Phantom design to enable harmonisation of quantitative PET imaging for lung disease

Francesca Alice Allen Leek

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
of
University College London.

Department of Medical Physics and Biomedical Engineering
University College London

July 31, 2025

I, Francesca Alice Allen Leek, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

Idiopathic pulmonary fibrosis (IPF) and chronic obstructive pulmonary disease (COPD) are characterised by distortions in normal lung parenchyma, resulting in regions of dense scar tissue and/or large air pockets. ^{18}F-fluorodeoxyglucose (FDG) positron emission tomography (PET)-computed tomography (CT) can be used to assess lung inflammation and fibrosis-related processes due to increased macrophage and neutrophil activity. However, accurate quantification of radiopharmaceutical uptake from lung PET-CT is challenging, due, in part, to large variations in fractions of air. This project developed digital test objects to validate pre-existing attenuation and air fraction correction (AFC) methods and explored the need for resolution matching between PET and CT. Resolution dependence on radiopharmaceutical uptake, tissue structure, and reconstruction convergence were assessed.

Current harmonisation efforts in PET-CT focus on oncology and brain studies; however, there is a similar need for studies of diffuse lung disease. Practical implementation of a resolution matching method was investigated with ^{18}F-FDG scans of a modified commercially available thorax phantom on clinical PET-CT scanners. Novel physical phantoms were then designed, and refined, to better approximate diseased lung density and structure, and to simplify parameter determination and validation. The designed phantoms comply with the requirements of being easy to prepare, quick to process, while having features relevant to patient data. The latter was ensured through analytic and Monte Carlo simulations.

This work sets the groundwork for the assessment of the accuracy and reproducibility of lung PET imaging in a multi-centre setting.

Impact statement

^{18}F-FDG PET has been increasingly applied as a molecular biomarker of pulmonary inflammation, which is commonly associated with respiratory diseases, and can play an important role in diagnosis and assessment of disease progression and treatment response. Quantitative PET imaging can provide more accurate, less observer-dependent, metrics than visual interpretation alone.

Consensus recommendations compiled by academic, clinical, and industrial partners in 2020 [1], highlighted the disparate approaches to image acquisition, reconstruction, and data analysis of ^{18}F-FDG PET-CT in lung imaging. This lack of standardisation and harmonisation limits data comparison in the multi-centre clinical trials required for robust biomarker development. The work in this thesis aimed to address some of the challenges highlighted in these recommendations.

The recommendations propose correcting for the variable air content within the lung to improve lung tissue-specific ^{18}F-FDG quantification. The air fraction correction (AFC) has previously been shown to alter image interpretation in diffuse lung disease imaging, and is vital for quantitative comparison of uptake between different density regions in a single patient, between patients and/or between patients and controls. The work in this thesis validated, and optimised, the pre-existing methodology for AFC. The groundwork was laid for further investigation of partial volume correction (PVC) techniques, which are necessary for accurate quantification in the case of non-uniform lung tissue uptake. The results on image reconstruction convergence for diffuse activity distributions and low count data pave the way for standardisation of reconstruction parameters for PET lung imaging.

The consensus recommendations stated the need for phantoms, that model relevant aspects of lung physiology, in order to establish harmonisation standards for multi-centre studies. This thesis developed a patient-realistic digital phantom and iterated the design of a physical phantom that mimics the density and uptake patterns seen in diffuse lung disease. These phantoms were used to validate the proposed AFC methodology and conducted preliminary investigations into the optimal reconstruction parameters for diffuse lung disease.

The variability in the quantitative accuracy between PET-CT scanners of different makes and models for lung imaging must be understood and characterised if data from different scanners is to be collated and compared. This work assessed the variability across three PET-CT scanners in a single centre with a novel physical phantom design that mimicked diffuse lung disease density and ^{18}F-FDG uptake patterns – this is the first step on the long road to harmonisation.

UCL Research Paper Declaration Form: referencing the doctoral candidate’s own published work(s)

1. **For a research manuscript that has already been published** (if not yet published, please skip to section 2)**:**

(a) **What is the title of the manuscript?**

Optimisation of the air fraction correction for lung PET/CT: addressing resolution mismatch.

(b) **Please include a link to or doi for the work:**

https://ejnmmiphys.springeropen.com/articles/10.1186/s40658-023-00595-y

(c) **Where was the work published?**

European Journal of Nuclear Medicine and Molecular Imaging Physics

(d) **Who published the work?**

SpringerOpen

(e) **When was the work published?**

5th December 2023

(f) **List the manuscript’s authors in the order they appear on the publication:**

Francesca Leek, Cameron Anderson, Andrew P. Robinson, Robert M. Moss, Joanna C. Porter, Helen S. Garthwaite, Ashley M. Groves, Brian F. Hutton, Kris Thielemans.

(g) **Was the work peer reviewed?**

Yes.

(h) **Have you retained the copyright?**

Yes.

(i) **Was an earlier form of the manuscript uploaded to a preprint server (e.g. medRxiv)? If 'Yes', please give a link or doi**

If 'No', please seek permission from the relevant publisher and check the box next to the below statement:

☒ *I acknowledge permission of the publisher named under 1d t o include in this thesis portions of the publication named as included in 1c.*

2. **For a research manuscript prepared for publication but that has not yet been published** (if already published, please skip to section 3)**:**

(a) **What is the current title of the manuscript?**

(b) **Has the manuscript been uploaded to a preprint server e.g. medRxiv'? If 'Yes', please please give a link or doi:**

(c) **Where is the work intended to be published?**

(d) **List the manuscript's authors in the intended authorship order:**

(e) **Stage of publication:**

3. **For multi-authored work, please give a statement of contribution covering all authors** (if single-author, please skip to section 4)**:**

FL carried out the data generation and analyses and prepared the manuscript. CA assisted in the measured data acquisition. APR contributed to the project funding. FL, CA, APR, RMM, BFH and KT participated in the method development and interpretation of the data. JCP, HSG and AMG provided patient data and clinical context. All authors read and approved the final manuscript.

4. **In which chapter(s) of your thesis can this material be found?**

Chapter 4 and Appendices B, C and D.

e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)**:**

Candidate: FRANCESCA LEEK
Date: 01/03/2025

Supervisor/Senior Author signature (where appropriate)**: KRIS THIELEMANS**
Date: 1 Mar 2025

Acknowledgements

I acknowledge funding support from EPSRC Industrial CASE award (EP/T517628/1), the UK National Physical Laboratory (NPL) through the National Measurement System Programmes Unit of the UK's Department of Business, Energy and Industrial Strategy, GlaxoSmithKline (BIDS3000035300), and the University College London (UCL) EPSRC Centre for Doctoral Training in Intelligent, Integrated Imaging in Healthcare (i4Health; EP/S021930/1).

Firstly, I wish to thank my primary supervisor, Kris Thielemans, for his endless support, thought-provoking discussions and advice. I would like to thank my secondary supervisor Robert Moss for his encouragement throughout. Thank you to my industrial supervisors Andrew Robinson, Andrew Patterson and Fred Wilson for their vital contributions and guidance. Thank you also to Brian Hutton for his direction and invaluable feedback.

I would also like to thank members of the clinical physics team at the Institute of Nuclear Medicine, UCL, for their time – Cameron Anderson for the many hours he spent filling and scanning phantoms with me and John Dickson for his feedback and encouragement. Thank you to Ashley Groves, Joanna Porter and Helen Garthwaite for providing clinical data.

Contents

List of Figures

List of Tables

Acronyms

6MWT 6-minute walking test.

ABS acrylonitrile butadiene styrene.

AC attenuation correction.

ACD annihilation coincidence detection.

ACF attenuation correction factor.

AF air fracions.

AFC air fraction correction.

BGO bismuth germanate.

BSREM block sequential regularised expectation maximisation.

CIs confidence intervals.

CLE centrilobular emphysema.

COPD chronic obstructive pulmonary disease.

CoV coefficient of variation.

CSF cerebrospinal fluid.

CT computed tomography.

CTAC CT-based attenuation correction.

D710 Discovery 710.

DLCO carbon monoxide diffusion.

EANM European Association of Nuclear Medicine.

EARL EANM Research Ltd..

EM expectation-maximisation.

EPF expanding polyurethane foam.

EPS expanded polystyrene.

FBP filtered backprojection.

FDG fluorodeoxyglucose.

FOV field of view.

FWHM full width half maximum.

GATE Geant4 Application for Tomographic Emission.

GM grey matter.

GT ground truth.

HL healthy lung.

HPDL high precision deep learning.

HRCT high resolution CT.

HU Hounsfield units.

ILC interstitial lung change.

ILD interstitial lung disease.

IPF idiopathic pulmonary fibrosis.

kBq kilobecquerel.

LAC linear attenuation coefficient.

LDIs low density inserts.

LIR local impulse response.

LM list-mode.

LOR line of response.

LSO lutetium oxyorthosilicate.

LYSO lutetium–yttrium oxyorthosilicate.

MFs material fractions.

MG Müller-Gärtner.

MKA matched kernel approach.

ML maximum likelihood.

ml millilitre.

MRI magnetic resonance imaging.

NECR noise equivalent count rate.

NIST National Institute of Standards and Technology.

nWMSE normalised weighted mean square error.

OSEM ordered subsets expectation-maximisation.

PDL pre-trained neural network.

PET positron emission tomography.

PF post-filter.

PLA polylactic acid.

PMMA polymethylmethacrylate.

PMT photomultiplier tube.

PoC proof of concept.

PSF point spread function.

PVC partial volume correction.

PVE partial volume effect.

QC quality control.

QIBA quantitative imaging biomarker alliance.

RD relative difference.

RM resolution modelling.

RMSE root mean square error.

ROI region of interest.

SiPM silicon photomultiplier.

SIRF Synergistic Image Reconstruction Framework.

SLA stereolithography.

SNR signal to noise ratio.

SPECT single photon emission computed tomography.

SR singles rate.

STIR Software for Tomographic Image Reconstruction.

SUV standardised uptake value.

TACs time-activity curves.

TBR target-to-background ratio.

TFE tissue fraction effect.

TMMs tissue mimicking materials.

TOF time-of-flight.

UCLH University College London Hospitals.

VOI volume of interest.

WM white matter.

WMSE weighted mean square error.

XCAT 4D extended cardiac-torso.

.

Chapter 1

Introductory material

Respiratory disease affects one in five people and is the third biggest cause of death in England, after cancer and cardiovascular disease [2]. Inflammation characterises several lung diseases including pneumonia, asthma and cystic fibrosis. Positron emission tomography (PET) is an imaging technique offering functional information, for instance on metabolism. This work focuses on methodology for more reliable quantification of PET images of interstitial lung disease (ILD) and chronic obstructive pulmonary disease (COPD).

ILD, is a group of diseases that affects the tissue around the alveoli (”interstitium”). Idiopathic pulmonary fibrosis (IPF) is the most common ILD pathology but has a poorly understood aetiology and pathogenesis. It is characterised by an increase in lung scar tissue, which distorts the normal parenchyma architecture, leading to reduction in air and alterations in blood volume, depending on the stage of fibrosis [3].

COPD is a spectrum of obstructive airway diseases, including emphysema. It is irreversible and progressive and is characterised by destruction of alveolar septa and pulmonary capillaries, leading to decreased elastic recoil and results in air trapping. Pulmonary bullae, focal regions of emphysema, can result in compression of adjacent lung [4].

Treatment options for these diseases remain limited due to the pathogenesis being poorly understood and the number of clinical biomarkers to identify aggressive disease phenotypes is limited. Furthermore, 50 % of drugs fail in phase III due to lack of demonstrable efficacy. As inflammation is commonly associated with respiratory diseases, robust molecular biomarkers of inflammation can play an important role in diagnosing lung abnormalities and assessing disease progression and treatment response [3]. The increase in activated neutrophils, which have been found to have high glucose metabolism, makes ^{18}F-fluorodeoxyglucose (FDG) PET-computed tomography (CT) an appropriate imaging technique for both diseases [5].

Quantitative PET imaging can provide more accurate, less observer-dependent, metrics than visual interpretation alone. However, PET suffers from multiple technical, physical, and biological factors that can significantly affect quantification [6]. Accurate quantification in the lung is particularly challenging due to the presence of large variations in fractions of tissue, air, blood and water, as well as respiratory motion. Image quality is further degraded by positron range and affected by reconstruction and registration approaches.

Patients may be scanned on different PET-CT systems, whether in a single-centre, monitoring disease progression and treatment response, or when participating in multi-centre trials. In clinical practice, a wide range of PET systems are installed, some with state-of-the-art acquisition, e.g. time-of-flight (TOF) and digital detectors, and reconstruction technologies, e.g. resolution modelling (RM). These technical factors impose a significant source of variability in PET quantification. Accurate cross-calibration of systems and standardisation of methodology is crucial, and the resultant variability, reproducibility, and accuracy of PET quantification must be understood. Chen et al. highlighted the lack of widely accepted standard protocols for pulmonary ^{18}F-FDG PET-CT and the need to develop test objects that model relevant aspects of lung physiology to support the establishment of harmonisation standards [1].

The objective of this project is to develop test objects to validate pre-existing attenuation and air fraction correction methods. This will aid in the assessment of the accuracy and reproducibility of PET lung imaging in a multi-centre setting.

Chapter 2 provides a general overview of the physics of both PET and CT.

A brief overview of the reconstruction principles in PET is given. Current efforts to standardise and harmonise PET-CT imaging are described. Finally, the specific challenges of PET-CT in the lungs, specifically diffuse lung disease, are outlined.

Chapter 3 investigates the effect of mismatched resolution between PET and CT and its effect on attenuation correction (AC) and air fraction correction (AFC) in lung imaging. Simple noise-fee analytical simulations and non-TOF reconstructions are used to investigate the smoothing kernel that should be applied to the CT for AC and AFC. The extent to which the kernel for AFC varies with respect to iterative reconstruction algorithm convergence is assessed. The limitations of the AFC method for non-uniform tissue uptake are explored. The interplay of the AFC with a partial volume correction (PVC) method to correct for the spill-in / spill-out component of the partial volume effect (PVE) is investigated.

Chapter 4 investigates three different methodologies by which the reconstructed image resolution, as a function of reconstruction algorithm and convergence, can be determined for smoothing of the CT for AFC. Simulation studies were used to compare the point source insertion-and-subtraction method to two variations on the matched kernel approach (MKA). The feasibility of using a MKA methodology on measured data was assessed through phantom acquisitions.

Chapter 5 begins with a literature review of novel lung phantom designs. This is followed by the construction of a preliminary novel phantom design, that approximates diffuse lung disease. The preliminary phantom is assessed for its ability to determine and validate the kernel for AFC on a clinical PET-CT scanner. The design was iterated, constructed, and reassessed on three clinical PET-CT scanners. Finally, a preliminary study was conducted to assess if the final phantom design is suitable for harmonisation of lung studies across PET-CT scanners.

Chapter 2

Positron Emission Tomography of the lung

2.1 Introduction

Respiratory diseases are a significant health burden that affect millions of people globally [3]. Treatment options remain limited as the pathogenic mechanisms are poorly understood. Clinical biomarkers, to identify aggressive disease phenotypes with accelerated progression, are limited. This contributes to 50 % of drugs failing in phase III due to lack of demonstrable efficacy [3]. Robust molecular biomarkers of pulmonary inflammation, which is commonly associated with respiratory diseases, might be applied to aid diagnosis, treatment planning, drug discovery and therapy monitoring.

Inflammation characterises several lung diseases, including pneumonia, cystic fibrosis, COPD, asthma, and IPF [3]. Lungs contain varying amounts of air depending on the disease and its severity – COPD affected lungs have a higher air content compared to healthy controls; IPF is characterised by an increase in lung scar tissue which can result in regions of dense tissue adjacent to air pockets. Without correction, these differences in air content can cause significant variation in the quantified PET signal [3]. This work will focus on these diseases.

Inflammatory cell recruitment leads to increased glucose utilisation in the lungs, ^{18}F-FDG has therefore been widely explored as a biomarker of pulmonary inflammation [7, 5, 8]. COPD results in increased numbers of lung neutrophils and macrophages, which would be expected to increase the ^{18}F-FDG signal, despite a reduction in measured signal due to larger air fractions and reduced blood volumes as a result of emphysema [3]. IPF is characterised by interstitial pneumonia along with fibrosis. Depending on the stage of fibrosis this can lead to reduced air content, increased fibrosis, and alterations in blood volume [3]. The result is that quantification of pulmonary PET imaging remains challenging due to these variations of tissue fractions. Correction methodologies have been proposed to address these variabilities but a standardised approach has not been adopted to date [3].

2.2 PET-CT

PET-CT (or PET-magnetic resonance imaging (MRI)) combines functional and morphological modalities into single devices allowing a single-study evaluation of patients in clinical practice. The synergism of CT (or MRI) with PET provides not only the precise anatomical localisation of regions identified by the uptake of radionuclide tracers, but also the information required for correction methodologies so that the uptake value can be correctly quantified [9]. Only PET-CT is considered as part of this work due to difficulties of estimating attenuation in the lung in PET-MRI [10].

For PET imaging, patients are injected with (or inhale) a radiotracer – a biologically active molecule to target a specific biological process, combined with a positron-emitting radionuclide. ^{18}F-FDG is by far the most commonly used radiotracer for clinical studies; as a glucose analogue, the uptake of FDG reflects the glucose metabolism in tissue which allows the detection of regional alterations in glucose metabolism. ^{18}F has a half-life ($\tau_{1/2}$) of 1.82890 (23) h [11]. Other common PET radionuclides are ^{15}O ($\tau_{1/2}$ = 2.041 (6) min), ^{11}C ($\tau_{1/2}$ = 20.361 (23) min), all of which are produced in a cyclotron, and ^{68}Ga ($\tau_{1/2}$ = 67.83 (20) min), which is produced in a generator [11].

For clinical evaluation, data from a PET scanner is reconstructed into a 3D representation of the radiotracer concentration; when quantitative reconstruction methods are used, the resultant voxel values have units of kilobecquerel (kBq) per millilitre (ml). The image is often normalised to the activity concentration in the whole body, this is known as the standardised uptake value (SUV), where SUV is the activity concentration in the region of interest divided by the injected activity normalised to patient weight.

2.2.1 Physics of PET

2.2.1.1 Positron decay

In radioactive decay by positron emission ($\beta+$ decay), a proton (p^+) in the nucleus is transformed into a neutron (n); a positively charged electron, known as a positron (e^+), and a neutrino (ν) are ejected from the nucleus [12]:

$$p^+ \rightarrow n + e^+ + \nu + energy \qquad (2.1)$$

After ejection from the nucleus, the positron loses its kinetic energy in collisions with atoms of the surrounding matter until it forms a positronium with an electron. The positronium has a lifetime of approximately 10^{-10} seconds before the positron combines with the electron in an annihilation reaction, in which their masses are converted into energy. The mass-energy equivalent of each particle is 511 keV, which is released in the form of two 511 keV annihilation photons. The annihilation photons leave the site of the annihilation event in near opposing directions, this is discussed further in Section 2.5.

As the rest-mass energy of an electron, or a positron, is 511 keV, a total transition energy of 1022 keV is required. The excess transition energy above 1022 keV is shared between the positron, as kinetic energy, and the neutrino. Since the probability of annihilation increases as the kinetic energy of the positron decreases due to collisions with electrons, the positron range depends on the initial kinetic energy of the positron at emission and the density and composition of the media crossed by the positron before it annihilates. The positron range effect on the overall image resolution is relatively small for ^{18}F [13], however, the resolution degradation is greater for radionuclides emitting higher energy positrons, and in tissues of low density, such as the lung. Due to the varying composition of the lung, spatially-varying positron range effects may be seen [14]. This is discussed further in Section 2.7.3.

2.2.1.2 Photon interactions

Transmission of photons through material can be characterised by the linear attenuation coefficient (LAC), μ. The LAC is defined as the probability per unit path-length that the photon will interact with the absorber. For an ideal narrow beam of mono-energetic photons, the resulting fractional reduction of the beam intensity $-\mathrm{dI/I}$ through an absorber is proportional to the absorber thickness dx, expressed in units of inverse centimetres (cm^{-1}):

$$I(t) = I_0.exp(-\int_0^x \mu\, dx') \tag{2.2}$$

where I_0 is the incident beam intensity [15]. LACs are proportional to the density, ρ, of the absorber, it is therefore common to express the attenuation property of a material in terms of its mass attenuation coefficient, μ/ρ, in units of cm^2/g. The total mass attenuation coefficient for an interaction is given by the sum of the possible photon interaction mechanisms

$$\frac{\mu}{\rho} = \tau + \sigma + \kappa \tag{2.3}$$

where τ is the part of μ/ρ caused by the photoelectric effect, σ is the part caused by Compton scattering, and κ the part caused by pair production [12]. The relative magnitudes of τ, σ, and κ vary with photon energy (E) and the atomic number (Z) of the material. The photoelectric component also increases abruptly at energies corresponding to orbital electron binding energies of the absorber, due to the fact that the photon energy must exceed the K-shell binding energy for photoelectric absorption to occur. These abrupt changes are referred to as “K absorption edges” [12]. The attenuation coefficient for photoelectric absorption varies approximately as:

$$\tau \propto \frac{Z^3}{E^3} \tag{2.4}$$

τ decreases rapidly with increasing photon energy and increases rapidly with increasing atomic number, it is therefore the dominant effect for low photon energies in dense material i.e., X-ray imaging of cortical bone.

Compton scattering is the dominating interaction for tissues in the body over most of the nuclear medicine energy range [12]. It decreases slowly with increasing E and Z but the changes are so small that σ is usually considered to be invariant with E and Z.

The pair-production component is zero for photon energies less than the threshold energy of 1.02 MeV.

2.2.1.3 Photon detection

Near-simultaneous detection of the two annihilation photons by a ring of detectors allows their origin to be localised along a line of response (LOR) between the two detectors. This mechanism is called annihilation coincidence detection (ACD) and its ability to localise events on the basis of coincidence timing is known as electronic collimation. High detection efficiency is essential to be able to carry out ACD. Desirable properties for scintillators in PET imaging are:

- high stopping power for 511 keV photons. The stopping power is the mean distance the photon travels before it deposits all its energy. The high energy of the annihilation gammas requires high atomic number and density detectors.
- low decay time. The decay time is the time needed to process each pulse. A low decay time is desirable at high count rates, to reduce the number of random events, and enables TOF.
- high light output. Detector light output is the number of scintillation photons produced per unit of deposited energy.

Detectors are typically arranged in banks of discrete elements around the scanned object, e.g. the GE Discovery 710 (D710) uses 288 block detectors to form the ring, with 48 lutetium–yttrium oxyorthosilicate (LYSO) crystals per quad-anode photomultiplier tube (PMT). Multiple detector rings (24 for the D710) are employed and data are typically acquired in 3D acquisition mode, meaning it can be obtained for all possible LORs. The sensitivity is greater at the centre of the field of view (FOV), structures of interest are placed as close to the centre of the axial FOV as possible and bed positions are overlapped so as to reduce noise at the periphery.

Table 2.1: Scanner design for the three clinical PET-CTs investigated in this project; lutetium oxyorthosilicate (LSO); bismuth germanate (BGO); silicon photomultiplier (SiPM)

Scanner	**Crystals**	**Crystal dim. [mm^3]**	**Photo-detection**	**Axial FOV [cm]**
GE Discovery 710	LYSO	4.2 x 6.3 x 25	PMTs	15.7
Siemens Biograph Vision	LSO	3.2 x 3.2 x 20	SiPM	26.1
GE Omni Legend	BGO	4.1 x 4.1 x 30	SiPM	32.0

An energy window is used to discriminate photons that have originated from an annihilation event e.g., the GE D710 uses an energy window of 425 – 650 keV. A coincident event is assumed to have occurred when a pair of events are recorded within a specified coincidence timing window, which is typically 4 – 5 ns.

There are three types of coincidence events:

- True: a pair of photons produced from the same annihilation event are detected by opposing detectors, without any prior interaction with the medium, verifying that the annihilation event occurred somewhere within the coincident volume between the two detectors.

- Scattered: one or both of the photons from an annihilation event outside the sensitive volume for true coincidence events undergoes Compton scattering and is detected in a detector other than the one that would be appropriate for a true coincidence event.

- Random: photons from two unrelated positron annihilation events are detected in two different detectors, within the coincidence timing window, and recorded as a single coincidence event. Random coincidences are not rare events because the volume of tissue from which the photons for a random coincidence event could arise generally is much larger than the potential volume for true coincidence events.

Scattered and random events yield incorrect positional information and contribute a relatively uniform background to the image, resulting in a loss of contrast in the image.

2.2.1.4 Time-of-flight

In TOF PET, each detected photon is tagged with a detection time; if the detection time difference between two photons is smaller than a set coincidence window, the two events are considered physically correlated with the same annihilation event [16]. The difference in arrival time between the two photons depends on the position of the annihilation event along a LOR; TOF uses the difference to better locate the annihilation position.

TOF requires an accurate evaluation of the arrival time of the photons and so requires a “fast” scintillator crystal i.e., LSO or LYSO. Historically, BGO has been considered inappropriate for TOF PET due to its moderate light yield and relatively long decay time. More recent work has shown that prompt Cerenkov photons arising from 511 keV gamma interactions in BGO can be detected [17]. However, TOF PET detectors utilising Cherenkov photons in BGO are not currently in clinical use. Higher time resolutions can also be achieved through the use of SiPMs instead of PMTs e.g., the Siemens Vision PET-CT has a time resolution FWHM of 210 ps (3.15 cm) [18], whereas the GE D710 has a time resolution FWHM of 550 ps (8.25 cm) [19]. In place of TOF, the GE Omni Legend employs software corrections through a pre-trained neural network (PDL) (trained on non-TOF to TOF) that takes Bayesian penalized likelihood reconstruction (Q.Clear) images as input [20].

2.2.2 Physics of CT

In x-ray CT, a narrow beam of x-rays is rotated along the long axis of the patient and the transmitted photons are detected at each angular position. The recorded photons at each position are reconstructed to give a 3D map of the attenuation coefficient of the imaged anatomy. Tube voltages of 120 – 140 kVp produce photons with mean energies of around 40 kV. At these energies, Compton scatter becomes more probable than photoelectric absorption in soft tissue [9]. CT images are displayed as CT numbers (known as Hounsfield units (HU)), which are defined with reference to water:

$$CTnumber = \frac{\mu_{tissue} - \mu_{water}}{\mu_{water}} \times 1000 \tag{2.5}$$

where μ_{water} and μ_{tissue} are the LACs of water and tissue, respectively.

2.2.3 Corrections

Quantitative PET imaging requires that the intensity of the reconstructed image should be proportional to the amount or concentration of activity at the corresponding location in the object. Several physical effects can perturb this, the most significant are photon attenuation, scattered and random coincidences, detector efficiency variations, and scanner dead-time [15]. Corrections for each of these factors are discussed below.

Attenuation: At 511 keV, a large fraction of the emitted photons will interact in the subject before they exit the body. Only about 5 % of the 511 keV photons from the centre of the body are detected [21]. When a photon fails to travel along a straight line, due to scattering or other interactions, it is attenuated, representing the largest degradation to PET data. The probability of attenuation for a given pair of annihilation photons is independent of the position of the annihilation along the LOR, making it possible to pre-correct for this effect using an external source. PET-CT scanners employ a CT-based attenuation correction (CTAC). The CT data can be incorporated directly into the image reconstruction process [22]. In regions of non-uniform density, such as the thorax, a lack of AC and AFC, as discussed in Section 2.7.2.2, can mask the appearance of solid lesions as the tracer uptake will be elevated in the lungs [15].

PET uses mono-energetic 511 keV annihilation photons, whereas the x-ray source in CT emits photons with a broad energy spectrum from 30 to 140 keV. This presents two potential difficulties with scaling CT attenuation factors for the use with PET data. The first problem is the large difference in photon energies between PET and CT, while the second is the difference between monochromatic and wide-band energy spectra. There is no unique transformation from CT energies to 511 keV due to the possibility of independent variations in density and atomic number. Two materials with different values of Z may have similar CT values but different LACs at 511 keV, and vice versa. At 511 keV the contribution of photoelectric absorption is essentially negligible for all biological materials. In the CT energy range, however, photoelectric absorption has a larger contribution, Table 2.2.

Table 2.2: Mass attenuation coefficients, in $cm^2.g^{-1}$, for biological materials [23].

	80 keV			500 keV			Ratio of totals
Material	τ	σ	Total	τ	σ	Total	80 keV:500 keV
Air	0.006	0.161	0.167	<0.001	0.087	0.087	1.92
Water	0.006	0.178	0.184	<0.001	0.097	0.097	1.90
Muscle	0.006	0.176	0.182	<0.001	0.096	0.096	1.90
Bone	0.034	0.175	0.209	<0.001	0.093	0.093	2.26

Multi-linear scaling methods are the most commonly implemented on clinical PET-CT scanners; they convert attenuation coefficients measured with CT, typically in the range 30 - 140 keV, to appropriate values at 511 keV [24, 25]. Burger et al. [25] defined a simple bi-linear function to transform CT values into LACs at 511 keV, appropriate for human tissue. The function employed the apparent attenuation coefficient of water, as determined from ^{68}Ge/^{68}Ga transmission scans (0.093 cm^{-1}), rather than the value derived from narrow-beam experiments (0.096 cm^{-1}).

Bi-linear transformation has been shown to be sufficiently accurate for human tissue [25, 26]. In the presence of high-Z materials, such as contrast agents, quantitative errors have been noted [27, 28, 29]. It has been shown that biases in the CTAC image propagate to errors in the non-TOF PET image in the same general location [30]. Kinahan et al. demonstrated that both quantitative errors and complex artefacts occur when LACs estimated by the standard bi-linear method are incorrect [31]. Dual-energy CT methods were shown to reduce errors in the estimated LAC image of high-Z materials but the effects of noise propagation, if low-dose CT scans are used, was not investigated. Due to increased patient dose, dual-energy CT is not routinely implemented clinically. Bi-linear or tri-linear scaling of single-energy CT is most commonly used.

The attenuation correction factor (ACF) for an individual sinogram element is calculated by integrating the attenuation coefficients along the corresponding LOR [22]. Once ACFs have been determined for each sinogram element, they are applied as multiplicative correction factors either before or during iterative reconstruction, see Section 2.4.

Randoms: two approaches have been adopted to estimate the number of randoms in each LOR: (i) the “delayed window” method, which contains purely random events and is an estimate of the random events in the prompt window [32]; (ii) estimate the randoms rate from the singles rate (SR) [33]. The SR method uses the singles count rates of two detectors to infer the randoms rate in the corresponding LOR using Equation 2.6

$$R_{ij}^{SR} = 2\tau S_i S_j \tag{2.6}$$

where R_{ij}^{SR} is the estimated randoms rate for the LOR defined by detectors ij and τ is the value of the coincidence window. This method relies on the assumption that the singles rates are large compared to the trues rate [34].

Scatter: scattered events lead to the mispositioning of events which reduces the image contrast. Although scattered events produce a fairly uniform signal across the FOV, their contribution needs to be corrected for accurate quantification. Most systems use energy thresholding to discriminate heavily scattered photons from the 511 keV unimpeded photons. However, additional correction techniques are required, such as simulated scatter models. These use tissue attenuation coefficients derived from the CT image to model the underlying distribution of scattered events and their contribution to individual profiles by assuming that at 511 keV virtually all attenuation will be due to Compton scatter. This estimated scatter model can be added to the forward model in the reconstruction. This method works well when all the sources of radioactivity are contained within the FOV of the scanner but not if large amounts of activity are outside the FOV.

Both scatter and randoms are additive effects and are taken into account in-reconstruction.

Normalisation: in an ideal scanner, each detector pair would record the same number of counts for a uniformly distributed activity in a cylindrical phantom, within statistical limits; however, in practice, due to minor variations in efficiency and geometry, some detector pairs record more counts. The correction factors can be determined by collecting data from uniformly distributed activity in the scanner FOV, obtaining sensitivity values for each detector [35]. The detector efficiencies can also be factorised as the product of the individual detector efficiencies and geometrical factors, this is known as the component-based method [36].

Dead time: after receiving a photon, the detector has a time period in which it will not record new events. At high counting rates, the dead time can significantly limit the detection efficiency and the true system count rate will not increase linearly with activity in the FOV. Correction for dead time typically involves a model for system dead time behaviour at different count rate levels [37].

2.3 Data output

During acquisition, each coincidence event can be written sequentially, along with a time stamp, information about whether the event is prompt or delayed, and the detector pair address, known as list-mode (LM) acquisition, or histogrammed into a sinogram. Detector singles rates are also stored. Difference in photon arrival time information, when it can be measured, is stored in a number of "TOF bins", the TOF bin index being an additional coordinate to an event in TOF PET. The "TOF bin size" is equal to the coincidence window divided by the number of TOF bins.

Sinograms require significantly less storage space and so are commonly used for clinical examinations. The benefit of acquiring in LM is that events can be integrated over any time interval, allowing the number and duration of frames to be selected after a scan has been completed.

2.4 PET reconstruction

The goal of PET image reconstruction is to provide a cross-sectional image of the activity distribution using the coincidence events detected by the scanner.

Filtered backprojection (FBP) is the most common method for analytic image reconstruction due to its accuracy, speed of computation, and simplicity of implementation [38]. After the corrections discussed in Section 2.2.3 have been applied, the counts from each detector pair are projected back into an image array along the LOR between the detectors. Counts are assumed to be evenly distributed in each pixel along that LOR. By repeating this for all detector pairs, a linear superposition of backprojections is achieved. The resultant image resembles the true activity distribution but it is blurred due to the assumption that counts are evenly distributed along each LOR. To compensate for this blurring, the 1D Fourier transform of each projection is filtered with a ramp filter, before taking the inverse Fourier transform to obtain the filtered projection, which is then backprojected. The ramp filter amplifies high-frequency components, which are dominated by noise. Therefore, in practice, the ramp filter is modified with a low-pass filter to control the noise level in the images. By varying the cut-off frequency of the filters, a trade-off between noise level and spatial resolution can be reached for each clinical task.

The reconstruction accuracy of FBP is limited by (i) degrading factors in PET acquisition, i.e. intercrystal scatter, positron range, and acolinearity, which are not modelled; (ii) the stochastic variability in photon detection is not taken into account.

Iterative image reconstruction

By modelling the statistical noise of PET data, and the physical effects of the imaging model, improved performance over the analytical methods can be achieved. Due to the resultant complexity of the reconstruction problem, it is solved iteratively in a “forward project, compare, backproject, adjust” loop.

Image reconstruction can be formulated as an affine inverse problem:

$$E[p] = Hf + b \tag{2.7}$$

where f is a discrete representation of the continuous object (in voxels); matrix H

describes the imaging system; b is a background term accounting for scatter and randoms; projections are arranged into a vector p with $E[.]$ denoting the mathematical expectation due to the statistical nature of the uncertainties associated with the positron decay process, the attenuation effects, the additive scatter and random events, and the photon detection process, as described above.

In Equation 2.7, each element H_{ij} contains the probability that image element f_j contributes to data element p_i, excluding the background. H can model only the geometrical mapping from the object to the data, or it can include other physical effects such as attenuation and detector blurring i.e. :

$$H = H_{sensitivity}.H_{blur}.H_{attenuation}.H_{geometry}.H_{positron} \tag{2.8}$$

Iterative methods are characterised by (i) a criterion that defines the "best" image, i.e. a cost function, which measures the similarity between estimated data and measured data plus a penalty term; (ii) a numerical algorithm to determine how the image estimate should be updated at each iteration based on the criterion. The most widely used criterion is the maximum likelihood (ML) approach and the expectation-maximisation (EM) algorithm is used to find the ML image estimate. This can be done using the MLEM algorithm.

For image reconstruction with a Poisson likelihood, the MLEM method is a simple iterative equation:

$$\hat{f}_j^{(n+1)} = \frac{\hat{f}_j^{(n)}}{\sum_{i'} H_{i'j}} \sum_i H_{ij} \frac{p_i}{\sum_k H_{ik} \hat{f}_k^{(n)} + b_i} \tag{2.9}$$

where $\hat{\mathrm{f}}_{\mathrm{j}}^{(\mathrm{n})}$ is the image estimate at voxel j at iteration n. This method yields very noisy images due to the ill-conditioning of the problem [39]. Common solutions are stopping the algorithm before convergence or applying a smoothing filter to the reconstructed image for noise suppression. These solutions reduce noise at the expense of increasing bias [38].

To comply with the assumption of a Poisson distribution, which is the basis for EM-based reconstructions, a truncation (to zero) of negative sinogram values is conducted. This is known as the non-negativity constraint and can lead to bias in the reconstructed image; this is especially an issue in low count data.

MLEM is slow to converge. To address this, ordered subsets expectation-maximisation (OSEM) was proposed, which only uses part of the data at each update [40]. It partitions the projection data into B subsets and uses only one subset of data for each update. Each pass of the entire data set involves a greater number of updates, leading to significant acceleration compared to MLEM (roughly B times faster). It is however, not guaranteed to converge to the ML solution.

To address the effects of convergence of MLEM/OSEM algorithms, regularised algorithms have been introduced e.g. GE's "Q.Clear" algorithm. Q.Clear is a Bayesian penalised likelihood reconstruction algorithm (specifically, block sequential regularised expectation maximisation (BSREM)), which incorporates an additional term in the objective function discouraging difference in neighbouring voxel values. This term increases as image noise increases, reducing the objective function, which has the effect of steering the optimisation away from noisier images [41]. The user can determine the strength of the regularising term by specifying a β value, which is included as a weighting factor in the objective function. Q.Clear is run to "effective" convergence, therefore, iterations and subsets are no longer required inputs. As noise is controlled within the algorithm, it is not necessary to post-filter the image.

The Q.Clear algorithm includes resolution modelling (RM; the H_{blur} component in equation 2.8). Detector point spread function (PSF) in projection space can be obtained through analytical derivations, Monte Carlo simulations, or experimental measurements [38]. Including the PSF model in the reconstruction has been shown to improve spatial resolution in the reconstructed images [42, 43] as well as leading to improved contrast recovery and lesion detectability [44, 45]. It is however recognised that it may lead to errors in quantification, e.g. overshoots at object edges, referred to as the Gibbs effect [46].

In TOF reconstruction, the backprojection operator includes a probability function that weights the data along each LOR [47]. The centre of the probability function is the nominal position derived from the TOF information, and the Gaussian full width half maximum (FWHM) of the TOF kernel is related to the time resolution: $c\Delta t/2$. The incorporation of the TOF information in the reconstruction algorithm has the advantage that true coincidence events tend to gather around the “true” position, consistent with the TOF information, and random coincidence events affect only a limited section of the line of response [48]. The overall effect is lower noise and higher contrast recovery. TOF reconstruction brings a gain in the signal to noise ratio (SNR) proportional to the inverse of the square root of the time resolution. Faster convergence is an additional advantage of iterative TOF algorithms.

A deep learning (DL) algorithm is available on the GE Omni Legend PET-CT; it aims to transform non-TOF PET images to TOF-like images. PET data were reconstructed using the Q.Clear algorithm with and without TOF. Three DL-TOF models were trained on thousands of reconstructions to transform the non-TOF Q.Clear images to their target TOF images, with different levels of contrast-enhancement-to-noise trade off (low, medium or high). The least smooth DL-TOF model is referred to as “high precision deep learning (HPDL)” [49]. The deep neural network aims to achieve “better contrast-to-noise ratio and contrast recovery” compared to non-TOF Q.Clear reconstructions and “increased small, low-contrast lesion detectability” compared to TOF Q.Clear reconstructions (on GE’s Discovery MI) [50].

2.5 PET system characterisation

2.5.1 Resolution

In clinical PET imaging, for FBP reconstructions, the intrinsic spatial resolution of the system is around 3.5 – 5.0 mm at the centre of the FOV and is limited by:

- the detector width (d) - this is the dominant factor and is due to the solid angle coverage and the fact that the position of interaction within the crystal is not determined [13].
- positron range (s) - the distance the emitted positron travels before it thermalises, captures an electron and forms a positronium. The resultant effective

blurring is isotope and material dependent. Positron range is non-Gaussian, with a sharp central cusp with relatively broad tails [13].

- acolinearity - the annihilation photons are rarely emitted at exactly 180° directions from each other due to a small residual momentum of the positron when it reaches the end of its range. This gives rise to the photons deviating from the 180° by 0.25° FWHM; this deviation results in an overall resolution blurring related to the detector separation (R).
- Anger logic (b) - most PET scanners have more crystals than photo-detector elements; the decoding is often imperfect. The magnitude of the effect is linked to the width of the detector element ($\approx$ d/3 FWHM) [13].
- crystal penetration - the 511 keV photons often penetrate some distance into the detector ring before they interact and are detected. They may interact in a crystal other than the one that they impinge upon, which results in asymmetric blurring in the radial direction. The magnitude of blurring depends on the detector material, for BGO or LSO scintillators, the penetration is described by a Gaussian with FWHM $12.5\mathrm{r}/\sqrt{\mathrm{r}^2+\mathrm{R}^2}$ [13].
- sampling error - the sampling in the scanner FOV is not uniform, some voxels have a larger number of LORs intersecting them. The effect is especially pronounced at the centre of the scanner and has been empirically observed to multiply all the other contributions by a factor of 1.25.

Although some of the response function are not well-described by a Gaussian function i.e., the detector response and the positron range, it is often assumed that all effects add in quadrature for FBP reconstructions [13]:

$$G = 1.25\sqrt{(d/2)^2 + s^2 + (0.0044R)^2 + b^2 + \frac{(12.5r)^2}{r^2+R^2}} \tag{2.10}$$

where r is the radius of the source from the centre of the detector ring.

As recommended by the NEMA standards, the reconstructed image resolution of a PET system is characterised through point source in air measurements at different locations in the FOV [51]. FBP reconstruction is undertaken and the FWHM calculated from the profiles along the radial and tangential directions. As the use of iterative reconstruction algorithms have surpassed the use of FBP in clinic this test no longer provides a representative measure of reconstructed PET image resolution.

More recently, Gong et al. investigated the effect of the non-negativity constraint, as well as the nonlinear nature of iterative algorithms, on the determination of reconstructed spatial resolution using point sources in hot backgrounds [52]. They recommend that, to minimise the variation in the measured FWHM, the reconstructed point source contrast should be less than 0.1. As the measured FWHM also depends on the reconstruction algorithm and number of iterations, the same algorithm with a fixed number of iterations should be used.

Meikle et al. demonstrated that differences in the intrinsic resolution between the emission and transmission measurements have a significant effect on the accuracy of the AC [53]. They conducted noiseless simulations of digital phantoms reconstructed by FBP. When the transmission resolution was better than the emission resolution, counts on the high-density side of the interface were artificially increased. When transmission resolution was worse than emission resolution, the effect is reversed and counts in the high-density regions were underestimated. When resolutions matched, the resultant PET distribution was homogeneous, as simulated. The authors acknowledged that, for simulated data, the blurring function required to achieve matched resolution was known but for real data, the FWHM of the required blurring function may be different to the intrinsic detector resolution used in the simulated case. To support their conclusions for scanner data, a ^{68}Ge/^{68}Ga line source was both measured and simulated. The difference in reconstructed resolution was negligible indicating that a blurring function equal to the intrinsic detector resolution is appropriate for both simulated and real data.

An investigation into the blurring required for AC when reconstructing with iterative algorithms will be investigated in Chapter 3 of this thesis.

2.5.2 Noise characteristics

The sensitivity of the system is determined primarily by the absorption efficiency of the detector system i.e. the stopping power of the crystal, as well as the geometric efficiency (its solid angle of coverage of the imaging object).

The noise equivalent count rate (NECR) is a performance parameter that accounts for the additional statistical noise introduced by the correction for random and scattered coincidences. The NECR is defined as the equivalent counting rate that gives rise to the same statistical noise level as the observed counting rate after random and scattered coincidences have been accounted for [12]. True and scattered events are proportional to activity in the FOV, whereas random events are proportional to the square of the activity; this causes a decrease in NECR at higher activities (as do dead time losses).

The performance characteristics of the three clinical PET-CT scanners used in this work are described in Table 2.3.

Table 2.3: Published values for scanner performance for the three clinical PET-CTs investigated in this project.

Scanner	**Spatial res. @ 1 cm (xy (z)) [mm]**	**Sensitivity [cps.kBq^{-1}]**	**Peak NECR [kcps]**	**TOF res. [ps]**
GE D710 [19]	4.70 (4.74)	7.5	139.1 @ 29 kBq.mL^{-1}	550
Siemens Vision [18]	3.6 (3.5)	16.4	306 @ 32 kBq.mL^{-1}	210
GE Omni Legend [54]	3.81 (3.61)	47.3	501.7 @ 15.7 kBq.mL^{-1}	-

2.6 Standardisation and harmonisation

Multi-centre imaging studies are critical to obtaining adequate statistical power in research studies but, despite calibration and routine quality assurance, PET-CT scanners from different vendors, located at different sites, can generate differing images. This is not only a major barrier to conducting multi-centre studies but also the translation of advanced correction methods into clinical practice.

The European Association of Nuclear Medicine (EANM) procedure guidelines for tumour imaging highlighted the importance of all institutions participating, in ^{18}F-FDG oncology studies to employ similar methodologies [55]. In order to ensure the comparability of SUVs between centres, recommendations were made regarding patient preparation, radiopharmaceutical administration, imaging protocol, image analysis, and reporting. A minimum set of quality control (QC) procedures that must be performed were defined. The guidelines also state that reconstruction settings should be chosen so as to achieve matching convergence and reconstructed spatial resolution across various systems and studies, especially within a multi-centre setting.

In 2006, the EANM launched EANM Research Ltd. (EARL) as an initiative to promote multi-centre nuclear medicine and research. From 2010 until July 2016, the EARL FDG-PET-CT accreditation program collected over 2500 phantom datasets from approximately 200 systems and 150 imaging sites worldwide [56]. To obtain and maintain EARL accredited status, sites are required submit two phantom scans: (i) a uniform cylindrical phantom to assess calibration quality control and (ii) a NEMA NU2-2007 body phantom to assess image quality. SUV bias in the uniform cylindrical phantom must be within ± 10 % for a system to become accredited. Sphere SUV recovery coefficients (RC) must lie within specified ranges for each of the six sphere diameters in the NEMA NU2-2007 body phantom. In 2018 EARL published the results of their study assessing the feasibility of harmonising the performance of PET-CT systems equipped with TOF and RM technologies [57]. Four PET-CT systems from three major vendors were used to acquire data. 15 reconstruction parameter settings were used to compare quantitative performance of the systems. They found a significant increase in SUV$_{mean}$ and SUV$_{max}$ contrast recoveries compared with the original “EARL1” specifications. The RC ranges were updated to accommodate the higher recoveries, this is referred to as the “EARL2” standard [58]. These updated criteria were validated on 18 additional scanners from 16 sites.

EARL harmonisation concentrates on quantification of hot tumour-like objects. Special considerations are needed when applying quantitative imaging approaches to PET-CT lung imaging as the lungs contain varying amounts of air depending on the disease, are the source of respiratory motion, and have a relatively high blood volume. In 2020 Chen et al. recognised the lack of widely accepted standard protocols for pulmonary ^{18}F-FDG PET-CT and convened representatives from industry, academia and the clinic to form consensus recommendations on patient preparation, image acquisition, reconstruction, analysis and reporting [1]. Recommendations for patient preparation and positioning were in-line with the Radiological Society of North America's quantitative imaging biomarker alliance (QIBA) [59] and EANM recommendations [55] for ^{18}F-FDG for oncological applications. Noteworthy considerations from the authors were: i) corrections for the effects of air and blood feature prominently in the quantitative analysis methods proposed to improve lung tissue-specific ^{18}F-FDG quantification; ii) minimising respiratory motion effects on quantification remains challenging, which novel reconstruction methods may help to overcome. They also recognised the need to develop specific phantoms that model relevant aspects of lung physiology and support the establishment of harmonisation standards to support multi-centre studies.

A recent review of harmonisation strategies for quantitative PET by Akamatsu et al. [60] did not include any harmonisation studies for quantitative imaging of diffuse lung disease. There is therefore still a need for standards to be put into place to validate pre-existing correction methods to enable quantification in multi-centre PET lung imaging.

2.7 Challenges of PET-CT in the lung

In 2017 Chen et al. published a position statement on ^{18}F-FDG PET quantification approaches in lung diseases [3]. The authors noted that interpretation and quantification of ^{18}F-FDG lung imaging is confounded by the substantial proportions of air and blood in the lungs and highlighted the importance of implementing methods to measure lung cell metabolism specifically, in order to better characterise the role of lung cells in promoting disease activity and progression. Reconstruction algorithms have a significant impact on quantification accuracy, including issues with non-linearity and under-convergence when using iterative algorithms; the authors recommended further investigation into reconstruction optimisation for the diffusely distributed, relatively low count activity, typically seen in the lungs. The general lack of standardised quantification approaches was highlighted.

2.7.1 Attenuation correction

As discussed in Section 2.2.3, a large literature exists discussing the effects of errors in the attenuation image, which propagate into PET image reconstruction, subsequently affecting PET quantification [61]. Artefacts due to errors in the conversion of the CT numbers to LACs appropriate for 511 keV annihilation photons, truncation of the CT, beam hardening, and resolution mismatches between the two modalities are well documented. CT-based AC for PET is particularly challenging in the thorax due to respiratory motion.

2.7.1.1 Motion

Mismatched data due to patient movement between the PET and CT scan used for AC, whether global, or due to cardiac or respiratory motion, are common and well documented in reviews by Nehmeh and Erdi and Sun and Mok [27, 62]. The artefacts are well recognised, such as a curvilinear cold artefact on the dome of the liver at the lung-diaphragm interface. The loss of contrast and lesion dispersion result in misidentification of lesions, as well as underestimate of SUVs. Inaccurate SUVs of lung nodules of up to 30 % have been demonstrated [63]. The attenuation mismatch due to respiratory motion is not solely the result of tissue displacement caused by compression and dilation during the breath cycle, but also the change in lung volume, and therefore tissue density, while doing so [64].

