which corresponds to STIR scatter estimation.

B6. Discussion and Conclusion

Visual interpretation of the images in figure B3 shows differences following multiple scatter compensation, either by single scatter scaling, or directly by multiple scatter correction. Quantification accuracy has improved by scatter correction (figure B5). In particular, heart SUV was 3 in the ideal, 4 before scatter correction, 3.3 after applying single scatter correction and 3 after applying single and double scatter correction. Double scatter has high impact in quantification; kidneys SUV after single scatter correction was 4.1, while following single plus double scatter correction was 3.7.

Furthermore, the kidneys SUV after correction by scaling single scatter with tail-fit approaches 3.8, whereas it is 3.6 with total-fit. However, tail-fit slightly underestimates the total scatter in some places (e.g. kidneys and ribcage sternum), whereas this is not observed when scaled with total-fit.

Compensation for double scattered coincident events is required for the accurate estimation of activity distribution. In addition, this investigation has verified that scaling the single scatter distribution is an efficient method to compensate for multiple scattering; thus, separate modelling double or higher order scatter seems unnecessary for human PET scanners with BGO crystals, and consequently they are expected to be even less important for higher energy resolution scanners (e.g. LSO/LYSO crystals systems). However, results might have been different with significant out of FOV scatter events, longer axial FOV scanners, or cases of exceptionally higher multiple scatter.

Evaluating the SSS implementation in STIR, figure B3 shows very good agreement between tail-scaled or total-scaled SSS scatter correction and SimSET true events. Intensity profiles in figure B4 show also good agreement between SSS correction and SimSET true events. Quantitative investigation in figure B5 also verifies good agreement between SSS correction and SimSET true events. Similarly with the SimSET scaled single scatter results discussed above, total- fit SSS correction may slightly outperform tail-fit SSS correction in some regions. In particular, for the spinal cord, true events SUV was 1.4, whereas 1.6 after tail-scaled SSS correction and 1.5 after total- scaled SSS correction. For the region of the heart, the true events SUV was 3.0, whereas after tail-scaled SSS correction it was 3.1 and after total-scaled SSS correction it was 3.0.

This work has illustrated that the scatter correction methodology implemented in STIR successfully recovers scatter-free images and shows good agreement with Monte Carlo simulations.

Appendix C. Noise properties of RTA and MCIR motion correction methods including resolution blurring

The purpose of this appendix was to validate the results of chapter 4 by using more realistic simulations including resolution blurring.

C1. Methods

A numerical phantom was created by adding spherical lesions to the phantom created in section 3.2.2 with 12 *mm* diameter and lesion to background ratio equal to 6:1. The 3D FDG distribution was convolved with a Gaussian kernel of 6 *mm* FWHM in x, y, z to simulate a current resolution scanner. A real respiratory signal for type C breathing pattern of 20 *mm* maximum value was combined with an MRI-derived motion model to simulate a dynamic thorax PET dataset acquired from a continuously moving subject as described in section 3.2.2.

The dynamic PET data were gated based on amplitude to create 8 gates and each gated image was used as the input to an analytic simulation of a PET acquisition to create projection data as described in section 3.2.3. The projection data accounted for photon attenuation, scatter (approximately 33% of the total counts) but not for random coincidences. Statistical (Poisson) noise was added to the projection data such that the overall counts of the whole dataset correspond to 50 million unscattered coincidences (i.e. 3 to 5 *min* 3D PET clinical thorax acquisition). To add noise in the gated projection data, based on the corresponding amplitude based histogram (i.e. distribution of the motion amplitudes), each of the eight sinograms was weighted according to the dwell time at each displacement (i.e. histogram value). Each histogram value shows the duration that patient spends in the specific breathing position (i.e. how many times the specific amplitude was repeated in a respiratory signal). Sixty noise realisations were created.

The eight respiratory positions were reconstructed using both RTA and MCIR motion correction methods with MRI-derived motion fields. The transformation operators describing the motion between the gated data were estimated using voxel-wise deformation vectors, which were produced by co-registration of MRI gated datasets with a non-linear registration algorithm (Buerger *et al.* 2011). The transpose motion matrix operation of MCIR was approximated by warping the image with the inverted motion fields which were numerically estimated from the same motion vectors (Crum *et al.* 2007). All iterative reconstructions were performed with 23 subsets and for 2 iterations and for β_g equal to zero (i.e. OSEM). Each slice consisted of 250×250 pixels with size 2×2 mm each, and the entire volume consisted of 87 slices with 2 mm thickness.

Figures of merit were computed from all the reconstructed images produced from each numerical experiment. The following figures of merit were calculated:

Bias: An image of the bias with respect to the true simulated radioactivity distribution was determined as follows:

$$Bias = \frac{1}{N}\sum_{i=1}^{N}\Lambda_i - \Lambda_{ideal} = \bar{\Lambda} - \Lambda_{ideal} \tag{C1.1}$$

where Λ_{ideal} is the reference image, Λ_i is the reconstructed image of the i^{th} noise realization and $\bar{\Lambda}$ is the mean image calculated as the average of the sequence of N realisations of reconstructed images.

Noise: The noise was evaluated through the use of the standard deviation (std) calculated over the sequence of N reconstructed images:

$$std = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(\Lambda_i - \bar{\Lambda})^2} \tag{C1.2}$$

C2. Results

Figure C1 shows coronal planes of the bias and standard deviation images for the RTA and MCIR reconstructed images. The results are displayed at 2, 4, 6, 8 and 10 iterations.

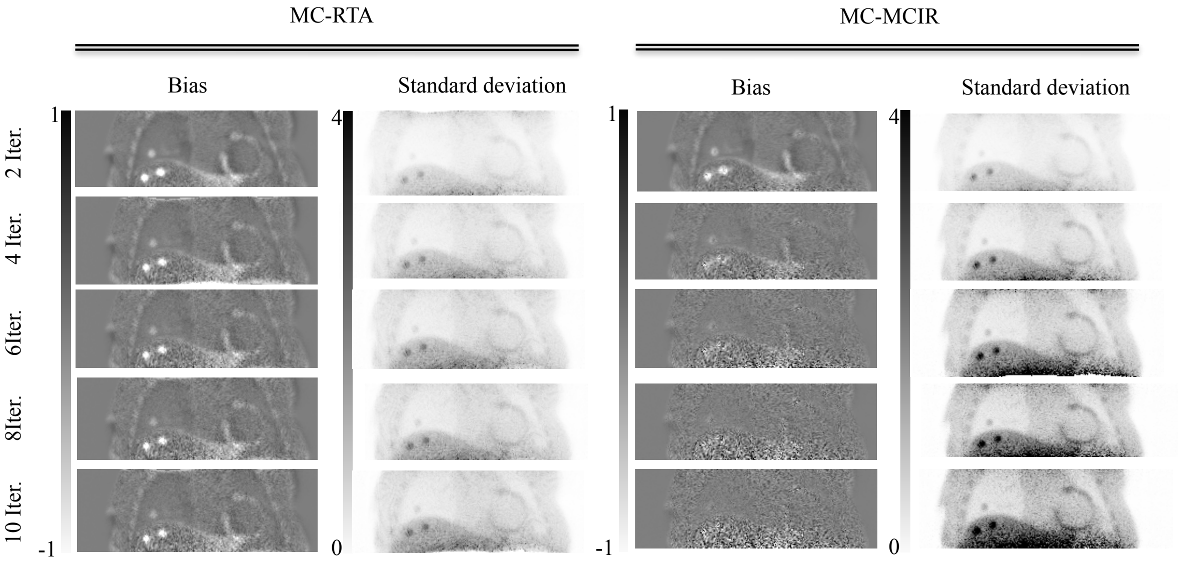


Figure C1: Coronal planes of the bias and standard deviation images with motion correction using both RTA (MC-RTA) and MCIR (MC-MCIR). For 2 iterations the MCIR bias is localized around edges and approaches a uniform small value at 10 iterations. On the other hand, RTA has high bias for all iterations especially at the edges of the lesions and the myocardium independent of the iteration number. Standard deviation is stabilized to a constant value for RTA, while for MCIR it progressively increases with the number of iterations.

C3. Discussion and Conclusion

The results are similar with chapter 4 including the reduced noise in RTA compared to MCIR. This confirms that the noted effects are indeed a characteristic of the RTA and MCIR methods and not a detail of the simulations used in this thesis and there is no effects of aliasing that may cause variance magnification which are attenuated with RTA via image interpolation.

Bibliography

Accorsi R, Adam L E, Werner M E and Karp J S 2004. Optimization of a fully 3D single scatter simulation algorithm for 3D PET. *Physics in Medicine and Biology* **49**(12), 2577–2598.

Adam L E, Karp J S and Freifelder R 2000. Energy-based scatter correction for 3-D PET scanners using NaI(T1) detectors. *IEEE Transactions on Medical Imaging* **19**(5), 513–521.

Adam R, Karp J S, Daube-Witherspoon M E and Smith R J 2001. Performance of a whole body PET scanner using curve plate NaI(Tl) detectors. *Journal of Nuclear Medicine* **42**(12), 1821–1830.

Adam R and Vinet E 2004. Regional treatment of metastasis: Surgery of colorectal liver metastases. *Annals of Oncology* **15**(SUPPL. 4), iv103–iv106.

Ahn S and Fessler J A 2004. Emission image reconstruction for randoms-precorrected PET allowing negative sinogram values. *IEEE Transactions on Medical Imaging* **23**(5), 591–601.

Alavi A, Dann R, Chawluk J, Alavi J, Kushner M and Reivich M 1986. Positron emission tomography imaging of regional cerebral glucose metabolism. *Seminars in Nuclear Medicine* **16**(1), 2–34.

Alenius S and Ruotsalainen U 1997. Bayesian image reconstruction for emission tomography based on median root prior. *European Journal of Nuclear Medicine* **24**(3), 258–265.

Alenius S and Ruotsalainen U 2002. Generalization of median root prior reconstruction. *IEEE Transactions on Medical Imaging* **21**(11), 1413–1420.

Alessio A M, Stearns C W, Tong S, Ross S G, Kohlmyer S, Ganin A and Kinahan P E 2010. Application and evaluation of a measured spatially variant system model for PET image reconstruction. *IEEE Transactions on Medical Imaging* **29**(3), 938–949.

Andronache A, von Siebenthal M, Székely G and Cattin P 2008. Non-rigid registration of multi-modal images using both mutual information and cross-correlation. *Medical Image Analysis* **12**(1), 3–15.

Antoch G, Stattaus J, Nemat A T, Marnitz S, Beyer T, Kuehl H, Bockisch A, Debatin J F and Freudenberg L S 2003. Non-small cell lung cancer: dual-modality PET/CT in preoperative staging. *Radiology* **229**(2), 526–533.

Asma E, Manjeshwar R, and Thielemans K 2006. Theoretical comparison of motion correction techniques for PET image reconstruction. *in* 'IEEE Nuclear Science Symposium and Medical Imaging Conference' pp. 1762–1767.

Badawi R D and Dahlbom M 2005. NEC: Some coincidences are more equivalent than others. *Journal of Nuclear Medicine* **46**(11), 1767–1768.

Badawi R D, Lodge M A and Marsden P K 1998. Algorithms for calculating detector efficiency normalization coefficients for true coincidences in 3D PET. *Physics in Medicine and Biology* **43**(1), 189–205.

Badawi R D and Marsden P K 1999. Developments in component-based normalization for 3D PET. *Physics in Medicine and Biology* **44**(2), 571–594.

Bai C, Shao L, Silva A J D and Zhao Z 2003. A generalized model for the conversion from CT numbers to linear attenuation coefficients. *IEEE Transactions on Nuclear Science* **50**(5), 1510–1515.

Bai W and Brady M 2009. Regularized B-spline deformable registration for respiratory motion correction in PET images. *Physics in Medicine and Biology* **54**(9), 2719–2736.

Bai W and Brady M 2011. Motion correction and attenuation correction for respiratory gated PET images. *IEEE Transactions on Medical Imaging* **30**(2), 351–365.

Bailey D L, Karp J S and Surti S 2005 *Physics and Instrumentation in PET. In: D. L. Bailey, D. W. Townsend, P. E. Valk and M. N. Maisey (eds) Positron emission tomography: Basic sciences* Springer London.

Bailey D L and Meikle S R 1994. A convolution-subtraction scatter correction method for 3D PET. *Physics in Medicine and Biology* **39**(3), 411–424.

Balfour D, Polycarpou I, Marsden P K and King A P 2013. Respiratory motion correction of PET imaging using a combination of PET and MR data. *in* 'IEEE Nuclear Science Symposium and Medical Imaging Conference'.

Barrett H H and Myers K J 2003 *Foundations of Image Science* Willey.

Barrett H H, Wilson D W and Tsui B M W 1994. Noise properties of the EM algorithm: I. Theory. *Physics in Medicine and Biology* **39**(5), 833–846.

Barrett H H, Yao J, Rolland J P and Myers K J 1993. Model observers for assessment of image quality. *Proceedings of the National Academy of Sciences of the United States of America* **90**(21), 9758–9765.

Beiderwellen K, Gomez B, Buchbender C, Hartung V, Poeppel T D, Nensa F, Kuehl H, Bockisch A and Lauenstein T C 2013. Depiction and characterization of liver lesions in whole body [(18)F]-FDG PET/MRI. *European Journal of Radiology* **82**(11), e669–e675.

Bentourkia M, Msaki P, Cadorette J, Héon M and Lecomte R 1993. Assessment of scatter components in multispectral PET imaging. *in* 'IEEE Nuclear Science Symposium and Medical Imaging Conference' Vol. 3 pp. 1505–1509.

Bettinardi V, Pagani E, Gilardi M, Alenius S, Thielemans K, Teras M and Fazio F 2002. Implementation and evaluation of a 3D one-step late reconstruction algorithm for 3D positron emission tomography brain studies using median root prior. *European Journal of Nuclear Medicine and Molecular Imaging* **29**(1), 7–18.

Beyer T, Antoch G, Blodgett T, Freudenberg L F, Akhurst T and Müeller S 2003. Dual-modality PET/CT imaging: The effect of respiratory motion on combined image quality in clinical oncology. *European Journal of Nuclear Medicine and Molecular Imaging* **30**(4), 588–596.

Beyer T, Antoch G, Müller S, Egelhof T, Freudenberg L S, Debatin J and Bockisch A 2004. Acquisition protocol considerations for combined PET/CT imaging. *Journal of Nuclear Medicine* **45**(Suppl 1), 25S–35S.

Beyer T, Townsend D W, Brun T, Kinahan P E, Charron M, Roddy R, Jerin J, Young J, Byars L and Nutt R 2000. A combined PET/CT scanner for clinical oncology. *Journal of Nuclear Medicine* **41**(8), 1369–1379.

Bloomfield P M, Spinks T J, Reed J, Schnorr L, Westrip A M, Livieratos L, Fulton R and Jones T 2003. The design and implementation of a motion correction scheme for neurological PET. *Physics in Medicine and Biology* **48**(8), 959–978.

Boellaard R, Lingen A V and Lammertsma A A 2001. Experimental and clinical evaluation of iterative reconstruction (OSEM) in dynamic PET: Quantitative characteristics and effects on kinetic modeling. *Journal of Nuclear Medicine* **42**(5), 808–817.

Boellaard R, Oyen W J G, Hoekstra C J, Hoekstra O S, Visser E P, Willemsen A T, Arends B, Verzijlbergen F J, Zijlstra J, Paans A M, Comans E F I and Pruim J 2008. The Netherlands protocol for standardisation and quantification of FDG whole body PET studies in multi-centre trials. *European Journal of Nuclear Medicine and Molecular Imaging* **35**(12), 2320–2333.

Boss A, Bisdas S, Kolb A, Hofmann M, Ernemann U, Claussen C D, Pfannenberg C, Pichler B J, Reimold M and Stegger L 2010. Hybrid PET/MRI of intracranial masses: Initial experiences and comparison to PET/CT. *Journal of Nuclear Medicine* **51**(8), 1198–1205.

Bouchelouche K, Turkbey B, Choyke P and Capala J 2010. Imaging prostate cancer: An update on positron emission tomography and magnetic resonance imaging. *Current Urology Reports* **11**(3), 180–190.

Boucher L, Rodrigue S, Lecomte R and Bénard F 2004. Respiratory gating for 3-dimensional PET of the thorax: Feasibility and initial results. *Journal of Nuclear Medicine* **45**(2), 214–219.

Brownell G L and Sweet W H 1953. Localization of brain tumors with positron emitters. *Nucleonics* **11**, 40–45.

Budinger T F 1983. Time-of-flight positron emission tomography: status relative to conventional PET. *Journal of Nuclear Medicine* **24**(1), 73–78.

Budinger T F 1998. PET instrumentation: what are the limits? *Seminars in Nuclear Medicine* **28**(3), 247–267.

Buerger C, Aitken A, Tsoumpas C, King A P, Schulz V, Marsden P K and Schaeffter T 2012. Investigation of MR-based attenuation correction and motion compensation for hybrid PET/MR. *IEEE Transactions on Nuclear Science* **59**(5), 1967–1976.