Holman et al. investigated both the tissue displacement and density change component of motion on attenuation map mismatch in the lungs in non-TOF data [65]. Simulations of the 4D extended cardiac-torso (XCAT) phantom were validated with PET data from five patients suffering from IPF, who underwent PET-cine-CT. Maximum errors due to density mismatch were found to be as high as 14 % for patient-derived static concentrations and 31 % for kinetic parameter estimates, for non-TOF reconstruction. The results indicated that the respiratory induced density errors in the attenuation map affect quantification throughout the lung, not just near the boundaries. The effect of attenuation map mismatch was demonstrated to be: i) mostly local to the mismatch; ii) dependent on the tracer distribution in the thorax at the time of imaging (and hence on the tracer and the time of acquisition); iii) dependent on the magnitude of density variation and volume change of the perturbation area. This was in agreement with previous findings by Thielemans et al. [66].

Ahn et al. considered the effect of mismatched attenuation maps to be local for TOF reconstruction [67]. An estimation close to that in Thielemans et al. was found, now depending on the time resolution instead of the size of the emission object. Emond et al. demonstrated that, although at the centre of the perturbation results were in agreement with Ahn et al. , errors were observed outside the location of the mismatch [68]. It was demonstrated that, for MLEM and OSEM reconstructions, the errors at a distance from the mismatch increased with improved timing resolution (e.g., from 2.8 % in non-TOF up to 8.2 % at 550 ps and 20.9 % at 200 ps in the left ventricle).

Methods for respiratory motion correction were primarily developed for oncological applications and so focus on addressing the issue of spatial displacement of tissue but few of them consider changes in density or tracer concentration throughout the breath cycle [27, 62].

Techniques such as shallow breathing or breath-holds at gentle end-expiration during CT acquisition [69, 70] have been shown to improve PET accuracy, but do not address the degrading effects of breathing motion in PET and involve some form of patient training and compliance, which can be difficult in patients with respiratory conditions [62].

The use of cine-CT has been suggested [71, 72, 42, 73]; these methods are not widely used clinically due to the higher patient dose. Cine-CT does have the advantage that both location and density will be more closely matched to the PET data.

The position statement by Chen et al. found that, within the seven participating institutions, the most used strategy to minimise PET and CT misregistration was breath-hold at gentle end-expiration or mid-expiration for the CT scan [1]. They recommend that, since many of the advanced techniques described are not widely available and require post-imaging offline processing, for research studies, list-mode acquisitions should be acquired and stored for future reprocessing when respiratory gating and other advanced techniques to minimise respiratory motion become readily available.

2.7.2 Partial volume effects and correction methods

The finite intrinsic spatial resolution of the PET system results in a 3D blurring throughout the PET image reducing PET quantification accuracy. PET has a relatively poor intrinsic resolution of between 5 and 6 mm FWHM for PMT-based photodetection systems to between 3 and 4 mm FWHM for newer solid-state detector systems. This poor spatial resolution can cause partial volume effect (PVE), a term used to describe two phenomena: (i) the blurring due to the intrinsic scanner resolution causing apparent displacement of activity [74]; (ii) the sampling of the image resulting in each voxel containing multiple tissue types, the voxel value represents the average signal from these fractional contributions; this is known as the tissue fraction effect (TFE).

PVC methods have been proposed which compensate for the effects of image system blurring on the PET data either within image reconstruction with RM and/or the introduction of anatomical priors or post-reconstruction. The TFE can be accounted for post-reconstruction with the inclusion of structural information from high resolution modalities such as CT or MRI. These methods have been extensively covered in the literature [75, 74].

2.7.2.1 Partial volume correction for imaging system blurring

As previously discussed, PVC approaches can be included in-reconstruction or applied post-reconstruction. These techniques can be further sub-divided into region of interest (ROI)-based or voxel-based methods. In addition, some approaches are purely data-driven, while others utilise information from structural imaging modalities.

RM incorporates an estimate of the PSF or detector response function into the system matrix used for forward and backprojection [76, 77], as discussed in Section 2.4. Comtat et al. demonstrated that modelling the resolution reduced the convergence rate of OSEM and increased activity in small structures [76]. However, taking the resolution into account during reconstruction also makes the reconstruction problem underdetermined, i.e. different images exist that agree equally well with the acquired data. OSEM tends to select a solution that suffers from Gibbs artefacts [78]. The Gibbs phenomenon is due to the step-function being represented using a Fourier series with finite number of terms, resulting in an overshoot at the point of discontinuity. This over-shoot is also accompanied by ringing, related to the sinc function, which is the impulse response of a perfect low-pass filter [79].

Reconstruction algorithms that incorporate an anatomical prior use the structural information to control noise and enhance edges between functional structures. Penalised iterative reconstruction algorithms often use a smoothness prior to encourage solutions with a low degree of variability between adjacent image voxels within regions, while allowing for large inter-voxel variations across boundaries [80, 81]. Some algorithms do not require segmentation of the anatomical image and instead promote greater similarity among nearby voxels that have a similar intensity in the anatomical image [82]. Erlandsson et al. stated that more work was needed to better understand the sensitivity of these methods to mismatched anatomical information [75].

Post-reconstruction methods can be sub-divided into purely data-driven, e.g. iterative deconvolution and those that utilise information from structural imaging modalities.

The anatomically based PVC methods, e.g. Müller-Gärtner (MG), multi-target correction, region-based voxel-wise correction, iterative Yang, correct for PVEs that correspond to spill-over of data from one region to another (inter-regional PVEs). They do not correct for PVEs between voxels within the same region (intra-regional PVEs). Deconvolution can be used for intra-regional PVC as an extension to voxel-based inter-regional PVC algorithms [83]. Richardson-Lucy [84, 85] and van Cittert [86] are common deconvolution techniques. These approaches suffer from noise amplification issues, which often results in premature termination and incomplete recovery. Noise suppression, such as the inclusion of priors and regularisation have been applied during deconvolution [87, 88].

These PVC techniques have mostly been developed for, and tested on, neurological applications. As noted in Erlandsson et al., PVC will be much more difficult to implement in patients with multiple diffuse lesions or a heterogeneous lesion than in patients with a small single lung tumour, easily detected on a CT scan [75]. For this reason, PVC methods have been studied for well-defined lesions in the lung [89, 90, 91, 92, 93] but not in diffuse lung diseases.

2.7.2.2 Partial volume correction for image sampling

The finite size of an imaging voxel, being significantly larger than the average alveolus size, results in a single voxel containing air and multiple tissue elements, this is known as the TFE. Much less work has been undertaken to address the tissue fraction problem [75].

Air fraction correction

The importance of correcting for the air component in quantitative PET-CT in the lung has been presented for patients with IPF [94, 95, 96]. To isolate metabolic changes from those associated with the degree of aeration, Lambrou et al. proposed a tissue fraction correction method to account for the variability of density in the lung [94]. A simplified model of the composition of the lung was used, assuming lung parenchyma and air only, see equations 3.1 and 3.2 in Chapter 3. The developed method was validated in a phantom study and applied to a group of subjects (n = 12) with no evidence of pulmonary disease and a small group of patients (n = 5) with confirmed IPF. It was observed that, once aeration of tissue had been taken into account, the variation in SUV was reduced in non-IPF subjects. In the small number of IPF patients studied, conducting AFC resulted in larger SUVs than in non-IPF subjects, but reduced SUVs in regions of interstitial lung change (ILC) compared to non-ILC regions; this is contrary to the uncorrected situation reported by Groves et al. [8]. The authors stated that the “CT data utilized for attenuation correction were already pre-smoothed so as to have a resolution similar to that of the PET study and identical voxel size. This dataset was therefore used to derive the TF correction factors”.

Win et al. conducted an additional study to investigate the ^{18}F-FDG PET signal at sites of normal-appearing lung on high resolution CT (HRCT) in IPF [96]. It was demonstrated that IPF patients (n = 25) have increased pulmonary uptake on ^{18}F-FDG PET in areas of lung with normal morphological appearance on HRCT, compared to controls (n = 25), with or without correction for air fraction. This finding suggests that the parenchyma is globally affected in IPF and that PET may have a higher sensitivity in the detection of early disease than HRCT. The application of AFC in this study increased the difference between the disease and control groups, while reducing intra-group variability.

Torigian et al. investigated the effect of AFC in 49 ^{18}F-FDG PET studies of COPD with emphysema [97]. A significant correlation between lung metabolism and (CT-derived) degree of emphysema was only observed when AFC was implemented. However, the results were not validated with tissue-based or clinical outcomes.

Castiaux et al. studied the relationship between severity of IPF and uptake of ^{18}F-FDG in the lung [98]. SUV values, uncorrected or corrected for lung density [94], were correlated with lung function tests, carbon monoxide diffusion (DLCO) and 6-minute walking test (6MWT) at baseline and at one year. Uncorrected SUV was shown to be related to lung volume measurements but normalisation for lung density eliminated correlations between SUV_{mean} and respiratory function variables. The AF-corrected measurements were more directly related to cellular and tissue changes in the parenchyma itself. AF-corrected SUV_{mean} was found to be correlated with the DLCO. The significance of this correlation was supported by the correlation with the desaturation post-exercise during the 6MWT. This indicates that an increase in glucose metabolism estimated by AF-corrected SUV_{mean} indicates a reduction in gas diffusion in the lungs and an increase in oxygen desaturation of the blood after exercise.

Blood fraction correction

The lung contains a large fraction of blood ($\approx$ 20 %) per unit whole lung volume [99]. Most methods for calculating the fractional volume of blood, such as the pulmonary function tests and cardiovascular MRI methods, result in global estimates. In lung diseases, such as IPF, where the quantity of blood reaching regions of diseased lung parenchyma in comparison to normal appearing tissue is poorly understood, determination of blood volume on a regional or voxel-wise level would be beneficial.

Regional blood volume can be measured from a ^{11}C or ^{15}O PET scan in combination with venous blood samples [100]. This method does, however, require an onsite cyclotron facility. Holman et al. utilised kinetic analysis of dynamic PET-CT to determine voxel-wise fractional blood volumes in six patients with IPF [101]. The advantage of this method is that almost any PET tracer can be used. It was shown that, if TFEs were not corrected for, regions of high density (fibrosis) appeared to have a higher uptake than lower density (normal appearing tissue), as demonstrated by Groves et al. [8]. However, this was reversed after air and blood correction for both uptake (SUV) and influx rate (K_i). The authors note the reliance of this technique on PET-CT registration, which is a challenge in the lung. They state that CT data was filtered to the resolution of the PET image and downsampled. How PET image resolution was determined was not described. Patient data was reconstructed using FBP to avoid issues with quantification when using iterative reconstruction methods such as variable convergence rates and bias in regions of low tracer uptake. Comparisons to healthy controls were not performed in this paper.

In more diffuse diseases, like COPD, quantification of ^{18}F-FDG has been analysed using a "slope-intercept" Patlak graphical approach. A metabolic activity outcome measure is estimated where the Patlak intercept is used as a surrogate measure to correct for air volume [5]. The slopes were normalised for variations in distribution volume (determined from the t = 0 intercept). Previous studies employing the Patlak method did not specifically include any correction for blood activity;

the Patlak intercept method to correct for the amount of air in the lung is only an approximate method [102].

Coello et al. corrected ^{18}F-FDG PET scans from COPD subjects (n = 10) and matched healthy volunteers (n = 10) for the presence of air and blood in the lung. The fractional air volume was estimated from CT and a weighted least squares fitting was used to estimate kinetic parameters from measured activity in the ROI and blood, assuming an irreversible two-tissue compartmental model [102]. The inclusion of the estimation of the fractional volume of air in the fitting differs from Holman et al. [101] by avoiding division of the parameters of interest by a correction factor that approaches zero, which may be of particular problem in emphysema. They compared their technique with the normalised slope-intercept Patlak approach described above [5]. The authors noted no difference between healthy volunteers and COPD patients using their method and thus inferred that whole lung FDG may be insufficiently sensitive to detect the increased inflammation associated with lung parenchyma in COPD. The differences observed between groups when using the normalised Patlak approach may therefore be driven by other quantities, i.e. blood or air volume, to which it is sensitive.

2.7.3 Positron range

Incorporation of positron range effects as a component of PVC algorithms has received limited attention to date. Several groups have worked on (partly) correcting for positron range within iterative reconstruction by adding an element to the system matrix in the reconstruction, as in equation 2.8, [103, 104]. This element is usually based on a Gaussian kernel constructed from attenuation values [14].

Emond et al. [14] used Monte Carlo simulation to investigate the effect of positron range in regions of small high-density structures within the healthy lung i.e. small tumours and localised fibrosis, for multiple radionuclides. The results demonstrated quantification biases, where the measured uptake of high-density regions can be increased by neighbouring radioactivity from regions of lower density. This indicates that, although positron range is usually perceived as a blurring effect, in heterogeneous media it should instead be considered as a more complex change in apparent radioactivity distribution. This effect was more noticeable for radionuclides with high-energy positron emission such as ^{68}Ga, ^{124}I or ^{82}Rb. Anisotropic kernels, as proposed by Bai et al. [30], Alessio and MacDonald [104], Rahmim et al. [105], Szirmay-Kalos et al. [106] and Cal-Gonzalez et al. [107], or Monte Carlo simulations investigated by Lehnert et al. [108], could reduce the bias seen in Emond et al. [14]. However, the challenge of determining spatially variant kernels in such an inhomogeneous medium as fibrotic lung, and the high computational cost once determined, might be a limiting factor in clinical settings.

2.7.4 Effects due to iterative reconstruction methods

For iterative reconstruction algorithms, the positivity bias and slower convergence rate in low-count regions could contribute to quantitative biases [109, 110]; this indicates that quantification in the lung may be affected. Chen et al. [1] recommended using a large number of iterative updates (number of iterations multiplied by the number of subsets), to ensure uniform convergence. However, it was recognised that more investigation is needed to define optimal reconstructions specifically for lung PET imaging.

TOF reconstructions are now common practice in the clinic, due to better performance in terms of SNR improvements and a greater robustness in the presence of model inconsistencies [16], such as the effects of attenuation mismatch, albeit at the expense of non-local effects outside the lung [68].

Chen et al. recommend that, for multi-centre studies, the data from scanners with a similar TOF time resolution are used, or that only non-TOF reconstructed images are used [1].

Chapter 3

The effect of PET-CT resolution mismatch in lung imaging

3.1 Introduction

The finite size of an imaging voxel, being significantly larger than the average alveolus size, results in a single voxel containing air and multiple tissue elements, this is known as the TFE, section 2.7.2. Correcting for this has been shown to alter image interpretation in patients with IPF [94, 96].

The AFC assumes a simple model of the lung where the observed activity concentration is considered to be the result of the radiotracer distributed throughout the tissue component (in this case, a combination of both parenchyma and blood) and a gas component, containing no activity. The CT acquired for AC is utilised to determine voxel-wise fractions of unit volume occupied by air, V_A:

$$HU_L = V_T HU_T + V_A HU_A \tag{3.1}$$

where V_T is the factional volume of tissue and HU_T, HU_A, HU_L are the CT numbers of tissue, air and in the lung voxel, respectively [101].

Given the linear scaling between HUs and LACs in the lung [25], the relationship between the fraction of air in each voxel, V_A, and lung density, can be instead expressed in terms of LACs for 511 keV photons.

$$V_A = \frac{\mu_L - \mu_T}{\mu_A - \mu_T} \tag{3.2}$$

where μ_L, μ_T and μ_A are the LACs for 511 keV photons in the mu-map lung voxel, soft tissue, and air, respectively. These air fracions (AF) can be used to account for the variable air content in the lung, providing an estimate of tracer uptake per gram of tissue.

In practice, the CT image is used for both AC (during reconstruction) and AFC (after reconstruction). However, the difference in resolution between PET and CT can cause artefacts in either of these corrections. It was demonstrated by Meikle et al. that for FBP the resolution of the transmission image should match that of the intrinsic PET scanner resolution for AC [53]. The extent of smoothing to be applied to the CT for determination of the AFs, to minimise artefacts in the AF-corrected image, has only partially been studied.

This chapter details preliminary findings from simplistic noise-free analytical simulations and non-TOF reconstructions to investigate the kernel that should be applied to the CT for AC and AFC. Throughout this chapter, non-TOF simulations and reconstructions were performed due to computational limitations of the STIR TOF implementation when this work was performed. Digital phantom design and simulation methodology is outlined in Section 3.2. In the first instance, these phantoms are designed such that lung tissue uptake is uniform post-AC/AFC, see Section 3.3. The extent to which the kernel for AFC varies with respect to iterative reconstruction algorithm convergence is assessed with simplistic noise-free simulations using the point source insertion-and-subtraction methodology in Section 3.4. Finally, a proof-of-concept PVC-AFC methodology is investigated in the case that lung tissue uptake is not uniform, see Section 3.5.

3.2 Phantom design and simulation methodology

When considering location and density mismatches of the emission and attenuation images, the effect on AC and AFC quantification is coupled [65]. The coupled effects of resolution mismatches on AC and AFC have not previously been investigated in detail. For ease of interpretation, the effects are considered separately in the first instance; the optimal kernels for AC and AFC were therefore determined independently.

Digital emission phantoms, paired with inhomogeneous attenuation (mu) maps, were constructed in Software for Tomographic Image Reconstruction (STIR) [111] via the Synergistic Image Reconstruction Framework (SIRF) [112]. These phantoms were used as input to an analytic model in STIR via SIRF (GE Discovery 710 (D710) [19] template); the projection data were reconstructed using the same GE D710 template in STIR via SIRF.

3.2.1 Digital phantom design

For determination of the optimal Gaussian kernel to be applied to the mu-map for AC, h_{AC}, a cylindrical test-object was designed (phantom 1 in Figure 3.1(a)(d), Table 3.1). The mu-map included two spherical inserts with densities approximating a region of IPF and one of emphysema, positioned on the central axial slice. A soft tissue / healthy lung horizontal boundary ran the length of the cylinder. This phantom was designed to produce a homogeneous attenuation corrected image when reconstructed with an optimally smoothed mu-map.

Phantom 2, Figure 3.1(b)(e), Table 3.1, was used in the determination of the appropriate kernel to apply to the mu-map for determination of the AFs, h_{AFC}. The inhomogeneous mu-map consisted of an elliptical cylinder with an LAC representative of soft tissue for 511 keV photons, containing two “lungs” with an LAC representative of healthy lung at full inspiration for 511 keV photons. The left lung contained a 2D high-resolution CT patch of centrilobular emphysema (CLE) [113], repeated on the three central slices. The right lung contained a spherical insert representative of IPF density. The simulated inhomogeneous emission was designed to produce a homogeneous AFC-SUV equal to 3 in the lungs, representative of AFC-SUVs in IPF lung for ^{18}F-FDG [94, 101], Figure 3.2(f). The kernel that resulted in the greatest homogeneity in the AC-AFC reconstructed lungs was deemed optimal.

A more complex test object was designed with multiple low-density inserts, phantom 3, Figure 3.1(c)(f) and Table 3.1. As with phantom 2, the emission object was designed such that the ground truth (GT) reconstructed AC-AFC PET distribution was homogeneous. Due to the different feature diameters and densities, the AC-AFC reconstructed image will be the most prone to artefacts due to resolution mismatches. This phantom that was therefore used to gain an understanding of the interaction between AC and AFC when PET and CT resolutions are mismatched.

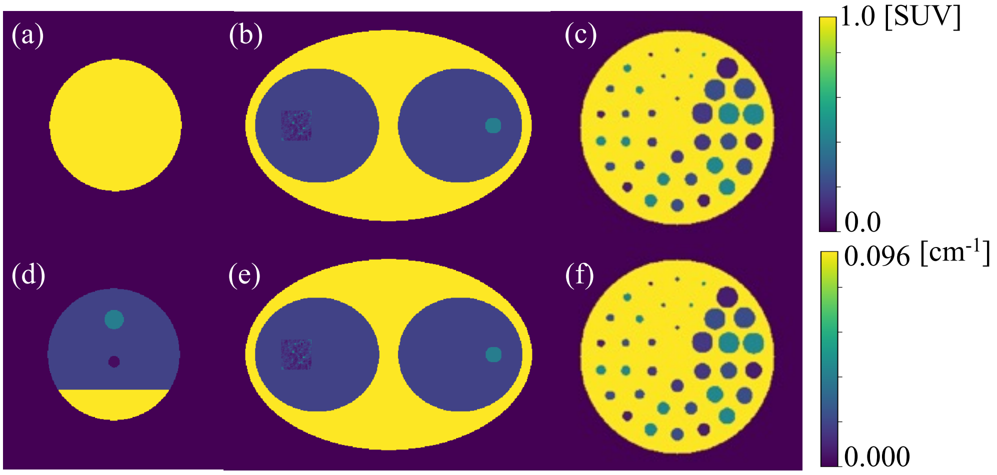


Figure 3.1: Digital phantom simulated emissions (top row) and simulated mu-maps (bottom row). **(a)(d) phantom 1**: mu-map includes two spherical inserts with densities approximating IPF and emphysema with a soft tissue-to-healthy lung boundary running horizontally along the length of the cylinder. The emission is homogeneous. **(b)(e) phantom 2**: mu-map and emission simulate a torso containing two lungs with healthy lung equivalent LACs, the left lung contains a 2D patch of CLE, the right lung a spherical insert of IPF density. The emission produces a homogeneous AF-corrected lung. **(c)(f) phantom 3**: mu-map consists of a "soft tissue" shell containing inserts of varying diameters and densities found in diseased lung. The emission resulted in a homogeneous AF-corrected image.

Table 3.1: Digital phantom specifications, the dimensions, simulated LACs of the mu-maps, and simulated SUVs of the emissions are given.

	Phantom 1	**Phantom 2**	**Phantom 3**
Shell dimensions [mm]	Cylinder 150 ϕ; length 78	Ellipse 300x200 ϕ; length 78 ’Lungs’ 130x120 ϕ; length 78	Cylinder 200 ϕ; length 40
Insert dimensions [mm]	13, 22 ϕ	2D patch CLE: 31x31 IPF: 22 ϕ	4 - 22 ϕ
Shell SUV	1.0	Ellipse: 1.0 “Lungs”: 0.597	1.0
Insert SUV	1.0	$3 \times (\frac{\mu_v - \mu_A}{\mu_T - \mu_A})$ CLE: 0.10 - 1.83 IPF: 1.27	$1 \times (\frac{\mu_v - \mu_A}{\mu_T - \mu_A})$ 0.20 - 0.46
Shell LACs [cm^{-1}]	Lower: 0.0960 Upper: 0.0191	Ellipse: 0.0960 “Lungs”: 0.0191	0.0960
Insert LACs [cm^{-1}]	13 mm: 0.0031 22 mm: 0.0406	CLE: 0.0031 - 0.0586 (mean = 0.0154) IPF: 0.0406	0.0191 - 0.0446

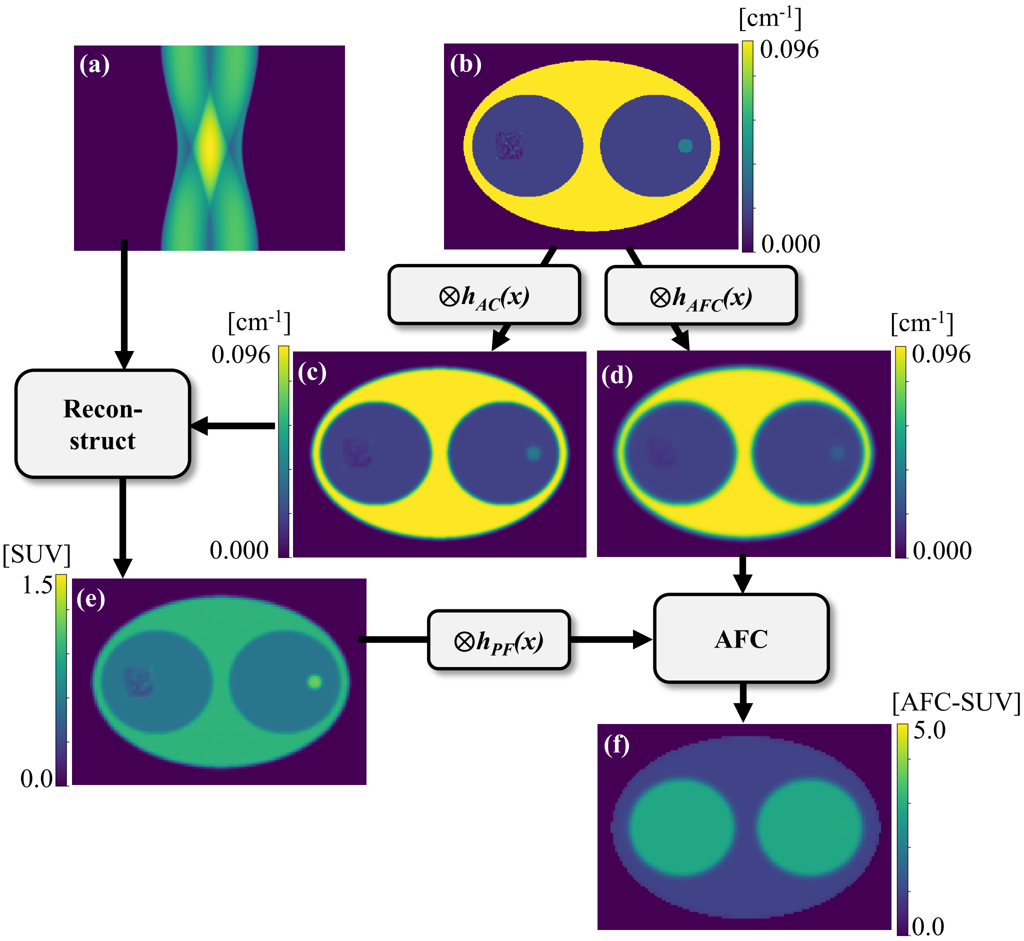


Figure 3.2: Pictoral representation of the AFC workflow as illustrated by a simulated noiseless scenario: (a) simulated emission projection data, smoothed by $h_{IR}(x)$ prior to forward projection; (b) simulated inhomogeneous mu-map; (c) mu-map $\otimes h_{AC}(x)$ for AC; (d) mu-map $\otimes h_{AFC}(x)$ for AFC; (e) AC reconstructed PET; (f) AC-AFC PET: AC PET $\otimes h_{PF}(x)$ and divided by (1 - V_A), voxel-wise.

3.2.2 Simulation methodology

The workflow for AC and AFC of PET images, depicted in Figure 3.2, was as follows:

1. Simulation of high resolution emission images and mu-maps (voxel size: 0.78 x 0.78 x 3.27 mm^3).

2. Convolution of the emission images and mu-maps with a 3D isotropic Gaussian kernel prior to forward-projection, $h_{IR}(x)$; this was to approximate the intrinsic spatial resolution of the D710 PET scanner [19], Table 3.2.

3. All data were noiseless to allow study of the attenuation and AF mismatch effect size.

4. OSEM reconstruction with 1000 iterative updates and 6 subsets (one update uses on subset of data). The FWHM of $h_{AC}(x)$ was varied from 0 to 10 mm in 1 mm increments. The reconstructed voxel size was 2.71 x 2.71 x 3.27 mm^3. All reconstructions were non-TOF.

5. Post-filter the AC PET image with an isotropic Gaussian post-filter (PF), $h_{PF}(x)$, of varying FWHM (0 to 8 mm, 2 mm increments), Table 3.4.. This was to mimic clinical practice, while also allowing the study of different resolutions in the reconstructed image.

Table 3.2: Analytic simulation parameters for phantoms 1, 2 and 3.

1. Simulated image parameters	Voxel size	0.78 x 0.78 x 3.27 mm^3
2. Acquisition model for forward projection	Acquisition model	4.7 mm^3 Gaussian filter, unless otherwise indicated, followed by ray-tracing
	Projection data template	GE D710, non-TOF (span=2; max. ring diff.=23)
	LORs per bin	12
	Sensitivity model	Includes attenuation factors

Table 3.3: Reconstruction parameters for phantoms 1, 2 and 3.

3. Acquisition model for reconstruction	Acquisition model	Ray-tracing
	Sensitivity model	Includes attenuation factors $h_{AC}(x)$ varied 0-10 mm in 1 mm increments
	Projection data template	GE D710, non-TOF (span=2, max. ring diff.=23)
	LORs per bin	12
4. Reconstruction	Reconstructor	OSEM
	N iterative updates	1000
	N subsets	6
	Voxel sizes	2.71 x 2.71 x 3.27 mm^3

Table 3.4: Post-reconstruction for phantoms 1, 2 and 3.

5. Post-reconstruction	Smoothing	$h_{PF}(x)$ varied 0-8 mm in 2 mm increments
	AF map computation	$h_{AFC}(x)$ varied 0-12 mm in 1 mm increments

Preliminary simulations suggested that the mu-map used in determination of the AFs should approximate the reconstructed PET image resolution. Reconstructed PET image resolution cannot be considered spatially invariant and is dependent on imaging conditions, including radionuclide administered, position in the FOV, the reconstruction model (the non-linearity of iterative reconstruction algorithms can result in a non-uniform reconstructed image resolution), and patient-specific factors such as activity distribution and attenuation, as discussed in Section 2.5. The point source insertion-and-subtraction (perturbation) method was first proposed by Stamos in 1988 to address the object and noise dependence of MLEM convergence rates [114]. The reconstructed spatial resolution of each smoothed attenuation-corrected PET image was determined at various positions on the central axial slice. Single-voxel point sources were positioned at multiple points in the phantoms in regions of differing densities and structure. The positions of the point sources are depicted in Figure 3.7. In each case, 1D Gaussian fits of the profiles in the radial, tangential, and axial direction were used to determine the reconstructed FWHM using curve fitting in the SciPy optimisation package in Python [115]. An estimate of the global measured reconstructed image resolution was calculated from the mean of the radial, tangential and axial FWHMs of all point sources. The non-negativity constraint inherent in the iterative reconstruction can artificially enhance the apparent reconstructed spatial resolution if a point source image is reconstructed without any background, Section 2.5. In our work, it was ensured that the point-source-to-background ratio was constant such that the point source contrast in the reconstructed image was less than 0.1, in line with recommendations by Gong et al. [52] for iterative reconstruction.

For investigation of the optimised kernel width for AFC, a 3D Gaussian filter of varying FWHM (0 to 12 mm, 1 mm increments), $h_{AFC}(x)$, was independently applied to the mu-map for voxel-wise AF, V_A, determination, Equation 3.2. The reconstructed, post-filtered, PET image was divided by $(1 - V_A)$, on a voxel-wise basis so as to represent uptake per gram of tissue rather than per unit volume.

3.3 Optimisation of AC and AFC kernels for uniform parenchymal uptake

3.3.1 Introduction

In the first instance, the optimal smoothing for the mu-map for AC and AFC was investigated for homogeneous tissue uptake, this simplification allows the effects of resolution mismatches on the AFC to be studied without having to consider the spill-in/spill-out component of the PVE.

3.3.2 Methods

Phantoms were designed to produce uniform activity distributions post-AC (phantom 1) and post-AC-AFC (phantoms 2 and 3), when associated kernels were optimised. Data were simulated and reconstructed according to Section 3.2.2.

The coefficient of variation (CoV) was calculated in the central slices of phantoms 1 and 3 and in the “lungs” for phantom 2. To remove edge-of-phantom artefacts from the analysis, the volume of interest (VOI) was eroded by 6 mm in both the transaxial and axial directions. The mean AC-AFC-SUV in the CLE and IPF regions were also determined for phantom 2.

For phantom 3, the FWHM of both $h_{AC}(x)$ and $h_{AFC}(x)$ were varied; the simulated emission and mu-map were convolved with an isotropic Gaussian kernel with FWHM of 5 mm^3 prior to forward-projection, followed by MLEM reconstruction with 200 iterations. The FWHM of $h_{PF(x)}$ was varied between 0 and 16 mm in 2 mm increments. The root mean square error (RMSE) between the homogeneous AC-AFC GT emission and the AC-AFC reconstructed images for each triple (h_{AC}, h_{AFC} and h_{PF}) of blurring kernels was calculated in a VOI that encompassed all features but avoided edge-of-phantom effects.

3.3.3 Results

Image artefacts were minimised (CoV = 0.17 %) in the reconstructed images of phantom 1 when the mu-map resolution for AC matched that of intrinsic PET scanner resolution, as simulated by STIR, Figure 3.3.

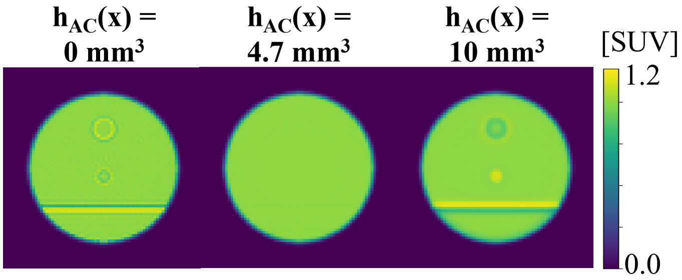


Figure 3.3: AC reconstructed image of phantom 1 (OSEM 1000 updates, 6 subsets) with various $h_{AC}(x)$. **Left**: under-smoothed mu-map used to attenuation correct; photopoenic regions on the low density side of the tissue boundaries are observed; **middle**: mu-map smoothed to resolution of simulated intrinsic PET scanner resolution; artefacts are minimised; **right**: over-smoothed mu-map used to attenuation correct; photopoenic regions on the high density side of the tissue boundaries are observed.

When there was a mismatch in mu-map and PET resolutions artefacts were seen at tissue density boundaries. When the mu-map resolution was higher than the intrinsic PET scanner resolution, counts on the low-density side of the tissue boundaries were artificially decreased. The converse was true for a lower mu-map resolution than the intrinsic PET scanner resolution.

The mu-map resolution for AC was therefore matched to the simulated intrinsic PET scanner resolution for determination of the optimal kernel for AFC. Figure 3.4 depicts the artefacts produced in phantom 2 at tissue density boundaries in the AF-corrected reconstructed images when there was a mismatch in the reconstructed image resolution and the mu-map used in the determination of the AFs. These artefacts were reduced (whole lung CoV = 4.08 %) when $h_{AFC}(x)$ approximated the measured global reconstructed image resolution (mean radial, tangential and axial FWHMs, FWHM(xyz), of point sources positioned in the lungs, shown in Figure 3.7(a)(b)), which in Figure 3.4 was 7.6 ± 0.27 mm^3.

The change in mean AFC-SUV in the CLE and IPF features as $h_{AFC}(x)$ is var-

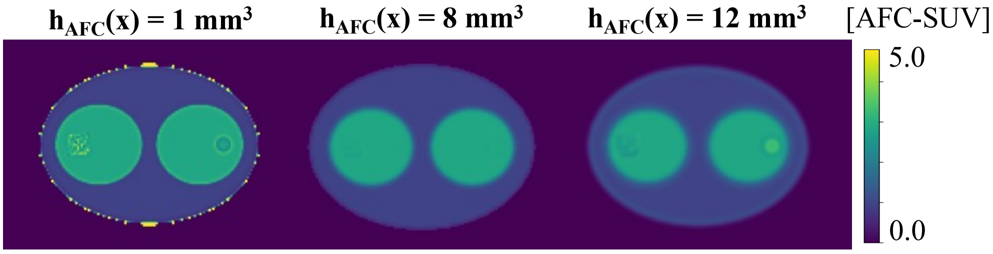
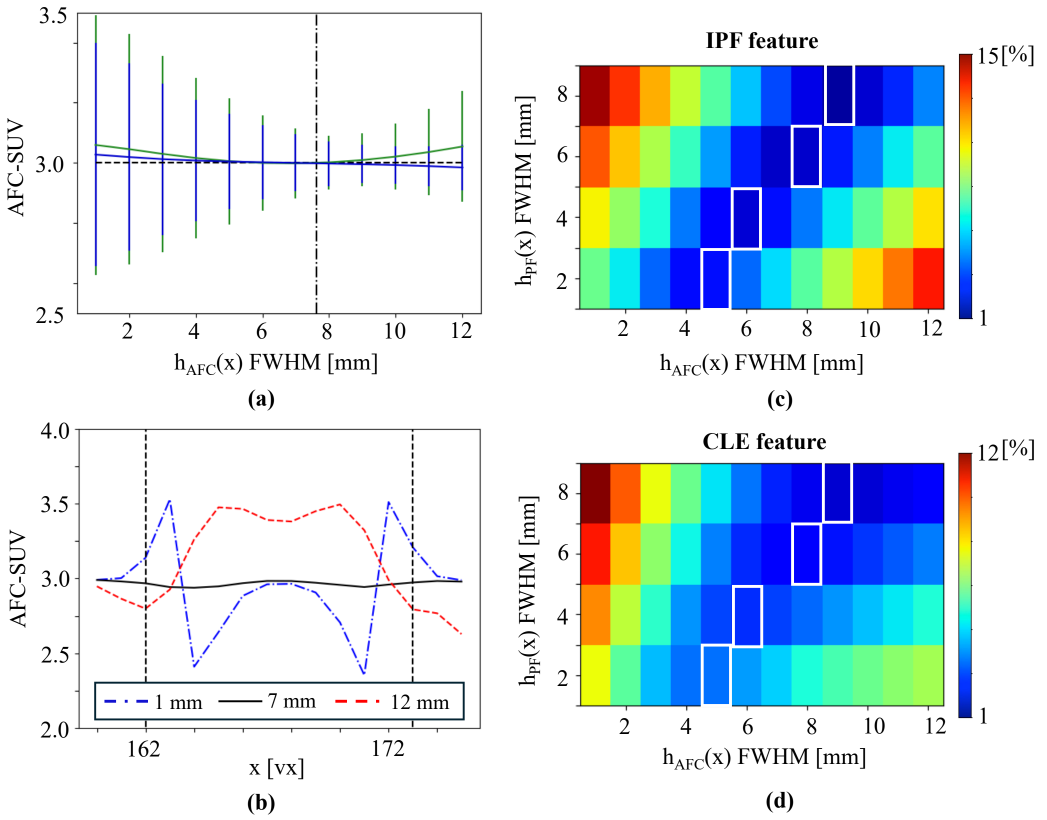


Figure 3.4: AC-AF-corrected reconstructed images of phantom 2 (OSEM 1000 updates, 6 subsets). AFs determined from mu-map $\otimes h_{AFC}(x)$ with various FWHM, $h_{IR}(x)$ = 4.7 mm^3, $h_{AC}(x)$ = 4.7 mm^3, $h_{PF}(x)$ = 6.0 mm^3. **Left**: mu-map used for AF determination less smooth than the reconstructed PET image results in photopoenic regions on the high density side of tissue boundaries; **middle**: mu-map smoothing approximated the mean measured reconstructed transaxial image resolution (7.6 ± 0.27 mm^3) resulting in a reduction of artefacts in the lungs; **right** mu-map used for AF determination smoother than the reconstructed PET image results in photopoenic regions on the low density side of tissue boundaries.

ied is shown in Figure 3.5(a) for $h_{PF}(x)$ = 6.0 mm^3. The optimal smoothing for AFC-SUV accuracy, for both the high- and low-density lung pathologies, approximates, to the nearest millimetre (only integer values of $h_{AFC}(x)$ were investigated), the measured reconstructed image resolution. This was observed for all PFs investigated, Figure 3.5(c)(d). When the mu-map resolution for AF determination is higher than the reconstructed PET image resolution, AFC-SUV in the low-density CLE feature is artificially increased (0.93 % at $h_{AFC}(x)$ = 1 mm^3; CoV = 10.7 %) but decreased when the mu-map resolution is lower than the reconstructed PET image resolution (-0.50 % for $h_{AFC}(x)$ = 12 mm^3; CoV = 3.72 %). For the higher density IPF region the AFC-SUV is overestimated by 2.0 % (CoV = 12.5 %) at $h_{AFC}(x)$ = 1 mm^3; and 1.8 % for $h_{AFC}(x)$ = 12 mm^3 (CoV = 7.4 %). Despite an error in quantification of $<$ 2 %, the artefacts that result from a mismatch in resolution can be seen for the AF-corrected IPF region in Figure 3.5(b). When the CT is under-smoothed for AFC, ringing artefacts are observed; when the CT is over-smoothed, the AFs are over-estimated, and the feature over-corrected, Figure 3.4.

The heatmaps in Figure 3.6 depict the interaction of $h_{AC}(x)$ and $h_{AFC}(x)$ for three PFs for phantom 3. It can be seen that $h_{AC}(x) = h_{IR}(x)$ and $h_{AFC}(x) \approx$ measured reconstructed image resolution resulted in low RMSEs for all PFs. However,



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

Figure 3.5: (a) Mean AFC-SUV in CLE (blue) and IPF (green) features in phantom 2 for changing $h_{AFC}(x)$, $h_{IR}(x) = 4.7$ mm^3, $h_{AC}(x) = 4.7$ mm^3, $h_{PF}(x) = 6.0$ mm^3. The dashed-dotted vertical line indicates the measured reconstructed image resolution, FWHM(xyz); the simulated AFC-SUV GT is represented by the horizontal dotted line. Error bars indicate 1σ; (b) x-direction profiles through the AF-corrected IPF feature for $h_{AFC}(x) = 1$ mm^3 FWHM (blue), 7 mm^3 (black), 12 mm^3 (red). The VOI boundaries are shown by the vertical dotted lines. (c)(d) heatmaps depicting the CoV in the IPF and CLE regions respectively for varying $h_{PF}(x)$ and $h_{AFC}(x)$; the $h_{AFC}(x)$ that best matches the reconstructed resolution as measured by the perturbation method are highlighted in white for each $h_{PF}(x)$.

for each PF, a diagonal “valley” in global RMSE minima was observed, indicating that there is not a clearly defined kernel pair solution.

3.3.4 Discussion

Artefacts observed in the AC-AFC PET images at tissue density boundaries were minimised when the mu-map resolution matched the intrinsic resolution of the simulation model for AC and approximated, to the nearest millimetre (only integer values of $h_{AFC}(x)$ were investigated), the reconstructed PET image resolution for

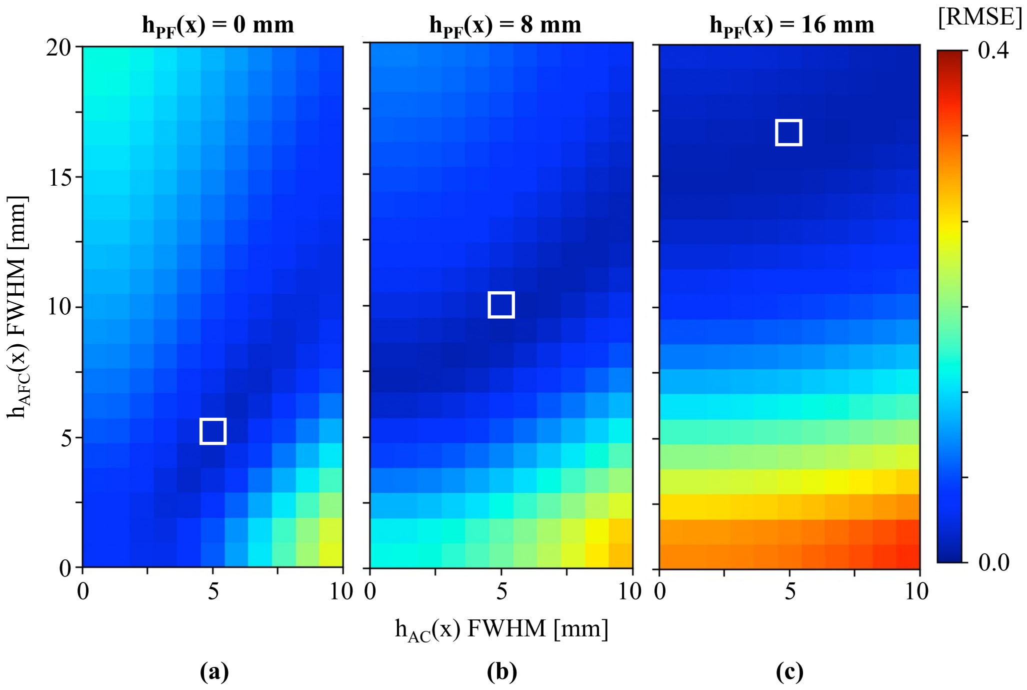


Figure 3.6: Global RMSE in the AC-AFC non-TOF reconstructed image (200 MLEM iterations) for various $h_{PF}(x)$ FWHMs for phantom 3. $h_{IR}(x)$ = 5 mm^3. Mean measured reconstructed image resolution was (a) 5.9 mm^3; (b) 9.5 mm^3; (c) 16.7 mm^3.

AFC. This was observed in all phantoms. The CT should therefore be smoothed to match intrinsic scanner resolution for AC but match the reconstructed PET image resolution for AFC. As the smoothness of the mu-map for CTAC is determined by the manufacturers and incorporated in-reconstruction, all simulations in the remainder of this thesis will assume a fixed kernel for AC.

Mismatched resolutions for AFC result in a greater error in image quantification when imaging larger-scale structures i.e. the 22 mm diameter IPF sphere in phantom 2, compared to the fine-scale structure seen in the CLE patch due to the magnitude of the blurring kernels implemented smoothing out fine-scale structures.

A mismatched mu-map resolution has opposing effects on the resultant activity distribution for AC and AFC. Therefore, if the same mu-map resolution is used for both AC and AFC, the effect of the mismatch for AC can be partially corrected for by the AFC. We believe this to be due to mismatches in PET-CT resolution for AC resulting in local errors in non-TOF PET [66], which can be (approximately) corrected for by using a mismatched resolution for AFC. When $h_{AC}(x)$ matched the simulated intrinsic PET scanner resolution ($h_{IR}(x)$ = 5 mm) and $h_{AFC}(x)$ matched the spatial resolution of the reconstructed PET image the RMSE falls within the "valley" of RMSE minima for all image spatial resolutions investigated. As discussed in section 2.7.1.1, errors in the PET image at a distance from the AC mismatch would increase with time resolution in TOF PET; a more unique kernel pair solution may therefore be expected in TOF PET and will be investigated with measured data in Chapter 4.

Residual image artefacts, despite optimised kernels, are likely due to spatially variant reconstructed image resolution; in these simulations, only spatially invariant global kernels at selected, isotropic, $h_{AC}(x)$ and $h_{AFC}(x)$ were investigated for each PF. Spatially variant anisotropic kernels are discussed in Section 3.4. The simulations were reconstructed with > 160 full passes of the data and may be prohibitively slow for a clinical workflow. The magnitude of AFC-SUV quantification error in relation to the number of iterative updates is discussed in Section 3.4.

3.4 AFC kernel dependence on reconstruction convergence and location

It has been demonstrated in Section 3.3 that the AFC requires an accurate knowledge of the resolution of the reconstructed PET images. The simulations discussed in Section 3.3 assumed a spatially invariant resolution in the reconstructed PET image, such that a global blurring kernel could be applied to the mu-map for determination of AFs. However, reconstructed PET image resolution cannot be considered spatially invariant, as discussed in Section 2.5.

The purpose of the following analytical simulations was to determine the magnitude of spatial variance in non-TOF reconstructed PET image resolution for the phantoms simulated in Section 3.3. The dependence on iterative reconstruction convergence was investigated with the point source insertion-and-subtraction method to determine localised reconstructed image resolution, see Section 2.5.

3.4.1 Methods

The two phantoms simulated to have homogeneous uptake post-AFC, phantoms 2 and 3, were used to investigate the dependence of spatial variance in measured reconstructed image resolution with respect to iterative algorithm convergence. The parameters listed in Section 3.2.2 were used for the simulation with the exceptions that: (i) $h_{AC}(x) = h_{IR}(x)$; (ii) 4000 iterative updates, 6 subsets (666 full iterations) were conducted; (iii) no PF was applied. Single-voxel point sources were positioned on the central axial slice in multiple positions within the phantoms / FOV, as shown in Figure 3.7. All data were noiseless.

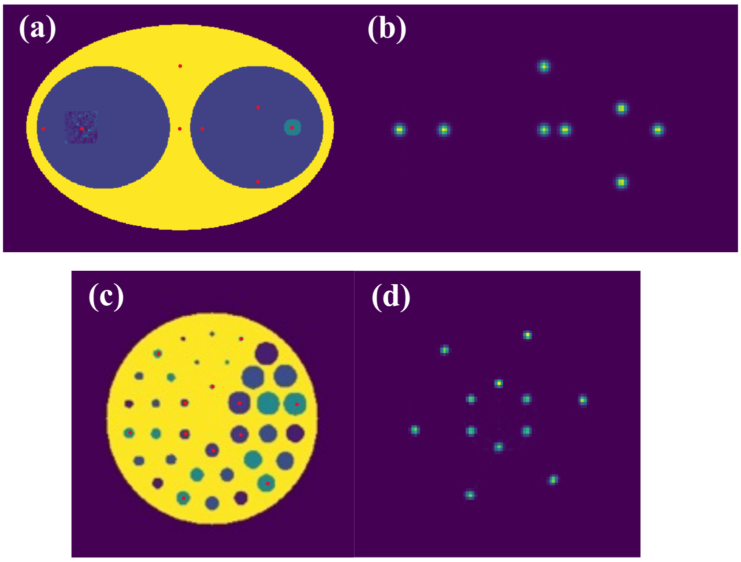


Figure 3.7: Point source insertion-and-subtraction (a) simulated emission of phantom 2 with embedded point sources shown in red; (b) difference image (image with point sources minus image without point sources) for OSEM 2000 iterative updates, 6 subsets; (c) simulated emission of phantom 3 with embedded point sources in red; (d) difference image for OSEM 400 iterative updates, 6 subsets.