Buerger C, Schaeffter T and King A P 2011. Hierarchical adaptive local affine registration for fast and robust respiratory motion estimation. *Medical Image Analysis* **15**(4), 551–564.

Bundschuh R A, Martínez-Moeller A, Essler M, Martínez M J, Nekolla S G, Ziegler S I and Schwaiger M 2007. Postacquisition detection of tumor motion in the lung and upper abdomen using list-mode PET data: a feasibility study. *Journal of Nuclear Medicine* **48**(5), 758–763.

Burger C, Goerres G, Schoenes S, Buck A, Lonn A and von Schulthess G 2002. PET attenuation coefficients from CT images: experimental evaluation of the transformation of CT into PET 511-keV attenuation coefficients. *European Journal of Nuclear Medicine and Molecular Imaging* **29**(7), 922–927.

Burgess A E 1994. Statistically defined backgrounds: performance of a modified nonprewhitening observer model. *Journal of Optical Society of America A* **11**(4), 1237–1242.

Büscher K, Judenhofer M S, Kuhlmann M T, Hermann S, Wehrl H F, Schäfers K P, Schäfers M, Pichler B J and Stegger L 2010. Isochronous assessment of cardiac metabolism and function in mice using hybrid PET/MRI. *Journal of Nuclear Medicine* **51**(8), 1277–1284.

Büther F, Dawood M, Stegger L, Wübbeling F, Schäfers M, Schober O and Schäfers K P 2009. List mode-driven cardiac and respiratory gating in PET. *Journal of Nuclear Medicine* **50**(5), 674–681.

Cao Z, Gilland D R, Mair B A and Jaszczak R J 2003. Three-dimensional motion estimation with image reconstruction for gated cardiac ECT. *IEEE Transactions on Nuclear Science* **50**(3), 384–388.

Casey M E, Gadagkar H and Newport D 1995. A component based method for normalization in volume PET. *in* 'International Meeting on Fully Three Dimensional Image Reconstruction in Radiology and Nuclear Medicine' pp. 67–71.

Casey M E and Nutt R 1986. A Multicrystal Two Dimensional BGO Detector System For Positron Emission Tomography. *IEEE Transcactions of Nuclear Science* **33**(1), 460–463.

Chakraborty D P 2000 *The FROC, AFROC and DROC variants of the ROC Analysis. In: R. L. Van Metter, J. Beutel and H. L. Kundel (eds) Handbook of Medical Imaging Volume 1. Physics and Psychophysics* Bellingham: SPIE Press.

Chan M T, Leahy R M, Mumcuoglu E U, Cherry S R, Czernin J and Chatziioannou A 1997. Comparing lesion detection performance for PET image reconstruction algorithms: A case study. *IEEE Transactions on Nuclear Science* **44**(4), 1558–1563.

Chang L T 1978. A Method for Attenuation Correction in Radionuclide Computed Tomography. *IEEE Transactions on Nuclear Science* **25**(1), 638–643.

Cherry S R, Dahlbom M and Hoffman E J 1991. 3D PET using a conventional multislice tomograph without septa. *Journal of Computer Assisted Tomography* **15**(4), 655–668.

Cherry S R, Sorenson J A and Phelps M E 2012 *Positron Emission Tomography. Physics in Nuclear Medicine* Springer London.

Chun S Y and Fessler J A 2012b. Spatial resolution properties of motion-compensated tomographic image reconstruction methods. *IEEE Transactions on Medical Imaging* **31**(7), 1413–1425.

Chun S Y and Fessler J A 2013. Noise properties of motion compensated tomographic image reconstruction methods. *IEEE Transactions on Medical Imaging* **32**(2), 141–152.

Chun S Y, Reese T G, Ouyang J, Guerin B, Catana C, Zhu X, Alpert N M and El Fakhri G 2012a. MRI-based nonrigid motion correction in simultaneous PET/MRI. *Journal of Nuclear Medicine* **53**(8), 1284–1291.

Commowick O, Arsigny V, Isambert A, Costa J, Dhermain F, Bidault F, Bondiau P Y, Ayache N and Malandain G 2008. An efficient locally affine framework for the smooth registration of anatomical structures. *Medical Image Analysis* **12**(4), 427–441.

Comtat C, Kinahan P, Defrise M, Michel C, Lartizien C and Townsend D W 1999. Simulating whole-body PET scanning with rapid analytical methods. *in* 'IEEE Nuclear Science Symposium and Medical Imaging Conference' Vol. 3 pp. 1260–1264.

Conti M, Hong I and Michel C 2012. Reconstruction of scattered and unscattered PET coincidences using TOF and energy information. *Physics in Medicine and Biology* **57**(15), N307–NN317.

Crum W R, Camara O and Hawkes D J 2007. Methods for inverting dense displacement fields: Evaluation in brain image registration. *International Conference on Medical Image Computing and Computer-Assisted Intervention* **10**(1), 900–907.

Crum W R, Hartkens T and Hill D L G 2004. Non-rigid image registration: Theory and practice. *British Journal of Radiology* **77**(SPEC. ISS. 2), S140–S153.

Daou D 2008. Respiratory motion handling is mandatory to accomplish the high-resolution PET destiny. *European Journal of Nuclear Medicine and Molecular Imaging* **35**(11), 1961–1970.

Daube-Witherspoon M, Yan Y, Green M, Carson R, Kempner K and Herscovitch P 1990. Correction for motion distortion in PET by dynamic monitoring of patient position. *Journal of Nuclear Medicine* **31**(5), 816.

Dawood M, Büther F, Jiang X and Schäfers K P 2008. Respiratory motion correction in 3D PET data with advanced optical flow algorithms. *IEEE Transactions on Medical Imaging* **27**(8), 1164–1175.

Dawood M, Büther F, Lang N, Schober O and Schäfers K P 2007. Respiratory gating in positron emission tomography: A quantitative comparison of different gating schemes. *Medical Physics* **34**(7), 3067–3076.

Dawood M, Büther F, Stegger L, Jiang X, Schober O, Schäfers M and Schäfers K P 2009. Optimal number of respiratory gates in positron emission tomography: A cardiac patient study. *Medical Physics* **36**(5), 1775–1784.

Dawood M, Lang N, Jiang X and Schäfers K P 2006. Lung motion correction on respiratory gated 3-D PET/CT images. *IEEE Transactions on Medical Imaging* **25**(4), 476–485.

DeBenedetti S, Cowan S, Konneker W R and Primakoff H 1950. On the angular distribution of two-photon annihilation radiation. *Physical Review* **77**(2), 205–212.

Delso G, Furst S, Jakoby B, Ladebeck R, Ganter C, Nekolla S G, Schwaiger M and Ziegler S I 2011. Performance measurements of the Siemens mMR integrated whole-body PET/MR scanner. *Journal of Nuclear Medicine* **52**(12), 1914–1922.

Detorie N and Dahlbom M 2007. Motion correction for gated PET using sinogram registration. *in* 'IEEE Nuclear Science Symposium and Medical Imaging Conference' Vol. 5 pp. 3577–3582.

Dey J and King M A 2009. Theoretical and numerical study of MLEM and OSEM reconstruction algorithms for motion correction in emission tomography. *IEEE Transactions on Nuclear Science* **56**(5), 2739–2749.

Dikaios N and Fryer T D 2009. Respiratory motion correction of PET using motion parameters from MR. *in* 'IEEE Nuclear Science Symposium and Medical Imaging Conference' pp. 2806–2808.

Dikaios N and Fryer T D 2011a. Improved motion-compensated image reconstruction for PET using sensitivity correction per respiratory gate and an approximate tube-of-response backprojector. *Medical Physics* **38**(9), 4958–4970.

Dikaios N and Fryer T D 2011b. Acceleration of motion-compensated PET reconstruction: ordered subsets-gates EM algorithms and a priori reference gate information. *Physics in Medicine and Biology* **56**(6), 1695–1715.

Dikaios N and Fryer T D 2012a. Registration weighted motion correction for PET. *Medical Physics* **39**(3), 1253–1264.

Dikaios N, Izquierdo-Garcia D, Graves M J, Mani V, Fayad Z A and Fryer T D 2012b. MRI-based motion correction of thoracic PET: Initial comparison of acquisition protocols and correction strategies suitable for simultaneous PET/MRI systems. *European Radiology* **22**(2), 439–446.