Image artefact magnitude dependence on iterative reconstruction convergence was assessed through determination of the CoV of AF-corrected reconstructions of phantom 2. A 6 mm^3 PF was applied, as an approximation, to the nearest millimetre, for a PF that may be applied clinically. The effect of using a global anisotropic kernel for AFC, measured from the point source insertion-and-subtraction method, to decrease image artefacts was investigated.

3.4.2 Results

The local reconstructed image resolution is assumed to approximate the measured radial, tangential, and axial FWHM of the reconstructed local impulse response (LIR). The extent to which the global transaxial reconstructed image resolution (mean radial and tangential FWHM of all point sources) varied with the number of iterative updates is shown in Figure 3.8. High numbers of iterative updates were needed to achieve measured reconstructed image resolution stability for both phantoms. As the number of iterative updates increases, the standard deviation on the mean, computed over all point sources in the FOV, decreases (CoV reduced from 13.2 % to 3.80 % for phantom 2 and from 10.7 % to 9.60 % for phantom 3). The more complex object, phantom 3, was slower to converge than phantom 2 (mean measured transaxial reconstructed image resolution at 4000 iterative updates being 4.67 ± 0.45 mm and 4.49 ± 0.17 mm, respectively).

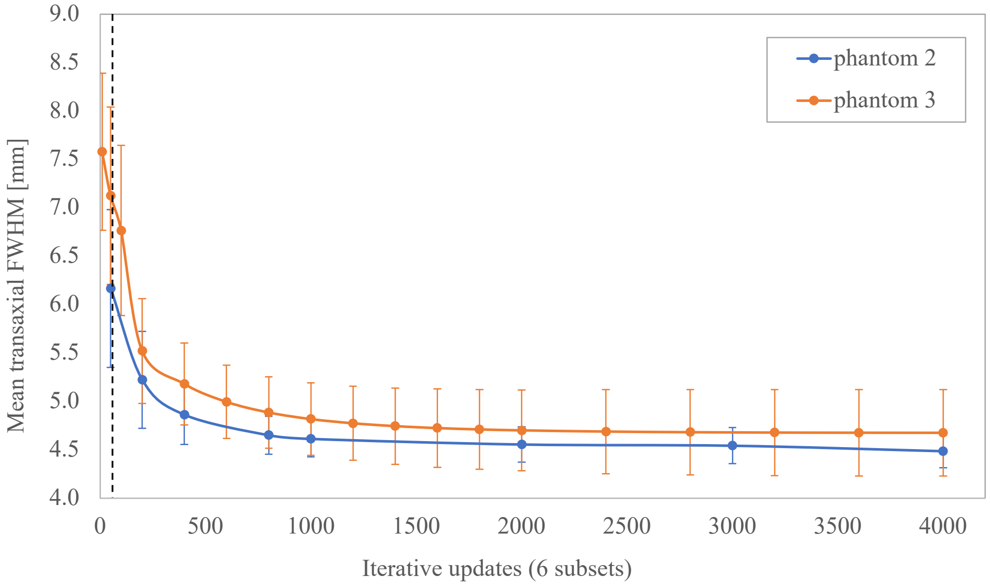


Figure 3.8: Mean measured transaxial image resolution for noiseless point sources at various positions in the phantom (/FOV) for increasing numbers of OSEM iterative updates, 6 subsets. $h_{PF}(x) = 0$ mm. Error bars depict standard deviation on the mean, computed over eight locations for phantom 2 and twelve locations for phantom 3. The black dashed vertical line indicates 50 MLEM iterations equivalence, which is representative of the number of iterative updates that may be used clinically.

Figure 3.9 depicts measured transaxial reconstructed image resolution in the “lungs” of phantom 2 for each point source ($h_{PF}(x)$ = 6 mm^3 as is typical in the clinic). It can be seen that there is spatial variation in the resolution, even after high numbers of iterative updates, depending on the position of the perturbation within the FOV, the contrast, and attenuation. At 1000 iterative updates the mean reconstructed image resolution across all point sources positioned in the lung was 7.36 ± 0.02 mm radially, 7.48 ± 0.02 mm tangentially, and 7.97 ± 0.01 mm axially. Measured reconstructed image resolution does not alter beyond 1000 updates for this phantom, ΔFWHM(xyz) $<$ 0.2 %.

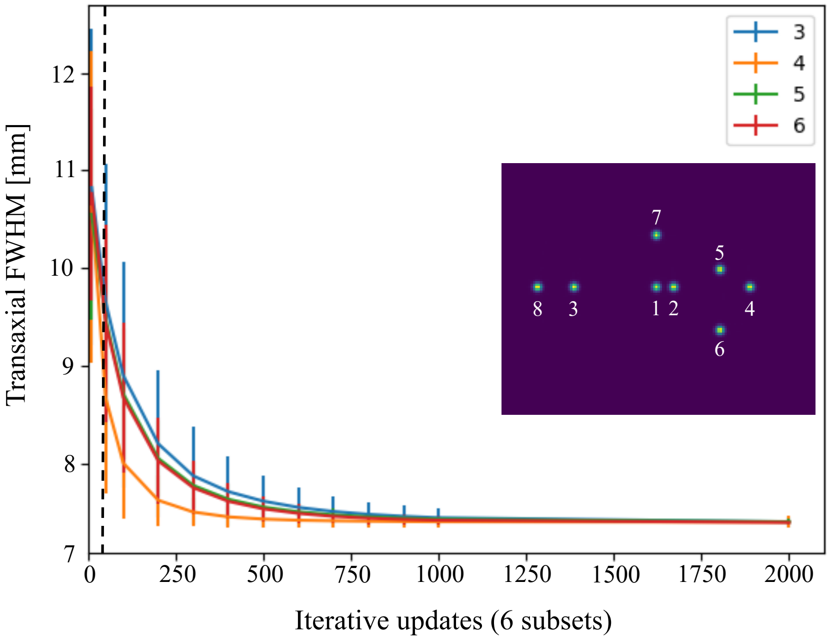


Figure 3.9: Measured transaxial reconstructed image resolution for point sources simulated in noiseless conditions at various positions in the lung for increasing numbers of non-TOF iterative updates, 6 subsets. $h_{AC}(x)$ = 4.7 mm, $h_{PF}(x)$ = 6.0 mm. The point source positioned in the centre of the CLE feature is shown in blue (point source 3), IPF in orange (point source 4), and healthy lung in green and red; error bars depict one standard deviation of FWHM(x) and FWHM(y). The black dashed vertical line indicates 50 MLEM iterations equivalence, which is representative of the number of iterative updates that may be used clinically.

In Section 3.3.3, it was demonstrated that image artefacts were minimised when $h_{AFC}(x)$ = 8 mm^3 (only integer values for each of the blurring kernels were investigated), Figure 3.4(b). Kernels were assumed to be isotropic, and the reconstructed image resolution was assumed to be spatially invariant. At lower numbers of iterative updates the magnitude of the image artefacts is greater, despite the same blurring kernels being applied, Figure 3.10.

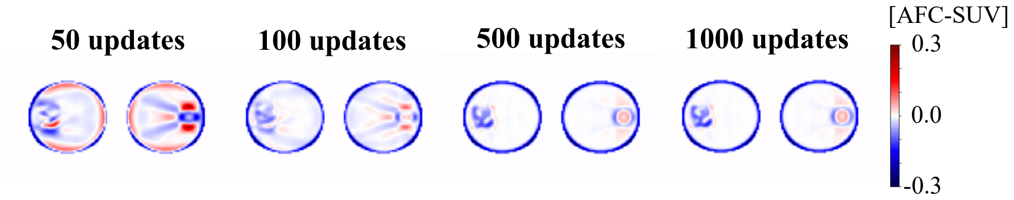


Figure 3.10: Central axial slice of difference images, AF-corrected reconstruction with respect to AF-corrected GT; $h_{IR}(x)$ = 4.7 mm^3, $h_{AC}(x)$ = 4.7 mm^3, $h_{PF}(x)$ = 6.0 mm^3, $h_{AFC}(x)$ = 8.0 mm^3. The magnitude of the artefacts reduces with increasing numbers of iterative updates; little change is seen beyond 500 iterative updates.

If a global anisotropic kernel, $h_{AFC}(x,y,z)$, that matches the mean measured reconstructed image resolution (determined from the image with the same number of iterative updates that is to be corrected) is applied, the CoV in the lungs is reduced with respect to the optimal isotropic kernel at all iterative updates investigated, Figure 3.11.

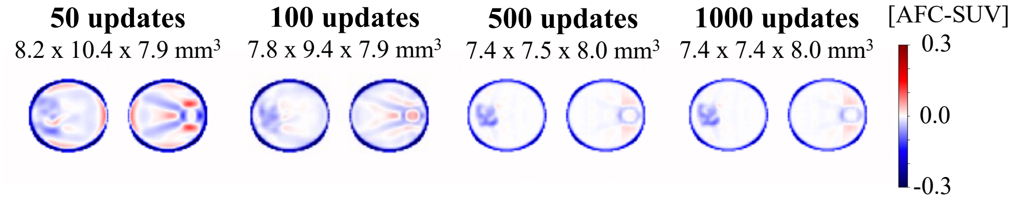


Figure 3.11: Central axial slice of difference images, AF-corrected reconstruction with respect to AF-corrected GT; $h_{IR}(x)$ = 4.7 mm^3, $h_{AC}(x)$ = 4.7 mm^3, $h_{PF}(x)$ = 6.0 mm^3; $h_{AFC}(x,y,z)$ equivalent to measured reconstructed image resolution in each case. The magnitude of the artefacts reduces with increasing numbers of iterative updates; little change is seen beyond 500 iterative updates.

3.4.3 Discussion

These results demonstrate that the more complex test object, with multiple regions of varying tracer concentration, converged slower than the object with smaller variations in activity distribution. This phenomenon has been well documented. Hutton et al. described the rate of convergence for OSEM reconstructions as complex and dependent on the signal and frequencies present at the point of measurement [116]. Iterative algorithms, such as MLEM and OSEM, are noted to be slower to converge in regions of low tracer uptake with respect to FBP [117]. In our comparison of convergence rates between phantoms no PF was applied; the difference in convergence rates would be reduced once a PF is applied, as would be done in clinical practice.

We looked at the implications of OSEM algorithm convergence on AFC. Artefacts were more evident in the AC-AFC images reconstructed with low numbers of iterative updates than in the images that had approached convergence, despite $h_{AFC}(x)$ being determined from the same image that was to be corrected. This suggests that the perturbation and underlying tissue structure converge at different rates. However, despite a lack of stability for measured reconstructed image resolution at low numbers of iterations, the error in AFC-SUV quantification is small for these simplistic phantoms, Figure 3.11.

The limitations of this work include that only non-TOF iterative reconstruction has been investigated on noiseless data. More work is needed to investigate the convergence rate of clinically relevant reconstructions for patient-realistic objects before conclusions can be drawn as to the number of iterative updates needed for accurate AF-corrected image quantification. Recommending that a large number of iterations be used for quantitative clinical lung imaging will increase reconstruction time and may not be practicable in a clinical workflow. In addition, increasing the number of iterations for patient data is known to amplify noise, potentially requiring additional post-reconstruction smoothing, thus reducing the reconstructed image resolution. Chapter 4 investigates the effects of non-TOF convergence on noiseless patient-realistic digital data, to gain a better understanding of the magnitude of quantitative error in highly heterogeneous media. In Chapter 5, these simulations are validated with scanned phantoms that mimic the shape, total activity, and changes in attenuation expected in a patient, allowing reconstruction recommendations to be made.

3.5 AFC for non-uniform parenchymal tissue uptake

3.5.1 Introduction

In Section 3.3 it has been demonstrated that, when parenchymal tissue uptake is uniform, artefacts in the AC-AFC image are minimised when the resolution of the CT for determination of AFs matches the PET reconstructed image resolution. This section investigates the limitations of the AFC method for non-uniform tissue uptake and conducts a preliminary exploration into a proof-of-concept PVC-AFC method. Although uniform tissue uptake can be assumed for ^{18}F-FDG IPF [101], there is the potential for non-uniform uptake in focal disease and lung diseases that result in regions of necrotic tissue, e.g. tuberculosis, and / or when alternative radiopharmaceuticals are used.

3.5.2 Limits of current AFC method

3.5.2.1 Simulations

In the first instance, a simplified elliptical phantom containing a single 22 mm diameter spherical feature, with a contrast of 2 with respect to background, was utilised to investigate non-uniform parenchymal uptake (phantom 4: Figure 3.12, Table 3.5). The GT AFC-SUV of the feature is 13.57. For ease of simulation, no reconstruction or down-sampling was performed; the GT was instead smoothed with a 3D Gaussian kernel. The details of the smoothing are listed in Table 3.6.

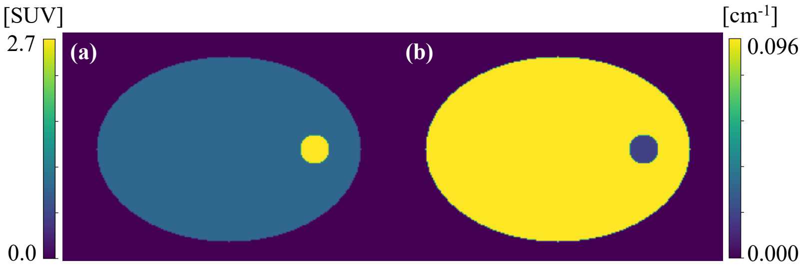


Figure 3.12: Phantom 4: (a) simulated emission; (b) simulated mu-map containing a 22 mm diameter sphere of healthy lung LAC in an ellipse representative of soft tissue LAC. The emission is designed to have a sphere-to-background contrast of 3.

Phantom 3 was modified to result in an inhomogeneous AC-AFC image; the

Table 3.5: Phantom 4 specifications, the dimensions, simulated LACs of the mu-map, and simulated SUV of the emission are given.

Shell dimensions [mm]	Ellipse: 300 x 200 ϕ
Insert dimensions[mm]	Sphere: 22 ϕ
Shell SUV (C_b)	0.9
Insert SUV (C_f)	2.7
Shell LACs [cm^{-1}]	Soft tissue: 0.0960
Insert LACs [cm^{-1}]	Healthy lung at expiration: 0.0191

Table 3.6: Smoothing parameters for phantom 4.

Simulated image voxel size	1.000 x 1.000 x 1.635 mm^3
Gaussian smoothing to approximate reconstruction, $h_{PET}(x)$	6 mm^3 (to approx. 3 mm^3 recon. voxels and a 4 mm^3 PF)
$h_{AFC}(x)$	6 mm^3

features in the emission image were simulated to have a feature-to-background-ratio of 2:1. Noiseless analytic simulations and reconstructions were then conducted using the same methodology as described in the workflow in Figure 3.2 and Tables 3.2 and 3.3, with the exception that $h_{AC}(x)$ was set to match $h_{IR}(x)$. The PF was fixed at 6 mm^3, as an approximation, to the nearest millimetre, for a PF that may be applied clinically. The feature of interest, ringed in red in Figure 3.13, is a 22 mm diameter sphere with an LAC of 0.0191 cm^{-1}, resulting in an expected AFC-SUV of 10.05. The reconstructed image resolution was measured using the point source insertion-and-subtraction method with a single voxel point source placed in the centre of a single feature of interest, as indicated in Figure 3.13(a); $h_{AFC}(x)$ was set to match the measured LIR.

3.5.2.2 Results

When $h_{AFC}(x)$ was set to match the simulated image resolution for phantom 4, $h_{PET}(x)$ = 6 mm^3, the resultant AF-corrected feature did not match the simulated AF-corrected GT, Figure 3.14. Near-recovery of the feature was seen at the centre of the feature (AFC-SUV$_{max}$ = 13.55), however, AFC-SUV$_{mean}$ in the feature was 6.46 ± 3.43, a bias of -74.7 %.

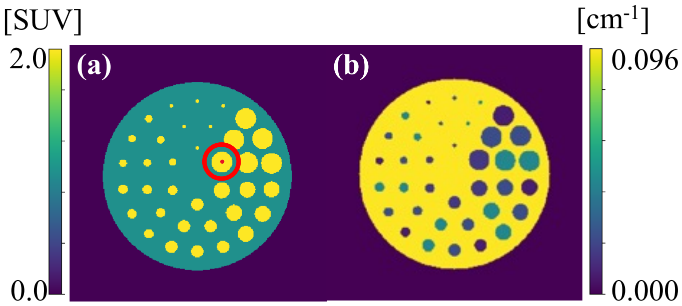


Figure 3.13: Modified phantom 3; (a) simulated emission, all features have a feature-to-background ratio of 2:1, the feature of interest is ringed in red with a point source in the centre; (b) simulated mu-map, LACs as listed in Table 3.1.

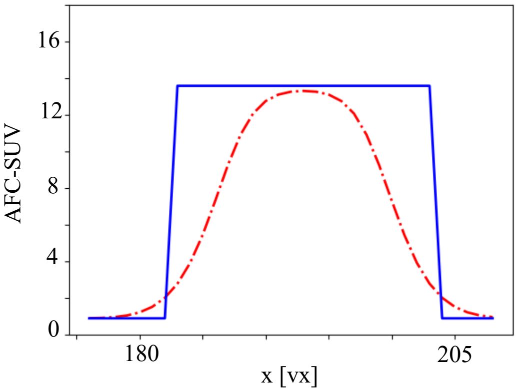


Figure 3.14: Profile in the x-direction through AF-corrected feature in phantom 4: smoothed emission, $h_{AFC}(x)$ = 6 mm^3 (red dash-dotted line); AF-corrected GT emission (blue solid line).

The reconstructed image resolution in the feature of interest in modified phantom 3 was estimated to be 7.43 x 7.46 x 7.19 mm^3; the mu-map was smoothed with this anisotropic kernel for determination of the AF map. As for phantom 4, full recovery was not achieved for the feature in the AC-AFC image (AFC-SUV$_{max}$ = 9.80). AFC-SUV$_{mean}$ in the feature was 5.40 ± 3.08, a bias of -46.3%.

3.5.2.3 Discussion

The lack of full feature recovery, and underestimation of AFC-SUV$_{mean}$, is due to count spill-out from the feature into the background, which had not needed to be considered in the uniform parenchymal uptake case. A correction for the spill-in/spill-out component, referred to as PVC for the remainder of this chapter, in addition to the fraction of air, is therefore required when parenchymal tissue uptake is not uniform.

3.5.3 PVC-AFC method

It is proposed that the reconstructed PET image be PV-corrected prior to AFC. A single post-reconstruction PVC method, utilising the high frequency information from the CT, was investigated in this proof-of-concept study. We hypothesise that the PVC methodology will enhance the emission resolution to match the CT resolution such that AFC can be conducted at CT resolution.

3.5.3.1 Background

PVC methods have been proposed which compensate for the effects of image system blurring on the PET data. PVC approaches can either be included in the reconstruction or applied as a post-reconstruction correction (ROI-based or voxel-based), as described in the extensive review by Erlandsson et al. in 2012 [75].

The Müller-Gärtner (MG) PVC method [118] is a post-reconstruction voxelwise method. It was developed for the brain and so does not take AFs into consideration. The MG method involves segmenting the region of interest (grey matter (GM)), the white matter (WM) and cerebrospinal fluid (CSF) on the anatomical image to generate binary masks. CSF is assumed to be devoid of activity. The binary WM map is scaled to estimate the true WM mean value, which is then convolved with the kernel and subtracted from the original PET image. The GM binary mask is also convolved by the system PSF to produce voxel-wise RCs. PVC is then performed by dividing the WM spill-in corrected PET image by the RC map.

The division by the RC factors in the MG method is similar to the division by (1-AF) in the AFC method. We therefore investigated if the MG method can be extended to the lung. For simplicity, we used a test case where the background region has strictly positive activity such that the CSF term can be removed. The feature can be considered as GM and the background as WM, resulting in equation 3.3.

$$f_f(x) = \frac{f_o(x) - [C_b p_b(x)] \otimes h(x)}{p_f(x) \otimes h(x)} \tag{3.3}$$

where: $f_o(x)$ is the original (reconstructed) image; $f_f(x)$ is the PV corrected feature;

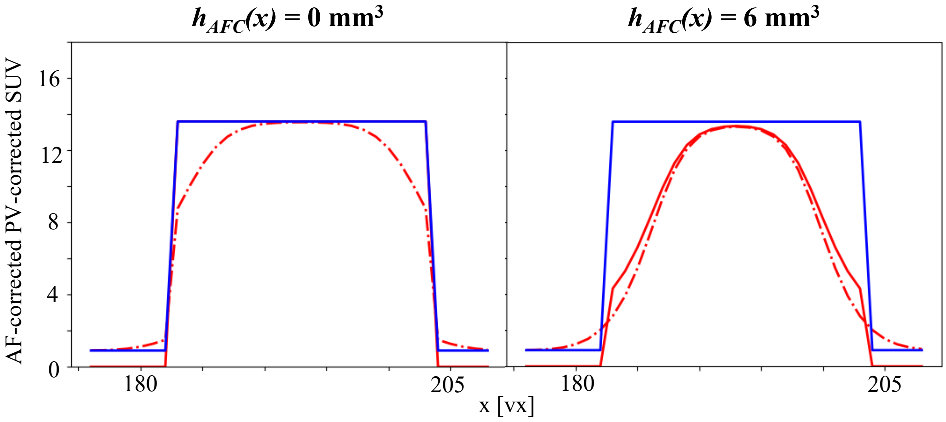
C_b is the estimated background mean value; $p_b(x)$ and $p_f(x)$ are the masks where $p(x) \in [0,1]$ of the background and feature respectively; $h(x)$ is the estimated PSF of the image.

3.5.3.2 Methods

The MG method of PVC followed by AFC was tested with phantom 4. The emission was again smoothed by a Gaussian kernel 6 mm^3 FWHM ($h_{PET}(x)$), to simulate reconstruction. The unsmoothed mu-map was used to segment the feature and background regions. The emission was PV-corrected with $h(x) = h_{PET}(x)$, before being AF-corrected, $h_{AFC}(x)$ = 0 - 12 mm^3.

3.5.3.3 Results

The profiles through the centre of the feature in the x-direction are shown in Figure 3.15 for (a) $h_{AFC}(x)$ set to match the CT resolution and (b) the unPV-corrected PET image resolution. Full recovery of the AF-corrected feature is demonstrated when the PET image was PV-corrected with a kernel that matched the simulated PET image resolution, $h(x) = h_{PET}(x)$, and then AF-corrected with an AF-map determined from the unsmoothed CT, $h_{AFC}(x)$ = 0 mm^3, Figure 3.15(a).



(a) (b)

Figure 3.15: Profiles in the x-direction through the centre of feature when the AF map has been matched to (a) the CT resolution, $h_{AFC}(x)$ = 0 mm^3 and (b) matched to the unPV-corrected PET image resolution, $h_{AFC}(x) = h_{PET}(x)$ = 6 mm^3. The GT AFC-SUV is shown in blue, the unPV-corrected AFC-SUV in red dot-dashed, and the AF-corrected, PV-corrected profile in solid red.

3.5.3.4 Discussion

When lung tissue uptake is non-uniform, the PET image must be PV-corrected prior to AFC; the AFC must use AFs determined from a CT resolution image. In this proof-of-concept PVC-AFC study, only a single PVC method was investigated. The results demonstrate that the hypothesis that the CT used for AF determination should match the reconstructed PET image resolution still holds for inhomogeneous tissue uptake, provided that an accurate PVC has been conducted.

3.6 Conclusions

Artefacts observed in the AC-AFC PET images at tissue density boundaries are minimised when the CT resolution matches the simulated intrinsic PET scanner resolution for AC and approximates, to the nearest millimetre (only integer values of $h_{AFC}(x)$ were investigated), the reconstructed PET image resolution for AFC. AF-corrected PET image quantification is therefore dependent on PET and CT being matched in terms of location, density, and resolution.

A mismatched PET and CT resolution have opposing effects on the resultant activity density for AC and AFC. Thielemans et al. demonstrated that, in non-TOF PET, the effects of attenuation mismatch are local [66]. Therefore, if the same mu-map resolution is used for both AC and AFC, the effect of the mismatch for AC can be partially corrected for by the AFC. A limitation of these simulations is that only non-TOF behaviours have been studied; Emond et al. demonstrated that in TOF PET, the errors at a distance from an attenuation mismatch increased with improved timing resolution [68]. The compensating effect is therefore likely to be reduced for TOF PET.

These simulations use simplistic digital phantoms and assume spatially invariant Gaussian kernels, the implications of using a single global kernel for AFC is investigated in patient-realistic simulations in Chapter 4. Positron range was not considered as part of these simulations; the magnitude of the effect on AFC kernel determination for ^{18}F is investigated in Chapter 5.

Large numbers of iterative updates were needed to achieve stability in the measured reconstructed image resolution, which has potential implications when designing a clinical lung imaging protocol that includes AFC. Our simulations have shown that reconstructed PET spatial resolution can be spatially variant, but these simulations were noiseless, simplistic, and used settings to investigate trends. Further investigation is needed to understand the impact in a clinical setting, see Chapters 4 and 5.

Our proof-of-concept simulation of non-uniform tissue uptake demonstrated that the spill-in/spill-out component of the PVE should be corrected for prior to AFC. The kernel for AFC should match the PV-corrected PET image resolution.

Chapter 4

Kernel determination methodologies

4.1 Introduction

As an alternative to the point source insertion-and-subtraction method, Joshi et al. [119] determined the reconstructed resolution of clinical scanners from scans of the Hoffman brain phantom. A digital Hoffman phantom was smoothed in all three dimensions with incremental FHWM Gaussian kernels to obtain a library of the digital phantom at various resolutions. The effective resolution was estimated by determining the smoothed digital phantom that was closest to the reconstructed image using a least squares approach. The advantage of this method over the point source insertion-and-subtraction method is that the former does not require simulation capabilities. It does, however, rely on an accurate digital representation of the measured object.

This chapter investigates three different methodologies by which the reconstructed image resolution, as a function of reconstruction algorithm and convergence, can be determined for smoothing of the CT for AFC. In simulation studies, the point source insertion-and-subtraction method was compared to two variations on the methodology similar to that proposed by Joshi et al.. A single methodology was then selected and the feasibility of its application to measured phantom data was assessed.

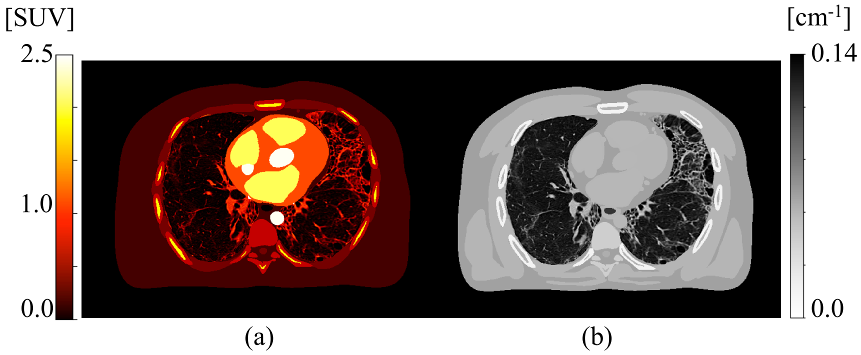
This chapter is based on Leek et al. [120].

4.2 Simulations

4.2.1 Methods

Three different methods for determining the optimal kernel with which to smooth the CT for AFC were investigated with noiseless simulations and non-TOF reconstructions of a digital patient-realistic phantom.

A patient-realistic mu-map GT phantom was constructed by substituting the lungs in a digital XCAT phantom [121] with a diagnostic CT from an IPF patient, Figure 4.1. This map was used to create an emission map of ^{18}F uptake. It was assumed that lung tissue uptake was uniform with a homogeneous AF-corrected SUV (SUV_{AFC}) equal to 0.996, see phantom construction details in Appendix B.



(a) (b)

Figure 4.1: Modified digital XCAT phantom used for simulations (voxel size = 0.61 x 0.61 x 1.50 mm^3); (a) GT emission (pre-AFC); (b) GT mu-map.

Data were simulated and reconstructed with a GE Discovery 710 PET template using STIR [111] via SIRF [112]. Both the emission and mu-maps were convolved with an isotropic Gaussian kernel with a 4.7 mm^3 FWHM prior to forward-projection, to approximate the intrinsic resolution of the GE Discovery 710 [19]. All data were noiseless.

Non-TOF MLEM reconstruction, with 1000 iterations (1000i), to investigate convergence, was performed into 2.71 x 2.71 x 3.27 mm^3 voxels. MLEM, rather than OSEM, was chosen to avoid extra complications caused by the effect of subsets on convergence. The mu-map was smoothed with a kernel that matched the simulated intrinsic resolution of the PET scanner for AC [122], as discussed in Chapter 3. A 6 mm^3 FWHM PF was applied to the reconstructed image, as an approximation, to the nearest millimetre, for a PF that may be applied clinically.

The localised optimal kernel with which to smooth the mu-map for AFC was determined in six 20 mm diameter VOIs – three in healthy lung (HL) tissue, and three in regions of IPF (Figure 4.2). A VOI diameter representative of that which might be quantitatively assessed clinically for interstitial lung disease was chosen; it was ensured the VOI was large enough (approximately twice the expected FWHM of the kernel) to reduce uncertainty on the kernel estimation.

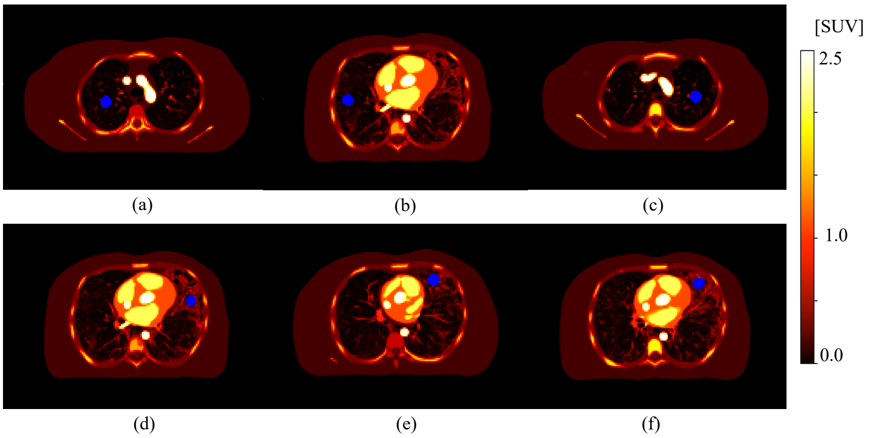


Figure 4.2: Axial slices through the modified XCAT depicting the six VOIs investigated in blue; top row shows VOIs in healthy lung and bottom row VOIs in fibrotic lung (a) $VOI1_{HL}$; (b) $VOI2_{HL}$; (c) $VOI3_{HL}$ (d) $VOI4_{IPF}$; (e) $VOI5_{IPF}$; (f) $VOI6_{IPF}$.

Three methods were investigated (as depicted in Figures 4.3 and 4.4):

1. h_{pts}: the point source insertion-and-subtraction method using a single voxel point-source in the centre of the VOI (point source intensity was set to 1.1 times the underlying voxel value, resulting in a minimum (maximum) reconstructed point source intensity with respect to reconstructed VOI of 0.12 in 0.15±0.07 (0.21 in 0.16±0.06)). The GT point source image was smoothed by kernels of varying width; h_{pts} represents the kernel that resulted in the lowest RMSE with respect to the difference image (reconstructed emission subtracted from the reconstructed emission plus point source) in the VOI.

2. h_{PVC}: the GT emission was smoothed by kernels of varying width. h_{PVC} represents the kernel that resulted in the lowest RMSE in the VOI with respect to the reconstructed emission image, similar to Joshi et al. [119].

3. h_{AFC}: the mu-map was smoothed by kernels of varying width and the reconstructed image AF-corrected with the smoothed mu-map; h_{AFC} denotes the kernel that resulted in the lowest RMSE with respect to the GT AF-corrected VOI.

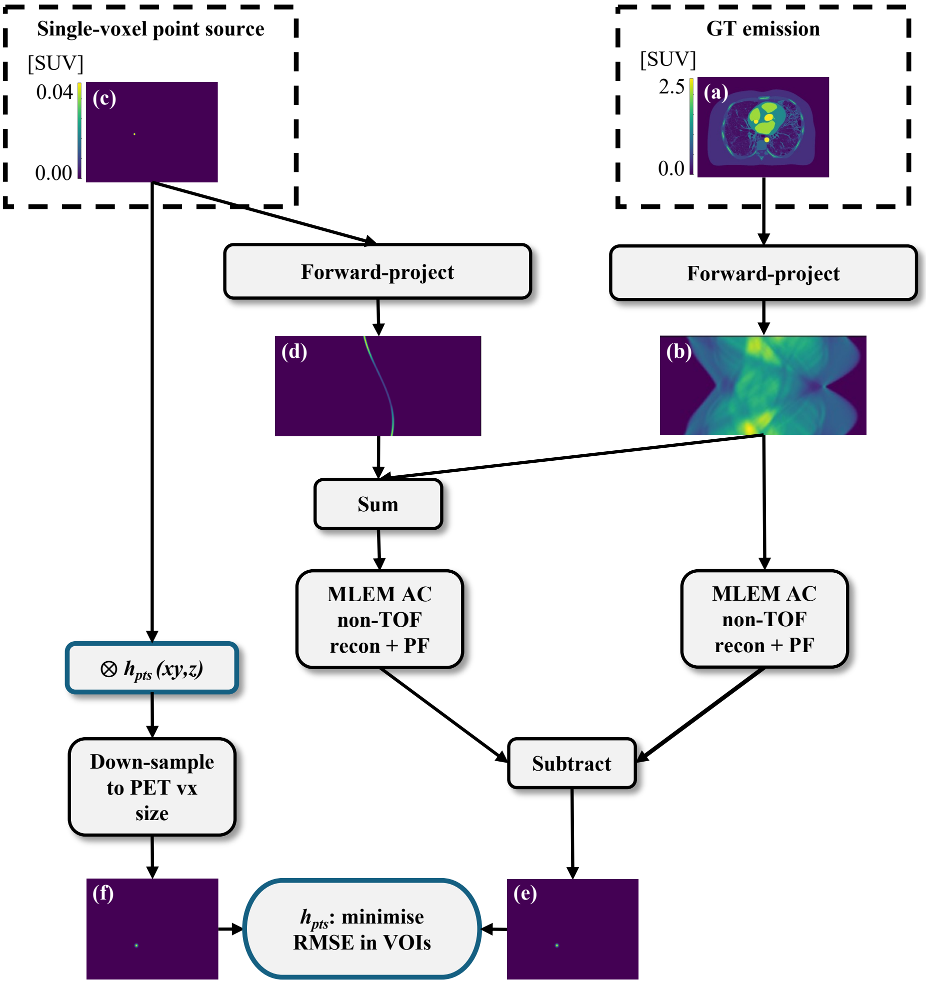


Figure 4.3: Workflow for the point source insertion-and-subtraction method for kernel determination; (a) simulated GT emission; (b) forward-projected and attenuated GT emission; (c) simulated single-voxel point source; (d) forward-projected and attenuated point source; (e) difference image between AC reconstructed and post-filtered (h_{PF} = 6 mm^3) PET with perturbation and without; (f) simulated single-voxel point source convolved with h_{pts} and down-sampled to PET voxel (vx) size. "minimise RMSE in VOIs" refers to the six VOIs in the lung.

Joshi et al. focusses on reducing inter-scanner variation and therefore examined resolution differences at a standardised very small voxel size; this methodology will be referred to as MKA_J and is adopted in Chapter 5 in a move towards harmonisa-

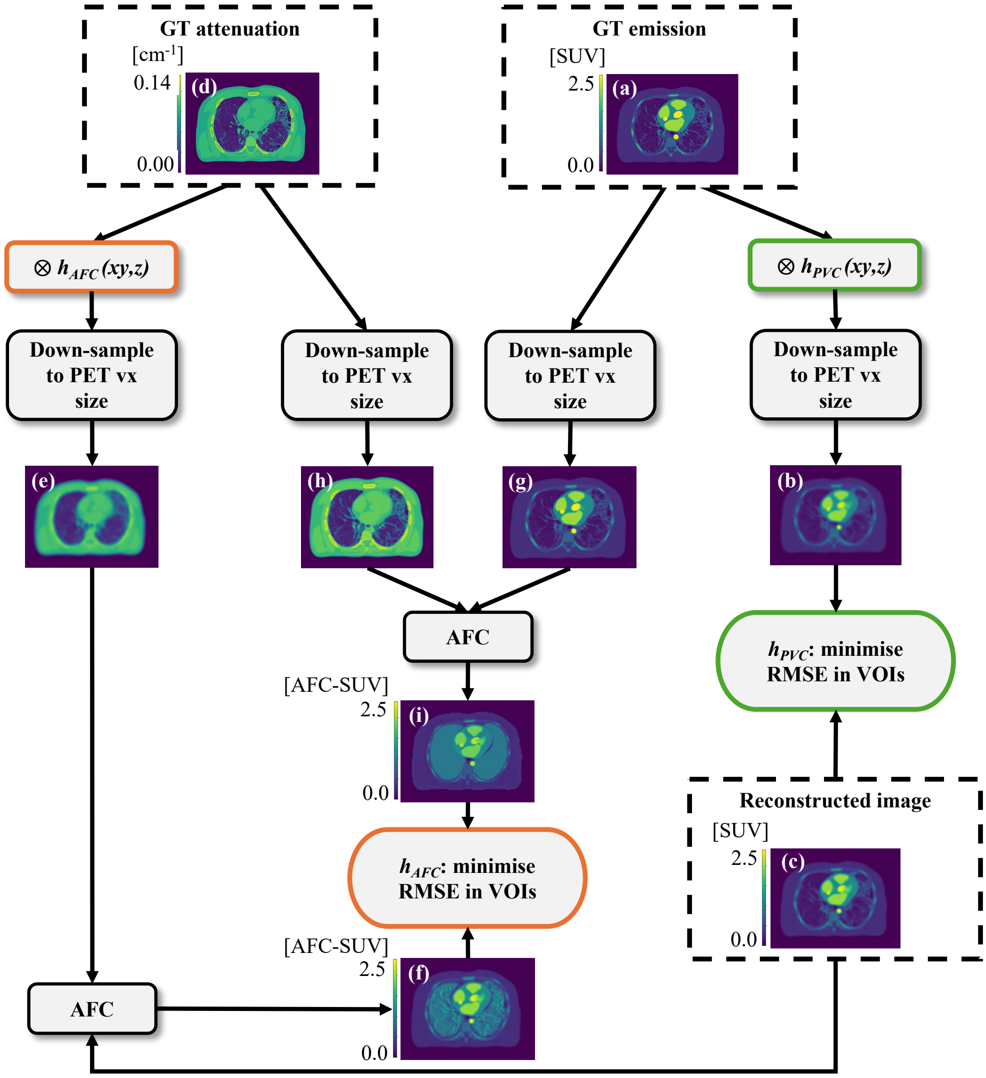


Figure 4.4: Workflow for the h_{PVC} and h_{AFC} methodologies for kernel determination; (a) simulated GT emission; (b) GT emission convolved with h_{PVC} and down-sampled to PET voxel (vx) size; (c) AC non-TOF reconstructed and post-filtered (h_{PF} = 6 mm^3) PET; (d) simulated GT mu-map; (e) mu-map convolved with h_{AFC} and down-sampled to PET vx size; (f) AFC-AC PET (AC non-TOF reconstructed and post-filtered (h_{PF} = 6 mm^3) PET divided by voxel-wise AFs, derived from (e)); (g) GT emission down-sampled to PET vx size; (h) GT mu-map down-sampled to PET vx size; (i) GT AF-corrected PET image. "minimise RMSE in VOIs" refers to the six VOIs in the lung.

tion. The method comparison and validation conducted in this chapter does not address inter-scanner variability and so is conducted at PET image resolution, this will be referred to as the matched kernel approach (MKA). For each method, the GT was convolved with 3D Gaussian kernels of decoupled in-plane and axial resolutions, of increasing FWHM (5 – 15 mm, 0.1 mm increments), before down-sampling to PET voxel size. For conciseness, only mean transaxial-axial kernel width (FWHM(xyz)) is reported in this chapter. The kernels estimated by each method were used to smooth the mu-map for voxel-wise AFC of the reconstructed data and the RMSE of each of the AF-corrected VOIs, with respect to the GT, was assessed. In addition, a whole lung volume (segmented on the HRCT), down-sampled to PET voxel size, and eroded by two voxels isotropically to avoid edge effects) was determined to quantify overall AFC performance.

4.2.2 Results

The simulated data demonstrated that for fewer than 200i the kernel width was dependent on iteration number and VOI position in the lung, Figures 4.5(a-c). For 200i or more, each VOI had converged to a stable (ΔFWHM$(xyz) < 0.5$ mm) estimate of kernel FWHM, independent of iteration number.

A large inter-VOI variation in kernel width estimates was observed between methodologies (range = 7.6 - 14.1 mm). The h_{pts} methodology estimated a lower, more uniform, FWHM(xyz) across all six VOIs, see Figures 4.5(a-c). Even at 200i or more, a large inter-VOI kernel width variation was still observed, Table 4.1. However, at these higher iterations, all three kernel determination methods result in approximately equivalent AF-corrected RMSE in their respective VOIs (6.3 ± 1.3 % for all VOIs; range = 4.6 - 8.8 %), Figures 4.5(d-f). This suggests that a single global kernel could be applied to AF-correct the whole lung. To assess this, the six VOIs were combined to determine a global kernel for each methodology. At 200i or more, the resultant FWHM ranged between 7.7 - 9.5 mm, depending on methodology, Figure 4.6(a). When this global kernel was applied in the voxel-wise AFC of the whole lung volume, the RMSE was consistent across all the methods at 200i or greater (range = 8.8 - 9.6 %), Figure 4.6(b). Using a kernel derived from a single VOI to AF-correct the whole lung resulted in a greater whole lung RMSE than when a global kernel was used, for all methodologies, as can be seen from the shaded areas in Figure 4.6(b). The RMSE of each AF-corrected individual VOI was less than 10 %, at as low as 30i, when the global kernel was used to AF-correct.

Reducing the VOI diameter increased the uncertainty on the estimated kernel

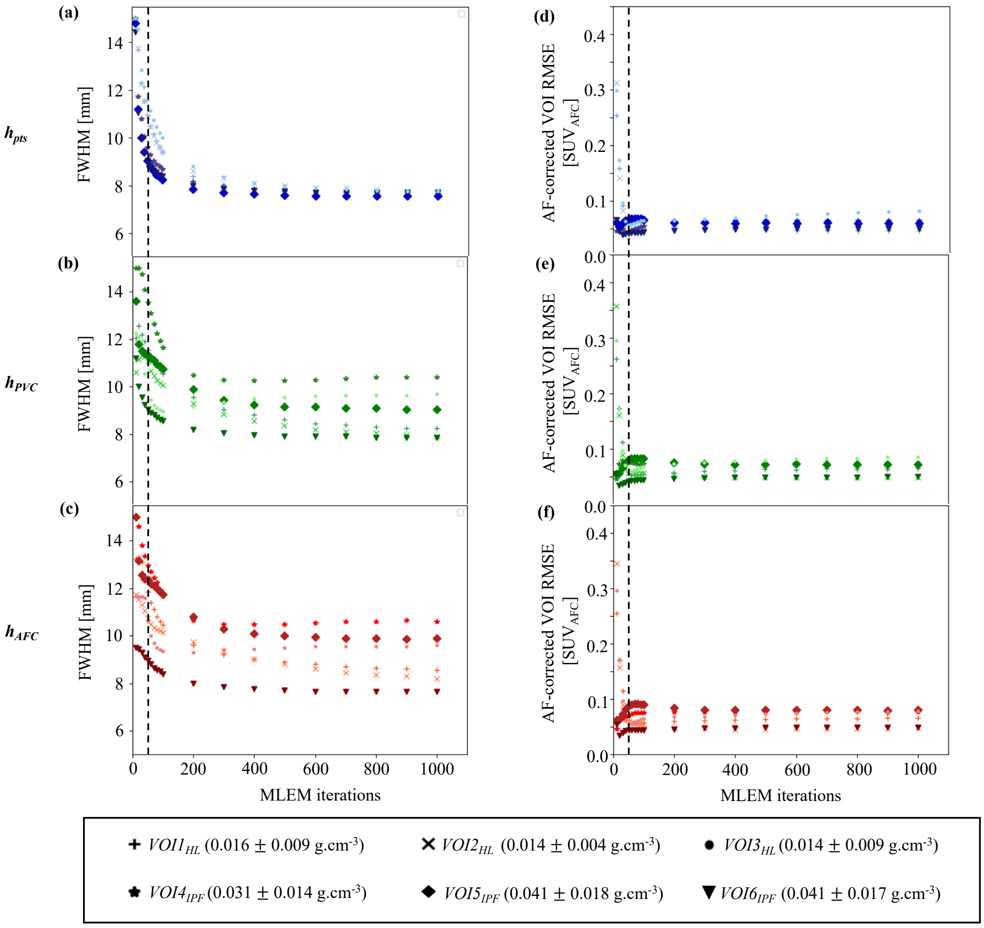


Figure 4.5: Simulated data: (a-c) estimated FHWM(xyz) for 10 to 1000 MLEM iterations for h_{pts}, h_{PVC} and h_{AFC} for each of the six VOIs, simulated on a GE D710 scanner in SIRF; (d-f) RMSE of the AF-corrected reconstructed VOIs at 10 to 1000 MLEM iterations; the mu-map was smoothed by the optimal kernel for each VOI determined for h_{pts}, h_{PVC} and h_{AFC}, respectively. Each marker shape represents a VOI density (mean and 1σ). The black dashed vertical lines indicate 50 MLEM iterations, which is representative of the number of iterative updates that may be used clinically.

width (individual VOI $h_{PVC}(xyz)$ = 8.6 ± 0.6 mm (mean ± 1σ) for 20 mm diameter VOI; $h_{PVC}(xyz)$ = 9.1 ± 1.0 mm for 4 mm diameter VOI).

For comparison, if a kernel is used that matched the intrinsic resolution of the simulated PET scanner, even at 200 or more MLEM iterations, the minimum VOI RMSE is 13.1 % (for VOI2$_{HL}$ at 1000i).

Table 4.1: Simulated data: mean estimated kernel FWHM(xyz) of the six VOIs for each methodology at 200i or greater, and the resultant AF-corrected VOI RMSE when these kernels are applied to the CT for AFC. In each case, the quoted uncertainty is one standard deviation across the six VOIs.

Method	**kernel** FWHM(xyz) **[mm]**	**AF-corrected VOI RMSE** [$\%SUV_{AFC}$]
h_{pts}	7.82 ± 0.25	6.07 ± 1.25
h_{PVC}	9.00 ± 0.87	6.51 ± 1.25
h_{AFC}	9.23 ± 0.99	6.56 ± 1.37

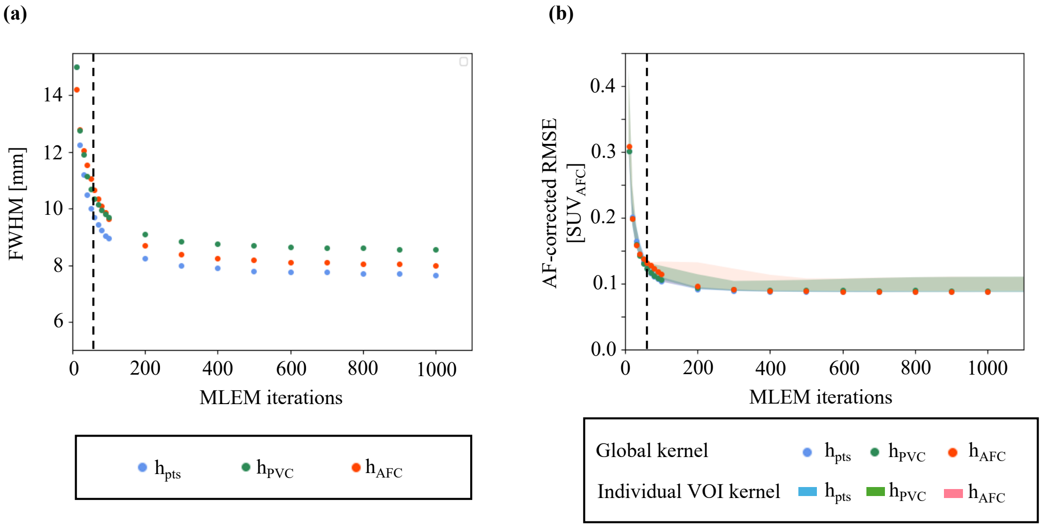


Figure 4.6: Simulated data: (a) estimated FWHM$_{xyz}$ for 10 to 1000 MLEM iterations for the global h_{pts}, h_{PVC} and h_{AFC} kernels determined from the six VOIs combined; (b) RMSE of the AF-corrected whole lung ($\rho = 0.024 \pm 0.018$ g.cm^{-3}) when the mu-map was smoothed by a global kernel derived from the six VOIs combined (dots), and kernels derived from individual VOIs (shaded regions depict range across six VOIs) at 10 to 1000 MLEM iterations. The black dashed vertical lines indicate 50 MLEM iterations, which is representative of the number of iterative updates that may be used clinically.

4.2.3 Discussion

The simulation study demonstrated that the FWHM estimates stabilised after approximately 200 non-TOF MLEM iterations, Figures 4.5(a-c) and 4.6(a).

The three different kernel determination methods estimate different kernel widths for CT smoothing for AFC, even on images from a large number of MLEM iterations, Figures 4.5(a-c) and 4.6(a). Despite this, an AF-corrected RMSE of less than 10 % was achieved for all three methods at 200 MLEM iterations or more, Figures 4.5(d-f) and 4.6(b). This suggests that, for this test data, AF-corrected quantification is not very sensitive to the smoothing applied to the CT for AFC. This could be due to the RMSE measures being insensitive to kernel widths in VOIs with little contrast, implying that determining h_{PVC} (*resp.* h_{AFC}) on VOIs within which there is a relatively uniform activity concentration before (*resp.* after) AFC is numerically unstable. Note that h_{pts} always fits on data with contrast by design of the method. A limitation of these simulations is the constant simulated point source intensity used, this resulted in a range of reconstructed point source contrasts with respect to the surrounding voxels due to the highly heterogeneous structure in the lungs. Reducing the VOI diameter increased the uncertainty on the estimated kernel width, as would be expected with the corresponding decrease in contrast within the VOI.

The RMSEs in the AF-corrected lung were comparable (RMSE range = 8.8 - 9.7 % at greater than 200i), regardless of the kernel determination method. The simulation results in Figure 6(b) provide confidence in the single global kernel approach, as the RMSE in the AF-corrected whole lung was lower with respect to using any of the kernels derived from an individual VOI. As noted above, the kernel width estimation derived from a single region is unstable, this instability is reduced by combining the VOIs. When the global kernel is used to AF-correct the individual VOIs, the RMSE was comparable to using the localised kernel for that VOI (global kernel RMSE range = 4.2 - 8.5 %; localised kernel RMSE range = 4.6 - 8.8 %, at 200i or more).

A faster convergence of the estimated kernel widths for the simulations would be expected with TOF reconstruction [123]. In the lung in particular, Emond et al. showed that 1600 non-TOF MLEM iterations were needed for the mean difference between the last two iterations to be less than 0.1 % overall, while only 240 MLEM iterations were needed for a 550 ps TOF FWHM system [68]. However, this was not investigated in this chapter due to computational limitations of the STIR TOF implementation at the time of these experiments. For the results of this chapter, 200 non-TOF MLEM iterations took 01:00:12 (hh:mm:ss) on an Apple MacBook Pro (2.4 GHz Quad-Core Intel Core i5; memory: 16 GB); a recent implementation of STIR allows GPU enabled reconstruction via the `parallelproj` library [124], which resulted in a 12-fold reduction in this reconstruction time on a AMD Ryzen 9 5900 12-Core Processor with GEForce RTX 3070 GPU.

The h_{PVC} method is the most practical to implement on measured data as h_{pts} requires accurate simulation capability and h_{AFC} needs VOIs that are uniform after AFC, which is difficult to achieve in phantom studies due to the presence of inacctive insert walls. It was therefore solely the h_{PVC} method that was implemented on physical phantoms.

4.3 PET-CT acquisitions

4.3.1 Methods

The feasibility of determining the kernel with which to smooth the CT for AFC from scans of a physical phantom on clinical scanners was assessed.

The h_{PVC} methodology was utilised to determine the optimal kernel for AFC on a GE Discovery 710 (D710) and a Siemens Biograph64 Vision 600 (Vision) at University College London Hospitals (UCLH), using a commercially available elliptical thorax phantom (ECT/LUNGSPINE/I) [125]. The phantom contains two lung inserts containing polystyrene beads; the phantom was modified to allow the insertion of a 4 ml hollow sphere into the left lung and an 8 ml sphere into the right lung. The phantom was filled with ^{18}F-FDG, with activity concentration ratios in the lung background and spheres such that the left lung was (approximately) homogeneous pre-AFC, and the right lung homogeneous post-AFC (sphere walls aside). A HRCT was acquired before the phantom was filled to provide a higher contrast between the Perspex walls and the fillable compartments, thereby allowing easier segmentation of the phantom components.

A GT emission was constructed from known activity concentrations, as measured in a radionuclide calibrator, and the segmented HRCT of the phantom (ITK-SNAP [126]), Figure 4.7 (d-f). Further details on GT construction are given in Appendix C.

The reconstructions used clinically for tumour imaging, i.e. 2 iterations, 24

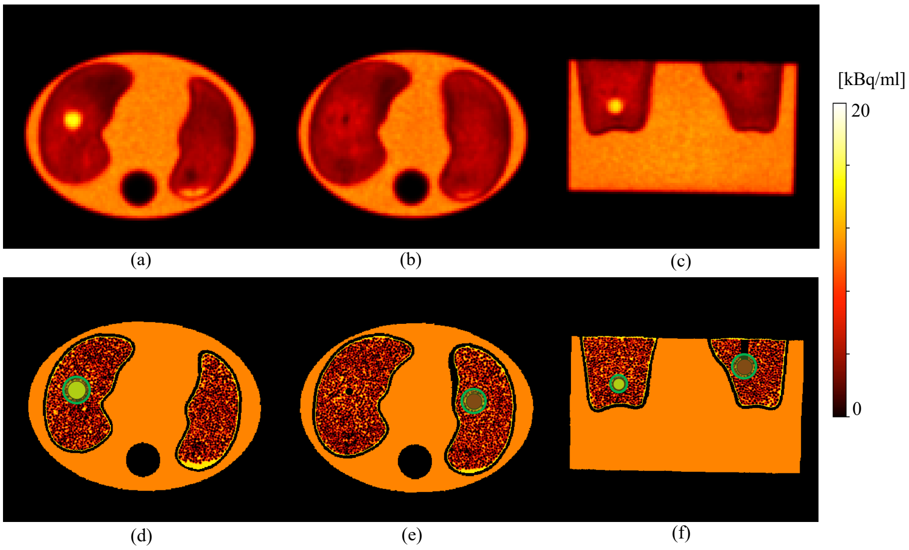


Figure 4.7: Siemens “TOF” reconstruction (top row); GT constructed from HRCT of the phantom and known activity concentrations (bottom row); (a)(d) axial slice through the centre of the sphere in the right lung; (b)(e) axial slice through centre of the sphere in the left lung; (c)(f) coronal slice through the phantom showing both spheres. The VOIs positioned over the sphere inserts, combined into a single VOI, for kernel determination, are shown in green.

subsets (2i24s) for the GE scanner and 4 iterations, 5 subsets (4i5s) for the Siemens scanner, were assessed using the h_{PVC} method, Table 4.2. GE’s “VPFX” (Vue Point FX) is a fully 3D TOF iterative reconstruction. This reconstruction is iterated to 2i (24 subsets (24s)) clinically, and produces images in-line with “EARL1” standards [56], described in Section 2.6. A 4i24s reconstruction was also conducted to assess the effect of iterating for longer on the required smoothing for AFC. A 6.4 mm^3 FWHM Gaussian PF in the transaxial plus a 1:4:1 ratio triangle filter in the axial was applied for both 2i24s and 4i24s. GE’s “QCFX” reconstruction is a regularised iterative reconstruction algorithm, otherwise known as “Q.Clear”. Q.Clear is always run to effective convergence, therefore iterations and subsets are not specified inputs. TOF reconstructions, one that complies with “EARL1” (PF FWHM = 6.0 mm^3) standards, and one with a 4.0 mm^3 FWHM PF were reconstructed on the Siemens scanner. All reconstructions were conducted using CT based AC.

The optimal kernel for AFC was determined via the h_{PVC} methodology. The spherical inserts were segmented on the HRCT, the mask was down-sampled to PET voxel size, and then dilated by two voxels isotropically, to ensure that arte-

Table 4.2: Reconstruction parameters for the clinical reconstructions assessed on the two clinical PET-CT scanners at UCLH. All PFs are Gaussian, except for the GE axial PF (a “standard” axial filter is a triangle filter with the ratio 1:4:1).

Scanner	**Reconstruction**	**i**	**sub.**	**Voxel size [mm^3]**	**PF FWHM [mm]**
D710	OSEM TOF (“EARL1”)	2	24	2.73x2.73x3.27	6.4 / “standard”
D710	OSEM TOF	4	24	2.73x2.73x3.27	6.4 / “standard”
D710	BSREM TOF+RM (β=400)	-	-	2.73x2.73x3.27	0.0
Vision	OSEM TOF	4	5	3.30x3.30x3.00	4.0
Vision	OSEM TOF+RM	4	5	1.65x1.65x3.00	0.0
Vision	OSEM TOF (“EARL1”)	4	5	3.30x3.30x3.00	6.0

facts caused by a mismatched kernel were analysed. While the simulations were noiseless, the acquired data were not. Therefore, a weighted mean square error (WMSE) of each of the activity concentrations in the smoothed GT activity VOI voxels ($A_{true,v}$) to the activity concentrations in the reconstructed image VOI voxels ($A_{obs,v}$) were calculated, equation 4.1. As a reasonable first approximation, it was assumed that the variance was proportional to the mean voxel value [127]. Each WMSE was normalised to the mean lesion value for each reconstruction (normalised weighted mean square error (nWMSE)), where N is the number of voxels in the VOI.

$$\text{nWMSE} = 1/N \sum_{v} (A_{obs,v} - A_{true,v})^2 / A_{obs,v} \tag{4.1}$$

The h_{PVC} that resulted in the smallest nWMSE in the VOI, with respect to the reconstructed image, was reported.

The GT emission was smoothed by the estimated kernel for each reconstruction and the RMSE from the reconstructed image was assessed within the VOI dilated by an additional two voxels isotropically. This larger VOI was used for analysis to be representative of a VOI used to assess diffuse disease e.g. IPF.

The uncertainty on the determined kernel FWHM was estimated by computing the standard deviation of h_{PVC} FWHMs derived by randomly sampling half of the VOI voxels. Details on the voxel sub-sampling methodology can be found in Appendix D.

4.3.2 Results

Heatmaps of the nWMSE in the VOI for each h_{PVC} pair are displayed for the clinical TOF reconstructions on each scanner in Figures 4.8(a) and 4.9(a). The nWMSE heatmap for the 4i24s reconstruction on the D710 is shown in Figure 4.8(b). The nWMSE was lower for the clinical Siemens TOF reconstruction than the clinical GE TOF reconstruction, Figures 4.9(a) and 4.8(a).

Profiles through each of the spheres, in all three dimensions, for the TOF reconstructed image, the unsmoothed GT, and the smoothed GT that was closest, in the least squares sense, to the reconstructed image, are shown in Figures 4.8(c) and 4.9(b).

The heatmaps show that the minimum nWMSE, and therefore the optimal kernel to smooth the CT for AFC, was ill-defined for all clinical reconstructions investigated. The transaxial-axial FWHM kernel pair that resulted in the nWMSE minimum for each reconstruction, are shown in table 4.3. The greatest uncertainty was associated with the kernel width determination in the axial direction, as demonstrated by the asymmetry of the heatmap gradients; this was seen for all reconstructions, table 4.3.

Table 4.3: Measured data results: transaxial and axial h_{PVC} FWHM that resulted in the smoothed GT emission image that best matched the reconstructed image. Uncertainty is one standard deviation of 100 voxel sub-sampling realisations, see Appendix D.

Scanner	**Reconstruction**	$h_{PVC}(xy)$ **FWHM [mm]**	$h_{PVC}(z)$ **FWHM [mm]**
D710	OSEM TOF 2i (“EARL1”)	10.7 ± 0.06	7.2 ± 0.39
D710	OSEM TOF 4i	9.6 ± 0.13	8.5 ± 0.25
D710	BSREM TOF+RM	8.9 ± 0.14	7.2 ± 0.18
Vision	OSEM TOF	8.1 ± 0.11	6.8 ± 0.15
Vision	OSEM TOF+PSF	6.7 ± 0.05	4.3 ± 0.18
Vision	OSEM TOF (“EARL1”)	9.3 ± 0.07	8.3 ± 0.64

The profiles through the spheres show that the GT emission, smoothed by the

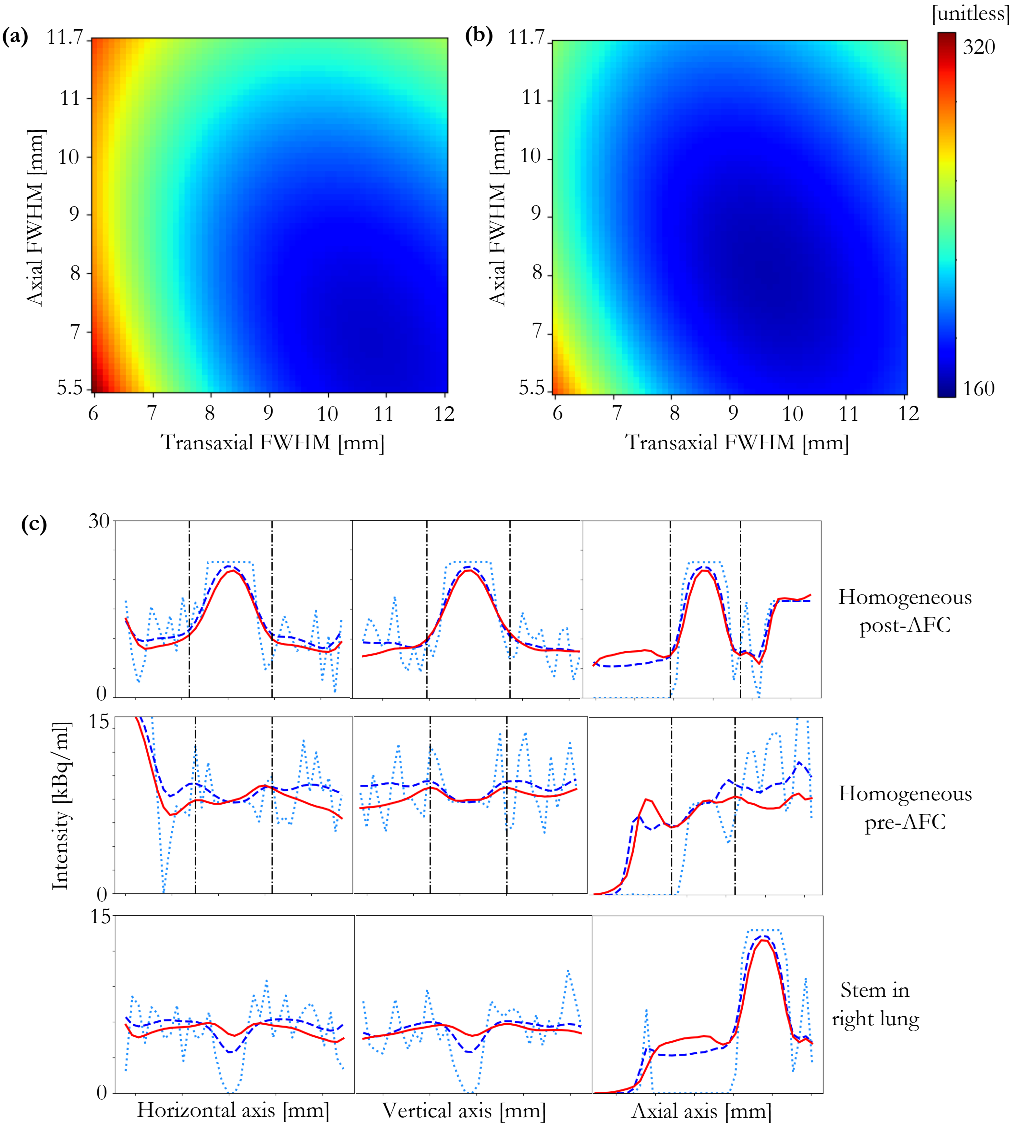


Figure 4.8: Measured data: (a) heatmap depicting nWMSE for each kernel combination investigated for GE D710 VPFX 2i reconstruction (minimum at h_{PVC} = 10.7 x 10.7 x 7.2 mm^3); (b) heatmap depicting nWMSE for each kernel combination investigated for GE D710 VPFX 4i24s reconstruction (minimum at h_{PVC} = 9.6 x 9.6 x 8.5 mm^3); (c) horizontal, vertical and axial profiles through the homogeneous post-AFC sphere (top row), the homogeneous pre-AFC sphere (middle row) and a cold Perspex stem (bottom row). The down-sampled GT emission (light blue, dotted), VPFX 2i reconstructed image (red, solid), GT emission smoothed by the h_{PVC} estimated kernel and down-sampled (royal blue, dashed) are depicted; the extent of the VOI used for kernel determination, in each orientation is shown by the black dash-dotted line.

estimated h_{PVC} kernel, matches the reconstructed image well, both within the VOI and in the lung background. As both spheres contained activity concentrations,

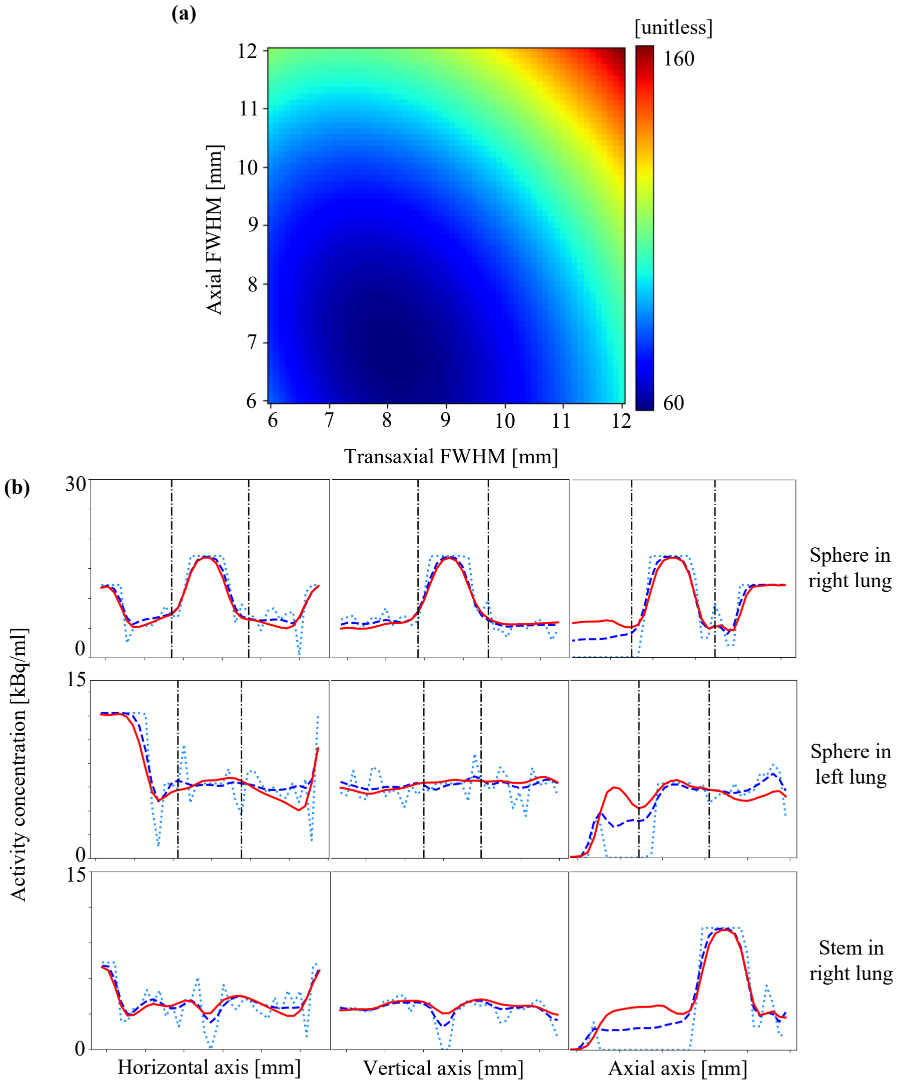


Figure 4.9: Measured data: (a) heatmap depicting nWMSE for each kernel combination investigated for Siemens Biograph64 Vision 600 Vision TOF reconstruction (minimum at 8.1 x 8.1 x 6.8 mm^3); (b) horizontal, vertical and axial profiles through the sphere filled to be homogeneous post-AFC (top row), the sphere filled to be homogeneous pre-AFC (middle row) and a cold Perspex stem (bottom row). The down-sampled GT emission (light blue, dotted), TOF reconstructed image (red, solid), GT emission smoothed by the kernel and down-sampled (royal blue, dashed); the extent of the VOI in each orientation is shown by the black dash-dotted line.

equivalent to, or greater than, lung background, a profile was drawn through the inactive Perspex stem of the right sphere; this was to assess the applicability of the smoothed CT to AF-correct the reconstructed PET within a low count region, as may be found in regions of diseased lung. Figures 4.8(c) and 4.9(b) show the h_{PVC}

smoothed GT emission matches the reconstruction within 10 %, even within these regions.

The mean transaxial-axial kernel width for “EARL1” reconstructions was consistent between scanners (D710 $h_{PVC}(xyz)$ FWHM = 9.40 ± 0.32 mm; Vision $h_{PVC}(xyz)$ FWHM = 8.97 ± 0.52mm). Increasing the number of iterations from two to four for the “EARL1” reconstruction on the D710 did not significantly alter the mean transaxial-axial kernel width estimation (2i $h_{PVC}(xyz)$ FWHM = 9.40 ± 0.32 mm; 4i $h_{PVC}(xyz)$ FWHM = 9.10 ± 0.24 mm).

The narrowest kernel width was obtained for the standard “TOF+PSF” reconstruction on the Siemens Vision, which has the highest reconstructed resolution of those investigated. This reconstruction also had the highest AF-corrected RMSE (% RMSE = 9.2), table 4.4. The lowest was for the “EARL1” reconstruction on the Vision (RMSE = 5.5 %), which employs the largest PF of those investigated. VOI RMSE was below 10 % for all reconstructions (range = 5.5 - 9.2 %).

Table 4.4: Measured data: RMSE in the reconstructed VOI with respect to the GT emission convolved with h_{PVC}. Uncertainty on RMSE is one standard deviation of 100 voxel sub-sampling realisations, see Appendix D. Uncertainty on GT VOI mean is one standard deviation over all voxels in the original VOI.

Scanner	**Reconstruction**	**GT VOI mean [kBq/ml]**	**VOI RMSE [kBq/ml]**	**VOI RMSE [%]**
D710	OSEM TOF 2i (“EARL1”)	10.1 ± 2.5	0.847 ± 0.009	8.4 ± 0.04
D710	OSEM TOF 4i	10.1 ± 2.5	0.829 ± 0.003	8.2 ± 0.01
D710	BSREM TOF+RM	10.1 ± 2.6	0.907 ± 0.004	9.1 ± 0.02
Vision	OSEM TOF	6.9 ± 2.1	0.435 ± 0.007	6.4 ± 0.11
Vision	OSEM TOF+PSF	7.3 ± 2.9	0.667 ± 0.003	9.2 ± 0.02
Vision	OSEM TOF (“EARL1”)	6.9 ± 2.0	0.385 ± 0.003	5.5 ± 0.05

4.3.3 Discussion

Figure 4.9(a) shows that the nWMSE was lower for the clinical Siemens TOF reconstruction than the clinical GE TOF reconstruction, Figure 4.8(a). This could be due to various differences between the two scanners and their processing methods, including (i) the slightly larger reconstructed voxel size used by Siemens (non-PSF reconstruction: 32.7 mm^3 versus 24.4 mm^3 for the GE reconstructions), resulting in higher counts per voxel, and therefore less uncertainty on the reconstructed images; (ii) a smaller crystal size, Vision: 3.2 x 3.2 x 20 mm^3; D710: 4.2 x 6.3 x 25 mm^3) resulting in less spatial uncertainty in the measured data.

The RMSE in the reconstructed VOI for the Siemens TOF+PSF reconstruction was the highest of all reconstructions, potentially due to the edge artefacts commonly observed with PSF-based image reconstruction, or the increased average variance in the smaller voxels. The smaller voxel size also contributes to the lowest $h_{PVC}(xyz)$ kernel FWHM.

It has been suggested that a major factor contributing to uncertainty in quantification is the uncertainty in the delineation of the VOI [128]. In this chapter, the VOI was drawn on the CT data set and copied to the registered PET image. The coordinates of the original boundary will therefore be rounded to the nearest voxel coordinates of the PET image, adding to uncertainty. The effect of VOI size was investigated for the GE VPFX 2i ("EARL1") reconstruction, by dilating the down-sampled spherical mask by 0, 1, 2, 3 and 4 voxels isotropically. The range of $h_{PVC}(xyz)$ FWHMs for varying VOI diameters was 8.75 - 9.55 mm, resulting in RMSE in the reconstructed VOI in the range 8.22 - 8.27 %. These results suggest that, for this experimental design, the AFC kernel was relatively insensitive to VOI size, provided that the edges of the lung were not included in the VOI.

The optimal kernel to smooth the CT for AFC was ill-defined for all clinical reconstructions. Non-spherical / non-symmetrical VOIs may aid in differentiation between the transaxial and axial kernels, and may potentially offer a more unique solution. Alternatively shaped inserts, which can be embedded within the lungs of the thorax phantom, are investigated in Section 5.3.

In this chapter a HRCT was used to create a GT for measured data. Clinically, an HRCT of patient data might not be available, thus patient PET data might need to be AF-corrected with a smoothed low-dose CT acquired for AC.

The ability to determine scanner- and reconstruction-specific kernels with which to smooth the CT for the AFC of patient data, relies on phantoms that approximate the characteristics of diseased lung. The phantoms must be easy to prepare and generate reproducible results. The conclusions from this chapter inform on the design of the phantoms in Chapter 5.

4.4 Conclusions

Simulations of an IPF patient-realistic digital phantom for a non-TOF PET scanner indicated that a large number of MLEM iterations (around 200) were needed to approach reconstructed image resolution stability. AF-corrected quantification was shown to not be very sensitive to the smoothing applied to the CT, indicating that a single global kernel could be applied to the CT for determination of AFs, with which to AF-correct the whole lung, on a voxel-wise basis. Moreover, a kernel derived on combined VOIs has been shown to be more numerically stable. The most practical method to determine the kernel for AFC is to smooth a GT emission image to match the reconstructed image, h_{PVC}. It has the potential to be used to determine the kernel for AFC for a clinical scanner from scans of physical phantoms, not just for ^{18}F-FDG, but for other PET tracers used to study the lung. However, the construction of the GT for phantoms containing lung-like structures is non-trivial due to registration and segmentation accuracy. This method is not limited to the determination of the AFC kernel but could also be utilised to determine the PSF for partial volume correction for both lung disease and tumour imaging. The applicability of the kernel to features that approximate diffuse lung disease, as well as solid tumours, is investigated in Section 5.4.3.

Chapter 5

Novel phantom design

5.1 Introduction

Current harmonisation efforts in PET-CT focus on oncology and brain studies [56, 58, 129]. However, there is a similar need for studies of diffuse lung disease. Patients may be scanned on different PET-CT systems, whether in a single centre, to monitor disease progression and treatment response, or when participating in multi-centre trials. In clinical practice, a wide range of PET systems are installed. Accurate cross-calibration of systems and standardisation of methodology is crucial; the resultant variability, reproducibility and accuracy of PET quantification must be understood. Chen et al. 2020 highlighted the lack of widely accepted standard protocols for pulmonary ^{18}F-FDG PET-CT, and the need to develop test objects that model relevant aspects of lung physiology, to support the establishment of harmonisation standards [1].

In this chapter, a literature review of novel lung phantom designs and material characterisations is conducted. The construction of a preliminary design of a novel phantom, that better approximates diffuse lung disease, follows. The design is assessed for its ability to determine and validate the kernel for AFC on a clinical PET-CT scanner. The design was iterated and reassessed on three clinical PET-CT scanners. Finally, additional smoothing was applied to selected reconstructions in an attempt to harmonise the three scanners for quantification of AF-corrected PET-CT lung studies.

A subsection of the literature review, relating to positron range, was published in "Tissue mimicking materials for imaging and therapy phantoms: a review" [130]. Preliminary results from the first and second generation phantoms were presented as posters at the European Association of Nuclear Medicine's Annual Meeting 2023 [120] and the Society of Nuclear Medicine and Molecular Imaging's Annual Meeting 2024 [131], respectively. The final phantom results from all three clinical PET-CT scanners were presented as a poster at IEEE's NSS/MIC/RTSD 2024 [132].

5.2 Lung phantoms in literature

Phantoms are designed to simulate the response and interactions of human tissues with the imaging device, thereby ensuring appropriate imaging quality in human subjects. Standard test objects allow for advanced calibration and permit comparisons between different systems participating in multi-centre trials.

Emphysematic lung is characterised by increased air trapping; IPF by an increase in the quantity of extra-cellular matrix, which distorts the parenchyma architecture, leading to regions of dense tissue with large air pockets. Therefore, to fully simulate the interactions of radiotracers in the lung with a test object, highly heterogeneous, low density structures, and in the case of nuclear medicine imaging, a radioactive component, are required.

Commercially available "lung" phantoms include the addition of inserts filled with a low atomic number material, typically polystyrene beads, with an average density of 0.3 ± 0.1 kg.m^{-3} to simulate the attenuation of lung; this set-up is utilised in the NEMA NU 2-2007 image quality phantom [51], elliptical Jaszczak and torso phantoms (Data Spectrum Corporation Products). These densities are greater than the lung density at 50 % vital capacity measured in 153 healthy subjects, which were 0.154 - 0.186 g.cm^{-3}, depending on age [133]. A limitation of commercially available multi-chamber fillable phantoms is the non-zero thickness of chamber walls, resulting in an artificial boundary effect between regions of activity in the phantom. This is especially problematic for small-scale structures. The effect of positron range on quantification in low density regions needs to be considered for isotopes with higher positron energies; the influence of positron range can result in spatially-variant, absorber material dependent, resolution degradation and image artefacts.

There has been a sharp increase in the number of publications concerning phantoms for medical imaging, including several reviews [134, 135, 130]. Tino et al. reviewed 53 papers on additive manufacturing for radiotherapy planning applications. Eleven of these papers modelled the lung but only one of those considered nuclear medicine, in which the lungs were hollow compartments [136]. Many of the materials used in 3D printed phantoms for nuclear medicine have not been characterised and tested to the same extent as commercial tissue-equivalent materials; half of the research articles reviewed in Filippou and Tsoumpas do not mention the properties of the materials and none have fully characterised low-density tissue mimicking materials (TMMs) suitable for use in nuclear medicine imaging test objects [130].

Heikkinen et al. designed lung-shaped shells, which they filled with 1-4 mm diameter plastic pellets to produce low density lung defects [137]. Molding wax and plastic pellets were mixed and shaped to simulate defects at various positions in the lungs. The lungs were filled with a $^{99\text{m}}$Tc water solution and scanned on gamma cameras in 18 Finnish Nuclear Medicine departments. The CT number of the pellets in water was -616 HU and -488 HU for the pellet and molding wax combination. The defects were therefore higher density than healthy lung and inactive.

Lehnert et al. investigated the use of $^{99\text{m}}$Tc-laced expanding polyurethane foam (EPF) for creating a more realistic lung phantom [138]. The range of CT numbers achieved with EPF (-980 to -750 HU) closely matched that of lung tissue (-950 to -800 HU). The EPF phantom demonstrated an increased homogeneity, in both material density and activity concentration, with respect to a polystyrene-bead-and-water phantom used for comparison. As the construction of the foam phantom takes less than an hour, once the mould has been prepared, the authors propose that this technique could be used to create lung phantoms for PET studies. Only $^{99\text{m}}$Tc-laced EPF was investigated. Material stopping power was not investigated.

Robinson et al. noted that the effective density of a 3D printed material depends on the model geometry and properties of the printer [139]. The acrylonitrile butadiene styrene (ABS) inserts produced in their work had a mean value of -54±13 HU, thus providing a closer analogue to water (-7.4±8.7 HU) than Perspex (116±9 HU), making it suitable for nuclear medicine phantom production assessing radionuclide distribution in soft tissue. However, as these CT numbers do not represent those found in the lung, commonly used 3D printing materials may not be appropriate for test objects designed to validate corrections such as AFC and positron range in the lung.

A wall-less model was proposed by DiFilippo et al. Small holes were machined in polystyrene to manufacture a "multi-attenuation porous phantom" [140]. The phantom was split into quadrants, three of which included additives to alter the mass density; one being glass microbubbles, resulting in an effective mass density of 0.96 g.cm^{-3} when filled with radiotracer solution. Hunt et al. 3D printed a phantom designed around the "porous phantom"; the density of the cured material was 1.15 g.cm^{-3} [141]. The phantom consisted of a cylindrical matrix of columns of decreasing width and separation to adjust contrast; when filled with a radioactive solution, changes in the sub-resolution columns produced a phantom with sphere inserts of differing size and contrast. Both the "porous phantom" described by DiFilippo et al. and that 3D printed by Hunt et al. are prone to air bubbles, which was partly remedied by adopting a syphon filling technique. Wollenweber et al. offered an alternative design to reduce the effect of air bubbles; the design uses hollow 3D printed dodecahedral features, which create voids in small, solid non-porous acrylic spheres [142]. Neither Hunt et al. nor Wollenweber et al. investigated low density features.

Madamesila et al. [143] and Dancewicz et al. [144] characterised materials at different infill percentages, attempting to match them with various anatomic regions, including the lung. The phantoms were designed for external beam purposes and so do not attempt to incorporate radionuclides into the prints. These studies examined single samples of each material or infill density, it has, however, been demonstrated that using a different printer, a filament from a different manufacturer, or even different rolls of the same filament from the same manufacturer, can result in large differences in the photon attenuation properties [145, 146]. Stopping and scattering power of the materials were not considered in either paper.

There are only a few reports of printing radioactive materials, mostly limited to 2D printing on paper sheets. In this technique, papers are stacked at fixed distances, separated by thin layers of polystyrene, 3D printed polylactic acid (PLA) or polymethylmethacrylate (PMMA) to build 3D sandwich phantoms for single photon emission computed tomography (SPECT), using $^{99\mathrm{m}}$Tc [147, 148], or PET, ^{18}F [149]. Negus et al. estimated the narrow beam attenuation coefficient of the 3D printed 3.8 mm thick PLA slabs (85 % infill) at 140 keV by measuring transmission with a $^{99\mathrm{m}}$Tc source. Measurements of CT number showed attenuation varied linearly with print density [148].

Miller and Hutchins used powder deposition printing with a cellulose-based powder and incorporated ^{18}F into the binder [150]. The density of the cellulose-based powder was 0.5 g.cm^{-3} with a CT number of -600 HU. A lung phantom, including lung nodule and bone, was produced; the bound powder had a CT number of -500 HU. The use of minimal binding to reduce the density was discussed, as was the inclusion of additives to alter the density of the powder, noting the need to take into account the change in viscosity and particulate size.

Läppchen et al. [151], Gear et al. [152], and Gillett et al. [153] incorporated $^{99\mathrm{m}}$Tc directly into photo-polymer resin. Gear et al. confirmed print consistency (CoV = 0.08 %), accuracy of activity concentration within ±2 %, better than ^{57}Co flood source uniformity and heterogeneous reproducibility with radionuclide calibrator and gamma camera acquisitions. They noted that attenuation properties of both the support and build materials were similar (density = 1.17 g.cm^{-3}). A slightly higher CT number was measured within the cubic arrays, designed to reduce activity concentration, than the solid rods.

A low density 3D printing material was developed by Wirth et al. for use with stereolithography (SLA) 3D printers [154]. The foaming prepolymer resin is capable of post-printing expansion of up to 40 times and with the right combination of layer thickness, exposure time, and blowing agent concentration, the density is reduced to that of Styrofoam®. It was demonstrated that if the blowing agent mass fraction was kept below a 10 wt % threshold, the structure remained closed cell. The authors found that the expansion ratios of the 3D printed parts was reproducible (within ±15 % of the mean value across triplicate parts).

In PET imaging (and Bremsstrahlung SPECT) secondary radiations are the principal emissions detected for imaging. Positron range depends on both the electron and physical density of the surrounding material. The majority of work studying this effect has focused on Monte Carlo simulation studies to determine the positron range effects of different radionuclides in water and other biological materials. These simulations, and subsequent phantom measurements, confirmed that positron range depends on both electron and physical density of the surrounding material. To date, there is a scarcity of publications that have experimentally measured positron range in lung-equivalent materials for different radionuclides [130].

Kemerink et al. measured the effect of positron range on spatial resolution and activity quantification of ^{18}F, ^{68}Ga and ^{124}I in homogeneous lung-like tissues [155]. Point sources, lines source and agar gel samples were measured in water, air and in cellular polyethylene (PE foam) with densities ranging between 0.037 g.cm^{-3} and 0.164 g.cm^{-3}, to simulate various lung pathologies. Although ^{68}Ga and ^{124}I sources were visualised nearly as sharp as ^{18}F in air, a large fraction of all annihilations were contained in the “tails”. It was recommended that lung lesions containing ^{68}Ga or ^{124}I be quantified with oversized regions of interest for complete activity recovery. Alva-Sánchez et al. demonstrated degradation of both spatial resolution and activity concentration quantification due to positron range for ^{18}F, ^{13}N and ^{68}Ga line sources in commercially available tissue-equivalent materials ranging in physical density from 0.20 g.cm^{-3}, to simulate lung tissue at full inspiration, to 1.93 g.cm^{-3}, cortical bone equivalent [156]. Neither paper investigated heterogeneous media. With the emergence of high-energy positron emitting radionuclides in PET imaging, the importance of characterising the degradation in spatial resolution and activity quantification due to positron range becomes paramount. Without this, PET images, traditionally considered to reflect activity concentration in the body, may only depict positron annihilation in the different tissues [156], which can be compounded by TF corrections. Tissue equivalent materials for nuclear medicine should therefore replicate the interaction cross sections for the production of secondary radiations and not just the bulk attenuation and scattering properties [130].

This review has demonstrated that there is a need to develop a lung phantom for PET-CT lung imaging standardisation and harmonisation purposes. The use of additive manufacturing in the construction of a fillable phantom is challenging due to low density materials not being readily available. The addition of radionuclides into the printing process presents its own challenges with respect to radioactive contamination. Requirements for a 3D printed phantom for multi-centre scanner harmonisation would be either i) a long half-life radionuclide with a similar positron range to the radionuclide of interest for transportation between centres or ii) a relatively simple design with features that do not exceed each printer's resolution specifications so that each centre could print their own single-use phantom; this would require rigorous print quality control to be conducted at each centre. For these reasons, additive manufacturing was not considered further. Instead, low density, permeable-walled, features to be embedded in the lungs of commercially available thorax phantoms are investigated in the following sections.

5.3 Feature shape simulations

A simple simulation study was used to select the shape that gives the highest precision for kernel determination. Features with this shape were then constructed from low density materials.

It was hypothesised in Chapter 4 that the non-uniqueness of the optimal kernel solution could, in part, be due to using a symmetrical feature making it harder to distinguish between the transaxial and axial kernels. It is possible that an oversmooth in one direction could be partially compensated by an under-smooth in the other, regardless of feature orientation, due to the symmetry of a sphere. To test this hypothesis, simple shapes were simulated analytically and smoothed to approximate reconstructed resolution. The h_{PVC} methodology was used to determine the smoothing kernel.

5.3.1 Methods

Uniform features, with volumes approximately equivalent to a 15.4 mm diameter sphere (as measured in Chapter 4), were modelled in SIRF with a 4:1 feature-to-background ratio. Features were centred in a noiseless GT image (voxel size = 1 x 1 x 1 mm^3) and convolved with an isotropic 3D Gaussian with 5 mm^3 FWHM. The GT image was blurred by 3D isotropic Gaussian kernels (FWHMs varied 3 – 12 mm, in 1 mm increments), decoupled in the transaxial and axial directions. The RMSE was determined in a VOI the physical dimensions of the GT shape. The uniqueness of the kernel was quantified with the condition number, κ, with a large κ indicating an ill-conditioned problem. The condition number of a matrix A is defined as its largest singular value, $\sigma_{max}(A)$, divided by its smallest non-zero singular value, $\sigma_{min}(A)$:

$$\kappa(A) = \frac{\sigma_{max}(A)}{\sigma_{min}(A)} \tag{5.1}$$

The condition number for each shape was obtained by fitting an ellipse to the contours in the region close to the minimum and computing the aspect ratio.

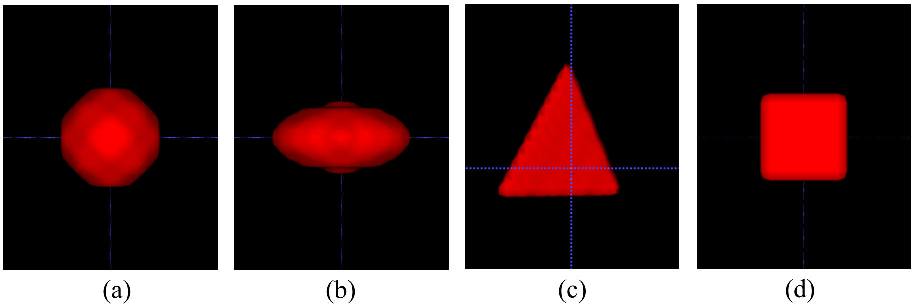

(a) (b) (c) (d)

Figure 5.1: 3D rendered simulated feature shapes (a) sphere (V = 1935 mm^3); (b) ellipsoid (V = 1931 mm^3); (c) tetrahedron (V = 2057 mm^3); (d) cube (V = 2197 mm^3).

5.3.2 Results

The minimum RMSE was found for an isotropic Gaussian smoothing with 5 mm^3 FWHM for all feature shapes, Figure 5.2.

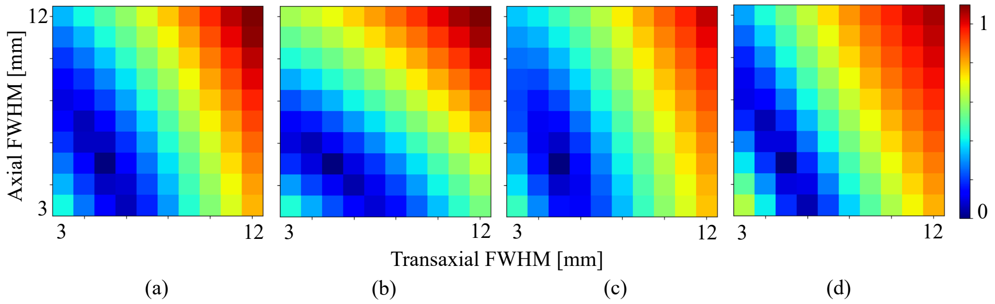


Figure 5.2: Heatmaps depicting RMSE for each kernel combination investigated for each feature shape (a) sphere; (b) ellipsoid; (c) tetrahedron; (d) cube.

The tetrahedron resulted in a better defined minimum, as well as the smallest condition number, κ, indicating that the transaxial and axial kernels are more uniquely identified, Figure 5.3.

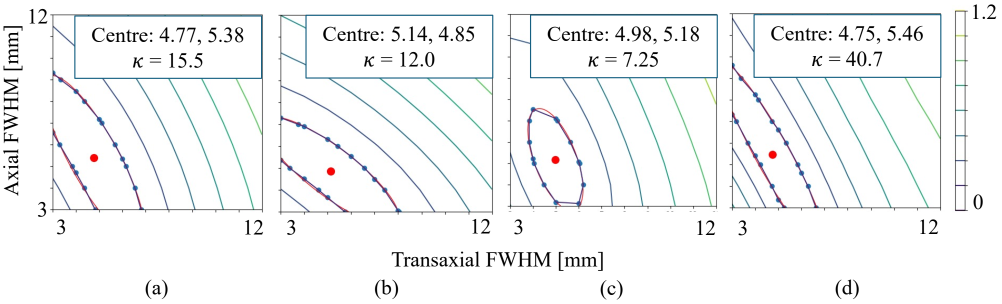


Figure 5.3: Contour plots depicting ellipse fitting for (a) sphere; (b) ellipsoid; (c) tetrahedron; (d) cube.

5.3.3 Discussion and conclusions

This simple simulation study suggests that irregular-shaped features result in more uniquely identifiable transaxial and axial kernel solutions. Tetrahedral features are relatively easy to manufacture and will be investigated further in physical phantoms.

5.4 Phantom construction

The phantoms in this section were designed to address some of the limitations of using the thorax phantom containing fillable spheres to determine the kernel for AFC. These aims were to:

- reduce the number of segmentations required for GT construction.

- utilise a non-symmetric shape to determine the kernel for AFC, which will potentially aid in the differentiation between the transaxial and axial kernels.

- make use of low density materials to approximate diffuse lung disease density and structure to determine and validate the reconstruction-dependent kernel for AFC.

The initial phantom design was used to validate the MKA methodology employed in Chapter 4. The reconstruction-dependent kernels for AFC were determined on a single low-density feature and verified on additional features of varying density and activity concentrations. The kernels determined from this phantom were compared to kernels determined from the fillable spheres for matched reconstructions. The results from the first generation phantom informed a second iteration of phantom design.

5.4.1 First generation design

A commercially available elliptical thorax phantom [125] was modified to include irregular-shaped, non-uniform, permeable-walled features. Six tetrahedral inserts were constructed from permeable polyester veil ($\rho \approx 0.21$ g/cm^3) sealed with cotton thread, and filled with different low density materials, Table 5.1 and Figure 5.4. A seventh tetrahedral feature was formed from solid balsa wood. The low density inserts (LDIs) were embedded in the polystyrene beads of the lung of the thorax phantom, ensuring at least 15 mm between each feature and the lung walls. The lungs were then repacked with polystyrene beads.

5.4.1.1 Methods

The optimal kernel for AFC was determined according to the MKA methodology from Section 4.2.1, on a single PET-CT scanner at UCLH (GE D710).

Phantom preparation and acquisition

The lungs and phantom background were filled with ^{18}F-FDG-laced water, with a lung-to-background activity concentration ratio of 1.7:1. The following acquisitons were undertaken on the D710 PET-CT:

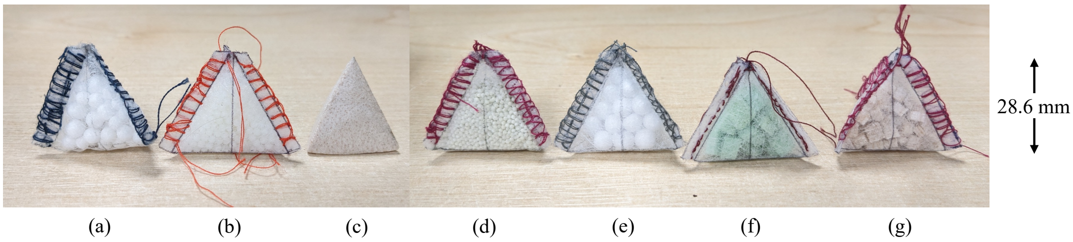


Figure 5.4: Low density features to be embedded in the lungs of the thorax phantom; (a) 2 – 5 mm diameter polystyrene beads; (b) 2 mm diameter polypropylene beads; (c) solid balsa wood; (d) 0.25 – 0.5 mm diameter expanded glass beads; (e) 5 mm diameter polystyrene beads; (f)˜5 mm cubes closed-cell foam; (g)˜3 mm cubes balsa wood.

Table 5.1: Material specifications for each insert and lung background.

Material	**Published density [g.cm^{-3}]**
polypropylene beads	0.91 [157]
polystyrene beads 2 – 5 mm ϕ	0.12-0.50 [158]
solid balsa	0.06-0.38 [159]
expanded glass	0.37 [160]
chipped balsa	0.06-0.38 [159]
closed-cell foam	0.12 [161]
polystyrene beads 5 mm ϕ	0.12-0.50 [158]
polystyrene beads 1 – 4 mm ϕ	0.12-0.50 [158]
healthy lung at full inspiration	0.20 [156]
IPF lung at full expiration	0.58 [162]

- HRCT (120 kVp, 200 mAs, 0.9 s, “Lung” kernel, 0.5 pitch, 0.90 x 0.90 x 0.63 mm^3 voxels).
- 20 minute per bed LM PET.
- CTAC (120 kVp, 32 mAs, 0.8 s, “Q AC” kernel, 1.375 pitch, 1.37 x 1.37 x 3.27 mm^3 voxels).

The PET acquisition was binned into one minute frames to assess if there was a delay in ^{18}F-FDG permeating into the features. The acquisition was commenced 34 minutes post-completion of lung filling; our filling protocol recommends filling the lungs before the phantom background to reduce dose to the operator and so this delay to imaging is considered representative.

Reconstructions

The same clinical reconstructions for the GE D710 that were investigated in Chapter 4 were investigated for this phantom, Table 4.2. In addition to the OSEM 2i24s reconstruction used clinically (PF = 6.4 mm^2 / “standard”; “EARL1”), an OSEM 2i24s reconstruction with no PF, “PF0”, was also reconstructed.

GT construction

A GT emission was constructed from image-derived activity concentrations and the segmented HRCT of the phantom (ITK-SNAP [126]), Figure 5.5, Appendix C. The phantom background, spine insert and lungs were semi-automatically segmented. The tetrahedral shapes were not rigid and so did not retain their shape once embedded in the closely packed polystyrene beads of the lung background. Therefore, each insert was manually segmented on the HRCT, approximating the known physical volume of the tetrahedron. When material CT numbers could be determined from the HRCT, these were used to determine material fractions (MFs) within each feature. For feature fragments smaller than the resolution of HRCT, a random packing fraction of equal diameter spheres was assumed. To determine image-derived activity concentrations, large regions were drawn in the phantom and the lung backgrounds on the reconstructed image, ensuring that feature and lung edges were avoided. The same region was transposed to the HRCT and the mean CT number in the region converted to a MF with which to MF-correct the mean activity concentration in the reconstructed image. The mean phantom background activity concentration and the mean MF-corrected lung background values were used to create image-derived GT activity concentrations.