Donati O F, Hany T F, Reiner C S, von Schulthess G K, Marincek B, Seifert B and Weishaupt D 2010. Value of retrospective fusion of PET and MR images in detection of hepatic metastases: comparison with 18F- FDG PET/CT and Gd-EOB-DTPA-enhanced MRI. *Journal of Nuclear Medicine* **51**(5), 692–699.

Drzezga A, Souvatzoglou M, Eiber M, Beer A J, Fürst S, Martinez-Möller A, Nekolla S G, Ziegler S, Ganter C, Rummeny E J and Schwaiger M 2012. First clinical experience with integrated whole-body PET/MR: Comparison to PET/CT in patients with oncologic diagnoses. *Journal of Nuclear Medicine* **53**(6), 845–855.

Du J, Hamilton G, Takahashi A, Bydder M and Chung C B 2007. Ultrashort echo time spectroscopic imaging (UTESI) of cortical bone. *Magnetic Resonance in Medicine* **58**(5), 1001–1009.

El Fakhri G, Surti S, Trott C M, Scheuermann J and Karp J S 2011. Improvement in lesion detection with whole-body oncologic time-of-flight PET. *Journal of Nuclear Medicine* **52**(3), 347–353.

Erdi Y E, Macapinlac H, Rosenzweig K E, Humm J L, Larson S M, Erdi A K and Yorke E D 2000. Use of PET to monitor the response of lung cancer to radiation treatment. *European Journal of Nuclear Medicine* **27**(7), 861–866.

Erdi Y E, Nehmeh S A, Pan T, Pevsner A, Rosenzweig K E, Mageras G, Yorke E D, Schoder H, Hsiao W, Squire O D, Vernon P, Ashman J B, Mostafavi H, Larson S M and Humm J L 2004. The CT motion quantitation of lung lesions and its impact on PET-measured SUVs. *Journal of Nuclear Medicine* **45**(8), 1287–1292.

Erlandsson K, Buvat I, Pretorius P H, Thomas B A and Hutton B F 2012. A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology. *Physics in Medicine and Biology* **57**(21), R119–R159.

Fahey F H 2002. Data acquisition in PET imaging. *Journal of Nuclear Medicine Technology* **30**(2), 39–49.

Farquhar T H, Llacer J, Hoh C K, Czernin J, Gambhir S S, Seltzer M A, Silverman D H S, Qi J, Hsu C and Hoffman E J 1999. ROC and localization ROC analyses of lesion detection in whole-body FDG PET: Effects of acquisition mode, attenuation correction and reconstruction algorithm. *Journal of Nuclear Medicine* **40**(12), 2043–2052.

Fayad H J, Buerger C, Tsoumpas C, Cheze-Le Rest C and Visvikis D 2012. A generic respiratory motion model based on 4D MRI imaging and 2D image navigators. *in* 'IEEE Nuclear Science Symposium and Medical Imaging Conference' pp. 4058–4061.

Ferlay J, Shin H R, Bray F, Forman D, Mathers C and Parkin D M 2010. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *International Journal of Cancer* **127**(12), 2893–2917.

Fessler J A 1994. Penalized weighted least-squares image reconstruction for positron emission tomography. *IEEE Transactions on Medical Imaging* **13**(2), 290–300.

Floyd C E J, Jaszczak R J, Greer K L and Coleman R E 1985. Deconvolution of Compton scatter in SPECT. *Journal of Nuclear Medicine* **26**(4), 403–408.

Frey E C, Gilland K L and Tsui B M W 2002. Application of task-based measures of image quality to optimization and evaluation of three-dimensional reconstruction-based compensation methods in myocardial perfusion SPECT. *IEEE Transactions on Medical Imaging* **21**(9), 1040–1050.

Gallas B D and Barrett H H 2003. Validating the use of channels to estimate the ideal linear observer. *Journal of the Optical Society of America A: Optics and Image Science, and Vision* **20**(9), 1725–1738.

Gifford H C, King M A, Vries D J D and Soares E J 2000. Channelized hotelling and human observer correlation for lesion detection in hepatic SPECT imaging. *Journal of Nuclear Medicine* **41**(3), 514–521.

Gilland D R, Mair B A, Bowsher J E and Jaszczak R J 2002. Simultaneous reconstruction and motion estimation for gated cardiac ECT. *IEEE Transactions on Nuclear Science* **49**(5), 2344–2349.

Goerres G W, Burger C, Schwitter M R, Heidelberg T N H, Seifert B and von Schulthess G K 2003. PET/CT of the abdomen: Optimizing the patient breathing pattern. *European Radiology* **13**(4), 734–739.

Goerres G W, Kamel E, Heidelberg T N H, Schwitter M R, Burger C and von Schulthess G K 2002. PET-CT image co-registration in the thorax: Influence of respiration. *European Journal of Nuclear Medicine* **29**(3), 351–360.

Gravier E and Yang Y 2005. Motion-compensated reconstruction of tomographic image sequences. *IEEE Transactions on Nuclear Science* **52**(1), 51–56.

Green P J 1990. Bayesian reconstructions from emission tomography data using a modified EM algorithm. *IEEE Transactions on Medical Imaging* **9**(1), 84–93.

Grezes-Besset L, Nuyts J, Boellard R, Buvat I, Michel C, Pierre C, Costes N and Reilhac A 2007. Simulation-based evaluation of NEG-ML iterative reconstruction of low count PET data. *in* 'IEEE Nuclear Science Symposium and Medical Imaging Conference' pp. 3009–3014.

Grootoonk S, Spinks T J, Sashin D, Spyrou N M and Jones T 1996. Correction for scatter in 3D brain PET using a dual energy window method. *Physics in Medicine and Biology* **41**(12), 2757–2774.

Guérin B, Cho S, Chun S Y, Zhu X, Alpert N M, El Fakhri G, Reese T and Catana C 2011. Nonrigid PET motion compensation in the lower abdomen using simultaneous tagged-MRI and PET imaging. *Medical Physics* **38**(6), 3025–3038.

Guérin B and El Fakhri G 2008. Realistic PET Monte Carlo simulation with pixellated block detectors, light sharing, random coincidences and dead-time modeling. *IEEE Transactions on Nuclear Science* **55**(3), 942–952.

Guivarch O, Turzo A, Bizais Y and Visvikis D 2004. Synchronisation of pulmonary scintigraphy by respiratory flow and impedance plethysmography. *in* 'Proceedings of SPIE in Medical Imaging' Vol. 5370 pp. 1166–1175.

Hanley J A and McNeil B J 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **143**(1), 29–36.

Hebert T and Leahy R M 1989. Generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors. *IEEE Transactions on Medical Imaging* **8**(2), 194–202.

Henschke C I, Wisnivesky J P, Yankelevitz D F and Miettinen O S 2003. Small stage I cancers of the lung: Genuineness and curability. *Lung Cancer* **39**(3), 327–330.

Higashi K, Ueda Y, Sakuma T, Seki H, Oguchi M, Taniguchi M, Taki S, Tonami H, Katsuda S and Yamamoto I 2001. Comparison of [18-F]FDG PET and 201-Tl SPECT in Evaluation of Pulmonary Nodules. *Journal of Nuclear Medicine* **42**(10), 1489–1496.

Hillner B, Siegel B, Liu D, Shields A, Gareen I, Hanna L, Stine S and Coleman R 2008. Impact of positron emission tomography/computed tomography and positron emission tomography (PET) alone on expected management of patients with cancer: initial results from the National Oncologic PET Registry. *Journal of Clinical Oncology* **26**(13), 2155–2161.

Holdsworth C H, Levin C S, Janecek M, Dalhbom M, and Hoffman E J 2002. Performance analysis of an improved 3-D PET Monte Carlo simulation and scatter correction. *IEEE Transactions on Nuclear Science* **49**(1), 83–89.

Hutton B F, Braun M, Thurfjell L and Lau D Y 2002. Image registration: an essential tool for nuclear medicine. *European Journal of Nuclear Medicine and Molecular Imaging* **29**(4), 559–577.

Jacobson M, Levkovitz R, Ben-Tal A, Thielemans K, Spinks T, Belluzzo D, Pagani E, Bettinardi V, Gilardi M C, Zverovich A and Mitra G 2000. Enhanced 3D PET OSEM reconstruction using inter-update Metz filters. *Physics in Medicine and Biology* **45**(8), 2417–2439.

Jacobson M W and Fessler J A 2003. Joint estimation of image and deformation parameters in motion corrected PET. *in* 'IEEE Nuclear Science Symposium and Medical Imaging Conference' Vol. 5 pp. 3290–3294.

Jan S, Santin G, Strul D, Staelens S, Assié K, Autret D, Avner S, Barbier R, Bardiés M, Bloomfield P M, Brasse D, Breton V, Bruyndonckx P, Buvat I, Chatziioannou A F, Choi Y, Chung Y H, Comtat C, Donnarieix D, Ferrer L, Glick S J, Groiselle C J, Guez D, Honore P F, Kerhoas-Cavata S, Kirov A S, Kohli V, Koole M, Krieguer M, van der Laan D J, Lamare F, Largeron G, Lartizien C, Lazaro D, Maas M C, Maigne L, Mayet F, Melot F, Merheb C, Pennacchio E, Perez J, Pietrzyk U, Rannou F R, Rey M, Schaart D R, Schmidtlein C R, Simon L, Song T Y, Vieira J M, Visvikis D, Van de Walle R, Wieërs E and Morel C 2004. GATE: a simulation toolkit for PET and SPECT. *Physics in Medicine and Biology* **49**(19), 4543–4561.

Jones T 2002. Molecular imaging with PET, the future challenges. *British Journal of Radiology* **75**(SPEC. ISS.), S6–S15.

Judenhofer M S, Catana C, Swann B K, Siegel S B, Jung W I, Nutt R E, Cherry S R, Claussen C D and Pichler B J 2007. PET/MR images acquired with a compact MR-compatible PET detector in a 7-T magnet. *Radiology* **244**(3), 807–814.

Kadrmas D J, Casey M E, Black N F, Hamill J J, Panin V Y and Conti M 2009. Experimental comparison of lesion detectability for four fully-3D PET reconstruction schemes. *IEEE Transactions on Medical Imaging* **28**(4), 523–534.

Karakatsanis N A, Lodge M A, Tahari A K, Zhou Y, Wahl R L and Rahmim A 2013. Dynamic whole-body PET parametric imaging: I. Concept, acquisition protocol optimization and clinical application. *Physics in Medicine and Biology* **58**(20), 7391–7418.

Khalaf M, Abdel-Nabi H, Baker J, Shao Y, Lamonica D and Gona J 2008. Relation between nodule size and 18F-FDG-PET SUV for malignant and benign pulmonary nodules. *Journal of Hematology and Oncology* **1**(13).

Killoran J H, Gerbaudo V H, Mamede M, Ionascu D, Park S J and Berbeco R 2011. Motion artifacts occurring at the lung/diaphragm interface using 4D CT attenuation correction of 4D PET scans. *Journal of Applied Clinical Medical Physics* **12**(4), 261–274.

Kim J H, Czernin J, Allen-Auerbach M S, Halpern B S, Fueger B J, Hecht J R, Ratib O, Phelps M E and Weber W A 2005. Comparison between 18F-FDG PET, in-line PET/CT, and software fusion for restaging of recurrent colorectal cancer. *Journal of Nuclear Medicine* **46**(4), 587–595.

Kinahan P E, Hasegawa B H and Beyer T 2003. X-ray-based attenuation correction for positron emission tomography/computed tomography scanners. *Seminars in nuclear medicine* **33**(3), 166–179.

Kinahan P E, Townsend D W, Beyer T and Sashin D 1998. Attenuation correction for a combined 3D PET/CT scanner. *Medical Physics* **25**(10), 2046–2063.

King A P, Buerger C, Tsoumpas C, Marsden P K and Schaeffter T 2012. Thoracic respiratory motion estimation from MRI using a statistical model and a 2D image navigator. *Medical Image Analysis* **16**(1), 252–264.

King A P, Rhode K S, Razavi R S and Schaeffter T 2009. An adaptive and predictive respiratory motion model for image-guided interventions: theory and first clinical application. *IEEE Transactions on Medical Imaging* **28**(12), 2020–2032.

King A P, Tsoumpas C, Buerger C, Schulz V, Marsden P and Schaeffter T 2011. Real-time respiratory motion correction for simultaneous PET-MR using an MR-derived motion model. *in* ‘IEEE Nuclear Science Symposium and Medical Imaging Conference’ pp. 3589–3594.

Klein G J, Reutter B W, Ho M W, Reed J H and Huesman R H 1998. Real-time system for respiratory-cardiac gating in positron tomography. *IEEE Transactions on Nuclear Science* **45**(4), 2139–2143.

Kolbitsch C, Prieto C, Tsoumpas C, Buerger C and Schaeffter T 2012. T1- And T2-Weighted MR Acquisition for Bulk Motion Correction for Simultaneous PET-MR. *in* ‘ISMRM Annual Meeting’.

Lalush D S, Cui L and Tsui B M W 1996. A priori motion models for four-dimensional reconstruction in gated cardiac SPECT. *in* ‘IEEE Nuclear Science Symposium and Medical Imaging Conference’ Vol. 3 pp. 1923–1927.

Lamare F, Carbayo M J, Cresson T, Kontaxakis G, Santos A, Cheze-Le Rest C, Reader A J and Visvikis D 2007b. List mode based reconstruction for respiratory motion correction in PET using non rigid body transformations. *Physics in Medicine and Biology* **52**(17), 5187–5204.

Lamare F, Cresson T, Savean J, Cheze-Le Rest C, Reader A J and Visvikis D 2007a. Respiratory motion correction for PET oncology applications using affine transformation of list mode data. *Physics in Medicine and Biology* **52**(1), 121–140.

Langen K M and Jones D T L 2001. Organ motion and its management. *International Journal of Radiation Oncology Biology Physics* **50**(1), 265–278.

Lartizien C, Kinahan P E and Comtat C 2004. A lesion detection observer study comparing 2-dimensional versus fully 3-dimensional whole-body PET imaging protocols. *Journal of Nuclear Medicine* **45**(4), 714–723.

Levin C S and Hoffman E J 1999. Calculation of positron range and its effect on the fundamental limit of positron emission tomography system spatial resolution. *Physics in Medicine and Biology* **44**(3), 781–799.

Lewellen T, Harrison R and Vannoy S 1998 *The SIMSET program. In: M. Ljungberg, S. E. Strand and M. King (eds) Monte Carlo calculations in nuclear medicine* IOP Publishers.

Li T, Thorndyke B, Schreibmann E, Yang Y and Xing L 2006. Model-based image reconstruction for four-dimensional PET. *Medical Physics* **33**(5), 1288–1298.

Likar B and Pernus F 2001. A hierarchical approach to elastic registration based on mutual information. *Image and Vision Computing* **19**, 33–44.

Little J A, Hill D L G and Hawkes D J 1996. Deformations incorporating rigid structures. *Computer Vision and Image Understanding* **66**, 223–232.

Liu C, Pierce II L A, Alessio A M and Kinahan P E 2009. The impact of respiratory motion on tumor quantification and delineation in static PET CT imaging. *Physics in Medicine and Biology* **54**(24), 7345–7362.

Livieratos L, Stegger L, Bloomfield P M, Schafers K, Bailey D L and Camici P G 2005. Rigid-body transformation of list-mode projection data for respiratory motion correction in cardiac PET. *Physics in Medicine and Biology* **50**(14), 3313–3322.

Loening M A and Gambhir S S 2001. AMIDE: A Free Software Tool for Multimodality Medical Image Analysis. *Molecular Imaging* **2**(3), 131–137.

Lowe V J, Fletcher J W, Gobar L, Lawson M, Kirchner P, Valk P, Karis J, Hubner K, Delbeke D, Heiberg E V, Patz E F and Coleman R E 1998. Prospective investigation of positron emission tomography in lung nodules. *Journal of Clinical Oncology* **16**(3), 1075–1084.

MacDonald L and Dahlbom M 1998. Parallax correction in PET using depth of interaction information. *IEEE Transactions on Nuclear Science* **45**(4), 2232–2237.

Marache-Francisco S, Lamare F, Fayad H, Visvikis D, Prost R, Rouet J M and Lartizien C 2010. Impact of respiratory motion correction on the detection of small lesions in whole-body PET imaging: A simulation study. *in* 'IEEE Nuclear Science Symposium and Medical Imaging Conference' pp. 3531–3533.

Martinez-Möller A, Souvatzoglou M, Delso G, Bundschuh R A, Chefd'hotel C, Ziegler S I, Navab N, Schwaiger M and Nekolla S G 2009. Tissue classification as a potential approach for attenuation correction in whole-body PET/MRI: Evaluation with PET/CT data. *Journal of Nuclear Medicine* **50**(4), 520–526.

McClelland J R, Hawkes D J, Schaeffter T and King A P 2013. Respiratory motion models: A review. *Medical Image Analysis* **17**(1), 19–42.

McRobbie D W, Moore E A, Graves M J and Prince M R 2006 *Getting in tune: resonance and relaxation. In: D. W. McRobbie, E. A. Moore, M. J. Graves and M. R. Prince (eds) MRI from Picture to Proton* Cambridge.