Kernel determination

The optimal kernel for AFC (or MFC in a phantom) was determined via the MKA, see Section 4.2.1. The feature masks were dilated by four HRCT voxels isotropically (3.6 mm transaxially, 2.5 mm axially) to capture edge effects and increase the PET inhomogeneity within the VOI, providing structure on which to fit the kernel. In addition to the lung background, two of the low density features con-

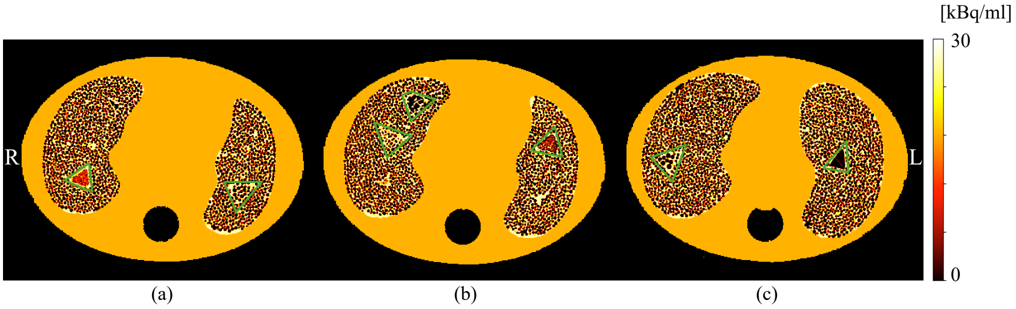


Figure 5.5: GT constructed from HRCT of the phantom and image-derived activity concentrations, the feature VOIs are highlighted in green; (a) axial slice through expanded glass (R lung) and 2 – 5 mm diameter polystyrene bead (L lung) features); (b) axial slice through closed-cell foam (anterior R lung) and chipped balsa (medial R lung) and polypropylene bead (L lung) features; (c) axial slice through 5 mm diameter polystyrene bead (R lung) and solid balsa (L lung) features.

tain polystyrene beads – one with mixed diameter beads (2 – 5 mm ϕ) and one with 5 mm ϕ beads. One was to be assigned as the feature on which to determine the kernel, and the other used as a kernel validation feature. In the first instance, the kernel was determined on each of the polystyrene bead features for the unsmoothed reconstructed image, "PF0". A nWMSE with respect to the reconstructed image, as described in Section 4.3.1, was computed; the h_{PVC} that resulted in the smallest nWMSE was reported as the optimal kernel. The uncertainty on the determined kernel FWHM was estimated as described in Appendix D; the standard deviation across 20 voxel sub-sampling realisations is reported.

The optimal kernels from each of the polystyrene features were used to smooth the GT emission and the relative difference (RD) in mean activity concentration, with respect to the reconstruction, RD_{Bqml}, were recorded for all seven of the features. The polystyrene bead feature that resulted in the lowest RD_{Bqml} across all features was to be assigned as the kernel determination feature.

The MKA methodology was validated by comparing the optimal kernel width for PF0 to the PF0 reconstruction smoothed with an isotropic 6 mm^3 Gaussian kernel.

Testing of kernel accuracy

For each reconstruction, the kernel determined from the polystyrene bead feature designated as the kernel determination feature was used to smooth the GT emission and RD_{Bqml} was recorded for all seven of the features individually, all features combined, and the whole lung. For comparison, the kernel determined from the sphere phantom in Chapter 4 for each reconstruction was also used to smooth the GT and RD_{Bqml} was calculated.

5.4.1.2 Results

LDI properties

Six of the seven inserts had mean CT numbers in the range of expected in a region of IPF in the patient diagnosed with severe IPF, Figure 5.6. The polypropylene bead feature (published $\rho = 0.91$ g.cm^{-3} [157]) had a CT number of 171 ± 41 HU and so will not be included in the analysis or in future phantom designs. The estimated MFs for each feature are listed in Table 5.2.

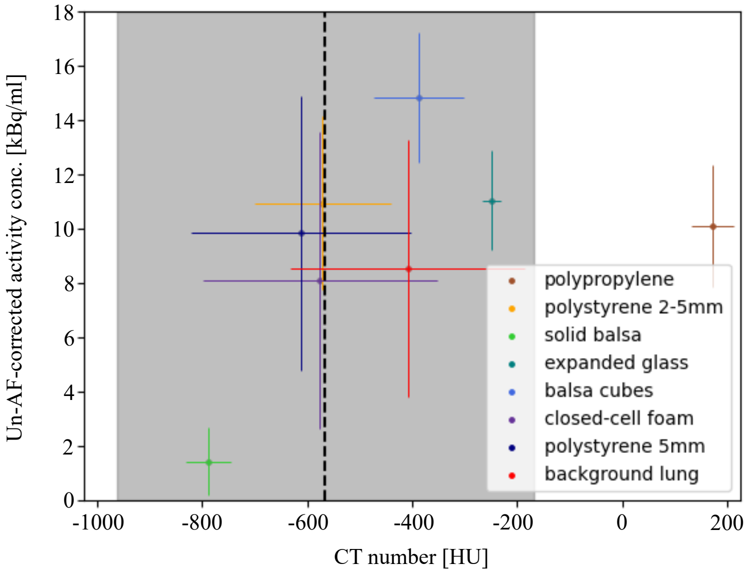


Figure 5.6: CT number versus unMF-corrected activity concentrations in unsmoothed GT for the seven features and background lung; error bars indicate one standard deviation. Black dotted line represents the mean CT number in a region of IPF in the lungs of the patient data; the shaded grey region is the range of CT numbers.

The activity concentration measured in each of the LDIs was constant over the course of the 20 minute dynamic acquisition (1 minute frames), Figure 5.7.

Kernel determination

Table 5.2: Estimated material CT numbers for each insert and lung background, see “GT construction” section; *mean HU from material fragments segmented on HRCT; **random packing fraction of equal diameter spheres [163].

Material	**Segmented VOI CT number [HU]**	**Estimated material CT number [HU]**	**Estimated material fraction**
polypropylene beads	171±41	269	0.635**
poly. beads 2-5mm ϕ	-570±129	-918*	0.621
solid balsa	-788±42	-805*	0.979
expanded glass	-249±8	-393	0.635**
chipped balsa	-387±85	-805*	0.481
closed-cell foam	-575±222	-921*	0.624
poly. beads 5mm ϕ	-611±209	-918*	0.666
poly. beads 1-4mm ϕ	-409±222	-918*	0.446

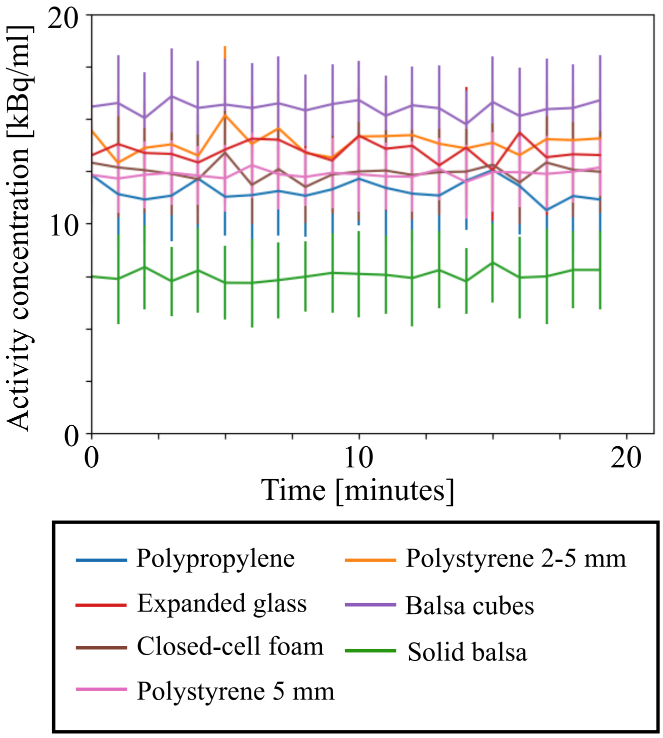


Figure 5.7: Activity concentration in each of the features (segmented on the HRCT and undilated) over the course of a 20 minute dynamic acquisition with 1 minute frames. The error bars indicate one standard deviation within the VOIs.

For the PF0 reconstruction, the kernel FWHM, $h_{PVC}(xyz)$, derived from the 2 – 5 mm ϕ polystyrene bead feature was 6.8 ± 0.24 mm^3 and 8.2 ± 0.40 mm^3 for the

5 mm ϕ polystyrene bead feature. Despite this difference in kernel FWHM, there was little effect on how well the optimally smoothed GT matched the reconstructed image within each feature ($|\Delta RD_{Bqml}| < 0.9$ %), Figure 5.8. For determination of the kernel for clinical reconstructions, which do employ PFs, the larger diameter beads will increase the PET inhomogenity, thereby providing structure on which to fit the kernel. The 5 mm ϕ bead feature was therefore used for kernel determination for the remaining results in this section. Increasing the diameter of the beads further is investigated in Section 5.4.2. The 2 – 5 mm ϕ bead feature was used as a validation of kernel applicability on a region of closely matched density and activity concentration.

A slight underestimation of optimally smoothed GT activity concentration, for this PF0 reconstruction, is observed within every feature but the 5 mm ϕ polystyrene beads and the solid balsa, Figure 5.8. With the exception of the very low density, low uptake, solid balsa feature, the difference is < 10 % of the lung background for all features and within the whole lung.

To validate the MKA, the optimal kernel width for PF0, h_{PVC_0}, was compared to that when the reconstructed image is smoothed by a 6 mm^3 isotropic Gaussian kernel. The predicted kernel for the smoothed image has a FWHM equal to $\sqrt{h^2_{PVC_0} + 6^2}$. The optimal kernel width for the unsmoothed image was found to be 8.1 x 8.1 x 8.3 mm^3 resulting in a predicted kernel FWHM of 10.1 x 10.1 x 10.2 mm^3 for the smoothed image. The computed kernel width for the smoothed image was 10.3 x 10.3 x 10.5 mm^3, Figure 5.9, which is within the typical uncertainty range for these measurements, as discussed in Chapter 4.

The optimal kernels determined from the 5 mm ϕ polystyrene bead feature and the kernel determined from the fillable spherical inserts described in Chapter 4 are shown in Table 5.3, for each reconstruction (OSEM reconstructions PF = 6.4 mm^2 FWHM Gaussian in-plane and “standard” axially).

Figure 5.10 depicts a mean negative bias of the reconstructed image with respect to the optimally smoothed GT for all reconstructions. The solid balsa is the only feature that demonstrates a greater activity concentration in the reconstructed

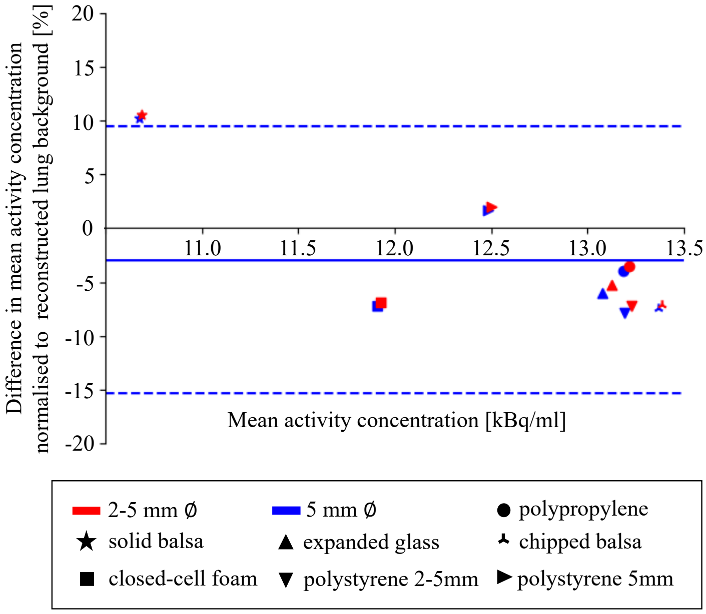
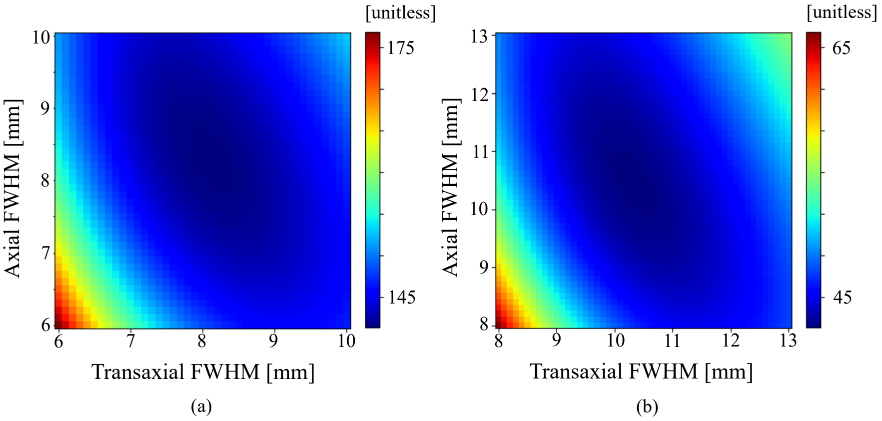


Figure 5.8: Bland-Altman plot of activity concentration in the optimally smoothed GT versus the reconstruction with no PF, normalised to the mean in the lung background, in all features. Features smoothed by the kernel derived from the 2 – 5 mm ϕ polystyrene bead VOI are shown in red and the 5 mm ϕ polystyrene bead VOI in blue. Feature VOIs are delineated on the HRCT and dilated by 4 voxels isotropically. The mean VOI difference and 95 % confidence intervals (CIs) are shown by the solid and dotted lines respectively for the kernel derived from the 5 mm ϕ polystyrene bead VOI.

Table 5.3: Mean transaxial and axial h_{PVC} FWHM that resulted in the smoothed GT emission image that best matched the reconstructed image. Uncertainty is one standard deviation of 20 voxel sub-sampling realisations, see Appendix D.

Reconstruction	**5 mm ϕ poly. bead $h_{PVC}(xyz)$ FWHM [mm^3]**	**spheres $h_{PVC}(xyz)$ FWHM [mm^3]**
OSEM TOF 2i (“EARL1”)	10.2 ± 0.39	9.5 ± 0.32
OSEM TOF 4i	9.5 ± 0.40	9.2 ± 0.46
BSREM TOF+RM (“Q.Clear”; β=400)	10.1 ± 0.48	8.3 ± 0.38

image than in the optimally smoothed GT. As would be expected, the feature on which the kernel was determined results in the closest match to the reconstructed image.



(a)

(b)

Figure 5.9: (a) Heatmap depicting nWMSE for each kernel combination investigated for OSEM TOF 2i reconstruction, no PF (minimum at h_{PVC} = 8.1 x 8.1 x 8.3 mm^3); (b) heatmap depicting nWMSE for each kernel combination investigated for OSEM TOF 2i reconstruction with an isotropic 6 mm^3 FWHM Gaussian filter applied (minimum at h_{PVC} = 10.3 x 10.3 x 10.5 mm^3).

5.4.1.3 Discussion

The “wall-less” phantom design required fewer segmentations than the fillable spheres, enabling a quick GT construction. Estimated CT numbers for four of the five low density materials investigated (polystyrene, balsa wood, expanded glass and closed-cell foam) ranged from -921 to -393 HU, which is within the range of that expected for a region of IPF in the lungs of a patient (-565 ± 396 HU). The CT number for polypropylene was estimated to be 269 HU; this material is therefore unsuitable to be used as diseased lung tissue equivalent and will not be included in future iterations of phantom design. When the insert material fragments were smaller than the resolution of the HRCT, random packing of equal diameter spheres was assumed. For the next generation of the phantom, a CT acquisition will be carried out pre-fill (each voxel will contain air and material) and post-fill (each voxel will contain water and material), such that material density can be solved for and the material fractions more accurately modelled.

The time-activity curves (TACs) were constant for all seven of the VOIs; this provides confidence that the activity diffusion through the permeable walls, or into the solid balsa, had reached an equilibrium in the time between phantom filling and phantom scanning.

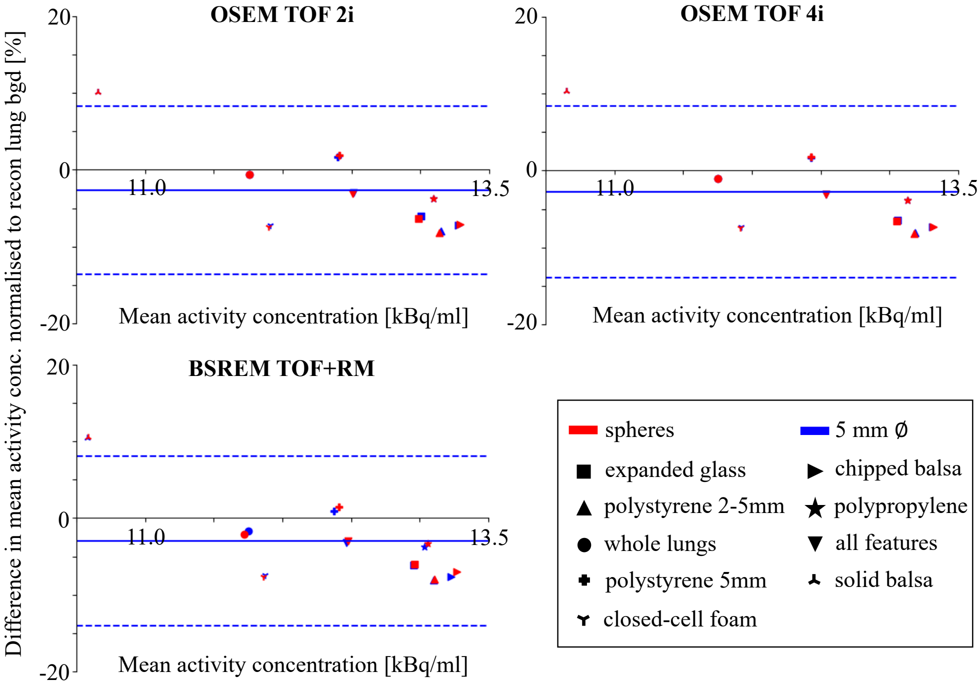


Figure 5.10: Bland-Altman plot of activity concentration in optimally smoothed GT versus reconstructed image, normalised to the mean in the lung background, for all features individually, all features combined, and the whole lung. Features smoothed by the kernel derived from the sphere phantom are shown in red and the 5 mm ϕ polystyrene bead VOI in blue. Feature VOIs are delineated on the HRCT and dilated by 4 voxels isotropically. The mean VOI difference and 95 % CIs are shown by the solid and dotted lines respectively.

The presence of the permeable-walled features not only results in a change in the packing fraction and/or material density contained within the feature, but also reduces the packing fraction of the polystyrene beads in the background lung at feature boundaries. This creates regions of higher activity concentration (and density) at feature boundaries, Figure 5.5. The additional structure in both the PET and CT images is potentially useful when determining and validating the kernel width for AFC.

The larger bead diameter feature was assigned as the kernel determination feature due to the beads being (i) much greater than the resolution of the HRCT, resulting in a more accurate GT construction, and (ii) approximate reconstructed PET resolution, which results in greater inhomogeneities in the reconstructed image. However, the non-uniqueness of the solution could be the result of the 5 mm diameter polystyrene beads not adding enough inhomogeneity into the PET image. This will be investigated with simulation studies in the next section.

Despite the change in feature shape from spherical to tetrahedral, the kernel solution is still not well defined, see Figure 5.9 and Table 5.3. This is potentially due to the non-rigidity of the tetrahedrons, which did not retain their shape when packed with the polystyrene spheres of the lung background. This will be addressed in Section 5.4.3.

The computed kernel for the $h_{PF}(x) = 6$ mm^3 reconstruction was within errors of that predicted from the $h_{PF}(x) = 0$ mm^3 reconstruction, providing confidence in our MKA methodology.

This LDI phantom estimated different kernel FWHMs with respect to those found with the sphere phantom in Chapter 4 for the same reconstruction on the same scanner, e.g. 10.2 ± 0.39 mm^3 (LDI) versus 9.5 ± 0.32 mm^3 (spheres) for the “EARL1” reconstruction. The difference was greatest for the Q.Clear reconstruction. This may be due to the edge-preserving prior used by Q.Clear being more apparent for the higher uptake spheres. The uncertainty on the optimal kernel width was consistent across reconstructions and kernel determination VOIs.

$|RD_{Bqml}| < 11$ % across all LDIs and reconstructions, regardless of whether the GT was smoothed by the kernel determined from the polystyrene bead feature or the spheres. This suggests that the fine-scale structures seen in this phantom (and diseased lung) are not that sensitive to the applied smoothing kernel; this supports the conclusion in Chapter 4 that a global kernel can be used for voxel-wise AFC in the lung.

The overestimation of activity in the solid balsa region highlights the potential for quantitative inaccuracies in low uptake regions due to iterative reconstruction (which can be slower to converge and is subject to the non-negativity constraint) and/or sub-optimal scatter correction. It is a useful feature to understand how airways, or air pockets within diseased lung, will be affected by the reconstruction used and the subsequent AFC.

5.4.1.4 Conclusions

This section has demonstrated the potential to construct a phantom that approximates diffuse lung disease. The permeable walled features increase the PET inhomogeneity both within the feature as well as in the surrounding region. Whether increasing the PET inhomogeneity further (in addition to the increase in rigidity of the irregularly shaped tetrahedrons) will result in a more unique kernel solution is investigated in Sections 5.4.2 and 5.4.3. The results indicate that fine-scale structures, as seen in diffuse lung disease, are not sensitive to the smoothing applied for AFC. For larger structures, such as air pockets or tumours, matching the kernel for AFC to the reconstructed PET image resolution (approaching convergence) will be necessary for accurate AF-corrected PET quantification.

5.4.2 PET inhomogeneity simulations

It is proposed that larger material fragments, in addition to permeable walls to create "edges", be used to increase inhomogeneity within the PET image for kernel determination. Increasing the diameter of the low density materials has the potential for greater positron range, leading to annihilation displacement from the site of emission, which may affect kernel determination. The mean and maximum positron ranges for ^{18}F in lung are 1.86 mm and 8.75 mm, respectively (0.48 mm and 2.27 mm in water) [108]. To investigate if the optimal kernel width is altered with an increase in polystyrene bead size, 8 mm diameter spheres were simulated on a hexagonal close-packing grid and a permeable wall was simulated in the x-direction by introducing a gap in the sphere lattice, Figure 5.11.

5.4.2.1 Methods

Kernel uncertainty: increased PET inhomogeneity (positron range not modelled)

To assess whether the increased inhomogeneity in the PET image provided by the gap between spheres resulted in a more unique kernel solution, than that determined from the sphere lattice only, noiseless SIRF simulations, Table 5.4, and reconstructions, Table 5.5, were conducted. "Cold" spheres in a "warm" background (29 Bq/voxel, equivalent to 18 kBq/ml, as in LDI phantom lung background) were simulated, Figure 5.11.

To obtain a LAC for polystyrene, the mu-map was simulated using Geant4 Application for Tomographic Emission (GATE) version 9.0 [164]. The GT emission map, as shown in Figure 5.11, was used as the voxelised emission input to GATE. The attenuation map was of identical geometry to the emission with the material composition of polystyrene taken from the National Institute of Standards and Technology (NIST) Standard Reference Database 126 [165], assuming expanded polystyrene (EPS) to have a 98 % air content [158], the composition of which was taken from NASA [166]. GATE is based on the GEANT4 toolkit (version 10.7.3), which simulates particle interactions through matter, therefore an accurate model of the path of an emitted positron before annihilation can be assessed. The GATE "MuMap" actor was enabled, which outputs the attenuation map, in units of cm^{-1}, Table 5.6. This mu-map output was utilised in all simulations. The LACs of polystyrene and water were 0.0018 cm^{-1} and 0.0960 cm^{-1}, respectively.

Table 5.4: Analytic simulation parameters for increased PET inhomogeneity simulations.

Simulated image parameters	Voxel size	1.000 x 1.000 x 1.635 mm^3
Acquisition model for forward projection	Acquisition model	$h_{IR}(x)$ = 4.7 mm^3 Gaussian filter followed by ray-tracing
	Projection data template	GE D710, non-TOF (span=2; max. ring diff.=23)
	LORs per bin	12
	Sensitivity model	Includes GATE attenuation factors

A 4 mm^3 FWHM Gaussian PF was applied to the reconstructed image, as would be done clinically. A 40 mm diameter spherical VOI was positioned over the "wall", shown in blue in Figure 5.11, and the resultant kernel compared to that determined from a VOI positioned centrally within the sphere lattice, shown in green. The MKA was utilised to determine kernel FWHM; 2 – 12 mm^3 (1 mm increments) isotropic Gaussian kernels were investigated.

Kernel uncertainty: increased PET inhomogeneity (positron range modelled)

To assess the potential for artefacts due to the increased positron range result-

Table 5.5: Reconstruction parameters for increased PET inhomogeneity simulations.

Acquisition model for reconstruction	Acquisition model	Ray-tracing
	Sensitivity model	Includes GATE attenuation factors $h_{AC}(x) = h_{IR}(\mathrm{x}) = 4.7$ mm^3
	Projection data template	GE D710, non-TOF (span=2, max. ring diff.=23)
	LORs per bin	12
Reconstruction	Reconstructor	MLEM
	N iterations	200
	Voxel sizes	2.71 x 2.71 x 3.27 mm^3

Table 5.6: GATE simulation parameters for increased PET inhomogeneity simulations (see Appendix A)

Physics and source specification	Physics model	“Livermore”
	Source	^{18}F ion source
Acquisition	Time	6 seconds
Random generator	Engine seed	“auto”
	Engine name	JamesRandom
Output	Actor	MuMap
	Recorded interactions	“initStep” and “annihil”

ing from a larger fractional volume of low density material, the recorded interactions in the GATE text file output were interrogated.

A six second acquisition time in GATE resulted in 221 x 10^6 annihilations being recorded. The locations of all emission and annihilation events was sampled into two 3D images (voxel size: 1.000 x 1.000 x 1.635 mm^3), one for emission events that resulted in an annihilation, and one for annihilation events, as previously described by Emond et al. 2019 [14], Figure 5.12(a)(b). The GT emission was scaled such that the mean Bq/voxel in the phantom background was equivalent to the mean annihilations in the same region. A 6 mm^3 FWHM isotropic Gaussian kernel was applied to the annihilation output from GATE to approximate reconstructing into 3 mm^3 voxels and applying a 4 mm^3 FWHM PF (this shall be referred to as the “reconstructed” annihilation). The optimal h_{PVC} was then determined for the GT emission in the blue VOI in Figure 5.11 by minimising the nWMSE (normalised

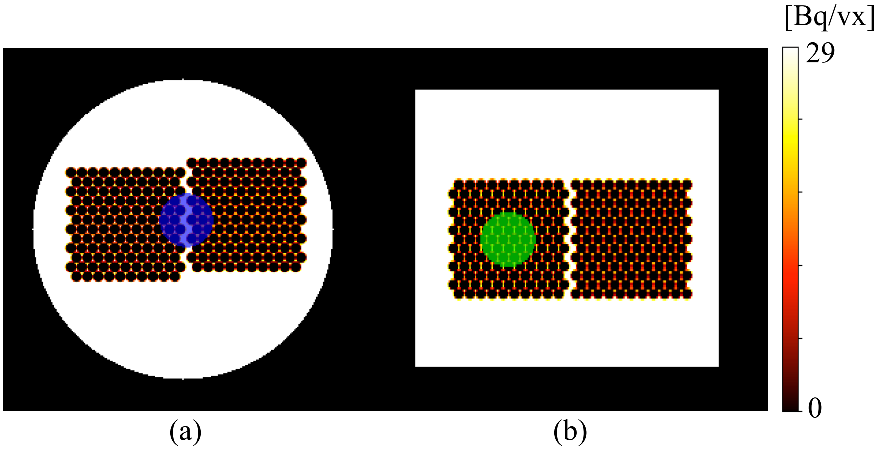


Figure 5.11: Simulated emission of 8 mm diameter polystyrene spheres arranged in a hexagonal close-packing grid within ^{18}F-laced water. A 4 mm gap between spheres has been simulated; (a) axial view depicting the 40 mm diameter VOI positioned over the "edge" in blue; (b) coronal view depicting the 40 mm diameter VOI positioned central to the spheres in green.

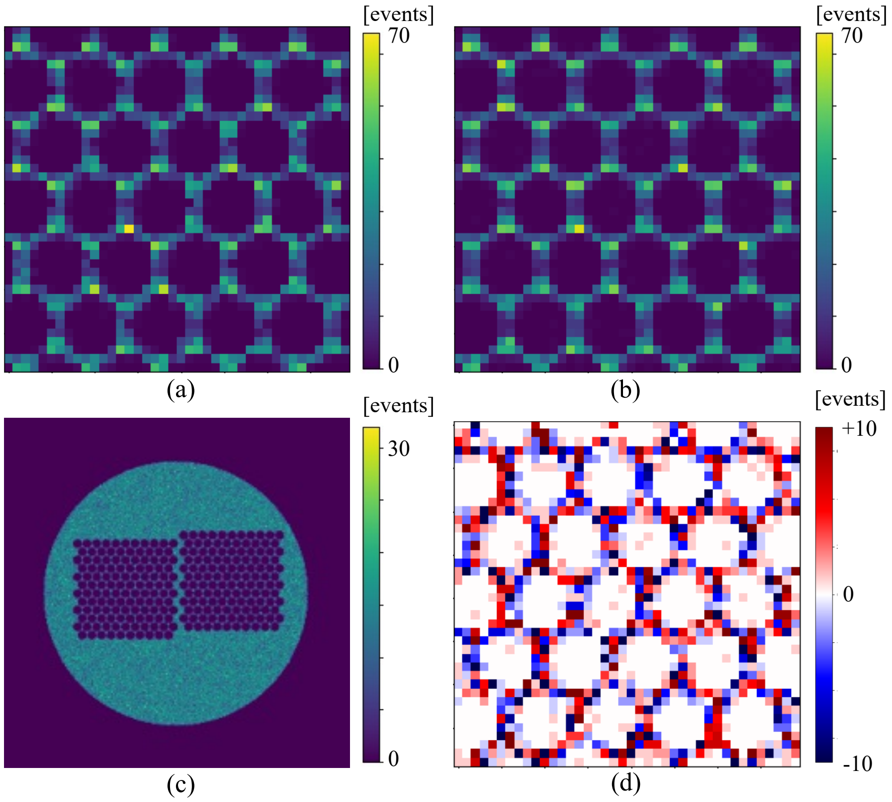
to the mean smoothed GT in the VOI), equation 4.1, with respect to the smoothed annihilation image. 20 voxel sub-sampling realisations, with a sub-sampling factor of 2, were conducted, as detailed in Appendix D. The emission-map output from GATE was also convolved with a 6 mm^3 FWHM isotropic Gaussian kernel and the RMSE within the blue VOI with respect to the GT smoothed by the optimal h_{PVC} determined. Comparison of the RMSEs for the annihilation and emission maps provides an indication of expected effect size from positron range in this experimental set-up for ^{18}F.

5.4.2.2 Results

Positron range not modelled

Analysis of the SIRF simulations and reconstructions showed that the optimal kernel width determined from the "wall" VOI and the VOI positioned centrally within the sphere lattice matched, at each iteration. The optimal kernel converges to 7 x 7 x 6 mm^3 at $\geq$160i, in both cases. The inhomogeneity in the PET image is increased by adding a "wall" in the x-direction, Figure 5.13(c-d), which results in a more unique kernel solution, Figure 5.13(a-b).

Positron range modelled



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

Figure 5.12: Central axial slice of maps formed from the output of GATE (a) magnified section of emission map; (b) magnified section of the annihilation map; (c) annihilation map; (d) magnified section of the difference in annihilation and emission positions (annihilation map - emission map).

The difference image in Figure 5.12(d) depicts some annihilations occurring in the polystyrene spheres, where no activity was simulated, due to positron range. The optimal kernel FWHM determined from the annihilation image, was found to be isotropic, 6.0 ± 0.0 mm (minimum nWMSE = 0.0119 ± 0.0001 % across voxel sub-sampling realisations), Figure 5.14. When the GT emission was smoothed with the optimal h_{PVC}, the RMSE in the VOI was 1.11 % with respect to the GATE annihilation map, and 1.17 % with respect to the GATE emission map.

5.4.2.3 Discussion

Profiles through the sphere lattice shows that 8 mm diameter spheres do not provide much inhomogeneity in the reconstructed PET image with $h_{PF} = 4$ mm^3 , even at high numbers of iterations (200i), Figure 5.13(d). While the closely-packed hexagonal lattice used in this simulation is the greatest achievable packing fraction for

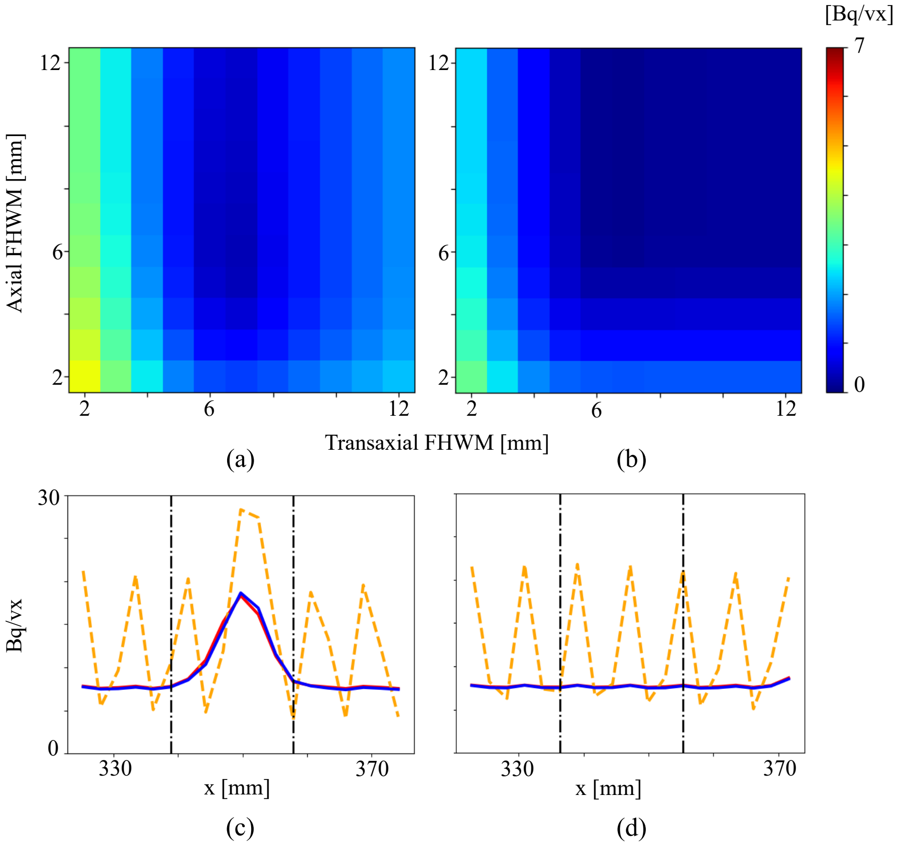


Figure 5.13: Heatmaps depicting RMSE for each kernel combination investigated for (a) VOI positioned over the "edge" and (b) VOI positioned centrally within the sphere lattice for 200i MLEM. The minimum RMSE is found at 7 x 7 x 6 mm^3 in both cases. Profiles in x-direction through the centre of the VOI positioned (c) over the "edge" and (d) centrally within the sphere lattice for 200i MLEM. The reconstructed image is depicted in blue, the unsmoothed GT in dotted orange, and the optimally smoothed GT in red.

equal diameter spheres, and so provides an underestimation of the inhomogeneity that would be seen in measured data using 8 mm beads, the kernel determination is still unstable. Inserting a gap between the spheres results in an unchanged, but more unique kernel solution, as can be seen from the heatmaps in Figures 5.13(a-b). Gaps between the spheres could be created by inserting a section of permeable polyester veil in each plane.

The kernel determined from the "reconstructed" annihilation (and emission) image of the sphere lattice edge was 6.0 mm^3 for all 20 voxel sub-sampling realisations. The RMSE in the VOI of the optimally smoothed GT emission with respect to the GATE emission and annihilation "reconstructed" images were comparable

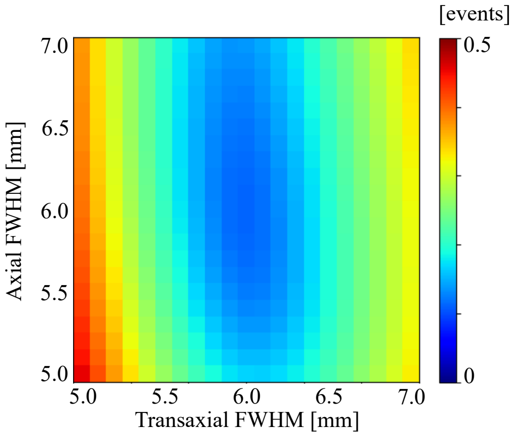


Figure 5.14: Heatmap depicting RMSE for each kernel combination investigated for a VOI positioned over the "edge" of the annihilation image.

within 0.06 % of the mean GT emission in the VOI. This suggests that the positron range of ^{18}F does not affect quantification for this simulation set-up.

5.4.2.4 Conclusions

An increase in inhomogeneity in the PET image results in a more unique kernel solution when using the MKA. AFC quantification was unaffected by the positron range of ^{18}F in 8 mm diameter polystyrene beads.

The applicability of a kernel determined from a region of high heterogeneity, as modelled in Section 5.4.2.1, to patient data has been investigated in Appendix E.

5.4.3 Second generation design

The purpose of this phantom redesign was to allow for (i) greater inhomogeneities in the PET image on which to fit the kernel, (ii) evaluate the applicability of the kernel to tumour imaging, and (iii) evaluate the applicability of the kernel to structures and densities approximating lung disease.

To increase PET inhomogeneity in the region used in the MKA, a permeable-walled structure was assembled from a 2 mm thick 3D printed frame and polyester veil walls. To assess the accuracy of the kernel (and MFC) determined from this feature, the low density materials investigated in Section 5.4.1, except for the polypropylene, were again arranged into tetrahedral shapes. These tetrahedra were embedded in the lung volumes separate to the kernel determination feature. To increase the rigidity of the tetrahedrons such that the irregular shape was maintained once embedded in the lungs, 2 mm thick frames were 3D printed. This also allows for easier segmentation. These kernel validation inserts were tetrahedral in shape to allow the in-plane and axial kernel applicability to be assessed, as discussed in Section 4.3.3. A fillable spherical insert was included to assess kernel applicability to tumour imaging.

To meet the aims of the second generation phantom, two thorax phantoms were required. The four lungs in the phantoms were used as follows:

- Lung 1 (kernel determination): packed with 8 mm ϕ polystyrene beads. A 50 mm sided frame, which intersects all three planes, was 3D printed and polyester veil walls glued to the frame, Figures 5.15(b) and 5.16(f) . The infill of 3D print material was set to 10 % to reduce density. The feature was not attached to the phantom, but the lung background was tightly packed with polystyrene beads to prevent movement. The feature was positioned such that one of the polyester veil walls was approximately orthogonal to the axial plane of the scanner.

- Lung 2 (applicability to tumour imaging): packed with standard 1 – 4 mm ϕ lung background polystyrene beads. A 24.8 mm diameter fillable spherical insert was attached via a stem to the lid of the phantom.

- Lungs 3 and 4 (applicability to diffuse lung disease): packed with standard 1 – 4 mm ϕ lung background polystyrene beads. 3D printed tetrahedral frames (35 mm sides), Figures 5.15(a) and 5.16(a-e), were filled with the low density materials from the LDI phantom and closed with polyester veil walls. The infill of 3D print material was again set to 10 % to reduce density. A sixth solid balsa $\approx$ 35 mm sided tetrahedral shape was also included. The tetrahedrons were not attached to the phantom but the lung background was adequately packed with polystyrene beads to prevent movement; this was confirmed with a HRCT pre- and post-PET. It was ensured that there was adequate spacing ($>$ 15 mm) between each feature and the lung walls.

To more accurately estimate the MFs and densities of each low density material, a HRCT was acquired before (each feature voxel containing material and air, equation 5.2) and after filling (each feature voxel containing material and tracer-laced water, equation 5.3).

$$\mu_{empty} = V_m\mu_m + (1 - V_m)\mu_a \tag{5.2}$$

$$\mu_{full} = V_m\mu_m + (1 - V_m)\mu_w \tag{5.3}$$

where μ_{empty} is the mean LAC in the VOI of the empty phantom, μ_{full} is the mean LAC in the VOI on the filled phantom, μ_w is the LAC of water, μ_a is the LAC of air, V_m is the fractional volume of the low density material and μ_m is the LAC of the material. The material fraction and LAC can be solved for:

$$V_m = 1 - \frac{\mu_{full} - \mu_{empty}}{\mu_w} \tag{5.4}$$

$$\mu_m = \frac{\mu_{empty}}{V_m} \tag{5.5}$$

All designs were produced in Autodesk Fusion 360 (v2.0.18961), prepared in Ultimaker Cura (v5.6.0), and printed in polylactic acid (PLA) on an Ultimaker S5.

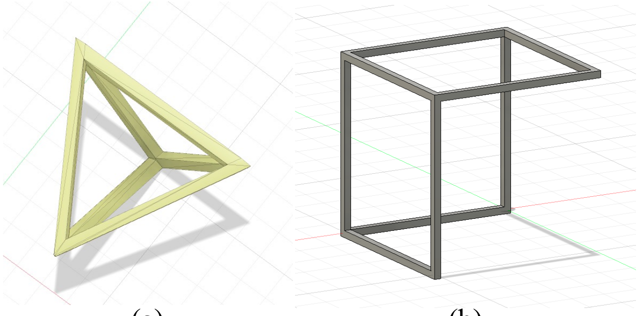

(a) (b)

Figure 5.15: CAD models of (a) the 35 mm sided tetrahedral frame, 2 mm thick; (b) the 50 mm sided square frame, with 2 mm thick sides.

5.4.3.1 Methods

^{18}F-FDG acquisitions of the second generation phantom were acquired on three clinical PET-CT scanners, of different makes and models at UCLH (GE Discovery

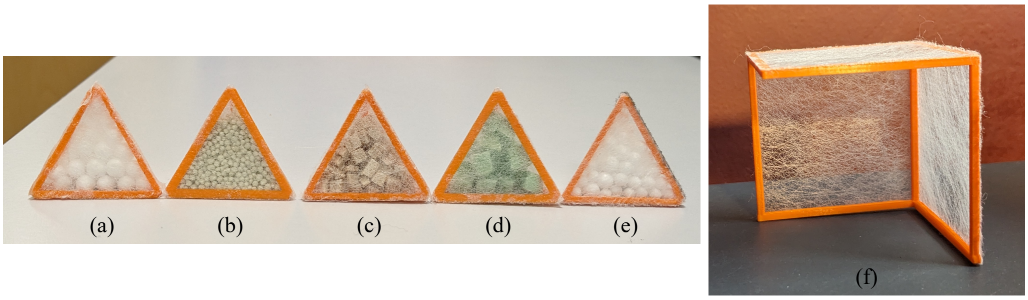


Figure 5.16: Photographs of 3D printed low density features before embedding in the lungs of the thorax phantom; (a) 5 mm diameter polystyrene beads; (b) 0.25 – 0.5 mm diameter expanded glass beads; (c)˜3 mm cubes balsa wood; (d)˜5 mm cubes closed-cell foam; (e) 2 – 5 mm diameter polystyrene beads; (f) kernel determination feature.

710, Siemens Biograph64 Vision 600, and GE Omni Legend). The optimal kernel for AFC was determined via the MKA_J, and its applicability assessed on a range of features of varying density, uptake and structures.

Phantom preparation and acquisitions

A HRCT of the two phantoms was acquired on the Siemens Vision before filling. This acquisition provided good feature wall contrast for segmentation, and enabled determination of MFs and CT numbers when voxel values are compared to the filled phantom using equations 5.2 and 5.3.

The phantoms were then filled with ^{18}F-FDG laced water, according to Table 5.7. A HRCT, CTAC, and LM PET were acquired on all three scanners, Table 5.8.

A HRCT was also conducted after the PET acquisition on the D710 to assess whether the inserts moved with respect to the phantom shell as the bed moved between the PET and CT acquisitions. The correlation of CT numbers in the feature VOIs pre- and post-PET acquisition were assessed.

Reconstructions

The reconstructions investigated for each scanner are listed in Table 5.9. A reconstruction that complies with “EARL2” standards, as described in Section 2.6, is included. OSEM reconstructions with an “EARL2” compliant PF were iterated for more updates to assess the effect on quantification. In addition, for the GE Omni Legend, images were also obtained using their PDL technology, described in Section 2.4. The parameters listed in Table 5.9 were being trialled clinically at

Table 5.7: ^{18}F-FDG activity concentrations, C, in second generation phantom at mid-point of scanning on three scanners. Desired phantom background activity concentration was 5.3 kBq/ml [51]. Lungs 1 and 2 were filled with an activity concentration ratio that resulted in a clinically representative unMF-corrected LBR $\approx$ 1 [101]. Lungs 3 and 4 were filled with an activity concentration that would result in an MF-corrected LBR $\approx$ 1 such that these VOIs could be MF-corrected without consideration of spill-in / spill-out from the lungs to the background.

Compartment	**Measured C [kBq/ml]**
Kernel determination and sphere phantom	
Background	4.83
Lungs 1 and 2	14.3
Sphere	46.1
LDI phantom	
Background	5.01
Lungs 3 and 4	5.05

Table 5.8: PET-CT acquisitions on each scanner. HRCT parameters: 120 kVp, 100 mAs, 0.8 pitch, "Edge" kernel.

Acquisitions		**D710**	**Vision**	**Omni**
Pre-fill	HRCT		x	
Post-fill	pre-PET HRCT	x	x	x
	CTAC	x	x	x
	LM PET (15 min/bed)	x	x	x
	Post-PET HRCT	x		

UCLH at the time of phantom acquisition. A clinical decision has since been made to not use the PDL reconstruction and as such the parameters (β value, level of contrast-enhancement-to-noise, and voxel size) have not been optimised for clinical practice.

GT construction

For determination of the MFs and LACs of each feature, the pre-fill HRCT on the Vision was registered (rigid only) to the post-fill HRCT on the Vision using NiftyReg [167]. The phantoms were registered independently due to their positioning not being fixed with respect to each other between acquisitions. The CT numbers were converted to LACs as in Carney et al. [24]. The voxel-wise MFs (V_{mv}) and

Table 5.9: Reconstruction parameters for the clinical reconstructions (in bold), reconstructions harmonised for tumour imaging ("EARL1" or "EARL2") and reconstructions with increased iterative updates for the three clinical PET-CT scanners at UCLH. The PDL algorithm, with "high" (H) contrast-enhancement-to-noise trade-off, was assessed on the GE Omni Legend. All PFs are Gaussian, except for the GE axial PF (a "standard" axial filter is a triangle filter with the ratio 1:4:1). All D710 and Vision reconstructions use TOF information; all Omni acquisitions are non-TOF.

Scanner	Reconstruction	i	sub.	Voxel size [mm^3]	PF FWHM [mm]
D710	OSEM ("EARL1")	2	24	2.73x2.73x3.27	6.4 / "standard"
D710	OSEM ("EARL2")	2	24	2.73x2.73x3.27	3.5 / "standard"
D710	OSEM	10	24	2.73x2.73x3.27	3.5 / "standard"
D710	OSEM	50	24	2.73x2.73x3.27	3.5 / "standard"
D710	**BSREM+RM (β=400)**	**-**	**-**	**2.73x2.73x3.27**	**0.0**
Vision	OSEM ("EARL1")	4	5	3.30x3.30x3.00	6.0
Vision	OSEM ("EARL2")	4	5	3.30x3.30x3.00	3.5
Vision	OSEM	24	5	3.30x3.30x3.00	3.5
Vision	OSEM	48	5	3.30x3.30x3.00	3.5
Vision	**OSEM+RM**	**4**	**5**	**1.65x1.65x3.00**	**0.0**
Omni	OSEM+RM ("EARL1")	3	22	2.73x2.73x2.07	8.0
Omni	**BSREM+RM(β=550; "EARL2")**	**-**	**-**	**2.73x2.73x2.07**	**0.0**
Omni	OSEM	3	22	2.73x2.73x2.07	3.5
Omni	OSEM	11	22	2.73x2.73x2.07	3.5
Omni	OSEM	55	22	2.73x2.73x2.07	3.5
Omni	HPDL+RM (β=450)	-	-	1.82x1.82x2.07	0.0

LACs (mu_{mv}) were determined from equations 5.2 and 5.3 on unsmoothed HRCTs. This methodology assumes:

- that there is a single material present in each voxel, in addition to water or air, hence the use of unsmoothed HRCTs.
- that the inserts do not move within the phantom during the filling process. This was checked through visual inspection of each feature on the pre-fill HRCT registered to the post-fill HRCT in ITK-SNAP [126].

A GT emission was constructed from the segmented HRCT of the empty phantom on the Vision (ITK-SNAP [126]). The phantom background, spine insert and lungs were semi-automatically segmented. The sphere and stem were modelled as outlined in Appendix C. To determine image-derived activity concentrations, C, large regions were drawn in the reconstructed phantom background, ensuring that phantom and lung edges were avoided. The mean phantom background activity concentrations were scaled for the lungs (C_{lung}) and sphere (C_{sphere}) according to the activity ratios measured in the radionuclide calibrator. The activity concentration in each lung background or feature voxel (C_v) is:

$$C_v = (1 - V_{mv})C_{lung} \tag{5.6}$$

which assumes that there is no air trapped in the filled phantom i.e., each voxel contains only material and tracer-laced water.

To ensure a GT activity concentration estimate in the sphere that is independent of any scanner calibration errors, a spherical VOI, large enough to ensure all counts originating in the sphere were included (34 mm diameter), was centred on the sphere. The activity concentration ratio between the “EARL2” reconstructed image with respect to the smoothed (isotropic Gaussian 6 mm^3 FWHM) GT was calculated for each scanner; the mean ratio for the three scanners was 0.90 ± 0.02; the activity concentration for the sphere was reduced by this factor. Axial slices of the unsmoothed GT are shown in Figure 5.17.

Kernel determination

The pre-processing methodology was conducted as in Joshi et al. (MKA$_J$) [119]. First, all reconstructions from the three scanners were registered to the GT image (voxel size: 0.74 x 0.74 x 0.60 mm^3; linear interpolation) [167]. The registered images were then normalised such that the mean of all voxels within both phantoms was one. The upsampling of all PET reconstructions to the GT voxel size allows comparison of different scanners.