Meikle S R and Badawi R D 2005 *Quantification techniques in PET. In: D. L. Bailey, D. W. Townsend, P. E. Valk and M. N. Maisey (eds) Positron emission tomography: Basic sciences* Springer London.

Meirelles G S P, Erdi Y E, Nehmeh S A, Squire O D, Larson S M, Humm J L and Schöder H 2007. Deep-inspiration breath-hold PET/CT: Clinical findings with a new technique for detection and characterization of thoracic lesions. *Journal of Nuclear Medicine* **48**(5), 712–719.

Melcher C L 2000. Scintillation crystals for PET. *Journal of Nuclear Medicine* **41**(6), 1051–1055.

Metz C E 1986. ROC methodology in radiologic imaging. *Investigative Radiology* **21**(9), 720–733.

Metz C E 2000 *Fundamental ROC analysis. In: R. L. Van Metter, J. Beutel and H. L. Kundel (eds) Handbook of Medical Imaging Volume 1. Physics and Psychophysics* Bellingham: SPIE Press.

Metz C E, Shen J, Wang P and Kronman H 1988. CLABROC, LABROC." Chicago: University of Chicago, Department of Radiology and the Franklin McLean Memorial Research Institute. *Medical Decision Making* .

Michel C, Sibomana M, Boi A, Bernard X, Lonneux M, Defrise M, Comtat C, Kinahan P E and Townsend D W 1998. Preserving Poisson characteristics of PET data with weighted OSEM reconstruction. *in* 'IEEE Nuclear Science Symposium Conference' Vol. 2 pp. 1323–1329.

Morey A M and Kadrmas D J 2013. Effect of varying number of OSEM subsets on PET lesion detectability. *Journal of Nuclear Medicine Technology* **41**(4), 268–273.

Myers K J and Barrett H H 1987. Addition of a channel mechanism to the ideal-observer model. *Journal of the Optical Society of America. A, Optics and image science* **4**(12), 2447–2457.

Nakazawa M, Ohi J, Tonami H, Yoshihiro Y, Furumiya T, Furuta M, Tomoaki T, Sato M, Yamakawa Y, Nobuya H, Akazawa A and Kitamura K 2010. Development of a prototype DOI-TOF-PET scanner. *in* 'IEEE Nuclear Science Symposium and Medical Imaging Conference' pp. 2077–2080.

Nehmeh S A 2013. Respiratory motion correction strategies in thoracic PET-CT imaging. *PET clinics* **8**(1), 29–36.

Nehmeh S A and Erdi Y E 2008. Respiratory motion in positron emission tomography computed tomography: A review. *Seminars in Nuclear Medicine* **38**(3), 167–176.

Nehmeh S A, Erdi Y E, Ling C C, Rosenzweig K E, Squire O D, Braban L E, Ford E, Sidhu K, Mageras G S, Larson S M and Humm J L 2002. Effect of respiratory gating on reducing lung motion artifacts in PET imaging of lung cancer. *Medical Physics* **29**(3), 366–371.

Nehmeh S A, Erdi Y E, Pan T, Yorke E, Mageras G S, Rosenzweig K E, Schoder H, Mostafavi H, Squire O, Pevsner A, Larson S M and Humm J L 2004. Quantitation of respiratory motion during 4D-PET/CT acquisition. *Medical Physics* **31**(6), 1333–1338.

Nekolla S, Souvatzoglou M, Schachoff S, Fuerst S, Martínez-Moeller A, Ziegler S and Schwaiger M 2011. MR/PET Attenuation correction using a multi tissue model for quantification of cardiac update: Initial comparison to PET/CT. *in* 'Journal of Nuclear Medicine' Vol. 52 p. Suppl 1:440.

Nomori H, Watanabe K, Ohtsuka T, Naruke T, Suemasu K and Uno K 2004. Evaluation of F-18 fluorodeoxyglucose (FDG) PET scanning for pulmonary nodules less than 3 cm diameter, with special reference to the CT images. *Lung Cancer* **45**(1), 19–27.

Nuyts J, Michel C, Brepoels L, Ceuninck L D, Deroose C, Goffin K, Mottaghy F M, Stroobants S, Riet J V and Verscuren R 2009. Performance of MAP reconstruction for hot lesion detection in whole-body PET/CT: An evaluation with human and numerical observers. *IEEE Transactions on Medical Imaging* **28**(1), 67–73.

Osman M M, Cohade C, Nakamoto Y, Marshall L T, Leal J P and Wahl R L 2003. Clinically significant inaccurate localization of lesions with PET/CT: frequency in 300 patients. *Journal of Nuclear Medicine* **44**(2), 240–243.

Ouyang J, Li Q and El Fakhri G 2013. Magnetic resonance-based motion correction for positron emission tomography imaging. *Seminars in Nuclear Medicine* **43**(1), 60–67.

Ozhasoglu C and Murphy M J 2002. Issues in respiratory motion compensation dunng external-beam radiotherapy. *International Journal of Radiation Oncology Biology Physics* **52**(5), 1389–1399.

Pan T, Lee T Y, Rietzel E and Chen G T Y 2004. 4D-CT imaging of a volume influenced by respiratory motion on multi-slice CT. *Medical Physics* **31**(2), 333–340.

Pan T, Mawlawi O, Nehmeh S A, Erdi Y E, Luo D, Liu H H, Castillo R, Mohan R, Liao Z and Macapinlac H A 2005. Attenuation correction of PET images with respiration-averaged CT images in PET/CT. *Journal of Nuclear Medicine* **46**(9), 1481–1487.

Panin V Y, Kehren F, Michel C and Casey M 2006. Fully 3-D PET reconstruction with system matrix derived from point source measurements. *IEEE Transactions on Medical Imaging* **25**(7), 907–921.

Park J, Lee T S, Chen S, Fung G and Tsui B M W 2012. Evaluation of 4D PET reconstruction with motion correction using Channelized Hottelling Observer for lesion detection. *Journal of Nuclear Medicine* **53**(Supplement 1), 2376.

Pevsner A, Nehmeh S A, Humm J L, Mageras G S and Erdi Y E 2005. Effect of motion on tracer activity determination in CT attenuation corrected PET images: A lung phantom study. *Medical Physics* **32**(7), 2358–2362.

Phelps M E, Hoffman E J, Mullani N A and Ter-Pogossian M 1975. Application of annihilation coincidence detection to transaxial reconstructed tomography. *Journal of Nuclear Medicine* **16**(3), 210–215.

Pichler B J, Kolb A, Nägele T and Schlemmer H P 2010. PET/MRI: Paving the way for the next generation of clinical multimodality imaging applications. *Journal of Nuclear Medicine* **51**(3), 333–336.

Polycarpou I, Thielemans K, Manjeshwar R, Aguiar P, Marsden P K and Tsoumpas C 2011. Comparative evaluation of scatter correction in 3D PET using different scatter-level approximations. *Annals of Nuclear Medicine* **25**(9), 643–649.

Polycarpou I, Tsoumpas C and Marsden P K 2012. Analysis and comparison of two methods for motion correction in PET imaging. *Medical Physics* **39**(10), 6474–6483.

Qi J and Huesman R H 2002. List mode reconstruction for PET with motion compensation a simulation study. *in* ‘IEEE International Symposium of Biology and Imaging’ pp. 413–416.

Qi J and Leahy R M 2006. Iterative reconstruction techniques in emission computed tomography. *Physics in Medicine and Biology* **51**(15), 541–578.

Qian H, Manjeshwar R M and Thielemans K 2010. A comparative study of multiple scatters in 3D PET. *in* ‘IEEE Nuclear Science Symposium and Medical Imaging Conference’ pp. 2700–2702.

Qiao F, Pan T, Clark Jr J W and Mawlawi O R 2006. A motion-incorporated reconstruction method for gated PET studies. *Physics in Medicine and Biology* **51**(15), 3769–3783.

Qiao F, Pan T, Clark Jr J W and Mawlawi O R 2007. Region of interest motion compensation for PET image reconstruction. *Physics in Medicine and Biology* **52**(10), 2675–2689.

Rahmim A, Bloomfield P, Houle S, Lenox M, Michel C, Buckley K R, Ruth T J and Sossi V 2004. Motion compensation in histogram-mode and list-mode EM reconstructions: Beyond the event-driven approach. *IEEE Transactions on Nuclear Science* **51**(5), 2588–2596.

Rahmim A, Rousset O G and Zaidi H 2007. Strategies for motion tracking and correction in PET. *PET clinics* **2**(2), 251–266.