The optimal kernel was determined via the MKA$_J$ from three 34 mm diameter spherical VOIs (two of which are depicted in Figure 5.17(a)), which intersect the centre of the permeable walls of the kernel determination feature, 5.16(f); the VOIs

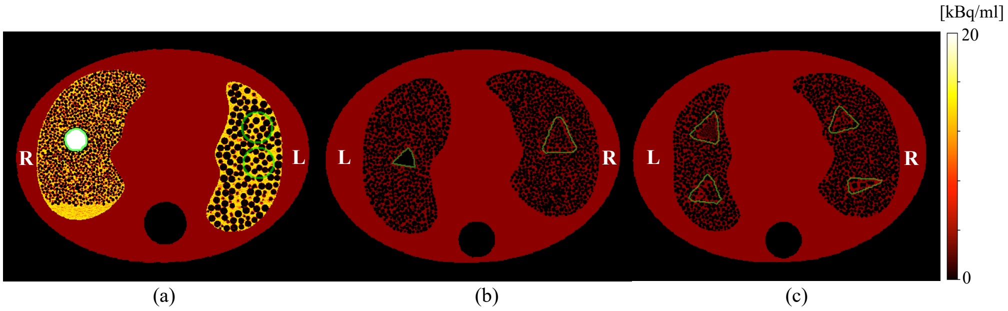


Figure 5.17: GT constructed from HRCT of the phantom and image-derived activity concentrations, the feature VOI edges are shown in green; (a) axial slice through the kernel determination feature (L lung) and hot sphere (R lung); (b) axial slice through solid balsa (L lung) and 2 – 5 mm polystyrene bead (R lung) feature; (c) axial slice through expanded glass (anterior L lung), 5 mm diameter polystyrene bead (posterior L lung), chipped balsa (anterior R lung), and closed-cell foam (posterior R lung) features.

were combined to form a single kernel determination VOI. nWMSE with respect to the upsampled reconstructed image, as described in Chapter 4, was computed; the kernel FWHM pair that resulted in the smallest nWMSE was reported as the optimal kernel. The uncertainty on the FWHM was estimated by computing the standard deviation of the FWHMs derived by randomly sub-sampling half of the VOI voxels, see Appendix D. 40 realisations were conducted (Omni reconstructions only) – the standard deviation of the first batch of 20 were equivalent to the second batch of 20. The uncertainty derived from 20 realisations are reported for all scanners.

Testing of MFC accuracy

Each VOI and lung background was MF-corrected with MFs determined from optimally smoothed HRCTs. The target-to-background ratio (TBR) of the LDIs compared to the MF-corrected GT TBRs ($RD_{MFC-TBR}$) were assessed in VOIs the physical dimensions of the insert.

Testing of kernel accuracy

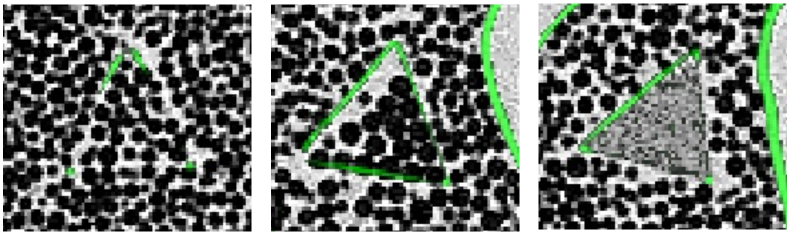
The relative difference between the optimally smoothed GT and reconstructed VOI values (RD_{Bqml}) was assessed in the six LDIs (normalised to reconstructed lung background; VOI: physical dimensions of the insert dilated by 4 HRCT voxels isotropically) and the sphere (normalised to reconstructed mean in sphere; VOI: physical dimension).

5.4.3.2 Results

The image derived unMF-corrected LBRs for each scanner were 1.05 ± 0.01 for lungs 1 and 2 and 0.35 ± 0.01 for lungs 3 and 4 (“EARL1” reconstructions).

LDI properties

Figure 5.18 depicts the HRCT of the 3D printed features on the Vision post-fill (greyscale) with the registered pre-fill phantom overlaid (green). It can be seen that the 3D printed tetrahedral frames in the empty phantom (green) align with the feature boundaries in the post-fill image. The Pearson’s correlation coefficient, r, of the solid balsa feature was 0.985.



2 – 5 mm ϕ poly. 5 mm ϕ poly. Expanded glass

Kernel det. feature

Closed-cell foam Balsa cubes

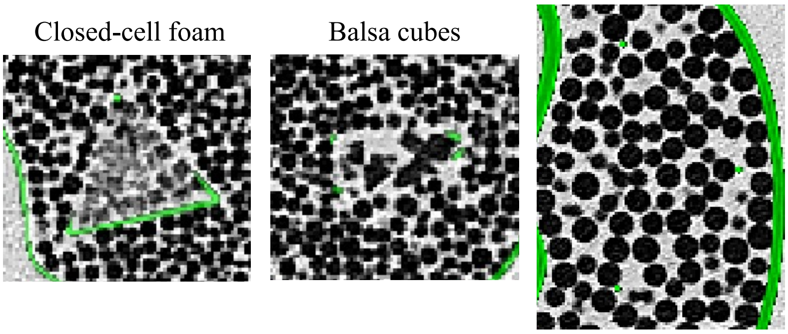


Figure 5.18: Pre-fill HRCT (black – green) overlaid on the post-fill HRCT (greyscale) for the five 3D printed LDIs and kernel determination feature (pre-fill window: -900 – 0; voxel size: 0.74 x 0.74 x 0.60 mm).

The polystyrene beads were not visible on the pre-fill HRCT due to the similarity of CT number to air; this negates the need to ensure that there was no movement of beads upon filling the phantom. Slices through the two non-polystyrene bead features that contained material fragments larger than the resolution of the HRCT (closed-cell foam and balsa wood cubes) are shown in Figure 5.19. Material fragments have not moved more than a couple of voxels (< 2 mm), which is acceptable in the context of MFC at reconstructed PET image resolution.

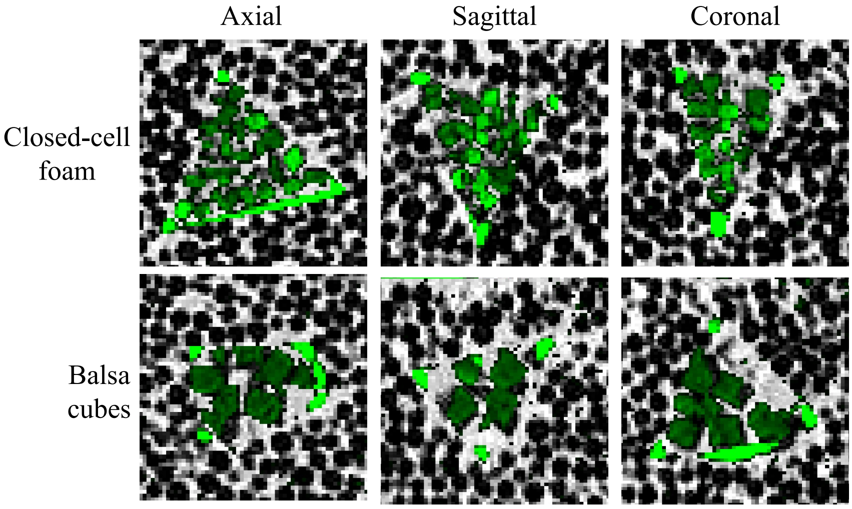


Figure 5.19: Axial, sagittal and coronal slices through the pre-fill HRCT (black – green) overlaid on the post-fill HRCT (greyscale) for the two LDIs with material fragments larger than the resolution of the HRCT (pre-fill window: -950 – -700; voxel size: 0.74 x 0.74 x 0.60 mm); top row: closed-cell foam; bottom row: balsa wood cubes.

The MFs and LACs estimated from equations 5.4 and 5.5 are listed in Table 5.10. The LAC for polystyrene beads ranged from 0.0020 - 0.0027 cm^{-1} (MF range: 0.611 - 0.677). The estimated LAC for the cubed balsa wood was 64.5 % greater than for the solid block.

Table 5.10: Estimated LACs and material fractions for each feature and lung backgrounds. Calculated from equations 5.4 and 5.5 and unsmoothed HRCTs.

Material	**Estimated LAC at 511 keV [cm^{-1}]**	**Material fraction**
polystyrene beads 2 – 5 mm ϕ	0.0020	0.677
solid balsa	0.0104	0.983
expanded glass	0.0489	0.685
balsa cubes	0.0171	0.504
closed-cell foam	0.0121	0.579
polystyrene beads 5 mm ϕ	0.0025	0.652
polystyrene beads 1 – 4 mm ϕ	0.0021	0.632
polystyrene beads 8 mm ϕ	0.0027	0.611

To assess whether the inserts moved with respect to the phantom shell when the bed moves between PET and CT acquisitions, the distribution of CT numbers in each LDI was assessed on the HRCT acquired pre- and post-PET. The HRCTs were first smoothed to approximate reconstructed PET image resolution (isotropic Gaussian kernel FWHM of 6 mm^3). It can be seen in Figure 5.20 that the CT numbers before and after bed movement correlate well on a voxel-by-voxel basis for all LDIs. The balsa cubes feature had the lowest correlation, potentially due to having the lowest MF of all of the inserts, Table 5.10, resulting in movement of the cubes within the tetrahedral structure. LDI CT numbers were in the range expected in a patient diagnosed with severe IPF, Figure 5.21.

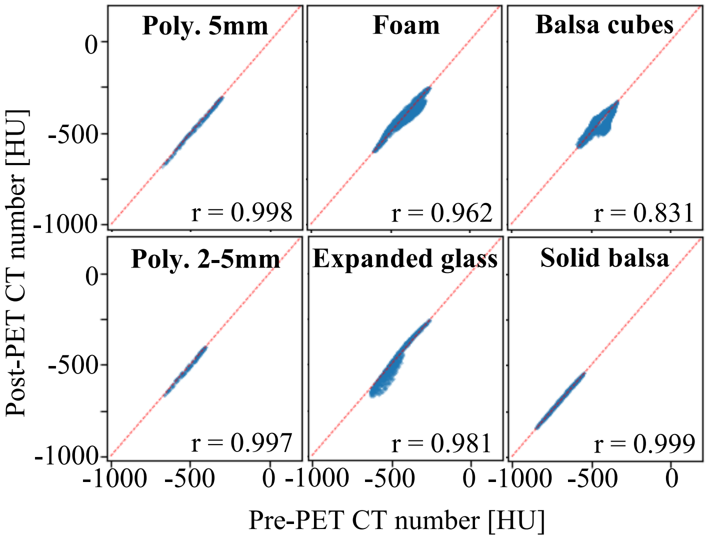


Figure 5.20: Correlation of LDI CT numbers for smoothed and down-sampled HRCTs acquired pre- and post-PET acquisition on the D710. Pearson's correlation coefficient is denoted by r.

Kernel determination

AFC kernels for each scanner/reconstruction pair were determined with the MKA$_J$ on a region of high PET inhomogeneity, Table 5.11. Agreement with the kernels determined from the fillable spheres in Chapter 4 was within uncertainties for "EARL1" reconstruction on the Vision, both in-plane and axially, and axially for "EARL1" and Q.Clear on the D710. The kernels determined from the RM reconstruction on the Vision were statistically significantly different ($p<0.0001$), Figure 5.22.

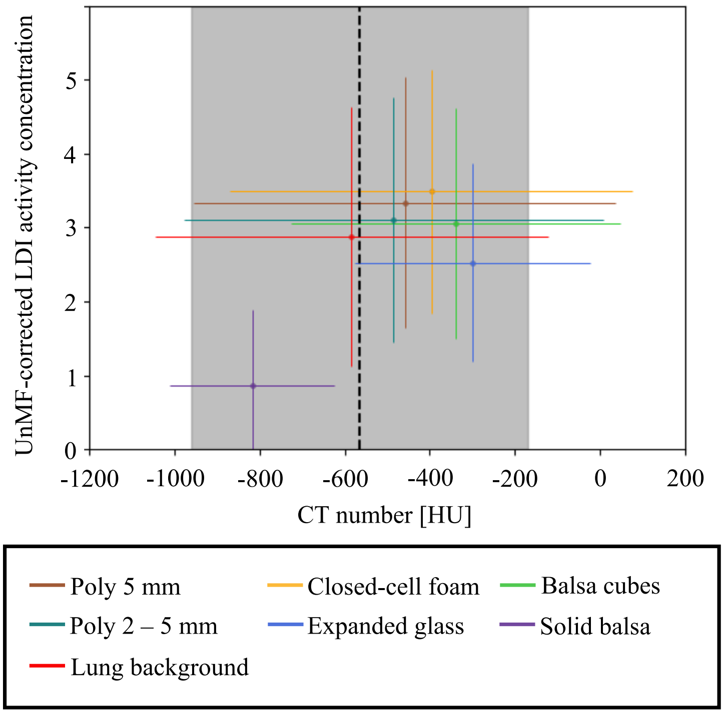


Figure 5.21: CT number versus unMF-corrected activity concentrations in unsmoothed GT for the six LDIs and lung background; error bars indicate one standard deviation within each VOI (physical dimensions of features). Black dotted line represents the mean CT number (grey region depicts range) in a VOI on the patient data used in Section 4.2.

The uncertainty on the optimal kernel width is reduced, Figure 5.22, compared to the non-rigid tetrahedron, Section 5.4.1, and the fillable spheres, Section 4.3. The width of the optimal kernel decreases with increasing iterations for all three scanners, as shown in Figure 5.23 for the Vision. It can be seen from the profiles through the permeable walls of the kernel determination feature, surrounded by 8 mm ϕ polystyrene beads, that structure in the activity distribution is more clearly defined as the reconstruction algorithm converges (in addition to increased noise), resulting in a less smooth GT being the optimal match, Figure 5.24.

MFC accuracy

The expanded glass LDI resulted in a mean MFC-TBR of 1.40 ± 0.16 across all scanners and reconstructions, Figure 5.25(b). The mean GT MFC-TBR was 1.35 ± 0.13, resulting in a mean $RD_{MFC-TBR}$ of 3.72 ± 5.78 % across scanners and reconstructions. Inspection of the water fraction (WF) maps (1 - MF) for this feature

Table 5.11: In-plane and axial h_{PVC} FWHM that resulted in the smoothed GT emission image that best matched the reconstructed image. Uncertainty is one standard deviation of 20 voxel sub-sampling realisations, see Appendix D. All D710 and Vision reconstructions are TOF; all Omni reconstructions are non-TOF.

Scanner	**Reconstruction**	**PF FWHM [mm]**	$h_{PVC}(xy)$ **FWHM [mm]**	$h_{PVC}(z)$ **FWHM [mm]**
D710	OSEM 2i24s ("EARL1")	6.4 / "standard"	12.2±0.00	7.5±0.00
D710	OSEM 2i24s ("EARL2")	3.5 / "standard"	9.2±0.00	7.7±0.00
D710	OSEM 10i24s	3.5 / "standard"	7.5±0.00	7.6±0.04
D710	OSEM 50i24s	3.5 / "standard"	7.2±0.00	6.0±0.02
D710	BSREM+RM (β=400)	0.0	12.2±0.00	7.6±0.00
Vision	OSEM 4i5s ("EARL1")	6.0	9.4±0.03	8.3±0.05
Vision	OSEM 4i5s ("EARL2")	3.5	8.0±0.00	6.7±0.04
Vision	OSEM 24i5s	3.5	6.5±0.00	6.5±0.00
Vision	OSEM 48i5s	3.5	6.3±0.00	6.5±0.00
Vision	OSEM+RM 4i5s	0.0	6.8±0.00	5.5±0.00
Omni	OSEM+RM 3i22s ("EARL1")	8.0	12.9±0.03	5.0±0.00
Omni	BSREM+RM (β=550; "EARL2")	0.0	12.0±0.04	5.6±0.00
Omni	OSEM 3i22s	3.5	10.3±0.00	5.2±0.00
Omni	OSEM 11i22s	3.5	9.0±0.05	5.1±0.00
Omni	OSEM 55i22s	3.5	8.8±0.02	5.1±0.05
Omni	HPDL+RM (β=450)	0.0	9.5±0.00	5.3±0.00

displayed large heterogeneity, despite the small glass bead diameters, Figure 5.26. This suggests that the assumption of each voxel containing only tracer-laced water and expanded glass had broken down in some voxels of the feature, most likely due to air trapping. A reliable MFC, which assumes only material and water in each voxel, can therefore not be performed for this feature and it will not be considered in further analysis.

The mean absolute $RD_{MFC-TBR}$ for all other 3D printed inserts was 2.64 ± 2.01 % for all reconstructions on all scanners; mean GT MFC-TBR of 1.00 ± 0.01, Figure 5.25(a). The maximum absolute $RD_{MFC-TBR}$ for 3D printed LDIs was 5.8 % across all scanners and reconstructions. The foam insert reconstructed with HPDL on the Omni was an outlier with an $RD_{MFC-TBR}$ = 12.5 %, Figure 5.25.

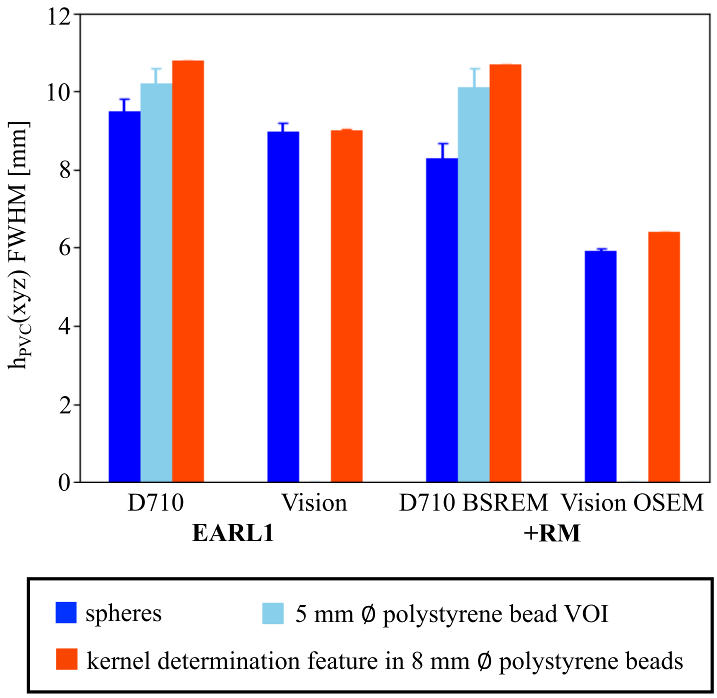


Figure 5.22: $h_{PVC}(xyx)$ FWHM for different VOIs, within which MSE was minimised to determine optimal kernel, for the reconstructions assessed in Chapter 4 ("EARL1" compliant and reconstructions that include RM) on the D710 and Vision. Error bars indicate one standard deviation of 20 voxel subsampling realisations (subsample factor = 2). NB: the 5 mm ϕ polystyrene bead VOI for kernel determination was not assessed on the Vision.

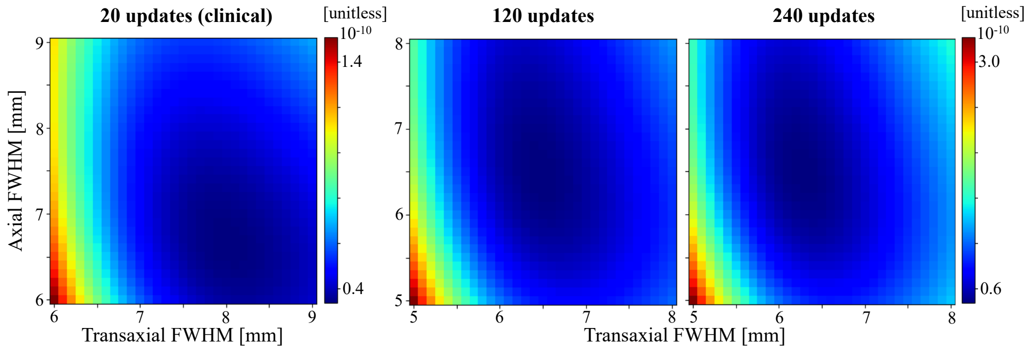


Figure 5.23: Heatmaps depicting nWMSE for each kernel combination investigated for the Vision OSEM reconstruction for increasing numbers of iterative updates (subsets = 5).

The solid balsa insert had a mean (maximum) $RD_{MFC-TBR}$ of 5.64 ± 3.61 % (12.2 %, RM recon on the Vision).

Kernel accuracy

The mean absolute RD_{Bqml} for all LDIs across all reconstructions on all scan-

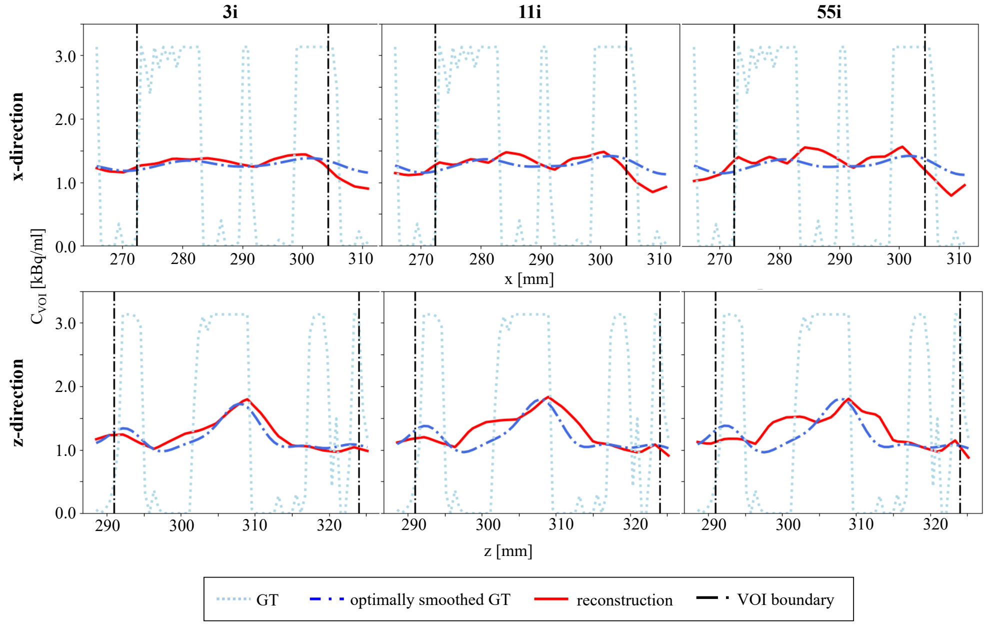


Figure 5.24: Profiles in x-direction through the centre of one of the kernel determination VOIs shown in Figure 5.17 (top row) and z-direction through the centre of a VOI intersecting the kernel determination feature wall orthogonal to the axial axis (bottom row) for non-TOF OSEM 3i, 11i and 55i reconstructions on the Omni Legend. The reconstructed image is depicted in red, the unsmoothed GT in dotted pale blue, and the optimally smoothed GT in dash-dotted royal blue.

ners was 13.4 ± 6.23 % (range: -0.46 – 32.1 %). The reconstructions that resulted in the lowest and highest mean absolute RD_{Bqml} across all inserts are shown for each scanner in Figure 5.28.

The 5 mm ϕ polystyrene bead insert on the Omni had the largest RD_{Bqml} across all reconstructions (mean=28.3 ± 2.00 %), Figure 5.28. Visual inspection of the registered "empty" HRCT acquired on the Vision to the "full" HRCT acquired on the Omni confirmed that the insert had not moved with respect to the phantom. The mean RD_{Bqml} for the 2 – 5 mm ϕ on the Omni was 15.7 ± 0.49 %, which is in-line with the other LDIs. Profiles through the unMF-corrected lung background, Figure 5.29, show a +8.43 % bias further supporting the hypothesis that the bias is not due to non-linearities in the polystyrene HU to LAC conversion, which could have led to an incorrect AC.

Sphere RD_{Bqml} was < 7.54 % for all reconstructions that did not include RM. On all three scanners, reconstructions that included RM resulted in the greatest

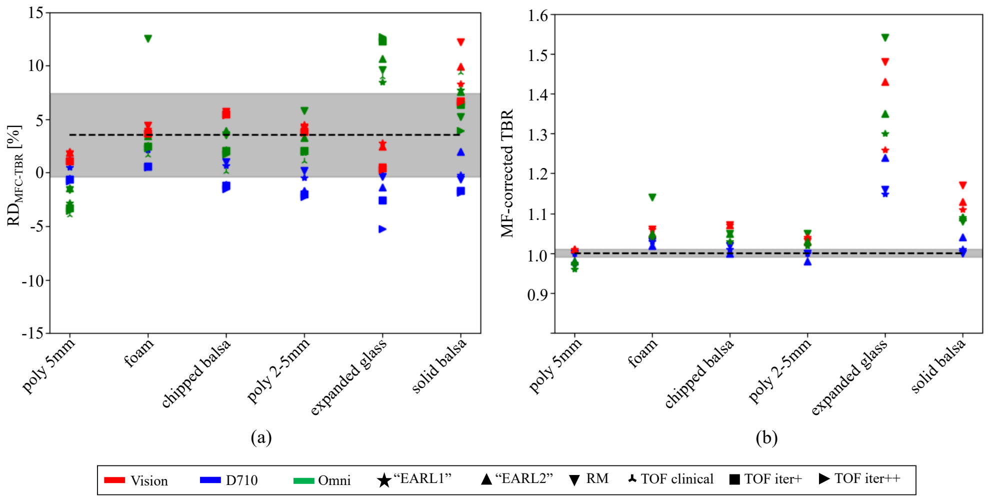


Figure 5.25: (a) $RD_{MFC-TBR}$ for LDIs for all reconstructions on all scanners; (b) LDI MFC-TBR for all reconstructions on all scanners; mean absolute values and one standard deviation are depicted by the black dotted line and shaded grey region, respectively.

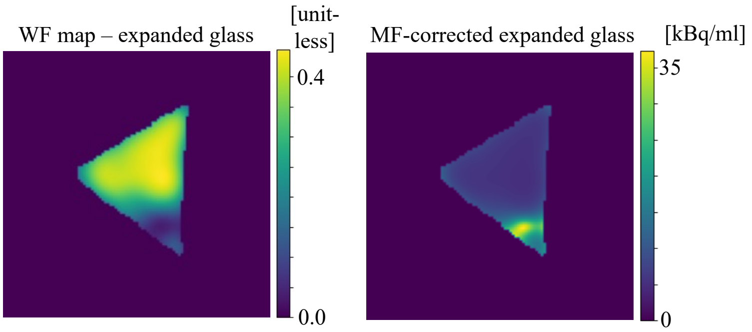


Figure 5.26: Left: WF map for expanded glass feature; right: MF-corrected expanded glass feature.

sphere RD_{Bqml}, and those with the most smoothing resulted in the smallest sphere RD_{Bqml}. Profiles through the sphere for the reconstruction with the smallest and greatest RD_{Bqml} on the Vision is shown in Figure 5.30. The edge artefacts observed in the RM reconstruction of the sphere are not as apparent for the kernel determination feature and solid balsa feature shown in Figure 5.31. The 3D printed LDIs were unaffected by edge artefacts as there was no observable sharp transition in image intensity due to the permeable walls.

For the TOF scanners, sphere RD_{Bqml} was negative for "EARL2" reconstructions (D710: -2.88 %; Vision: -3.05 %), i.e. the optimally smoothed GT was more

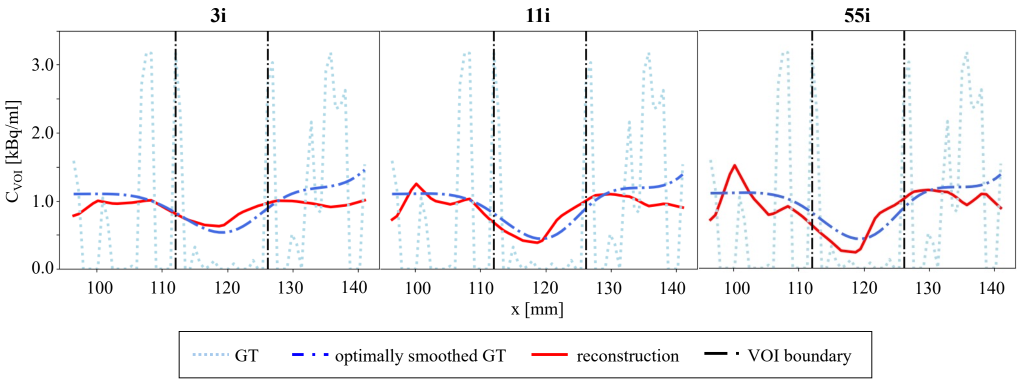


Figure 5.27: Profiles in x-direction through the centre of the unMF-corrected solid balsa insert reconstructed with non-TOF OSEM 3i, 11i and 55i on the Omni Legend. The reconstructed image is depicted in red, the unsmoothed GT in dotted pale blue, and the optimally smoothed GT in dash-dotted royal blue.

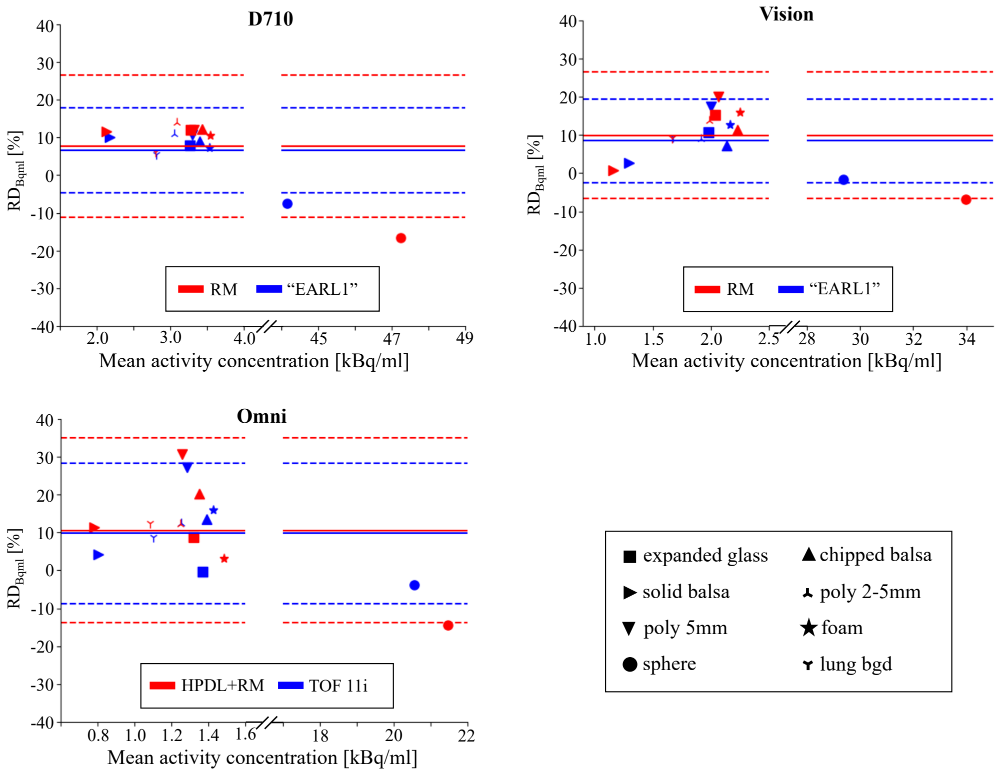


Figure 5.28: Bland-Altman plots for unMF-corrected C_{VOI} in optimally smoothed GT for the reconstructions with the highest (lowest) mean absolute RD_{Bqml} across all inserts in red (blue). Mean and 95 % CIs shown by solid and dotted lines, respectively.

smooth than the reconstruction. Increasing the number of iterations reduces the kernel width resulting in positive RD_{Bqml}s (D710: 1.82 %; Vision: 1.04 %), i.e. the optimally smoothed GT overestimates the activity in the sphere.

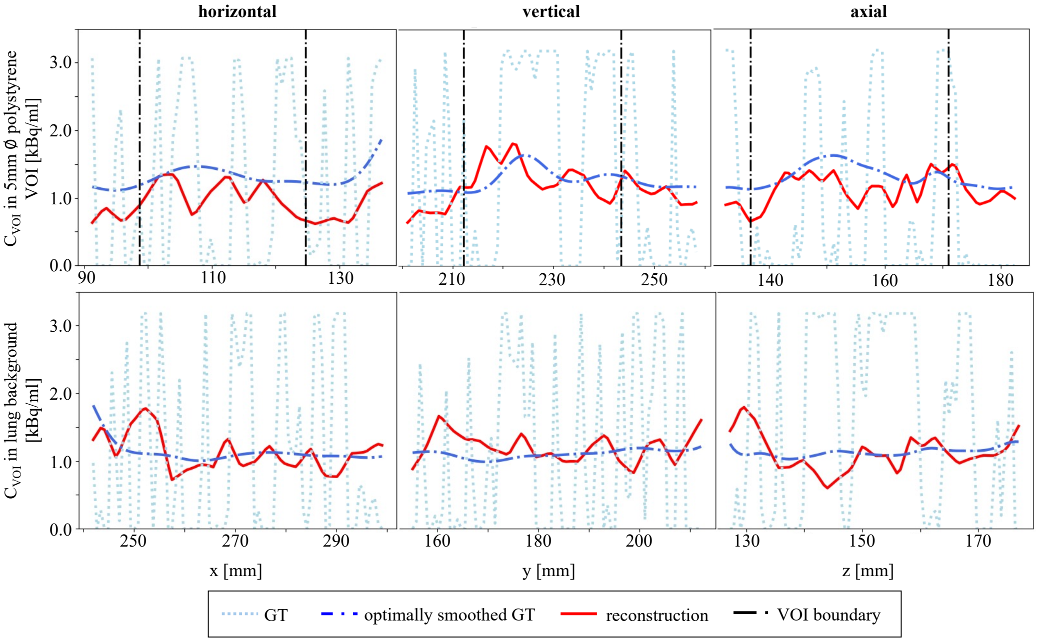


Figure 5.29: Horizontal, vertical and axial profiles through the unMF-corrected 5 mm ϕ polystyrene bead insert (top row) and lung background (bottom row) for the non-TOF OSEM 55i reconstruction on the Omni Legend. The reconstructed image is depicted in red, the unsmoothed GT in dotted pale blue, and the optimally smoothed GT in dash-dotted royal blue. VOI boundary shown in dash-dotted black.

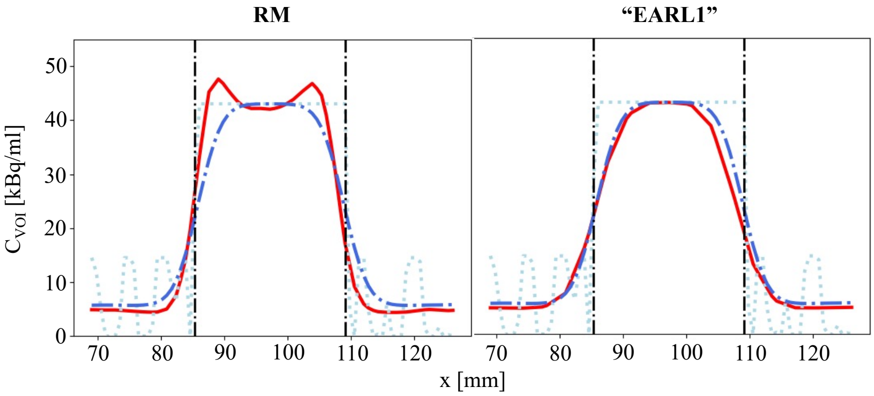


Figure 5.30: Vision acquisition: profiles in x-direction through the centre of the sphere for RM and "EARL1" reconstructions. The reconstructed image is depicted in red, the unsmoothed GT in dotted pale blue, and the optimally smoothed GT in dash-dotted royal blue. VOI boundary shown in dash-dotted black.

.

Increasing the number of iterations increased both the reconstructed contrast of the very low uptake solid balsa insert, and noise in the image, Figure 5.27 (RD_{Bqml} increased from -0.46 % to 7.01 %; $RD_{MFC-TBR}$ decreased from 9.40 % to 3.93 %

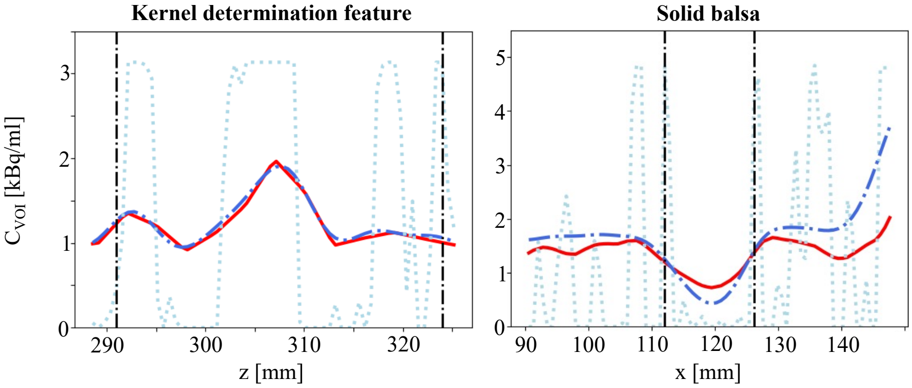


Figure 5.31: Vision acquisition: profiles in z-direction through the centre of the kernel determination feature, and x-direction through the centre of the solid balsa feature for the RM reconstruction. The reconstructed image is depicted in red, the unsmoothed GT in dotted pale blue, and the optimally smoothed GT in dash-dotted royal blue. VOI boundary shown in dash-dotted black.

.

for the non-TOF OSEM 3i to 55i reconstructions on the Omni Legend).

5.4.3.3 Discussion

This second generation of phantom design was easy to fill and required few segmentations to construction the GT. The method for determining MFs and LACs from equations 5.2 and 5.3 resulted in an estimated polystyrene bead LAC of 0.0023 ± 0.0003 cm^{-1} across the four different configurations. The two regions of balsa wood had different LAC estimates (0.0171 cm^{-1} for the balsa cubes and 0.0104 cm^{-1} for the solid balsa tetrahedron). Figure 5.19 indicates that the balsa cubes did not move substantially in the filling process. However, they did have the lowest correlation coefficient between bed positions, Figure 5.20, which was explained by the lowest packing fraction. The discrepancy could also be due to subdivision, or a different part of the wood being used, allowing more water to infiltrate the cubes than the solid block, resulting in a higher LAC for the cubes. This variance highlights the need to estimate MFs from measured data. The features did not move with respect to the phantom (measured at reconstructed PET image resolution) on filling, with bed motion, or between scanner acquisitions.

This phantom design enabled determination of reconstruction dependent kernels for AFC. The uncertainty on the optimal kernel width was reduced, Figure 5.22, compared to the non-rigid tetrahedron, Section 5.4.1, and the fillable spheres, Section 4.3, potentially due to the PET inhomogeneity introduced by the kernel determination feature. Increasing the number of iterations on the OSEM reconstruction reduced the h_{PVC} kernel width on all scanners, Table 5.11.

The D710 required the widest axial FWHM kernels to achieve optimally smoothed GTs, this is thought to be due to the larger crystal dimensions in the axial direction (6.3 mm versus 3.2 mm for the Vision and 4.1 mm for the Omni). The Omni also has a smaller voxel size (slice thickness: 2.07 mm versus 3.00 mm for the Vision and 3.27 mm for the D710). The in-plane crystal dimensions are roughly equivalent for the GE scanners, 4.2 mm for the D710 and 4.1 mm for the Omni, compared to 3.2 mm for the Vision.

The expanded glass LDI RD_{Bqml} and $RD_{MFC-TBR}$ were consistent with the other 3D printed LDIs, Figures 5.28 and 5.25(a). However, the MF-corrected GT and MF-corrected reconstruction were 1.35 ± 0.13 and 1.40 ± 0.16, respectively. Due to potential air trapping within this feature, made more likely by the very small diameter beads, it is not suitable for an easy-to-fill, reproducible, harmonisation phantom.

With the exception of the DL algorithm on the Omni applied to the foam insert, scanner and reconstruction specific AFC kernels resulted in $RD_{MFC-TBR} < 5.8$ % in all regions.

Increasing the number of iterations had little impact on the MF-corrected 3D printed, fine-scale structure, insert quantification, Figure 5.25(a). Improved matching was seen in the solid balsa insert and sphere at increased iterations. This suggests different reconstruction convergence rates in different parts of the image, supporting the need for lung-specific harmonisation. It also highlights the utility of validating the kernel with less fine-scale structures, which are more sensitive to changes in the kernel width.

The mean bias of +13.4 % in LDI RD_{Bqml} across all three scanners, Figure 5.28, could be due to:

1. Inaccuracies in the GT construction:

   (a) An overestimation of the lung background activity concentration, as seen for the sphere (where the C_{sphere} measured in the radionuclide calibrator was scaled by 0.90). The image derived MF-corrected activity concentrations for the lungs containing the LDIs was 5.0 % lower than the activity measured in the radionuclide calibrator and so could account for some of the observed bias. Absolute calibration was not conducted as part of this work due to difficulties with the radionuclide calibrator.

   (b) Inaccuracies in MF estimates – a more accurate estimate of MFs could be conducted with micro-CT acquisitions of each LDI prior to being a embedded in the phantom. As MF calculation relies on CT number linearity, calibration of the micro-CT would need to be ensured.

2. Systematic underestimation of kernel width – the permeable wall of each 3D printed LDI results in a higher activity concentration at the edges due to a lower packing fraction of the insert material and the polystyrene beads in the lung background, an underestimation of the kernel width would lead to the GT being less subject to the spill-out component of the PVE and therefore an overestimation of activity concentration in the VOI with respect to the reconstruction.

The solid balsa insert has a high MF and is therefore less sensitive to inaccuracies in activity concentration (it has roughly the same positive bias in the first and second generation of the phantom), see Section 5.4.1.3 for discussion. No evidence of overall bias, which would indicate a suboptimal scatter correction, was observed in the profiles in Figure 5.27.

On all three scanners, reconstructions that included RM resulted in the largest sphere RD_{Bqml}. This is due to the edge artefacts observed as overshoots at sharp intensity transitions. This effect is well documented in the PET and SPECT literature, as discussed in Tong et al. [168]. Due to the less sharp intensity transitions for the other features, edge artefacts were not observed.

5.4.3.4 Conclusions

The inclusion of a region of high PET heterogeneity in this phantom reduced the uncertainty on estimated reconstruction-dependent kernels with respect to previous generations of phantom. Improved quantification was observed for the solid balsa region, suggesting different convergence rates in different parts of the image, supporting the need for lung specific harmonisation. With the exception of a man-made material (closed-cell foam) and the reconstruction algorithm that employs DL, the MFC-TBRs of each feature were in good agreement with the MF-corrected GT. This suggests that harmonisation for application to patient PET-CT lung scans is feasible.

5.5 Harmonisation

The EANM's "EARL" accreditation scheme harmonises scanners for tumour imaging, see Section 2.6. It does this by ensuring that, regardless of scanner manufacturer or model, sphere RCs remain stable within ± 10 % for the four largest, and ± 20 % for the two smallest spheres in the NEMA/IEC body phantom (ECT/IECBODY/P NEMA) with a sphere-to-background ratio = 10:1 [56]. The specifications for the "EARL1" ^{18}F-FDG PET–CT accreditation program were developed during a pilot study performed during 2010 – 2011, involving 12 PET–CT systems. The performance of PET–CT systems significantly improved in the following years, due to new acquisition and reconstruction technologies. The harmonisation of the newer systems required an update to the "EARL1" standards to accommodate higher recoveries, this standard is known as "EARL2" [58]. The 24.8 mm sphere in the phantom in Section 5.4.3, with a 9.2 sphere-to-(inhomogeneous) background ratio, had an inter-scanner RC variation of $<$ 9.5 % for both "EARL1" and "EARL2" reconstructions.

Kernels have been determined for the "EARL1" reconstruction on homogeneous, relatively high density spheres, Section 4.3, and the low uptake, low density kernel determination feature, Section 5.4.3, for both the GE D710 and the Siemens Vision. The kernel widths estimated from the kernel determination feature agreed, within uncertainties, with the kernel determined from the spheres, for the "EARL1" reconstruction on the Vision, but were significantly different for the "EARL1" reconstruction on the D710. This suggests that harmonisation for lung disease imaging cannot be assumed despite EARL harmonisation for tumour imaging.

In this section, a preliminary study will be conducted to assess whether the phantom acquisitions in Section 5.4.3 allow harmonisation of the Siemens Vision, GE D710, and GE Omni Legend for lung studies. Whether inter-scanner variance on the quantification of low uptake, low density, fine-scale structures, as well as the sphere, can be reduced by matching image resolutions will be investigated. Matched reconstructed image resolution will be accomplished through additional smoothing.

5.5.1 Methods

In previous work, reconstructed image resolution was estimated via a Gaussian kernel, h_{PVC}. Extra post-filtering of these images with a Gaussian filter, h_{HPF}, will lead to a FWHM equal to $\sqrt{h_{PVC}^2 + h_{HPF}^2}$, as demonstrated in Section 5.4.1.2. We therefore propose to post-filter images such that final resolution matches the widest h_{PVC} of the scanners/reconstruction protocols tested.

A (non-RM) OSEM reconstruction, with the post-filter applied clinically, was investigated for each scanner with (i) the number of iterative updates (u) used in the clinic, (ii) 240u reconstruction (scanner subsets equivalent to those used clinically). The maximum in-plane and axial resolutions determined in Section 5.4.3 are listed in Table 5.12 as *"max. h_{PVC}"*. The additional in-plane and axial smoothing required to match image resolutions to the maximum h_{PVC} are listed for each scanner as $h_{HPF}(xy)$ and $h_{HPF}(z)$.

Each additionally smoothed reconstruction was assessed in terms of LDI unMF-corrected TBR and sphere RC. For the sphere, the EARL approach of RC values was adopted based on (i) 50 % background-corrected isocontour VOI (RC_{mean}), (ii) maximum voxel value included in the 50 % isocontour (RC_{max}), (iii) a spherical VOI with a 12 mm ϕ, positioned to yield the highest uptake (RC_{peak}). The RC was calculated:

$$RC = \frac{C_s/C_b - 1}{A_s/A_b - 1} \tag{5.7}$$

Table 5.12: Additional smoothing, h_{HPF}, required to ensure matched resolutions across all three scanners. h_{PF} is the post-filter applied as part of the reconstruction and max. h_{PVC} is the smoothest reconstructed image resolution estimated in Section 5.4.3.

	Scanner	**Recon.**	$h_{PF}(xy)$ **[mm^2]**	$h_{PF}(z)$ **[mm]**	**max.** $h_{PVC}(xy)$ **[mm^2]**	**max.** $h_{PVC}(z)$ **[mm^2]**	$h_{HPF}(xy)$ **[mm]**	$h_{HPF}(z)$ **[mm]**
OSEM clinical updates	D710	2i24s	3.5	1:4:1	12.0	7.7	7.7	0.0
	Vision	4i5s	3.5	3.5			8.9	3.8
	Omni	3i22s	3.5	3.5			0.0	5.3
OSEM 240u	D710	10i24s	3.5	1:4:1	9.0	7.6	5.0	0.0
	Vision	48i5s	3.5	3.5			6.4	3.9
	Omni	11i22s	3.5	3.5			0.0	5.6

where C_s is the measured activity concentration within the chosen VOI, C_b is the average activity concentration in a reconstructed lung background VOI the physical dimension of the sphere, A_s is the activity concentration in the unsmoothed sphere GT, and A_b is the activity concentration in the unsmoothed background VOI GT.

The expanded glass insert was excluded from harmonisation analysis, as discussed in Section 5.4.3.2.

5.5.2 Results

For the OSEM clinical updates reconstruction, harmonisation resulted in a reduction in inter-scanner variation in LDI TBRs greater than uncertainties (3σ) in all but the balsa cubes insert, Figure 5.32. Only the 2 – 5 mm ϕ polystyrene bead insert demonstrated a significant reduction in inter-scanner variation for the OSEM 240u reconstruction. Increasing the number of iterations to 240u resulted in a reduction in inter-scanner variation, without additional smoothing, for all 3D printed LDIs; this effect was within uncertainties for all but the 5 mm ϕ polystyrene bead insert.

The solid balsa region, being the lowest uptake of all of the regions is most sensitive to scanner differences, e.g. TOF versus non-TOF. This is highlighted by the increase in inter-scanner variance as the iterations increase – the TOF Vision and D710 TBR decreases to a greater extent than the non-TOF Omni. Harmonisation increased inter-scanner variation in sphere RC_{mean} for both clinical updates and

240u OSEM reconstructions. The additional smoothing decreased the mean RC_{mean} from 0.67 to 0.62 for clinical updates and from 0.69 to 0.65 for 240u.