Rahmim A and Zaidi H 2008. PET versus SPECT: Strengths, limitations and challenges. *Nuclear Medicine Communications* **29**(3), 193–207.

Ratib O and Beyer T 2011. Whole-body hybrid PET/MRI: ready for clinical use? *European Journal of Nuclear Medicine and Molecular Imaging* **38**(6), 992–995.

Raylman R R, Kison P V and Wahl R L 1999. Capabilities of two- and three-dimensional FDG-PET for detecting small lesions and lymph nodes in the upper torso: a dynamic phantom study. *European Journal of Nuclear Medicine* **26**(1), 39–45.

Reilhac A, Tomeï S, Buvat I, Michel C, Keheren F and Costes N 2008. Simulation-based evaluation of OSEM iterative reconstruction methods in dynamic brain PET studies. *NeuroImage* **39**(1), 359–368.

Reyes M, Malandain G, Koulibaly P M, González-Ballester M A and Darcourt J 2007. Model-based respiratory motion compensation for emission tomography image reconstruction. *Physics in Medicine and Biology* **52**(12), 3579–3600.

Rodriguez M, Liow J S, Thada S, Sibomana M, Chelikani S, Mulnix T, Johnson C A, Michel C, Barker W C and Carson R E 2007. Count-rate dependent component-based normalization for the HRRT. *IEEE Transactions on Nuclear Science* **54**(3), 486–495.

Rohren E M, Turkington T G and Coleman R E 2004. Clinical applications of PET in oncology. *Radiology* **231**(2), 305–332.

Rousset O G, Rahmim A, Alavi A and Zaidi H 2007. Partial volume correction strategies in PET. *PET clinics* **2**(2), 235–249.

Sánchez-Crespo A, Andreo P and Larsson S A 2004. Positron flight in human tissues and its influence on PET image spatial resolution. *European Journal of Nuclear Medicine and Molecular Imaging* **31**(1), 44–51.

Sánchez-Crespo A and Larsson S A 2006. The influence of photon depth of interaction and non-collinear spread of annihilation photons on PET image spatial resolution. *European Journal of Nuclear Medicine and Molecular Imaging* **33**(8), 940–947.

Sarikaya I, Yeung H W D, Erdi Y and Larson S M 2003. Respiratory artefact causing malpositioning of liver dome lesion in right lower lung. *Clinical Nuclear Medicine* **28**(11), 943–944.

Schäfers K P 2008a. The promise of nuclear medicine technology: Status and future perspective of high-resolution whole-body PET. *Physica Medica* **24**(2), 57–62.

Schäfers K P and Stegger L 2008b. Combined imaging of molecular function and morphology with PET/CT and SPECT/CT: Image fusion and motion correction. *Basic Research in Cardiology* **103**(2), 191–199.

Schlemmer H P, Pichler B J, Schmand M, Burbar Z, Michel C, Ladebeck R, Jattke K, Townsend D, Nahmias C, Jacob P K, Heiss W D and Claussen C D 2008. Simultaneous MR/PET imaging of the human brain: feasibility study. *Radiology* **248**(3), 1028–1035.

Schleyer P J, O'Doherty M J, Barrington S F and Marsden P K 2009. Retrospective data-driven respiratory gating for PET/CT. *Physics in Medicine and Biology* **54**(7), 1935–1950.

Schleyer P J, O'Doherty M J and Marsden P K 2011. Extension of a data-driven gating technique to 3D, whole body PET studies. *Physics in Medicine and Biology* **56**(13), 3953–3965.

Schreiter N F, Nogami M, Steffen I, Pape U F, Hamm B, Brenner W and Röttgen R 2012. Evaluation of the potential of PET-MRI fusion for detection of liver metastases in patients with neuroendocrine tumours. *European Radiology* **22**(2), 458–467.

Schwarz A J and Leach M O 2000. Implications of respiratory motion for the quantification of 2D MR spectroscopic imaging data in the abdomen. *Physics in Medicine and Biology* **45**(8), 2105–2116.

Schwenzer N F, Schmidt H and Claussen C D 2012. Whole-body MR/PET: Applications in abdominal imaging. *Abdominal Imaging* **37**(1), 20–28.

Segars W P and Tsui B M W 2009. MCAT to XCAT: The evolution of 4-D computerized phantoms for imaging research. *Proceedings of the IEEE* **97**(12), 1954–1968.

Seppenwoolde Y, Berbeco R I, Nishioka S, Shirato H and Heijmen B 2007. Accuracy of tumor motion compensation algorithm from a robotic respiratory tracking system: a simulation study. *Medical Physics* **34**(7), 2774–2784.

Seppenwoolde Y, Shirato H, Kitamura K, Shimizu S, Herk M V, Lebesque J V and Miyasaka K 2002. Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy. *International Journal of Radiation Oncology Biology Physics* **53**(4), 822–834.

Shao L, Freifelder R and Karp J S 1994. Triple energy window scatter correction technique in PET. *IEEE Transactions on Medical Imaging* **13**(4), 641–648.

Shepp L A and Vardi Y 1982. Maximum likelihood reconstruction for emission tomography. *IEEE Transactions on Medical Imaging* **1**(2), 113–122.

Shidahara M, Inoue K, Maruyama M, Watabe H, Taki Y, Goto R, Okada K, Kinomura S, Osawa S, Onishi Y, Ito H, Arai H and Fukuda H 2006. Predicting human performance by channelized Hotelling observer in discriminating between Alzheimer's dementia and controls using statistically processed brain perfusion SPECT. *Annals of Nuclear Medicine* **20**(9), 605–613.

Siebenthal M V, Székely G, Gamper U, Boesiger P, Lomax A and Cattin P 2007. 4D MR imaging of respiratory organ motion and its variability. *Physics in Medicine and Biology* **52**(6), 1547–1564.

Slates R B, Farahani K, Shao Y, Marsden P K, Taylor J, Summers P E, Williams S, Beech J and Cherry S R 1999. A study of artefacts in simultaneous PET and MR imaging using a prototype MR compatible PET scanner. *Physics in Medicine and Biology* **44**(8), 2015–2027.

Smith W E and Barrett H H 1986. Hotelling trace criterion as a figure of merit for the optimization of imaging systems. *Journal of the Optical Society of America. A, Optics and image science* **3**(5), 717–725.

Sonke J, Lebesque J and van Herk M 2008. Variability of Four-Dimensional Computed Tomography Patient Models. *International Journal of Radiation Oncology Biology Physics* **70**(2), 590–598.

Soret M, Bacharach S L and Buvat I 2007. Partial-volume effect in PET tumor imaging. *Journal of Nuclear Medicine* **48**(6), 932–945.

Surti S, Kuhn A, Werner M E, Perkins A E, Kolthammer J and Karp J S 2007. Performance of Philips Gemini TF PET/CT scanner with special consideration for its time of flight imaging capabilities. *Journal of Nuclear Medicine* **48**(3), 471–480.

Swensson R G 1996. Unified measurement of observer performance in detecting and localizing target objects on images. *Medical Physics* **23**(10), 1709–1726.

Swets J A and Pickett R M 1982 *Evaluation of diagnostic systems: Methods from signal detection theory* Academic Press, New York.

Ter-Pogossian M M, Phelps M E, Hoffman E J and Mullani N A 1975. A positron-emission transaxial tomograph for nuclear imaging (PET). *Radiology* **114**(1), 89–98.

Teräs M, Tolvanen T, Johansson J, Williams J and Knuuti J 2007. Performance of the new generation of whole-body PET/CT scanners: Discovery STE and Discovery VCT. *European Journal of Nuclear Medicine and Molecular Imaging* **34**(10), 1683–1692.

Thielemans K, Manjeshwar R M, Tao X and Asma E 2006. Lesion detectability in motion compensated image reconstruction of respiratory gated PET/CT. *in* 'IEEE Nuclear Science Symposium and Medical Imaging Conference' Vol. 6 pp. 3278–3282.

Thielemans K, Manjeshwar R M, Tsoumpas C and Jansen F P 2007. A new algorithm for scaling of PET scatter estimates using all coincidence events. *in* 'IEEE Nuclear Science Symposium and Medical Imaging Conference' Vol. 5 pp. 3586–3590.

Thielemans K, Mustafovic S and Schnorr L 2003. Image reconstruction of motion corrected sinograms. *in* 'IEEE Nuclear Science Symposium and Medical Imaging Conference' Vol. 4 pp. 2401–2406.

Thielemans K, Schleyer P, Marsden P K, Manjeshwar R M, Wollenweber S D and Ganin A 2013. Comparison of different methods for data-driven respiratory gating of PET data. *in* 'IEEE Nuclear Science Symposium and Medical Imaging Conference'.

Thielemans K, Tsoumpas C, Mustafovic S, Beisel T, Aguiar P, Dikaios N and Jacobson M W 2012. STIR: Software for tomographic image reconstruction release 2. *Physics in Medicine and Biology* **57**(4), 867–883.

Thoen H, Keereman V, Mollet P, Holen R V and Vandenberghe S 2013. Influence of detector pixel size, TOF resolution and DOI on image quality in MR-compatible whole-body PET. *Physics in Medicine and Biology* **58**(18), 6459–6479.

Tong S, Alessio A M and Kinahan P E 2010. Noise and signal properties in PSF-based fully 3D PET image reconstruction: An experimental evaluation. *Physics in Medicine and Biology* **55**(5), 1453–1473.

Townsend D W 2008. Dual-modality imaging: Combining anatomy and function. *Journal of Nuclear Medicine* **49**(6), 938–955.

Townsend D W, Spinks T, Jones T, Geissbuhler A, Defrise M, Gilardi M C and Heather J 1989. Three dimensional reconstruction of PET data from multi-ring camera. *IEEE Transcactions of Nuclear Science* **36**(1), 1056–1065.

Tsoumpas C, Aguiar P, Nikita K S, Ros D and Thielemans K 2004. Evaluation of the single scatter simulation algorithm implemented in the STIR library. *in* 'IEEE Nuclear Science Symposium and Medical Imaging Conference' Vol. 6 pp. 3361–3365.

Tsoumpas C, Aguiar P, Ros D, Dikaios N and Thielemans K 2005. Scatter simulation including double scatter. *in* 'IEEE Nuclear Science Symposium and Medical Imaging Conference' Vol. 3 pp. 1615–1619.

Tsoumpas C, Buerger C, King A P, Mollet P, Keereman V, Vandenberghe S, Schleyer P, Schaeffter T, Schulz V and Marsden P K 2011. Fast generation of 4D PET-MR data from real dynamic MR acquisitions. *Physics in Medicine and Biology* **56**(20), 6597–6613.

Tsoumpas C and Gaitanis A 2013a. Modeling and simulation of 4D PET-CT and PET-MR images. *PET clinics* **8**(1), 95–110.

Tsoumpas C, MacKewn J E, Halsted P, King A P, Buerger C, Totman J J, Schaeffter T and Marsden P K 2010. Simultaneous PET-MR acquisition and MR-derived motion fields for correction of non-rigid motion in PET. *Annals of Nuclear Medicine* **24**(10), 745–750.

Tsoumpas C, Polycarpou I, Thielemans K, Buerger C, King A P, Schaeffter T and Marsden P K 2013b. The effect of regularization in motion compensated PET image reconstruction: A realistic numerical 4D simulation study. *Physics in Medicine and Biology* **58**(6), 1759–1773.

Van Laere K, Koole M, Lemahieu I and Dierckx R 2001. Image filtering in single-photon emission computed tomography: Principles and applications. *Computerized Medical Imaging and Graphics* **25**(2), 127–133.

van Velden F H P, Kloet R W, van Berckel B N M, Molthoff C F M, de Jong H W, Lammertsma A A and Boellaard R 2008. Impact of attenuation correction strategies on the quantification of High Resolution Research Tomograph PET studies. *Physics in Medicine and Biology* **53**(1), 99–118.

Vanderhoek M, Perlman S B and Jeraj R 2012. Impact of the definition of peak standardized uptake value on quantification of treatment response. *Journal of Nuclear Medicine* **53**(1), 4–11.

Vansteenkiste J F and Stroobants S G 2001. The role of positron emission tomography with 18F-fluoro-2-deoxy-D-glucose in respiratory oncology. *European Respiratory Journal* **17**(4), 802–820.

Visvikis D, Barret O, Fryer T D, Lamare F, Turzo A, Bizais Y and Cheze-Le Rest C 2004. Evaluation of respiratory motion effects in comparison with other parameters affecting PET image quality. *in* 'IEEE Nuclear Science Symposium and Medical Imaging Conference' Vol. 6 pp. 3668–3672.

Visvikis D, Lamare F, Bruyant P, Boussion N and Cheze-Le Rest C 2006. Respiratory motion in positron emission tomography for oncology applications: Problems and solutions. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **569**(2 SPEC. ISS.), 453–457.

Wahl R L, Jacene H, Kasamon Y and Lodge M A 2009. From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. *Journal of Nuclear Medicine* **50**(1), S122–S150.

Walker M D, Asselin M C, Julyan P J, Feldmann M, Talbot P S, Jones T and Matthews J C 2011. Bias in iterative reconstruction of low statistics PET data: Benefits of a resolution model. *Physics in Medicine and Biology* **56**(4), 931–949.

Wallach D, Lamare F, Kontaxakis G and Visvikis D 2012. Super resolution in respiratory synchronized positron emission tomography. *IEEE Transactions on Medical Imaging* **31**(2), 438–448.

Wang G and Qi J 2012. Penalized likelihood PET image reconstruction using patch-based edge-preserving regularization. *IEEE Transactions on Medical Imaging* **31**(12), 2194–2204.

Watson C C 2000. New, faster, image-based scatter correction for 3D PET. *IEEE Transactions on Nuclear Science* **47**(4), 1587–1594.

Watson C C, Newport D and Casey M E 1996 *A single-scatter simulation technique for scatter correction in 3D PET. In: P. Grangeat and J. Amans (eds) Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine* Springer London.

Watson C C, Newport D, Casey M E, Dekemp R A, Beanlands R S and Schmand M 1997. Evaluation of simulation-based scatter correction for 3-D PET cardiac imaging. *IEEE Transactions on Nuclear Science* **44**(1), 90–97.

Weber W A, Petersen V, Schmidt B, Tyndale-Hines L, Link T, Peschel C and Schwaiger M 2003. Positron emission tomography in non-small-cell lung cancer: Prediction of response to chemotherapy by quantitative assessment of glucose use. *Journal of Clinical Oncology* **21**(14), 2651–2657.

Wells W M, Viola P, Atsumi H, Nakajima S and Kikinis R 1996. Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis* **1**(1), 35–51.

Wilson D W and Tsui B M W 1993. Noise properties of filtered -backprojection and ML-EM reconstructed emission tomographic images. *IEEE Transactions on Nuclear Science* **40**(4 pt 1), 1198–1203.

Wrenn F R, Good M L and Handler P 1951. The use of positron-emitting radioisotopes for the localization of brain tumors. *Science* **113**(2940), 525–527.

Xia T, Alessio A M, De Man B, Manjeshwar R, Asma E and Kinahan P E 2012. Ultra-low dose CT attenuation correction for PET/CT. *Physics in Medicine and Biology* **57**(2), 309–328.

Xu Q, Yuan K and Ye D 2011. Respiratory motion blur identification and reduction in ungated thoracic PET imaging. *Physics in Medicine and Biology* **56**(14), 4481–4498.

Zaidi H and Alavi A 2007a. Current trends in PET and combined (PET/CT and PET/MR) systems design. *PET clinics* **2**(2), 109–123.

Zaidi H and Del Guerra A 2011b. An outlook on future design of hybrid PET/MRI systems. *Medical Physics* **38**(10), 5667–5689.

Zaidi H and Koral K F 2004. Scatter modelling and compensation in emission tomography. *European Journal of Nuclear Medicine and Molecular Imaging* **31**(5), 761–782.

Zaidi H and Montandon M L 2007c. Scatter compensation techniques in PET. *PET clinics* **2**(2), 219–234.

Zaidi H, Ojha N, Morich M, Griesmer J, Hu Z, Maniawski P, Ratib O, Izquierdo-Garcia D, Fayad Z A and Shao L 2011a. Design and performance evaluation of a whole-body Ingenuity TF PET- MRI system. *Physics in Medicine and Biology* **56**(10), 3091–3106.

Zaidi H and Tsui B M W 2009. Review of computational anthropomorphic anatomical and physiological models. *Proceedings of the IEEE* **97**(12), 1938–1953.

Zaidi H and Xu X G 2007b. Computational anthropomorphic models of the human anatomy: the path to realistic Monte Carlo modeling in radiological sciences. *Annual Review of Biomedical Engineering* **9**, 471–500.

Zanzonico P 2004. Positron emission tomography: a review of basic principles, scanner design and performance, and current systems. *Seminars in Nuclear Medicine* **34**(2), 87–111.

Zubal I G, Harrell C R, Smith E O, Rattner Z, Gindi G and Hoffer P B 1994. Computerized three-dimensional segmented human anatomy. *Medical Physics* **21**(2), 299–302.