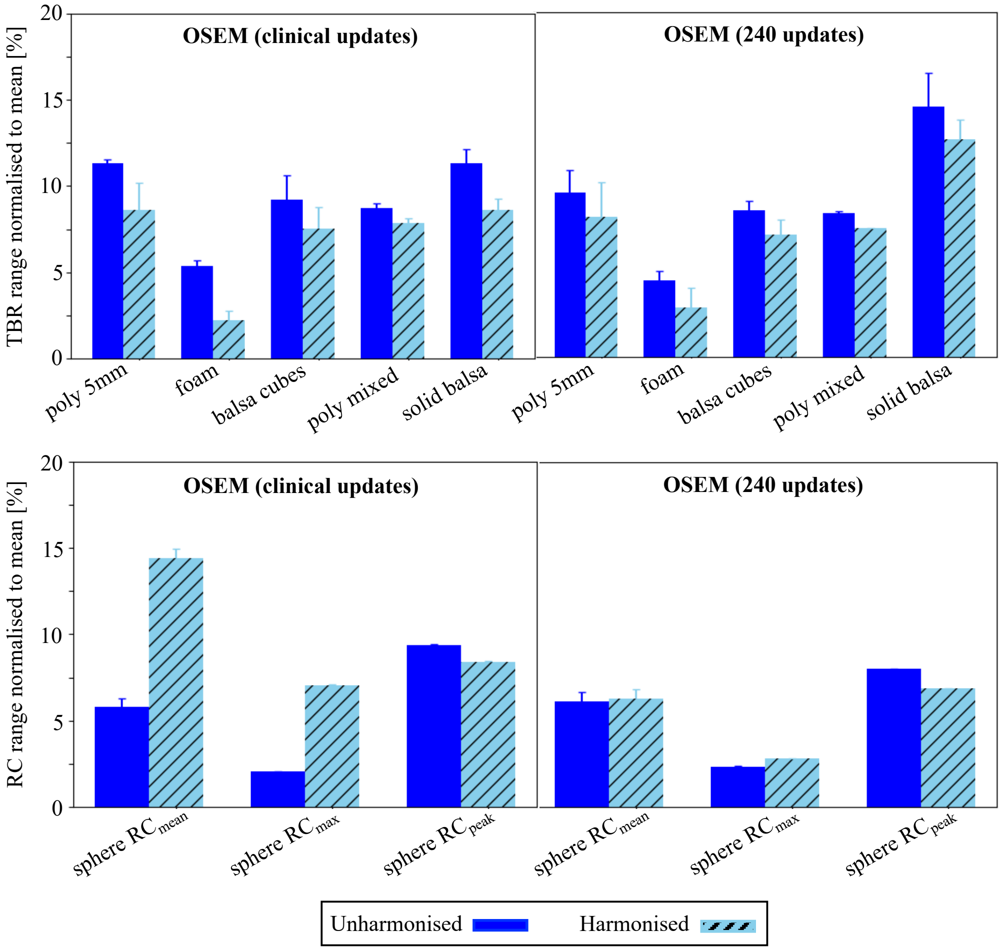


Figure 5.32: Range (as a percentage deviation from mean across three scanners) of unMF-corrected LDI TBRs (top row) and sphere RCs (bottom row) for OSEM clinical updates (left column) and OSEM 240u (right column) reconstructions for unharmonised and harmonised data. Error bars are 1σ on mean VOI voxels, propagated to calculation of range as percentage of mean.

All RC measures fulfilled the EARL requirement of RC remaining stable within ±10 % for this sphere diameter with the exception of the harmonised clinical updates reconstruction for RC_{mean}, Figure 5.32. The kernel was determined from a low density, low uptake, kernel determination feature; the variation in RC_{mean} suggests different rates of convergence in different parts of the image at the lower numbers of iterations. As noted by Kaalep et al. 2018b RC_{peak} is less sensitive to changes in reconstruction parameters than RC_{mean} or RC_{max} [57].

5.5.3 Discussion

Harmonisation resulted in a significantly reduced inter-scanner variation in LDI TBRs in all but the balsa cubes insert for the clinical update reconstructions. Only one of the LDIs demonstrated a decrease in inter-scanner variation greater than uncertainties at 240u. This is potentially due to increased quantitative accuracy of the low uptake regions as the reconstruction approaches convergence on each scanner.

Comparison of harmonised data demonstrated equivalent inter-scanner variability for the 3D printed (/fine-scale structure) features, and reduced variability for the sphere with increased iterations. This supports the conclusions of Chapter 3, that large numbers of iterative updates are needed to achieve stability in measured reconstructed image resolution. 240u reconstructions result in a superior harmonised reconstructed image resolution with respect to clinical numbers of iterations (clinical updates $h_{PVC}(x,y,z)$ = 12.0 x 12.0 x 7.7 mm^3; 240u $h_{PVC}(x,y,z)$ = 9.0 x 9.0 x 7.6 mm^3), Table 5.12, with comparable (or reduced in the case of the spheres) inter-scanner variability.

The solid balsa has the largest inter-scanner variance of the OSEM clinical updates and 240u reconstructions; with such low uptake (V_{mat} = 0.983), this VOI is particularly sensitive to differences in scanner technologies and convergence rates. Such low uptake regions may be seen clinically in lung diseases that give rise to pulmonary bullae and therefore highlights the importance of specific harmonisation for diseased lung imaging with a variety of PET scanner technologies and reconstructions currently available.

Too few scanners have been investigated as part of this PoC work to form definitive conclusions, especially given the difference in technologies between the three scanners.

5.6 Conclusions

A review of phantom literature suggested that there is a lack of low density phantoms suitable for optimisation and harmonisation of NM imaging. In general, publications mostly demonstrated that the materials were low density by quoting CT numbers. Tissue mimicking materials for PET imaging must reproduce the interaction cross sections for the production of secondary radiations as well as the bulk attenuation and scattering properties [130]. Identification of suitable lung tissue mimicking materials that would allow the production of lung-specific PET phantoms has not been widely adopted to-date.

Thorax phantom inserts were constructed to better approximate the density and activity distribution seen in diffuse lung disease. Designing rigid-framed, permeable-walled, low density features enabled quick and simple phantom filling, more accurate feature segmentation, and increased activity inhomogeneity within each feature, and in the surrounding volume.

The uncertainty in kernel width was reduced with increased PET inhomogeneity when using the matched kernel approach. Increasing the volume of inactive low density features to provide increased inhomogeneity in the PET image did not result in image artefacts due to positron range when imaging ^{18}F. There could however be implications for image quality when using radionuclides that emit higher energy positrons due to increased positron range. This would need to be investigated further with simulations.

The overall positive bias of the optimally smoothed GT with respect to the reconstructed image for all LDIs needs to be investigated further with improved radionuclide calibrator measurement accuracy and material fraction and density estimation.

The overestimation of activity in the solid balsa region highlights the potential for quantitative inaccuracies in low uptake regions due to iterative reconstruction and potentially sub-optimal scatter correction. Increasing the number of iterations improved the estimated reconstructed image resolution, and reduced the bias observed on the otherwise under-converged MF-corrected solid balsa region. The hot sphere RD_{Bqml} negative bias was reduced (to a positive bias of equal magnitude for the TOF capable scanners) with increased numbers of iterative updates. This suggests different reconstruction convergence rates in different parts of the image and supports the need for lung disease specific harmonisation projects.

The fine-scale structures in the 3D printed LDIs are not particularly sensitive to the blurring applied to the CT for MFC. This was also observed in regions of fibrosis in patient data in Chapter 4. However, both emphysematic and fibrotic lungs suffer from large air pockets. As demonstrated by the sphere and solid balsa inserts, these regions are likely to be more sensitive to kernel width accuracy. This underlines the need for inserts to be less fine-scale to assess the kernel validity. Alternative low density materials, within which activity concentration could be varied whilst remaining relatively homogeneous, could be investigated for their sensitivity to the effects of kernel width on quantification.

The effect of using a CTAC resolution CT for GT construction and kernel determination needs to be examined as HRCTs are not generally available for patient data.

A scanner specific AFC/PVC kernel would need to be determined for each radionuclide used to image lungs. Monte Carlo simulations, to ensure that the kernel determination feature design is appropriate for other radionuclides, e.g. Gallium-68, would need to be conducted.

The kernel determination design enables reconstruction dependent kernels to be determined for AFC. The phantom acquisitions could also be used to recommend scanner-specific PFs to promote harmonisation of different clinical scanners for application to patient PET-CT lung scans.

Chapter 6

General Conclusions

6.1 Summary and conclusions

Accurate quantification of radiotracer uptake from lung PET-CT is challenging due, in part, to large variations in the air fraction (AF). The AFC determines voxel-wise AFs from the CT acquired for AC to correct for this. The mismatch in resolution between PET and CT can cause artefacts in the attenuation- and AF-corrected image. The work in this thesis has concentrated on the development of digital and physical phantoms to (i) aid in matching the CT resolution to the PET image resolution and (ii) assess the impact of resolution mismatches on AFC quantification in clinically realistic densities and uptake patterns. Respiratory motion is not considered as part of this work. The feasibility of using a diffuse lung disease mimicking phantom to increase quantitative reproducibility in PET-CT lung imaging across scanners of different makes and models was assessed at a single site.

A summary of the chapters in this thesis and the main conclusions drawn from each are given below.

Chapter 3: The effect of PET-CT resolution mismatch in lung imaging

The first stage of this project was to validate existing correction techniques for attenuation and air fraction. Simulations and reconstructions of digital phantoms concluded that the CT should be smoothed to match the intrinsic PET scanner resolution for AC and the reconstructed PET image resolution for AFC.

For ease of implementation, the reconstructed PET spatial resolution is often assumed to spatially invariant; this is unlikely to be the case for patient datasets reconstructed with parameters that are practicable in the clinic. The magnitude of the variance in reconstructed PET spatial resolution for the digital phantoms was assessed in non-TOF MLEM and OSEM reconstructions of increasing iterative updates using the point source insertion-and-subtraction method. It was shown that artefacts were more evident at tissue boundaries in the AC-AFC images reconstructed with lower numbers of iterative updates. Despite this lack of stability in measured reconstructed image resolution at low iterative updates, the error in AFC quantification was small for the simple phantoms simulated. Reconstruction parameters with respect to bias on the kernel were investigated for patient realistic digital and physical phantoms in Chapter 4.

The interplay of the two aspects of the partial volume effect (PVE) – the tissue fraction effect (TFE) and the spill-in/spill-out component – was investigated with simulations of simple digital phantoms in a proof-of-concept study. It was concluded that the PET image must be corrected for the spill-in/spill-out component of the PVE prior to AFC. The CT resolution for determination of the AFs must match the PV-corrected PET image resolution.

Chapter 4: Kernel determination methodologies

This chapter investigated three different methodologies by which the reconstructed PET image resolution, as a function of reconstruction algorithm and convergence, can be determined for smoothing of the CT for AFC. In simulation studies, the point source insertion-and-subtraction method was compared to two variations of the matched kernel approach (MKA) (matching pre-AFC and matching post-AFC). Simulations of an IPF patient-realistic digital phantom for a non-TOF PET scanner indicated that a large number of MLEM iterations (around 200) were needed to approach reconstructed image resolution stability. AF-corrected quantification was shown to not be very sensitive to the smoothing applied to the CT for patient realistic data. This indicates that a single global kernel could be applied to the CT for determination of AFs, with which to AF-correct the lung on a voxel-wise basis. A kernel derived on VOIs positioned at multiple locations throughout the lung was shown to be more numerically stable than from a single VOI. It was concluded that the most practical method to determine the kernel for AFC is to smooth a GT emission image to match the reconstructed PET image. The applicability of the pre-AFC MKA, h_{PVC}, to measured phantom data was assessed.

^{18}F-FDG acquisitions of a thorax phantom with an additional fillable spherical insert in each lung were assessed on two clinical PET-CT scanners and a range of reconstructions used in the clinic. Kernels were successfully determined using the h_{PVC} methodology for each reconstruction, with less than 10 % bias on the unAF-corrected reconstruction with respect to the GT. The construction of the GT for this physical phantom was non-trivial due to registration and segmentation accuracy, simplification of this process would reduce workload and potentially improve kernel uncertainty (the uncertainty (3σ) on kernel FWHM ranged from 0.2 – 2.0 mm). Alternative phantom designs were investigated in Chapter 5 with these aims in mind.

Chapter 5: Novel phantom design

A review of phantom literature suggested that there is a lack of low-density phantoms suitable for optimisation and harmonisation of NM imaging. This chapter reports the development and application of a diffuse lung disease realistic phantom design to determine and validate the optimal kernel for AFC. The phantom was designed to be easy to fill, segment, and analyse. Lung inserts created variations in density and activity concentrations within the ranges expected in a patient with diffuse lung disease. A fillable sphere was also included to test kernel applicability to tumour imaging. An increase in PET inhomogeneity in the second generation of phantom reduced the uncertainty (3σ) on the estimated kernel FWHM to less than 0.2 mm.

As discussed in Chapter 4 for patient data, AF-corrected quantification was not shown to be very sensitive to the smoothing applied for the fine-scale features investigated. An overestimation of activity in the “cold” region was observed. This highlights the potential for quantitative inaccuracies in low uptake regions due to iterative reconstruction (and potentially sub-optimal scatter correction). Differences in quantification of the larger-scale "hot" and "cold" regions of uptake, with increased iterations, suggests different reconstruction convergence rates in different parts of the image. This is particularly important in both emphysematic and fibrotic lungs, which can be prone to large air pockets. This supports the need for lung disease specific harmonisation projects.

The final phantom design and kernel determination method investigated in this chapter are not limited to the determination of the AFC kernel but could also be utilised to determine the PSF for PVC for both lung disease and tumour imaging.

Finally, it was shown in a proof of concept (PoC) study, that these phantom acquisitions could also be used to recommend scanner-specific post-filters (PFs) to promote harmonisation of different clinical scanners for application to patient PET-CT lung scans.

6.2 Future directions

Several aspects of the research presented in this thesis would benefit from further investigation.

Phantom design

Detailed characterisation of the composition of materials contained in future phantom designs could be conducted with dual-energy micro-CT for material decomposition. This would provide confidence that the interaction cross section for annihilation, as well as the bulk attenuation and scattering properties for both the resultant 511 keV gammas and photons at CT energies, are being replicated in the low density materials selected [130].

To ensure accurate estimation of insert material fractions (MF), it is imperative that the lung inserts do not move in the phantom filling or acquisition process, see Section 5.4.3. Means of fixing the features in place, whether through (sub-PET resolution, so as to avoid "cold" regions) physical fixtures or tighter packing of lung background polystyrene beads, could be investigated.

The fine-scale structures were not shown to be very sensitive to the smoothing applied for AFC, either in patient data in Chapter 4 or the phantoms in Chapter 5.

The solid balsa insert provided a larger region (volume ≈ 5 cm^3) of low density and uptake and was shown to be more sensitive to differences in reconstruction parameters and AFC kernel widths. These larger regions of low density and uptake are clinically applicable for lung diseases that enlarge lung spaces, such as emphysema. The optimal dimensions of these larger scale features could be further investigated with simulations and reconstructions.

The applicability of the kernel determination feature in Section 5.4.3 for radionuclides with greater positron energies than ^{18}F could be investigated with Monte Carlo simulations, in the first instance.

Acquisition and reconstruction parameter recommendations

Throughout this work, i.e. the simulation studies in Sections 3.4 and 4.2 and phantom scans in Section 5.4.3, increased reconstructed image resolution stability has been demonstrated with increased numbers of iterative updates in the reconstruction. However, as the number of MLEM/OSEM updates increases, so does the noise in the image. Optimal reconstruction algorithms and parameters, with respect to the trade-off between noise and reconstructed image resolution stability, could be investigated by temporally rebinning a long duration phantom acquisition into realisations with equivalent statistical noise to a patient scan. The use of a regularised image reconstruction algorithm, which has the potential to control noise within the reconstruction process, could be investigated further. This work investigated GE's Bayesian penalised likelihood reconstruction algorithm (Q.Clear), however, this algorithm also includes resolution modelling (RM). Reconstructions that include RM, whether regularised or maximum likelihood based algorithms, demonstrated the greatest differences in AFC kernel width estimates from high uptake spheres (Chapter 4) and low uptake inserts (Chapter 5) and the greatest differences in both low uptake and high uptake insert quantification, with respect to the optimally smoothed GT. Further investigation is needed to determine the optimal reconstruction for PET lung imaging to optimise the trade-off between noise control and reconstructed image resolution uniformity. Optimal reconstructions for qualitative and quantitative assessment may differ, in which case, two images could be reconstructed, one for clinical reporting and one for quantitative analysis. The variability between vendor implementations of reconstruction algorithms would need to be considered in a multi-centre setting, as noted for tumour imaging by Kaalep et al. 2018b [57].

As in this work, the optimal width of the kernel for AFC should be determined on high count phantom data using the reconstruction parameters determined from patient-realistic count data; uncertainty on the kernel width could be determined from the different noise realisations. Our method of determining kernel width uncertainty was deemed to be an overestimate with respect to a ≈ 3 % noise level image, see Appendix D; this could be validated. Kernel uncertainty propagation to uncertainty in AF-corrected quantification could be investigated. Scanner-specific reconstruction parameters for patient data, and a Gaussian kernel width for AFC quantification accuracy to within a specified uncertainty could then be recommended.

The methodology for kernel determination in Chapter 5 requires a high resolution image on which to base the GT. The quoted optimal kernels for AFC are therefore the smoothing that should be applied to a high resolution CT (HRCT) to match reconstructed PET resolution. Due to radiation protection principles, patients undergoing a PET-CT scan will likely only undergo a low dose CT scan for AC (CTAC). As discussed in Chapter 3, this CTAC will have been smoothed to match the intrinsic resolution of PET scanner and will therefore require less smoothing than the HRCT to match reconstructed PET image resolution. An additional kernel, from HRCT resolution to CTAC resolution will need to be determined to be applicable to patient data.

Development of a partial volume correction (PVC) method, that performs well on very low uptake data and does not require segmentation of the very fine-scale structures of diseased lung on an anatomical image, could be investigated. The sensitivity of the chosen approach to mismatched anatomical information would need to be considered given that the data is subject to respiratory motion. In-reconstruction methods would be preferential from a noise perspective but increase the complexity for use in multi-centre studies.

Harmonisation

Standard operating procedures (SOPs) for phantom filling and acquisition would need to be produced to ensure a standardised approach to kernel determination across sites. Code written as part of this project could be developed into a robust GT construction, kernel determination and uncertainty quantification processing pipeline. Once validated, this could be used by the site conducting centralised analysis or rolled out to participating sites. To assess repeatability of kernel width determination, the phantom fill, acquisition, reconstruction, and processing should be repeated by the same operator on a scanner already evaluated using the SOPs and analysis pipeline developed. This test-retest study would give an overall measure of the variability introduced from activity measurement and dilution, slight (assuming fixed features in the final phantom) variations in feature positioning with respect to the detectors, reconstruction, GT activity concentration estimation and registration, and the MKA methodology.

This project examined variability across three scanners of different makes, models, and technologies in a single centre. Expanding to multiple sites and additional scanners would allow inter-scanner AF-corrected quantification variability, when working to SOPs, to be further characterised.

Recommendations for scanner-specific reconstruction parameters, and in some cases, additional post-filtering to enable comparable AF-corrected quantification between scanners of different makes and models could be made. Limits on AF-corrected metrics for scanner inclusion in the network would need to be determined from the clinical question. Cross-calibration between a centre's radionuclide calibrator, with proven traceability to a primary standard or equivalent, and PET scanner is paramount for reducing bias; this is verified through calibration QC scans in the EARL program [56]. For AF-corrected quantification, CT number accuracy and linearity would also need to be ensured. Once a standardised protocol has been established and implemented across participating sites, a system of regular verification could be introduced.

6.3 Publications and presentations

6.3.1 Peer-reviewed journal papers

- **F. Leek**, C. Anderson, A.P. Robinson, R.M. Moss, J.C. Porter, H.S. Garthwaite, A.M. Groves, B.F. Hutton and K. Thielemans. “Optimisation of the air fraction correction for lung PET/CT: addressing resolution mismatch”. Eur J Nucl Med Mol Imaging Physics 10:77, 2023.

- C.K. McGarry, L.J. Grattan, A.M. Ivory, **F. Leek**, G.P. Liney, Y. Liu, P. Miloro, R. Rai, A.P. Robinson, A.J. Shih, B. Zeqiri and C.H. Clark. “Tissue mimicking materials for imaging and therapy phantoms: a review”. Phys. Med. Biol. 65, 2020.

6.3.2 Conference proceedings

- **F. Leek**, C. Anderson, A.P. Robinson, R.M. Moss, B.F. Hutton and K. Thielemans. “Phantom design to enable harmonisation of quantitative PET imaging for lung disease”. Poster presentation. 2024 IEEE Nuclear Science Symposium (NSS), Medical Imaging Conference (MIC) and Room Temperature Semiconductor Detector Conference (RTSD), Tampa, FL, USA, 2024, pp. 1-1, doi: 10.1109/NSS/MIC/RTSD57108.2024.10657509.

- **F. Leek**, C. Anderson, A. Robinson, R. Moss, B. Hutton and K. Thielemans. “Preliminary results from a novel phantom design for harmonisation of lung PET/CT studies”. Journal of Nuclear Medicine 65 (suppl 2) 241445, 2024.

- **F. Leek**, C. Anderson, A. Robinson, R. Moss, B. Hutton and K. Thielemans. “Determination of the kernel with which to smooth the CT for air fraction correction in lung PET/CT studies”. Eur J Nucl Med Mol Imaging 50 (Suppl 1), 1–898, 2023. https://doi.org/10.1007/s00259-023-06333.

- **F. Leek**, A. P. Robinson, R. M. Moss, F. J. Wilson, B. F. Hutton and K. Thielemans. “Air Fraction Correction in PET Imaging of Lung Disease – Kernel Determination”. Poster presentation. 2022 IEEE Nuclear Science

Symposium and Medical Imaging Conference (NSS/MIC), Italy, 2022, pp. 1-3, doi: 10.1109/NSS/MIC44845.2022.10399046.

- R. Twyman, L. Brusaferri, E.C. Emond, **F. Leek**, S. Arridge, B.F. Hutton, V. Cuplov, K. Thielemans. “A Demonstration of STIR-GATE-Connection”. Poster presentation. 2021 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), Piscataway, NJ, USA, 2021, pp. 1-3, doi: 10.1109/NSS/MIC44867.2021.9875442.

- **F. Leek**, A. P. Robinson, R. M. Moss, F. J. Wilson, B. F. Hutton and K. Thielemans. “Air Fraction Correction Optimisation in PET Imaging of Lung Disease”. 2020 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), Boston, MA, USA, 2020, pp. 1-4, doi: 10.1109/NSS/MIC42677.2020.9507896.

6.3.3 Presentations

- “Improved quantification of PET imaging for lung disease - the long road to harmonisation”. Oral presentation. Virtual conference on applied radiation metrology (vCARM), National Physical Laboratory, 2021.

- “Harmonisation and quantification of PET imaging for lung diseases”. Oral presentation. Institute of Cancer Research, 2021.

Appendix A

GATE simulations

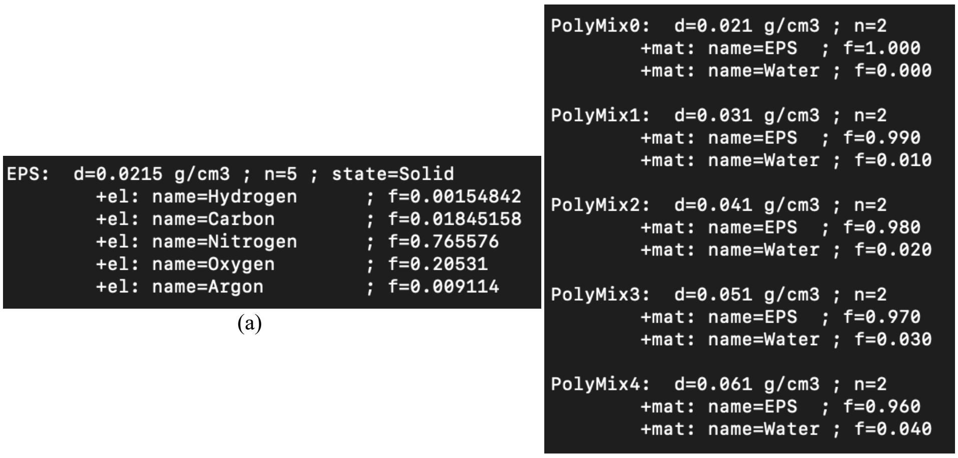
GATE is an open-source software package, developed by the OpenGATE collaboration (http://www.opengatecollaboration.org) [164]. GATE combines the advantages of GEANT4 i.e. well-validated physics models, sophisticated geometry description, and visualization and 3D rendering tools, with original features specific to emission tomography.

A.1 Defining a geometry

The imaging device and object to be scanned are described; particles are then tracked through the components of the geometry. The geometry of two PET scanners are described in the STIR-GATE-Connection (SGC; https://github.com/UCL/STIR-GATE-Connection), discussed in more detail in Appendix F. For the positron range simulations described in Chapter 5, it is not necessary to define a detector system as only the physical positions of positron emission and annihilation within the phantom are recorded. GATE compatible voxelised phantoms were created using STIR functionality, assigned activity or LAC values and saved and imported into GATE as interfiles.

A.2 Materials database generation

To read in voxelised geometry, GATE requires a text file for the description of materials as well as a 3D image. A text file provides a number of voxel value subdivisions, defines the intervals associated to each subdivision and specifies an associated material name. The properties of the materials used in GATE are defined in a materials database. Each material is specified in terms of its mass density in $g.cm^{-1}$, the elements comprising it, and their fractional contribution; this allows GATE to assign the nuclear properties from the GEANT4 datasets. The OpenGATE collaboration supplies a listing of materials in the database, however this needed to be modified to include expanded polystyrene (EPS), defined by it's elemental composition, fractions, and the corresponding density, Figure A.1(a). The boundary voxels in the spheres simulated in 5, are subject to the TFE i.e. they comprise of both water and EPS. Each attenuation voxel value was therefore assigned a "PolyMix" percentage within the voxel-value-to-material conversion text file, which points to a corresponding EPS and water fraction, Figure A.1(b), in the GATE materials database. "Nested parameterisation" was used to create a parametrised materials volume in Geant4, with a corresponding navigation algorithm.



```
EPS:  d=0.0215 g/cm3 ; n=5 ; state=Solid
        +el: name=Hydrogen      ; f=0.00154842
        +el: name=Carbon        ; f=0.01845158
        +el: name=Nitrogen      ; f=0.765576
        +el: name=Oxygen        ; f=0.20531
        +el: name=Argon         ; f=0.009114
```

(a)

```
PolyMix0:  d=0.021 g/cm3 ; n=2
        +mat: name=EPS   ; f=1.000
        +mat: name=Water ; f=0.000

PolyMix1:  d=0.031 g/cm3 ; n=2
        +mat: name=EPS   ; f=0.990
        +mat: name=Water ; f=0.010

PolyMix2:  d=0.041 g/cm3 ; n=2
        +mat: name=EPS   ; f=0.980
        +mat: name=Water ; f=0.020

PolyMix3:  d=0.051 g/cm3 ; n=2
        +mat: name=EPS   ; f=0.970
        +mat: name=Water ; f=0.030

PolyMix4:  d=0.061 g/cm3 ; n=2
        +mat: name=EPS   ; f=0.960
        +mat: name=Water ; f=0.040
```

(b)

Figure A.1: Excepts from GATE's material database, (a) EPS material specification; (b) fractional contributions of EPS and water.

A.3 Physics model and source specification

The "emlivermore" physics list, provided by the GEANT4 community was used. Voxelised activity phantoms were in interfile format, positioned in GATE space as described in the SGC. An ^{18}F ion source was utilised in the positron range simulations in Chapter 5. It was ensured that an ^{18}F ion source produced the expected mean and maximum positron ranges in water by first simulating a 10 kBq point source in a sphere of water. The positions of each annihilation event were recorded. The mean and maximum range were found to be 0.45 mm and 2.05 mm, respectively, which is within the range of values published in literature [108, 107].

Appendix B

Modified XCAT phantom generation

An IPF patient-realistic digital phantom was created by modifying the 4D extended cardiac-torso (XCAT) phantom. The near-homogeneous structure in the XCAT lungs were replaced by the lungs of a patient with severe IPF, Figure B.1. This allowed the effect of the disease structure on measures of resolution to be assessed whilst keeping the structure outside of the lungs relatively uniform. The simplicity of the phantom outside of the lungs aims to reduce computation time should the phantom be used as input to Monte Carlo modelling at a later date. The resolution of the phantom was maximised i.e., matched to the patient HRCT resolution, to preserve data in the construction of the GT.

The phantom construction steps were as follows:

1. **XCAT phantom simulation:** The XCAT phantom software [121] was used to create a 3D voxelised phantom. The parameters used are listed in table B.1. The voxel size was matched to the patient HRCT data. The resultant attenuation map was cropped to the thorax, to match the patient HRCT axial coverage, and the LAC units were converted from voxel^{-1} to cm^{-1}.

Table B.1: XCAT construction parameters

Parameter	Setting
Respiratory motion	Full inhale
Cardiac motion	None - end-diastole
Arm position	Not in the field-of-view
Voxel size (mm)	0.607 x 0.607 x 1.5

2. **Patient data acquisition:** A single patient with severe IPF was identified

from a study to assess the role of coagulation in IPF, conducted at University College London Hospital (UCLH, London, UK). The patient underwent a PET-CT acquisition on a GE Discovery 710 as part of a research study. An HRCT (voltage: 100 kVp, current: 149 mA, exposure time: 1.095 s, voxel size: 0.607 x 0.607 x 1.5 mm, pitch: 0.52) was acquired at end-expiration, in addition to the PET and CTAC scans. The HUs were converted to LACs [25].

3. **Segmentation of XCAT and patient lungs for registration:** The patient lungs were segmented on the HRCT using the "active contour segmentation mode" with thresholding, in ITK-SNAP (v3.8.0) [126]. Each XCAT lung was semi-automatically segmented using the same methodology as was used for the patient lungs. Mislabeling of vasculature and airways was manually corrected for such that they were included in the segmented lung. All segmentations were saved as NIfTIs.

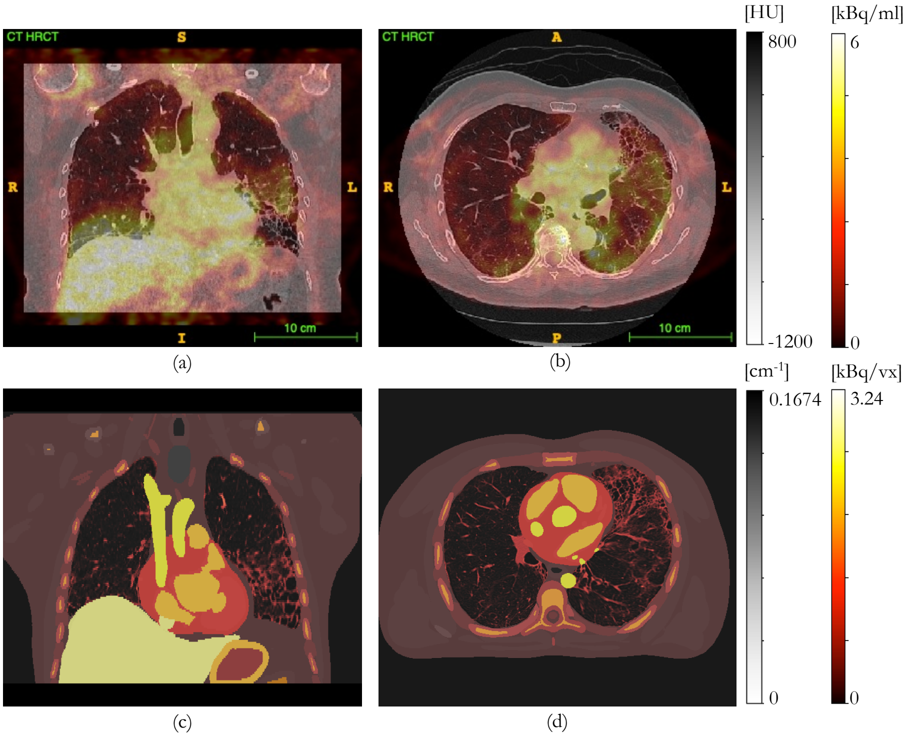
4. **Registration of patient lungs to XCAT lungs:** Both the XCAT and patient lung masks were binary dilated by a 7x7x1 box kernel using `fslmaths`; this ensured lung edges were included to assist registration. Patient lungs were registered to the XCAT attenuation map with Advanced Normalisation Tools (ANTs v2.3.5) [169] using symmetric diffeomorphic (differentiable map with differentiable inverse) transformation with the parameters listed in Table B.2. Registration was conducted only within the dilated lung masks and for each lung separately.

Table B.2: 3D ANTs registration parameters.

Parameter	Setting
Transformation model	Symmetric diffeomorphic (SyN)
Transformation	fast symmetric normalisation: affine + deformable transformation, with mutual information as optimisation metric. Gradient-descent step-size = 3
Similarity metric	Fast cross-correlation (window radius:4; weight: 1)
Optimisation	4 level resolution pyramid (25, 50, 25, 5 iterations)
Regularisation	12σ FWHM Gaussian on the gradient; 3σ FWHM on the deformation field

5. **Replacement of XCAT lungs with IPF patient lungs:** The XCAT lung values were replaced by the registered patient data lung values. In regions where the registration was not deemed acceptable i.e., too much warping of the patient lung, which occurred at the very base of one lung and a few voxels at the apex of both lungs, a value representative of the surrounding lung voxels was homogeneously assigned within that region.

6. **Creation of GT activity map:** Each organ outside of the lungs was segmented by their near-homogeneous XCAT voxel values. Each organ voxel was replaced by an activity concentration representative of IPF patient uptake. In the analytic simulations conducted in this work, only the ratios, not the units, of activity distribution are relevant; for convenience values are referred to in units of SUV throughout. The activity concentration in each lung voxel was weighted by the patient lung LACs, such that the AF-corrected lung was homogeneous. The value for AF-corrected lung uptake was derived as follows: normal physiologic skeletal muscle uptake varies between 0.5 - 2.2 $\mathrm{SUV}_{\mathrm{max}}$ [170]; a representative SUV of 1.0 was assumed. Holman et al. 2015 [101] demonstrates an approximate $\mathrm{SUV}_{\mathrm{AFC}}$ of 3 in the lungs, in this patient cohort. An AF-corrected uptake ratio of 3:1 between lungs and muscle was therefore assumed. The patient data used to construct the phantom had a mean uptake in muscle of 0.332, therefore, a homogeneous AF-corrected uptake, equal to 0.996, was set in the lungs.



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

Figure B.1: Fused PET-CT of patient with severe IPF (a) coronal slice; (b) axial slice; fused attenuation and GT activity map patient realistic XCAT phantom (c) coronal slice; (d) axial slice.

Appendix C

GT construction for scanner data

1. An HRCT of the empty (before filling with a solution of activity concentration) thorax phantom was acquired on a single scanner to improve contrast of insert and phantom edges to aid in semi-automatic segmentation. CT reconstruction with a sharp filter was used for clearer edge definition.

2. The phantom was filled with the solution of activity concentration, according to experimental design.

3. A HRCT, CTAC and LM PET scan were acquired on all clinical scanners.

4. All PET and CT scans of the phantom were converted from DICOM to interfile using the `stir_math` functionality.

5. The scanner beds were removed from both "full" and "empty" phantoms to ensure that the registration was constrained to phantom features only. In the case of multiple phantoms, each phantom was registered separately. The "empty" phantom HRCT (floating) was registered to the "full" phantom HRCT (reference); a rigid registration was conducted using NiftyReg `reg_aladin` [167].

6. The phantom components were segmented on the registered "empty" phantom in ITK-SNAP (v3.8.0) [126]. Each segmentation was saved as a NIfTI.

7. All segmentations were imported into STIR [111] as NIfTIs.

8. The centre-of-mass (COM) of sphere(s) and stem(s) was computed.
9. STIR objects were created for each of the spheres (value: -1), sphere walls (value: 1) and stems (value: 1), using the COMs and published / estimated values for dimensions. Each object was saved as a NIfTI.
10. STIR objects were imported into STIR as NIfTIs.
11. Sphere masks (-sphere), sphere wall masks (sphere wall + sphere) and stem masks (stem - sphere wall) were created. All were saved as NIfTIs.
12. The HRCT of the "full" phantom was converted from HUs to LACs using bilinear conversion detailed in Carney et al. [24] for a Siemens scanner and Burger et al. [25] for a GE scanner.
13. For lungs containing fillable Perspex spheres, a water fraction map was created within the lung background:

    $$\mu_v = V_p\mu_p + V_w(\mu_w - \mu_s) + \mu_s(1 - V_p) \tag{C.1}$$

    where μ_v, μ_p, μ_w and μ_s are the LACs for the voxel, Perspex, water and polystyrene respectively, and V_p and V_w are the fractional volume of Perspex and water respectively. In Chapter 4, μ_w and μ_s were 0.0960 cm^{-1} and 0.0079 cm^{-1} respectively. For lungs containing LDIs, a water fraction map was created within the lung background as in C.1, setting V_p=0. Insert water fraction maps substituted μ_s for μ_{mat}, where μ_{mat} (or HU_{mat}) were determined from the HRCT of the "full" and "empty" phantoms.
14. For lungs containing solid inserts (spheres and stems), these were subtracted from the lung background activity fraction map.
15. The component mask values were replaced with activity concentration values, as measured in the radionuclide calibrator (Chapter 4) or image-derived phantom background values and component ratios as measured in the radionuclide calibrator (Chapter 5).

16. The component masks were summed to form a high-resolution ground truth activity image.

An example workflow for GT construction is depicted in Figure C.1.

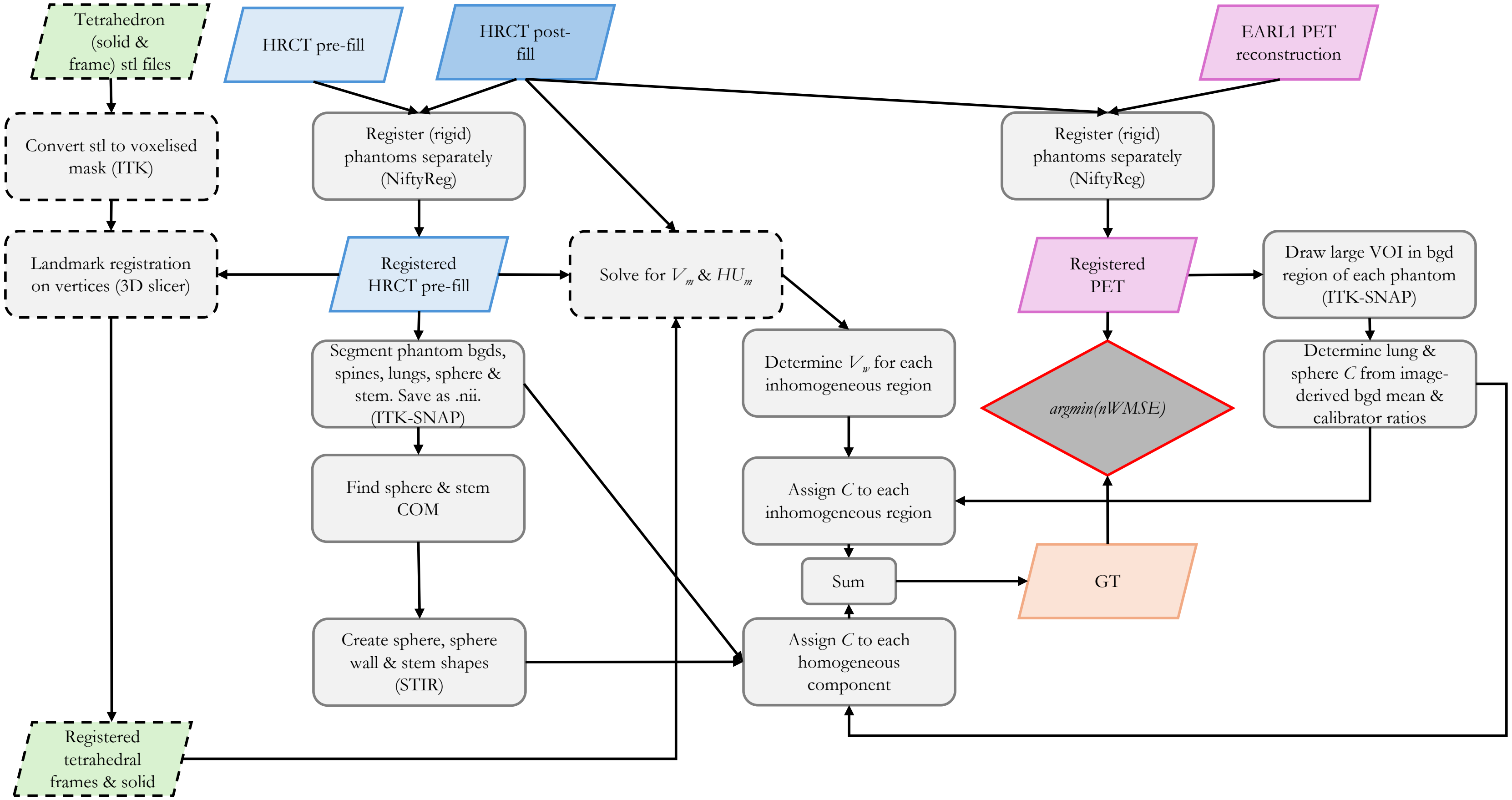


Figure C.1: Workflow for GT construction; boxes with dashed black boundaries indicate steps that were only conducted for the LDI phantom described in Chapter 5. C: activity concentration; V_m: material fraction; V_w: water fraction; HU_m: material CT number

Appendix D

Validation of voxel sub-sampling method for uncertainty estimation

The uncertainty on the determined kernel FWHM for measured phantom data was estimated by computing the standard deviation of h_{PVC} FWHMs derived by randomly sub-sampling half of the VOI voxels. Multiple voxel sub-sampling realisations were conducted and the standard deviation calculated across all realisations.

This approach was validated using a 1D array problem using the following methodology, as depicted in Figure D.1:

1. 1D array of N elements, where N = 100 – 700, was created. Each element randomly assigned an integer between 0 and 1000 (3.2% noise level, assuming a Poisson distribution).
2. Array smoothed with a Gaussian kernel ($\sigma = 1$) to form the ground truth (GT).
3. GT smoothed with a Gaussian kernel ($\sigma = 5$) to obtain "observed" data.
4. Poisson noise added to the "observed" data - "noisy observed".
5. GT smoothed with Gaussian kernel with varying σ (1 - 9; increments of 1) - "smoothed GT".
6. h_{PVC} computed by minimising the RMSE between smoothed GT and noisy observed.

$$RMSE = \sqrt{\sum_{v} (A_{obs,v} - A_{true,v})^2 / N_{sub}} \tag{D.1}$$

where $A_{obs,v}$ is the voxel value in the noisy observed array, $A_{true,v}$ is the voxel value in the smoothed GT array, and N_{sub} is the number of included elements.

7. Steps 4 - 6 repeated for 1000 noise realisations.
8. For a single noise realisation, a percentage (2.5 – 75%) of voxels were randomly sampled for h_{PVC} FWHM estimation. 1000 voxel sub-sampling realisations were conducted.
9. h_{PVC} was estimated by minimising the RMSE between sub-sampled smoothed GT and sub-sampled noisy observed.
10. Steps 8 - 9 was repeated for 1000 voxel sub-sampling realisations.
11. The standard deviation for the h_{PVC} FWHMs from all realisations was calculated for both the noise realisation and voxel sub-sampling methods.

1000 noise realisations returned the input kernel width ($\sigma = 5$) with zero standard deviation. The estimated kernel width from 1000 voxel sub-sampling realisations depended on the number of elements in the array and the sub-sample factor, Figure D.2.

As the number of elements in the array increased, the variance in the estimated kernel width decreased. As the sub-sample factor increased i.e., the number of voxels used to estimate the kernel decreased, for any given array size, the variance in the estimated kernel increased. At a sub-sample factor of 2 i.e., half of the voxels were sub-sampled, as was done for the measured data, the FWHM = 5.0 ± 0.0, for all array sizes investigated at 1,000 realisations.

The ratio of the two standard deviations was used to determine the scaling factor that should be applied to the voxel sub-sampling uncertainty. The scaling factor was computed as 1.4 for a sub-sample factor of 2. As the standard deviation from the sub-sampling realisations was shown to be an overestimate of the uncertainty that would have been estimated from noise realisations, it is this uncertainty that is quoted in Tables 4.3 and 4.4.

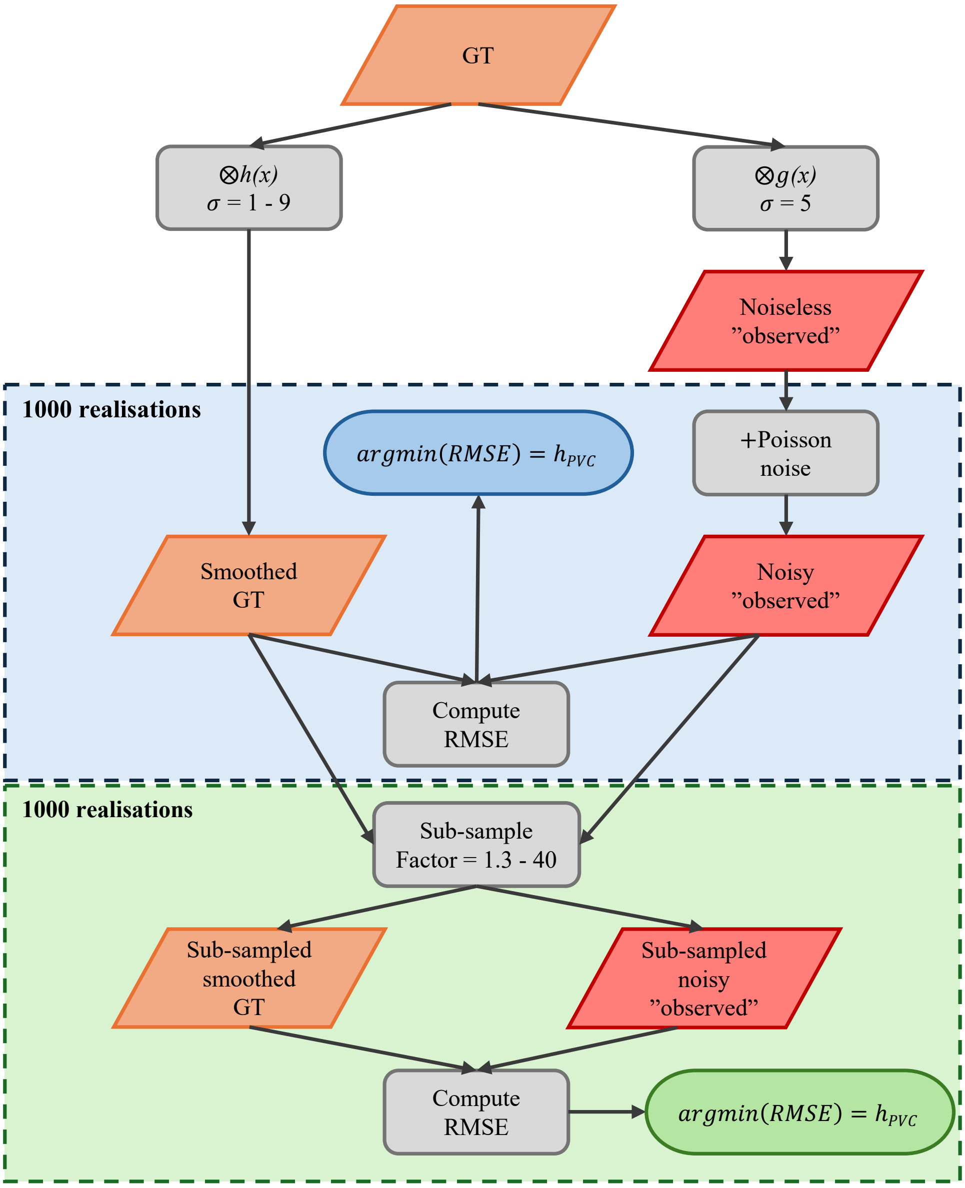


Figure D.1: Workflow for uncertainty estimation on h_{PVC} from 1000 sub-sampling realisations of a simulated 1D array (700 elements; values = 0 - 1000).

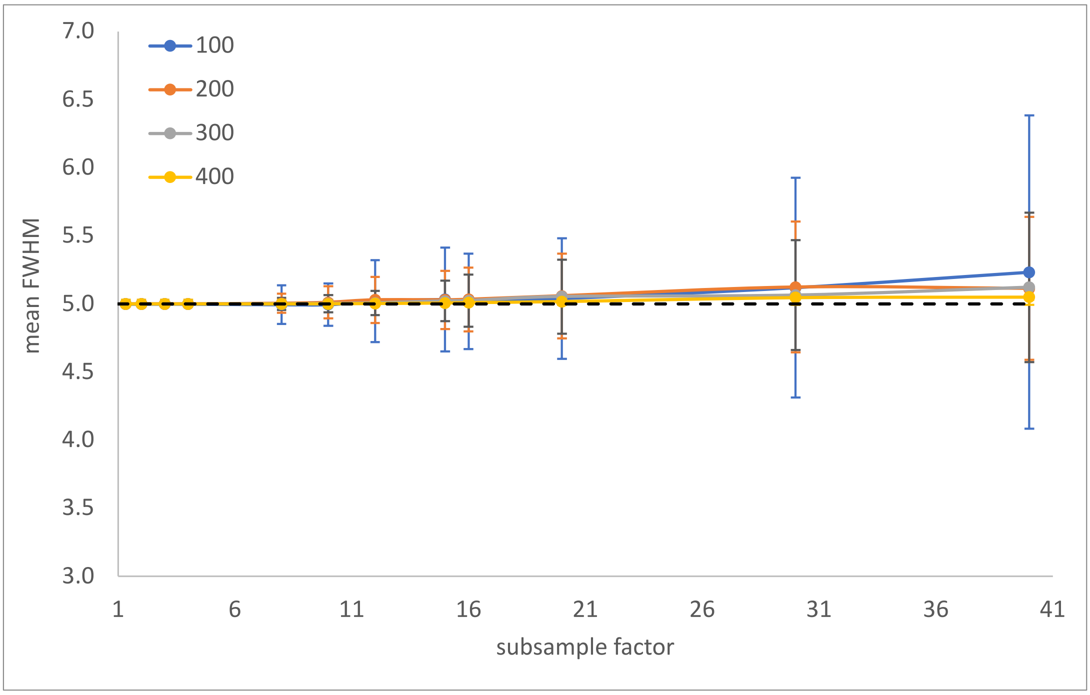


Figure D.2: Mean of estimated kernel width from 1000 voxel sub-sampling realisations at 3.2% noise level. Error bars indicate 1 standard deviation across the voxel sub-sampling realisations. Each colour represents the number of elements in the 1D array; 500+ elements resulted in a mean FWHM = 5.0 ± 0.0, and are therefore not depicted.

Appendix E

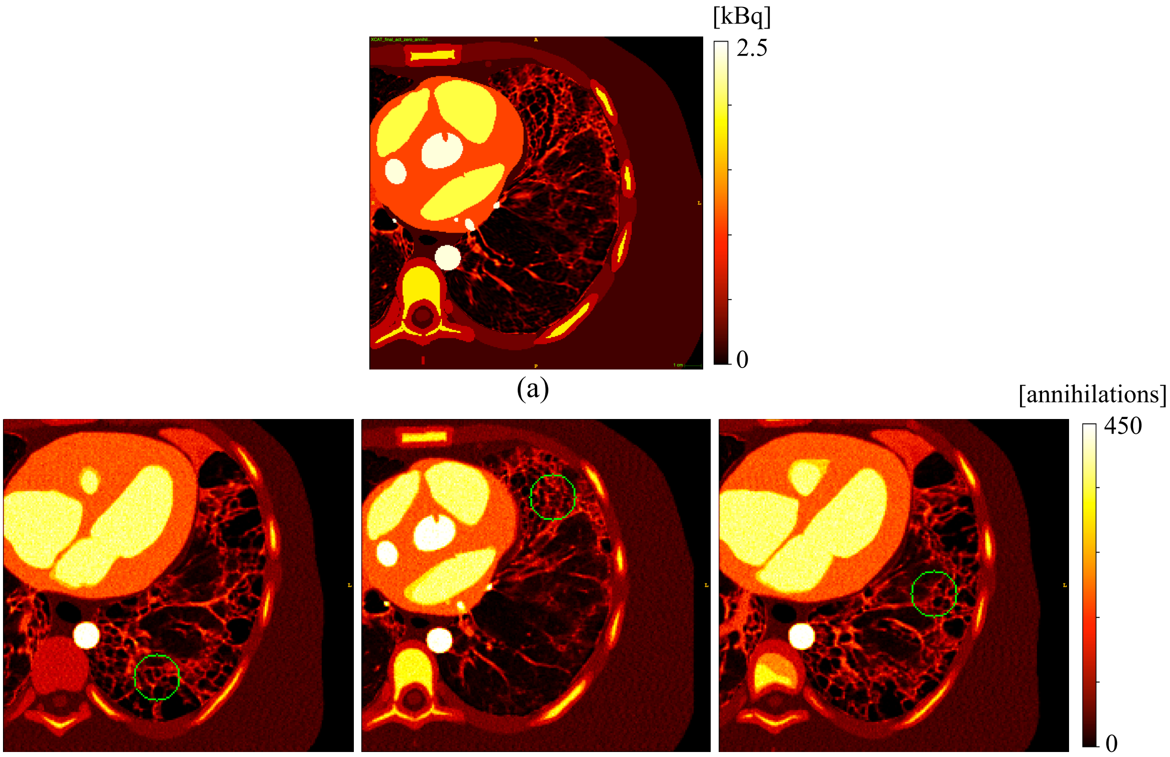
GATE simulations – patient data

To ensure that the kernel determined from the sphere lattice phantom in Section 5.4.2 is applicable to patient-realistic data, a patch of left lung of the modified XCAT phantom, Figure E.1(a), was simulated with the same GATE set-up as the sphere lattice phantom.

The emission that resulted in annihilations, and all annihilation events (642 x 10^6), were recorded and sampled into 3D images (voxel size: 1.000 x 1.000 x 1.635 mm^3). The kernel determined from the sphere phantom "reconstructed" annihilation image was used to smooth the downsampled XCAT GT activity map, which was scaled to the mean number of GATE emissions in a VOI in the homogeneous XCAT body. The RMSE was determined with respect to the smoothed (isotropic h_{PF} = 6 mm^3) "reconstructed" emission and annihilation images, in three 28 mm diameter VOIs positioned in regions of IPF, Figure E.1(b-d).

The optimal kernel width determined from the GATE simulation of the sphere lattice phantom (h_{PVC} = 6.0 mm^3) was used to smooth the downsampled XCAT GT emission. The RMSE in the three VOIs were in the range 0.81 – 1.00 % and 0.89 – 1.09 %, with respect to the smoothed emission map and smoothed annihilation map, respectively. When the optimal h_{PVC} was used to smooth the XCAT GT attenuation for AFC of the smoothed emission and annihilation maps, the AF-corrected RMSE in the three VOIs ranged from 0.84 – 1.10 % and 0.98 – 1.27 %, with respect to the GT AFC VOIs.

The XCAT "reconstructed" annihilation images had a slightly greater RMSE



(a)

(b) (c) (d)

Figure E.1: (a) Central axial slice of the unsmoothed modified XCAT activity map input to GATE (voxel size: 0.607 x 0.607 x 1.500 mm^3); (b-d) annihilation map formed from the output from GATE depicting the three VOIs used in green (voxel size: 1.000 x 1.000 x 1.635 mm^3.

than the “reconstructed” emission images (mean difference = 0.08 % of the GT activity map in each VOI). The RMSE in each VOI was comparable to those in the sphere lattice phantom.

These results provide confidence in the use of a kernel determined from a phantom containing 8 mm polystyrene beads with permeable barriers to create additional inhomogeneity in the PET image, to patient data acquired with ^{18}F.

Appendix F

STIR-GATE-Connection

The STIR-GATE-Connection is an open source research tool which aims to provide a simplified pipeline to quantitatively reconstruct GATE tomographic simulations with STIR. A public release of STIR-GATE-Connection, licensed under the Apache 2.0 License, can be downloaded at `https://github.com/UCL/STIR-GATE-Connection`. I assisted in the validation of this software prior to release, which resulted in co-authorship of a poster at IEEE NSS/MIC 2021 [171].

Specific contributions made to this project include checking alignment between scanners modelled in STIR and GATE and the effect misalignment has, not only on source positioning, but also normalisation.

Bibliography

[1] D.L. Chen, S. Ballout, L. Chen, J. Cheriyan, G. Choudhury, A.M. Denis-Bacelar, E. Emond, K. Erlandsson, M. Fisk, F. Fraioli, A.M. Groves, R.N. Gunn, J. Hatazawa, B.F. Holman, B.F. Hutton, H. Iida, S. Lee, W. MacNee, K. Matsunaga, D. Mohan, D. Parr, A. Rashidnasab, G. Rizzo, D. Subramanian, R. Tal-Singer, K. Thielemans, N. Tregay, E.J.R. van Beek, L. Vass, M.F. Vidal Merlo, J.W. Wellen, I. Wilkinson, F.J. Wilson, and T. Winkler. Consensus Recommendations on the Use of 18F-FDG PET/CT in Lung Disease. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*, 61(12):1701–1707, 2020.

[2] NHS England. Respiratory Disease. `https://www.england.nhs.uk/ourwork/clinical-policy/respiratory-disease/`. Accessed 11th December 2023.

[3] D.L. Chen, J. Cheriyan, E.R. Chilvers, G. Choudhury, C. Coello, M. Connell, M. Fisk, A.M. Groves, R.N. Gunn, B.F. Holman, B.F. Hutton, S. Lee, W. MacNee, D. Mohan, D. Parr, D. Subramanian, R. Tal-Singer, K. Thielemans, E.J.R. Van Beek, L. Vass, J.W. Wellen, I. Wilkinson, and F.J. Wilson. Quantification of lung PET images: Challenges and opportunities. *Journal of Nuclear Medicine*, 58(2):201–207, 2017.

[4] D.R. Subramanian, L. Jenkins, R. Edgar, N. Quraishi, R.A. Stockley, and D.G. Parr. Assessment of Pulmonary Neutrophilic Inflammation in Emphysema by Quantitative Positron Emission Tomography. *American Journal of Respiratory and Critical Care Medicine*, 186(11):1125–32, 2012.

[5] H.A. Jones, S. Sriskandan, A.M. Peters, N.B. Pride, T. Krausz, A.R. Boobis, and C. Haslett. Dissociation of neutrophil emigration and metabolic activity in lobar pneumonia and bronchiectasis. *European Respiratory Journal*, 10(4):795–803, 1997.

[6] R. Boellaard. Standards for PET image acquisition and quantitative data analysis. *Journal of Nuclear Medicine*, 50(SUPPL. 1):11–20, 2009.

[7] D.L. Chen and D.P. Schuster. Imaging pulmonary inflammation with positron emission tomography: A biomarker for drug development. *Molecular Pharmaceutics*, 3(5):488–495, 2006.

[8] A.M. Groves, T. Win, N.J. Screaton, M. Berovic, R. Endozo, H. Booth, I. Kayani, L.J. Menezes, J.C. Dickson, and P.J. Ell. Idiopathic pulmonary fibrosis and diffuse parenchymal lung disease: Implications from initial experience with 18F-FDG PET/CT. *Journal of Nuclear Medicine*, 50(4):538–545, 2009.

[9] T.C. Lee, A.M. Alessio, R.M. Miyaoka, and P.E. Kinahan. Morphology supporting function: attenuation correction for SPECT/CT, PET/CT, and PET/MR imaging. *The Quarterly Journal of Nuclear Medicine and Molecular Imaging*, 60(1):25–39, 2016.

[10] J. Lillington, L. Brusaferri, K. Kläser, K. Shmueli, R. Neji, B.F. Hutton, F. Fraioli, S. Arridge, M.J. Cardoso, S. Ourselin, K. Thielemans, and D. Atkinson. PET/MRI attenuation estimation in the lung: A review of past, present, and potential techniques. *Medical Physics*, 47(2):790–811, 2020.

[11] Laboratoire National Henri Becquerel. `http://www.lnhb.fr/nuclides/F-18.lara.txt`. Accessed 16 January 2025.

[12] S. Cherry and J. Sorenson. *Physics in Nuclear Medicine. 4th Edition.* Elsevier - OHCE, 2013.

[13] William W. Moses. Fundamental limits of spatial resolution in PET. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 648(SUPPL. 1):S236–S240, 2011.

[14] E.C. Emond, A.M. Groves, B.F. Hutton, and K. Thielemans. Effect of positron range on PET quantification in diseased and normal lungs. *Physics in Medicine and Biology*, 64(20), 2019.

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

[16] M. Conti. Why is TOF PET reconstruction a more robust method in the presence of inconsistent data? *Physics in Medicine and Biology*, 56(1):155–168, 11 2010.

[17] S. Il Kwon, A. Gola, A. Ferri, C. Piemonte, and S.R. Cherry. Bismuth germanate coupled to near ultraviolet silicon photomultipliers for time-of-flight PET. *Physics in Medicine and Biology*, 61(18), 2016.

[18] J. van Sluis, J. de Jong, J. Schaar, W. Noordzij, P. van Snick, R. Dierckx, R. Borra, A. Willemsen, and R. Boellaard. Performance characteristics of the digital biograph vision pet/ct system. *Journal of Nuclear Medicine*, 60:1031–1036, 2019.

[19] V. Bettinardi, L. Presotto, E. Rapisarda, M. Picchio, L. Gianolli, and M.C. Gilardi. Physical Performance of the new hybrid PETCT Discovery-690. *Medical Physics*, 38(10):5394–5411, 2011.

[20] R.L. Smith, L. Bartley, C. O'Callaghan, L. Haberska, and C. Marshall. Nema nu 2-2018 evaluation and image quality optimization of a new generation digital 32-cm axial field-of-view omni legend pet-ct using a genetic evolutionary algorithm. *Biomedical Physics and Engineering Express*, 10, 2024.

[21] R.J. Jaszczak and E.D. Hoffman. *Principles of nuclear medicine, 2nd edition; Positron emission tomography (PET): scatter and attenuation*. Philadelphia: Saunders, 1995.

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

[23] Hubbell, J.H. and Seltzer, S.M. Tables of x-ray mass attenuation coefficients and mass energy-absorption coefficients 1 keV to 20 meV for elements z = 1 to 92 and 48 additional substances of dosimetric interest. *U.S. Department of Energy Office of Scientific and Technical Information*, 1995.

[24] J.P.J. Carney, D.W. Townsend, V. Rappoport, and B. Bendriem. Method for transforming CT images for attenuation correction in PET/CT imaging. *Medical Physics*, 33(4):976–983, 2006.

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

[26] Y. Nakamoto, K.R. Zasadny, H. Minn, and R.L. Wahl. Reproducibility of common semi-quantitative parameters for evaluating lung cancer glucose metabolism with positron emission tomography using 2-deoxy-2-[18F]fluoro-D-glucose. *Molecular Imaging and Biology*, 4(2):171–178, 2002.

[27] S.A. Nehmeh and Y.E. Erdi. Respiratory Motion in Positron Emission Tomography/Computed Tomography: A Review. *Se*, 38, 2008.

[28] G.W. Goerres, C. Burger, E. Kamel, B. Seifert, A.H. Kaim, A. Buck, T.C. Buehler, and G.K. Von Schulthess. Respiration-induced Attenuation Artifact at PET/CT: Technical Considerations 1. *Radiology*, 226:906–910, 2003.

[29] D. Visvikis, C. Cheze-LeRest, D.C. Costa, J. Bomanji, S. Gacinovic, and P.J. Ell. Influence of OSEM and segmented attenuation correction in the calculation of standardised uptake values for [18F]FDG PT. *European Journal of Nuclear Medicine*, 28:1326–1335, 2001.

[30] B. Bai, A. Ruangma, R. Laforest, Y.C. Tai, and R.M. Leahy. Positron range modeling for statistical PET image reconstruction. *IEEE Nuclear Science Symposium Conference Record*, 4(2):2501–2505, 2003.

[31] P.E. Kinahan, A.M. Alessio, and J.A. Fessler. Dual Energy CT Attenuation Correction Methods for Quantitative Assessment of Response to Cancer Therapy with PET/CT Imaging. *Technology in Cancer Research and Treatment*, 5(4), 2006.

[32] D. Brasse, P.E. Kinahan, C. Lartizien, C. Comtat, M. Casey, and C. Michel. Correction methods for random coincidences in fully 3d whole-body pet: impact on data and image quality. *Journal of Nuclear Medicine*, 46(5):859–867, 2005.

[33] E.J. Hoffman, S. Huang, M.E. Phelps, and D.E. Kuhl. Quantitation in positron-emission computed tomography: 4. effect of accidental coincidences. *J Comput Assist Tomography*, 5, 1981.

[34] J.P. Oliver and M. Rafecas. Modelling random coincidences in positron emission tomography by using singles and prompts: A comparison study. *PLoS One*, 11(9), 2016.

[35] E.J. Hoffman, T.M. Guerrero, G. Germano, W.M. Digby, and M. Dahlbom. Pet system calibrations and corrections for quantitative and spatially accurate images. *IEEE Transactions on Nuclear Science*, 36(1):1108–1112, 1989.

[36] M. Defrise, D.W. Townsend, D. Bailey, A. Geissbuhler, C. Michel, and T. Jones. A normalization technique for 3d pet data. *Physics in medicine and biology*, 36(7):939–952, 1991.

[37] G. Germano and E.J. Hoffman. Investigation of count rate and deadtime characteristics of a high resolution pet system. *Journal of computer assisted tomography*, 12(5):836–846, 1988.

[38] S Tong and P.E. Alessio, A.M. abd Kinahan. Image reconstruction for PET/CT scanners: past achievements and future challenges. 2(5):529–545, 2011.

[39] L. A. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Transactions on Medical Imaging*, 1(2):113–122, 1982.

[40] H.M. Hudson and R.S. Larkin. Ordered Subsets of Projection Data. *IEEE transactions on medical imaging*, 13(4):601–609, 1994.

[41] S. Ross. Q.Clear. *GE Healthcare, White Paper*, page 1–9, 2014.

[42] A.M. Alessio, P.E. Kinahan, and T.K. Lewellen. Modeling and incorporation of system response functions in 3-D whole body PET. *IEEE Transactions on Medical Imaging*, 25(7):828–837, 2006.

[43] V.Y. Panin, F. Kehren, C. Michel, and M. Casey. Fully 3-d pet reconstruction with system matrix derived from point source measurements. *IEEE transactions on medical imaging*, 25(7):907–921, 2006.

[44] S. Tong, A. M. Alessio, and P. E. Kinahan. Noise and signal properties in psf-based fully 3d pet image reconstruction: an experimental evaluation. *Physics in medicine and biology*, 55(5):1453–1473, 2010.

[45] D.J. Kadrmas, M.E. Casey, N.F. Black, J.J. Hamill, V.Y. Panin, and Conti M. Experimental Comparison of Lesion Detectability for Four Fully-3D PET Reconstruction Schemes. *IEEE Transactions on Medical Imaging*, 28(4):523–534, 2009.

[46] A.M. Alessio. Application and Evaluation of a Measured Spatially Variant System Model for PET Reconstruction. *IEEE Transactions on Medical Imaging*, 29(3):938–949, 2010.

[47] M. Conti. Why is tof pet recontruction a more robust method in the presence of inconsitent data? *Phys. Med. Biol.*, 56:155–168, 2011.

[48] M. Conti. Effect of random reduction on signal-to-noise-ratio in tof pet. *IEEE Trans. Nucl. Sci.*, 53:1188–93, 2006.

[49] A. Mehranian, S.D. Wollenweber, M.D. Walker, K.M. Bradley, P.A. Fielding, M. Huellner, F. Kotasidis, K.-H. Su, R. Johnsen, F.P. Jansen, and D.R. McGowan. Deep learning-based time-of-flight (tof) image enhancement of non-tof pet scans. *EJNMMI*, 49, 2022.

[50] GE Healthcare. `https://www.gehealthcare.co.uk/products/molecular-imaging/pet-ct/omni-legend`. Accessed on 18 June 2024.

[51] NEMA. NEMA Standards Publication NU 2-2007 Performance Measurements of Positron Emission Tomographs. *National Electrical Manufactuers Association*, page 33, 2007.

[52] K. Gong, S.R. Cherry, and J. Qi. On the assessment of spatial resolution of PET systems with iterative image reconstruction. *Physics in Medicine and Biology*, 61(5):N193–N202, 2 2016.

[53] S.R. Meikle, M. Dahlbom, and S.R. Cherry. Attenuation correction using count-limited transmission data in positron emission tomography. *Journal of Nuclear Medicine*, 34(1):143–144, 1993.

[54] S. Yamagishi, K. Oguchi, K. Miwa, S. Kamitaki, K. Anraku, and T. Yamao. Performance characteristics of a new generation digital-bgo pet/ct system (omni legend 32) according to the nema standard. *Journal of Nuclear Medicine*, 64 (supplement 1):1365, 2023.

[55] R. Boellaard, R. Delgado-Bolton, W.J.G. Oyen, F. Giammarile, K. Tatsch, W. Eschner, F.J. Verzijlbergen, S.F. Barrington, L.C. Pike, W.A. Weber, S. Stroobants, D. Delbeke, K.J. Donohoe, S. Holbrook, M.M. Graham,

G. Testanera, O.S. Hoekstra, J. Zijlstra, E. Visser, C.J. Hoekstra, J. Pruim, A. Willemsen, B. Arends, J. Kotzerke, A. Bockisch, T. Beyer, A. Chiti, B.J. Krause, and European Associate of Nuclear Medicine. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. *EJNMMI*, 42:328–354, 2015.

[56] A. Kaalep, T. Sera, W. Oyen, B.J. Krause, A. Chiti, Y. Liu, and R. Boellaard. EANM/EARL FDG-PET/CT accreditation - summary results from the first 200 accredited imaging systems. *European Journal of Nuclear Medicine and Molecular Imaging*, 45(3):412–422, 3 2018.

[57] A. Kaalep, T. Sera, S. Rijnsdorp, M. Yaqub, A. Talsma, M.A. Lodge, and R. Boellaard. Feasibility of state of the art PET/CT systems performance harmonisation. *European Journal of Nuclear Medicine and Molecular Imaging*, 45(8):1344–1361, 7 2018.

[58] A. Kaalep, C.N. Burggraaff, S. Pieplenbosch, E.E. Verwer, T. Sera, J. Zijlstra, O.S. Hoekstra, D.E. Oprea-Lager, and R. Boellaard. Quantitative implications of the updated EARL 2019 PET–CT performance standards. *EJNMMI Physics*, 6(1), 12 2019.

[59] P.E. Kinahan, E.S. Perlman, J.J. Sunderland, R. Subramaniam, S.D. Wollenweber, T.G. Turkington, M.A. Lodge, R. Boellaard, N.A. Obuchowski, and R.L. Wahl. The QIBA profile for FDG PET/CT as an imaging biomarker measuring response to cancer therapy. *Radiology*, 294(2):647–657, 2020.

[60] G. Akamatsu, Y. Tsutsui, H. Daisaki, K. Mitsumoto, S. Baba, and M Sasaki. A review of harmonisation strategies for quantitative PET. *Annals of Nuclear Medicine*, 37:71–88, 2023.

[61] Y.J Tsai and C. Liu. Pitfalls on pet/ct due to artifacts and instrumentation. *Seminars in nuclear medicine*, 51(6):646–656, 2021.

[62] T. Sun and G.S.P. Mok. Techniques for respiration-induced artifacts reduc-

tions in thoracic PET/CT. *Quantitative Imaging in Medicine and Surgery*, 2(1), 2012.

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

[64] V. Cuplov, B.F. Holman, J. McClelland, M. Modat, B.F. Hutton, and K. Thielemans. Issues in quantification of registered respiratory gated PET/CT in the lung. *Physics in Medicine and Biology*, 63(1), 2018.

[65] B.F. Holman, V. Cuplov, B.F. Hutton, A.M. Groves, and K. Thielemans. The effect of respiratory induced density variations on non-TOF PET quantitation in the lung. *Physics in Medicine and Biology*, 61(8):3148–3163, 2016.

[66] K. Thielemans, E. Asma, R.M. Manjeshwar, A. Ganin, and J.T. Spinks. Image-based correction for mismatched attenuation in pet images. *IEEE Nuclear Science Symposium Conference Record*, pages 5292–5296, 2008.

[67] S. Ahn, L. Cheng, and R.M. Manjeshwar. Analysis of the effects of errors in attenuation maps on PET quantitation in TOF PET. *2014 IEEE Nuclear Science Symposium and Medical Imaging Conference, NSS/MIC 2014*, pages 1–4, 2016.

[68] E.C. Emond, A. Bousse, M. Machado, J.C. Porter, A.M. Groves, B.F. Hutton, and K. Thielemans. Effect of attenuation mismatches in time of flight PET reconstruction. *Physics in Medicine and Biology*, 65(8), 2020.

[69] M.D. Gilman, A.J. Fischman, V. Krishnasetty, E.F. Halpern, and S.L. Aquino. Optimal CT Breathing Protocol for Thoracic PET/CT Optimal CT Breathing Protocol for Combined Thoracic PET/CT. *AJR*, 187, 2006.

[70] M.J. Nyflot, T.C. Lee, A.M. Alessio, S.D. Wollenweber, C.W. Stearns, S.R. Bowen, and P.E. Kinahan. Impact of CT attenuation correction method on quantitative respiratory-correlated (4D) PET/CT imaging. *Medical Physics*, 42(1):110–120, 2015.

[71] T. Pan, O. Mawlawi, S.A. Nehmeh, Y.E. Erdi, D. Luo, H.H. Liu, R. Castillo, R. Mohan, Z. Liao, and H.A. Macapinlac. Attenuation Correction of PET Images with Respiration-Averaged CT Images in PET/CT. Technical report, The University of Texas M.D. Anderson Cancer Centre, 2005.

[72] C.C.A. Nagel, G. Bosmans, A.L.A.J. Dekker, M.C. Öllers, D.K.M. De Ruysscher, P. Lambin, A.W.H. Minken, N. Lang, and K.P. Schäfers. Phased attenuation correction in respiration correlated computed tomography/positron emitted tomography. *Medical Physics*, 33(6):1840–1847, 2006.

[73] B.F. Holman, V. Cuplov, L. Millner, R. Endozo, T.M. Maher, A.M. Groves, B.F. Hutton, and K. Thielemans. Improved quantitation and reproducibility in multi-PET/CT lung studies by combining CT information. *EJNMMI Physics*, 5(1):1–14, 2018.

[74] M. Soret, S.L. Bacharach, and I. Buvat. Partial-volume effect in PET tumor imaging. *Journal of Nuclear Medicine*, 48(6):932–945, 2007.

[75] K. Erlandsson, I. Buvat, P.H. Pretorius, B.A. Thomas, and B.F. Hutton. 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), 2012.

[76] C. Comtat, F.C. Sureau, M. Sibomana, I.K. Hong, N. Sjoholm, and R. Trebossen. Image based resolution modeling for the hrrt osem reconstructions software. In *2008 IEEE Nuclear Science Symposium Conference Record*, pages 4120–4123, 2008.

[77] A.J. Reader, P.J. Julyan, H. Williams, D.L. Hastings, and J. Zweit. EM Algorithm Resolution Modeling by Image-Space Convolution for PET Recon-

struction. *IEEE Nuclear Science Symposium and Medical Imaging Conference*, 2(5):1221–1225, 2002.

[78] J. Nuyts. Unconstrained image reconstruction with resolution modelling does not have a unique solution. *EJNMMI Physics*, 1, 2014.

[79] A. Rahmim, J. Qi, and V. Sossi. Resolution modeling in PET imaging: Theory, practice, benefits, and pitfalls. *Medical Physics*, 40(6):1–15, 2013.

[80] B. Lipinski, H. Herzog, E. Rota Kops, W. Oberschelp, and H.W. Müller-Gärtner. Expectation maximization reconstruction of positron emission tomography images using anatomical magnetic resonance information. *IEEE Transactions on Medical Imaging*, 16(2):129–136, 1997.

[81] Claude Comtat, Paul E. Kinahan, Jeffrey A. Fessler, Thomas Beyer, David W. Townsend, Michel Defrise, and Christian Michel. Clinically feasible reconstruction of 3D whole-body PET/CT data using blurred anatomical labels. *Physics in Medicine and Biology*, 47(1):1–20, 2002.

[82] J.E. Bowsher, H. Yuan, L.W. Hedlund, T.G. Turkington, G. Akabani, A. Badea, W.C. Kurylo, C.T. Wheeler, G.P. Cofer, M.W. Dewhirst, and G.A. Johnson. Utilizing MRI information to estimate F18-FDG distributions in rat flank tumors. *IEEE Nuclear Science Symposium Conference Record*, 4(C):2488–2492, 2004.

[83] B.A. Thomas, V. Cuplov, A. Bousse, A. Mendes, K. Thielemans, B.F. Hutton, and K. Erlandsson. PETPVC: A toolbox for performing partial volume correction techniques in positron emission tomography. *Physics in Medicine and Biology*, 61(22):7975–7993, 2016.

[84] L.B. Lucy. An iterative technique for the rectification of observed distributions. *The Astronomical Journal*, 79(6), 1974.

[85] W.H. Richardson. Bayesian-Based Iterative Method of Image Restoration. *J. Opt. Soc. Am.*, 62(1):55–59, 1972.

[86] Von P.H. van Cittert. Zum Einflu˜ der Spaltbreite auf die Intensit˜tsverteilung in Spektrallinien. II. *Zeitschrift fur Medizinische Physik*, 69(5-6):298–308, 1931.

[87] A.S. Kirov, J.Z. Piao, and C.R. Schmidtlein. Partial volume effect correction in PET using regularized iterative deconvolution with variance control based on local topology. *Physics in Medicine and Biology*, 53(10):2577–2591, 2008.

[88] N. Boussion, C. Cheze, L. Rest, M. Hatt, and D. Visvikis. Incorporation of wavelet-based denoising in iterative deconvolution for partial volume correction in whole-body PET imaging. *European Journal of Nuclear Medicine and Molecular Imaging*, 36:1064–1075, 2009.

[89] H. Vesselle, E. Turcotte, L. Wiens, R. Schmidt, J.E. Takasugi, T. Lalani, E. Valliè, and D.E. Wood. Relationship between Non-Small Cell Lung Cancer Fluorodeoxyglucose Uptake at Positron Emission Tomography and Surgical Stage with Relevance to Patient Prognosis. *Clinical Cancer Research*, 10:4709–4716, 2004.

[90] M. Hickson, M. Yun, A. Matthies, H. Zhuang, L.-E. Adam, L. Lacorte, and A. Alavi. Use of a corrected standardized uptake value based on the lesion size on CT permits accurate characterization of lung nodules on FDG-PET. *European Journal of Nuclear Medicine*, 29:1639–1647, 2002.

[91] Y. Menda, D.L. Bushnell, M.T. Madsen, K. McLaughlin, D. Kahn, and K.H. Kernstine. Evaluation of various corrections to the standardized uptake value for diagnosis of pulmonary malignancy. *Nuclear medicine communications*, 22(10):1077–1081, 2001.

[92] H. Vesselle, A. Salskov, E. Turcotte, L. Wiens, R. Schmidt, C.D. Jordan, E. Vallières, and D.E. Wood. Relationship between non-small cell lung cancer FDG uptake at PET, tumor histology, and Ki-67 proliferation index. *Journal of Thoracic Oncology*, 3(9):971–978, 2008.

[93] L. Thomas, T. Schultz, V. Prokic, M. Guckenberger, S. Tanadini-Lang, M. Hohberg, M. Wild, A. Drzezga, and R.A. Bundschuh. 4D-CT-based motion correction of PET images using 3D iterative deconvolution. *Oncotarget*, 2019.

[94] T. Lambrou, A.M. Groves, K. Erlandsson, N. Screaton, R. Endozo, T. Win, J.C. Porter, and B.F. Hutton. The importance of correction for tissue fraction effects in lung PET: Preliminary findings. *European Journal of Nuclear Medicine and Molecular Imaging*, 38(12):2238–2246, 2011.

[95] T. Win, T. Lambrou, B.F. Hutton, I. Kayani, N.J. Screaton, J.C. Porter, T.M. Maher, R. Endozo, R.I. Shortman, P. Lukey, and A.M. Groves. 18F-Fluorodeoxyglucose positron emission tomography pulmonary imaging in idiopathic pulmonary fibrosis is reproducible: Implications for future clinical trials. *European Journal of Nuclear Medicine and Molecular Imaging*, 39(3):521–528, 2012.

[96] T. Win, B.A. Thomas, T. Lambrou, B.F. Hutton, N.J. Screaton, J.C. Porter, T.M. Maher, R. Endozo, R.I. Shortman, A. Afaq, P. Lukey, P.J. Ell, and A.M. Groves. Areas of normal pulmonary parenchyma on HRCT exhibit increased FDG PET signal in IPF patients. *European Journal of Nuclear Medicine and Molecular Imaging*, 41(2):337–342, 2014.

[97] D.A. Torigian, V. Dam, X. Chen, B. Saboury, J.K. Udupa, A. Rashid, S. Moghadam-Kia, and A. Alavi. In vivo quantification of pulmonary inflammation in relation to emphysema severity via partial volume corrected18F-FDG-PET using computer-assisted analysis of diagnostic chest CT. *Hellenic Journal of Nuclear Medicine*, 16(1):12–18, 2013.

[98] A. Castiaux, G. Van Simaeys, S. Goldman, and B. Bondue. Assessment of 18F-FDG uptake in idiopathic pulmonary fibrosis: influence of lung density changes. *European Journal of Hybrid Imaging*, 2(27), 2018.

[99] P. Wollmer, C.G. Rhodes, and J. Hughes. Regional extravascular density and

fractional blood volume of the lung in interstitial disease. *Thorax*, 39:286–293, 1984.

[100] C.G. Rhodes and J.M.B. Hughes. Pulmonary studies using positron emission tomography. *European Respiratory Journal*, 8(6):1001–1017, 1995.

[101] B.F. Holman, V. Cuplov, L. Millner, B.F. Hutton, T.M. Maher, A.M. Groves, and K. Thielemans. Improved correction for the tissue fraction effect in lung PET/CT imaging. *Physics in Medicine and Biology*, 60(18):7387–7402, 2015.

[102] C. Coello, M. Fisk, D. Mohan, F.J. Wilson, A.P. Brown, M.I. Polkey, I. Wilkinson, R. Tal-Singer, P.S. Murphy, J. Cheriyan, and R.N. Gunn. Quantitative analysis of dynamic 18 F-FDG PET/CT for measurement of lung inflammation. *EJNMMI Research*, 7(47), 2017.

[103] L. Fu and J. Qi. A residual correction method for high-resolution PET reconstruction with application to on-the-fly Monte Carlo based model of positron range. *Medical Physics*, 37(2):704–713, 2010.

[104] A.M. Alessio and L. MacDonald. Spatially Variant Positron Range Modeling Derived from CT for PET Image Reconstruction. *IEEE Nuclear Science Symposium Conference Record*, 2008.

[105] A. Rahmim, J. Tang, M.A. Lodge, S. Lashkari, R.A. Mohammad, and F.M. Bengel. Resolution modeled PET image reconstruction incorporating space-variance of positron range: Rubidium-82 cardiac PET imaging. *IEEE Nuclear Science Symposium Conference Record*, pages 3643–3650, 2008.

[106] L. Szirmay-Kalos, M. Magdics, B. Tóth, B. Csébfalvi, T. Umenhoffer, J. Lantos, and G. Patay. Fast positron range calculation in heterogeneous media for 3d pet reconstruction. In *2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC)*, pages 2150–2155, 2012.

[107] J. Cal-González, J.L. Herraiz, S. España, P.M.G. Corzo, J.J. Vaquero, M. Desco, and J.M. Udias. Positron range estimations with PeneloPET. *Physics in Medicine and Biology*, 58(15):5127–5152, 2013.

[108] W. Lehnert. Analytical positron range modelling in heterogeneous media for PET Monte Carlo simulation. *Physics in Medicine and Biology*, 57:4075, 2012.

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

[110] R. Boellaard, A. Van Lingen, and A.A. Lammertsma. 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, 2001.

[111] K. Thielemans, C. Tsoumpas, S. Mustafovic, T. Beisel, P. Aguiar, N. Dikaios, and M.W. Jacobson. STIR: Software for tomographic image reconstruction release 2. *Physics in Medicine and Biology*, 57(4):867–883, 2012.

[112] E. Ovtchinnikov, R. Brown, C. Kolbitsch, E. Pasca, C. da Costa-Luis, A.G. Gillman, B.A. Thomas, N. Efthimiou, J. Mayer, P. Wadhwa, M.J. Ehrhardt, S. Ellis, J.S. Jørgensen, J. Matthews, C. Prieto, A.J. Reader, C. Tsoumpas, M. Turner, D. Atkinson, and K. Thielemans. SIRF: Synergistic Image Reconstruction Framework. *Computer Physics Communications*, 249:39–41, 2020.

[113] L. Sørensen, S.B. Shaker, and M. De Bruijne. Quantitative analysis of pulmonary emphysema using local binary patterns. *IEEE Transactions on Medical Imaging*, 29(2):559–569, 2010.

[114] J.A Stamos, W.L. Rogers, N.H. Clinthorne, and K.F Koral. Object-dependent performance comparison of two iterative reconstruction algorithms. *IEEE Transactions on Nuclear Science*, 35(1):611–614, 1988.

[115] P. Virtanen, R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S.J. van der Walt, M. Brett, J. Wilson, K. Jarrod Millman, N. Mayorov, A.R. Nelson, E. Jones, R. Kern, E. Larson, C.J. Carey, I. Polat, Y. Feng, E.W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E.A. Quintero, C.R. Harris, A.M. Archibald, A.H. Ribeiro, F. Pedregosa, and P. van Mulbregt. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17, 2020.

[116] B.F. Hutton, H.M. Hudson, and F.J. Beekman. A clinical perspective of accelerated statistical reconstruction. *European Journal of Nuclear Medicine*, 24(7):797–808, 1997.

[117] M. Defrise, P.E. Kinahan, and C.J. Michel. Image Reconstruction Algorithms in PET. In D.L. Bailey, D.W. Townsend, P.E. Valk, and M.N. Maisey, editors, *Positron Emission Tomography: Basic Sciences*, pages 63–91. Springer London, London, 2005.

[118] H.W. Müller-Gartner, J.M. Links, J.L. Prince, R.N. Bryan, E. McVeigh, J.P. Leal, C. Davatzikos, and J.J. Frost. Measurement of Radiotracer Concentration in Brain Gray Matter Using Positron Emission Tomography: MRI-Based Correction for Partial Volume Effects. *J Cereb Blood Flow Metab.*, 12(4):571–83, 1992.

[119] A. Joshi, R.A. Koeppe, and J.A. Fessler. Reducing between scanner differences in multi-center PET studies. *NeuroImage*, 46:154–159, 2009.

[120] F. Leek, C. Anderson, A.P. Robinson, R.M. Moss, J.C. Porter, H.S. Garthwaite, A.M. Groves, B.F Hutton, and K. Thielemans. Optimisation of the air fraction correction for lung PET/CT: addressing resolution mismatch. *EJNMMI Physics*, 10, 2023.

[121] W.P. Segars, G. Sturgeon, S. Mendonca, J. Grimes, and B.M.W. Tsui. 4D

XCAT phantom for multimodality imaging research. *Medical Physics*, 37(9):4902–4915, 2010.

[122] F. Leek, A.P. Robinson, R.M. Moss, F.J. Wilson, B.F. Hutton, and K. Thielemans. Air fraction correction optimisation in pet imaging of lung disease. In *2020 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, pages 1–4, 2020.

[123] S. Surti, J.S. Karp, L.A. Popescu, M.E. Daube-Witherspoon, and M. Werner. Investigation of time-of-flight benefit for fully 3-d pet. *IEEE Trans Med Imaging*, 25:529–538, 2006.

[124] G. Schramm. PARALLELPROJ – An open-source framework for fast calculation of projections in tomography. *Frontiers*, 2023 in press.

[125] DataSpectrum. Elliptical lung-spine body phantom lid insert™. `https://www.spect.com/wp-content/uploads/2020/04/Elliptical-Lung-Spine-Phantom-Lid-Insert.pdf`, 2020. Accessed 19 Aug 2023.

[126] P.A. Yushkevich, J. Piven, H.C. Hazlett, R.G. Smith, S. Ho, J.C. Gee, and G. Gerig. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *NeuroImage*, 31(3):1116–1128, 2006.

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

[128] J.I. Gear, M.G. Cox, J. Gustafsson, K. Sjögreen Gleisner, I. Murray, G. Glatting, M. Konijnenberg, and G.D. Flux. EANM practical guidance on uncertainty analysis for molecular radiotherapy absorbed dose calculations. *European Journal of Nuclear Medicine and Molecular Imaging*, 45:2456–2474, 2018.

[129] E.E. Verwer, S.S.V. Golla, A. Kaalep, M. Lubberink, F.H.P. van Velden, V. Bettinardi, M. Yaqub, T. Sera, S. Rijnsdorp, A.A. Lammertsma, and R. Boellaard. Harmonisation of PET/CT contrast recovery performance for brain studies. *European Journal of Nuclear Medicine and Molecular Imaging*, 48:2856–2870, 2021.

[130] C.K. McGarry, L.J. Grattan, A.M. Ivory, F. Leek, G.P. Liney, Y. Liu, P. Miloro, R. Rai, A. Robinson, A.J. Shih, B. Zeqiri, and C.H. Clark. Tissue mimicking materials for imaging and therapy phantoms: a review. *Physics in Medicine and Biology*, 2020.

[131] F. Leek, C. Anderson, A.P. Robinson, R.M. Moss, B.F. Hutton, and K. Thielemans. Preliminary results from a novel phantom design for harmonisation of lung pet/ct studies. In *Journal of Nuclear Medicine 65 (suppl 2) 241445*, 2024.

[132] F. Leek, C. Anderson, A.P. Robinson, R.M. Moss, B.F. Hutton, and K. Thielemans. Phantom design to enable harmonisation of quantitative pet imaging for lung disease. In *2024 IEEE Nuclear Science Symposium (NSS), Medical Imaging Conference (MIC) and Room Temperature Semiconductor Detector Conference (RTSD), Tampa, FL, USA*, 2024.

[133] W.A. Kalender, H. Fichte, W. Bautz, A. Zwick, R. Rienmüller, J. Behr, J.V. Engelshoven, R. Lamers, and P. Vock. Reference Values for Lung Density and Structure Measured by Quantitative CT. In *Advances in CT III*, 1994.

[134] V. Filippou and C. Tsoumpas. Recent advances on the development of phantoms using 3D printing for imaging with CT, MRI, PET, SPECT, and ultrasound. *Medical Physics*, 45(9):e740–e760, 2018.

[135] R. Tino, A. Yeo, M. Leary, M. Brandt, and T. Kron. A Systematic Review on 3D-Printed Imaging and Dosimetry Phantoms in Radiation Therapy. *Technology in Cancer Research and Treatment*, 18:1–14, 2019.

[136] J.I. Gear, C. Cummings, A.J. Craig, A. Divoli, C.D.C. Long, M. Tapner, and G.D. Flux. Abdo-Man: a 3D-printed anthropomorphic phantom for validating quantitative SIRT. *EJNMMI Physics*, 3(17), 2016.

[137] J.O. Heikkinen, J.T. Kuikka, and P.J. Rautio. Interdepartmental audit with an anatomically realistic lung phantom. *Journal of Nuclear Medicine Technology*, 34:34–42, 2006.

[138] A.L. Lehnert, W.C.J. Hunter, W.A. McDougald, R.L. Harrison, T.K. Lewellen, H.J. Vesselle, and R.S. Miyaoka. Development and testing of SPECT/CT lung phantoms made from expanding polyurethane foam. *Medical Physics*, 46(12):5593–5601, 2019.

[139] A.P. Robinson, J. Tipping, D.M. Cullen, D. Hamilton, R. Brown, A. Flynn, C. Oldfield, E. Page, E. Price, A. Smith, and R. Snee. EJNMMI Physics Organ-specific SPECT activity calibration using 3D printed phantoms for molecular radiotherapy dosimetry. *Physics*, 3:12, 2016.

[140] F.P. DiFilippo, J.P. Price, D.N. Kelsch, and R.F. Muzic. Porous phantoms for PET and SPECT performance evaluation and quality assurance. *Medical Physics*, 31(5):1183–1194, 2004.

[141] D.C. Hunt, H. Easton, and C.B. Caldwell. Design and construction of a quality control phantom for SPECT and PET imaging. *Medical Physics*, 36(12):5404–5411, 2009.

[142] S.D. Wollenweber, A.M. Alessio, and P.E. Kinahan. A phantom design for assessment of detectability using a lumpy background and 3D-printed features. *2015 IEEE Nuclear Science Symposium and Medical Imaging Conference, NSS/MIC 2015*, 75:3–5, 2016.

[143] J. Madamesila, P. Mcgeachy, Villarreal B.J.E., and R. Khan. Characterizing 3D printing in the fabrication of variable density phantoms for quality assurance of radiotherapy. *Physica Medica*, 32:242–247, 2016.

[144] O.L. Dancewicz, S.R. Sylvander, T.S. Markwell, S.B. Crowe, and J.V. Trapp. Radiological properties of 3D printed materials in kilovoltage and megavoltage photon beams. *Phys Med*, 2017.

[145] J. Solc, T. Vrba, and L. Burianova. Tissue-equivalence of 3D-printed plastics for medical phantoms in radiology. *Journal of Instrumentation*, 13(9), 2018.

[146] D.F. Craft, S.F. Kry, P. Balter, M. Salehpour, W. Woodward, and R.M. Howell. Material matters: Analysis of density uncertainty in 3D printing and its consequences for radiation oncology. *Medical Physics*, 45(4):1614–1621, 2018.

[147] S.A. Larsson, C. Jonsson, M. Pagani, L. Johansson, and H. Jacobsson. A novel phantom design for emission tomography enabling scatter- and attenuation-'free' single-photon emission tomography imaging. *European Journal of Nuclear Medicine*, 27(2):131–139, 2000.

[148] I.S. Negus, R.B. Holmes, K.C. Jordan, D.A. Nash, G.C. Thorne, and M. Saunders. Technical Note: Development of a 3D printed subresolution sandwich phantom for validation of brain SPECT analysis. *Medical Physics*, 43(9):5020–5027, 2016.

[149] B. Berthon, C. Marshall, R. Holmes, and E. Spezi. A novel phantom technique for evaluating the performance of PET auto-segmentation methods in delineating heterogeneous and irregular lesions. *EJNMMI Physics*, 2(1):1–17, 2015.

[150] M.A. Miller and G.D. Hutchins. Development of anatomically realistic PET and PET/CT phantoms with rapid prototyping technology. *IEEE Nuclear Science Symposium Conference Record*, 6:4252–4257, 2007.

[151] T. Läppchen, L.P. Meier, M. Fürstner, G.A. Prenosil, T. Krause, A. Rominger, B. Klaeser, and M. Hentschel. 3D printing of radioactive phantoms for nuclear medicine imaging. *EJNMMI Physics*, 7(1), 2020.

[152] J.I. Gear, C. Cummings, J. Sullivan, N. Cooper-Rayner, P. Downs, and I. Murray. Radioactive 3D printing for the production of molecular imaging phantoms. *Physics in Medicine and Biology*, 65(17), 2020.

[153] D. Gillett, D. Marsden, S. Ballout, B. Attili, N. Bird, S. Heard, M. Gurnell, I.A. Mendichovszky, and L. Aloj. 3D printing 18 F radioactive phantoms for PET imaging. *EJNMMI Physics*, 2021.

[154] D.M. Wirth, A. Jaquez, S. Gandarilla, J.D. Hochberg, D.C. Church, and J.K. Pokorski. Highly Expandable Foam for Lithographic 3D Printing. *ACS Applied Materials and Interfaces*, 2020.

[155] G.J. Kemerink, M.G.W. Visser, R. Franssen, E. Beijer, M. Zamburlini, S.G.E.A. Halders, B. Brans, F.M. Mottaghy, and G.J.J. Teule. Effect of the positron range of 18F, 68Ga and 124I on PET/CT in lung-equivalent materials. *European Journal of Nuclear Medicine and Molecular Imaging*, 38(5):940–948, 2011.

[156] H. Alva-Sánchez, C. Quintana-Bautista, A. Martínez-Dávalos, M.A. Ávila-Rodríguez, and M. Rodríguez-Villafuerte. Positron range in tissue-equivalent materials: Experimental microPET studies. *Physics in Medicine and Biology*, 61(17):6307–6321, 2016.

[157] Fastener Kingdom. 2mm to 40mm precision solid pp plastic balls polypropylene sealing rolling beads. `https://www.ebay.co.uk/itm/324698488572?hash=item4b9989befc:g:BW4AAOSwifVg3Vcf`. Accessed 23 Sept 2023.

[158] British Plastics Federation. Expanded polystyrene (eps). `https://www.bpf.co.uk/plastipedia/polymers/expanded-and-extruded-polystyrene-eps-xps.aspx`. Accessed 23 Sept 2023.

[159] E-S Jang and C-W Kang. Porosity analysis of three types of balsa (ochroma pyramidale) wood depending on density. *Journal of Wood Science*, 2022.

[160] PORAVER® GMBH. Technical data poraver expanded glass, https://poraver.com/en/poraver. `https://poraver.com/en/poraver/`. Accessed 13 Feb 2021.

[161] KingspanTM. Tarecpir technical data hd 120, cfc/hcfc free rigid polyisocyanurate insulation. Technical report, 2015.

[162] B.F Holman. *Improving Quantification of PET/CT Biomarkers for Evaluation of Disease Progression and Treatment Effectiveness in Pulmonary Fibrosis*. PhD thesis, University College London, 2016.

[163] Y. Wu, Z. Fan, and Y. Lu. Bulk and interior packing densities of random close packing of hard spheres. *Journal of Materials Science*, 38(9):2019–2025, 2003.

[164] S. Jan, D. Benoit, E. Becheva, H.-H. Lin, K.-S. Chuang, Y.-S. Lin, C. Pet, S. España, J.L. Herraiz, E. Vicente, D. Lazaro, I. Buvat, and G. Loudos. GATE : a simulation toolkit for PET and SPECT. *Physics in Medicine and Biology*, 49:4543–4561, 2004.

[165] Nist standard reference database 126 table 2: X-ray mass attenuation coefficients. `https://physics.nist.gov/PhysRefData/XrayMassCoef/tab2.html`. Accessed 01 March 2024.

[166] A. Buis. The atmosphere: Getting a handle on carbon dioxide; sizing up humanity's impacts on earth's changing atmosphere: A five-part series. `https://climate.nasa.gov/news/2915/the-atmosphere-getting-a-handle-on-carbon-dioxide/`. Accessed 01 March 2024.

[167] M. Modat, G.R. Ridgway, Z.A. Taylor, M. Lehmann, J. Barnes, D.J. Hawkes, N.C. Fox, and S. Ourselin. Fast free-form deformation using graphics processing units. *Computer Methods And Programs In Biomedicine*, 98:278–284, 2010.

[168] S. Tong, A.M. Alessio, K. Thielemans, C. Stearns, S. Ross, and P.E. Kinahan. Properties and mitigation of edge artifacts in psf-based pet reconstruction. *IEEE Transactions on Nuclear Science*, 58(5):2264–2275, 2011.

[169] B. Avants, N.J. Tustison, and G. Song. Advanced Normalization Tools: V1.0. *The Insight Journal*, pages 1–35, 2022.

[170] S. Karunanithi, R. Soundararajan, P. Sharma, N. Naswa, C. Bal, and R. Kumar. Spectrum of Physiologic and Pathologic Skeletal Muscle 18F-FDG Uptake on PET/CT. *American Journal of Roentgenology*, 205, 2015.

[171] R. Twyman, L. Brusaferri, E.C. Emond, F. Leek, B.F. Hutton, V. Cuplov, and K. Thielemans. A demonstration of stir-gate-connection. In *2021 IEEE Nuclear Science Symposium and Medical Imaging Conference*, 2021.