DYNAMIC FUNCTIONAL IMAGING OF THE RESPONSE OF LUNG TUMOURS TO A POTENTIAL HYPOXIA MODIFIER

Daniel McGowan

St Hugh's College

A thesis submitted for the degree of
Doctor of Philosophy at the University of Oxford

Hilary Term 2016

NHS
National Institute for Health Research

Abstract

Dynamic Functional Imaging of the Response of Lung Tumours to a Potential Hypoxia Modifier

Daniel McGowan, St Hugh's College

Submitted for the degree of Doctor of Philosophy
at the University of Oxford, HT 2016

In this thesis, the kinetics of fluoromisonidazole (FMISO) uptake by lung tumours in patients with advanced stage non-small cell lung cancer (NSCLC) is analysed. The uptake data was extracted from dynamic positron emission tomography (PET) images collected in the course of a clinical trial designed to investigate the impact of Buparlisib, a potentially hypoxia-modifying drug, on perfusion and hypoxia in NSCLC patients. The trial design is a 3+3 dose escalation study exploring the effect of daily doses of 50 mg (Cohort 1), 80 mg (Cohort 2), and 100 mg (Cohort 3) Buparlisib.

Many trials combining drug and RT fail in the clinical setting despite yielding promising pre-clinical data: the work described here uses and develops kinetics analysis techniques that can be applied to proof-of-principle dynamic PET imaging studies. For whole tumour FMISO kinetics analysis it has been shown that an irreversible three-compartment model is optimal, whereas for voxel-by-voxel analysis an irreversible two-compartment model is optimal. A clustering method is also presented to distil the information from the noisy voxel-by-voxel kinetics data into the minimum number of clusters that adequately describe the data.

Analysis of both PET kinetics and perfusion CT has shown significant decreases in the rate of tracer flow from blood into the tumour post-Buparlisib (in Cohorts 2 and 3). Voxel-by-voxel analysis found that this occurred throughout the tumour volume. The volume of blood within the tumour was shown to increase significantly after Buparlisib according to the PET kinetics analysis, whereas analysis of perfusion CT data showed a decrease post-Buparlisib. For Cohorts 2 and 3, significant decreases were seen in tumour volumes hypothesised to be hypoxic (tumour-to-blood ratio greater than 1.4) after administration of Buparlisib. The estimated FMISO intracellular binding rate-constant (associated with cellular oxygen concentration) according to the voxel-by-voxel kinetics analysis was found to significantly decrease at the tumour edge and significantly increase in the tumour centre, with an overall decrease for Cohort 3.

Dedicated to my wife

whose support was essential for this thesis to come to fruition

Acknowledgements

I am extremely grateful to my supervisors John Fenwick and Geoff Higgins, for their support, patience and invaluable guidance throughout this D.Phil and also for their instrumental assistance during the prior fellowship application. I am also thankful to my other supervisor Fergus Gleeson for all his advice and support along the way.

I have been fortunate to work alongside respected colleagues in the Buparlisib trial, and am indebted to Ruth for her many patient hours of outlining. I would also like to record my thanks to all those involved in the Buparlisib trial, in particular Naomi, Claire, Lainey, Finn, Pradeep, and all of the willing patients. I am also grateful to PET radiographers John and Paul for allowing me access to the scanner at all manner of hours.

Bartek and Jesper at the engineering department provided vital assistance with the perfusion data, and my colleagues at Hermes Medical and GE Healthcare have been generous with their advice on software and hardware. I would also like to thank my colleagues in the Radiation Research Institute, especially Tracy, Dan, Sara and Sam, and to Kevin in Radiology for his kind assistance.

I acknowledge my thanks to the National Institute for Health Research (NIHR) and Health Education England (HEE) who funded my doctoral fellowship and to the NHS who allowed me to go on secondment for three years. This thesis presents independent research funded by the NIHR/HEE. The views expressed are those of the author and not necessarily those of the NHS, the NIHR, HEE, or the Department of Health.

Finally I would like to thank my family for all their unending support and encouragement in particular my parents Ken and Jackie, my brother Dave, and my wife Rosie.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

σ_B^2	variance of bias values
σ_P^2	variance of parameter values
2C	two-compartment
3C	three-compartment
A	injected activity
AATH	Adiabatic approximation to the tissue homogeneity
AIC	Akaike information criteria
Akt	protein kinase B
AUC	area under the curve
AVH	activity volume histograms
BF	blood flow
BIC	Bayesian information criteria
BOLD	blood-oxygen-level dependent
BPL	Bayesian penalised likelihood
c	radioactivity concentration
CAIX	carbonic anhydrase-IX
CRT	chemoradiotherapy
CT	computed tomography
Cu	Copper
Cu-ATSM	Copper(II)-diacetyl-bis(N4-methylthiosemicarbazone)
CVH	concentration volume histogram
D	diffusivity
DNA	deoxyribonucleic acid

dPET	dynamic PET
DVH	dose volume histograms
E	extraction fraction
EGFR	epidermal growth factor receptor
ETM	Extended Toft's model
F	Fluorine
FAZA	fluoroazomycinarabinofuranoside
FBP	filtered back projection
FDG	fluorodeoxyglucose
FETNIM	fluoroerythronitroimidazole
FMISO	fluoromisonidazole
Gate-PM	gated phase-matched attenuation correction
Gate-SC	gated single CT attenuation correction
Ge	germanium
GLUT1	glucose transporter 1
HIF1	hypoxia inducible factor-1
HX4	flortanidazole
IDIF	image-derived input function
IR	infrared
IRF	impulse residue function
K	partition fraction
K_1	rate constant (3C models)
$K_{1\text{-}2C}$	rate constant (2C models)
k_{trans}	volume transfer constant
LOR	line of response

MB	mean bias
MRI	magnetic resonance imaging
MSEP	mean square error of prediction
MTD	maximum tolerated dose
mTOR	mammalian target of rapamycin
MTT	mean transit time
NSCLC	non-small cell lung cancer
OER	oxygen enhancement ratio
OSEM	ordered subset expectation maximization
P	permeability
pCT	perfusion CT
PET	Positron emission tomography
PI3K	phosphatidylinositol 3-kinase
PS	permeability surface area product
r	Pearson's correlation coefficient
ROC	receiver operating characteristic
RT	radiotherapy
S	surface
SABR	stereotactic ablative radiotherapy
SF	scale factor
SUV	standardised uptake value
SUV_{max}	maximum SUV
SUV_{mean}	mean SUV
TAC	time-activity curve
TBR	tumour-to-blood ratio

TBR_{mean}	mean TBR
TMR	tumour-to-muscle ratio
ToF	time of flight
uncorr	without correcting for movement
$V_{50\%}$	volumes with uptakes greater than 50% of SUV_{max}
v_B	fractional vascular volume (3C5K model)
$v_{B\text{-}2C}$	fractional vascular volume (2C model)
$v_{B\text{-}pCT\text{-}AATH}$	fractional vascular volume (AATH model)
$v_{B\text{-}pCT\text{-}ETM}$	fractional vascular volume (ETM model)
$v_{B\text{-}PET}$	fractional vascular volume (3C5K model)
v_e	interstitial volume
VOI	volume of interest
$V_{TBR1.4}$	hypothesised hypoxic volumes
w	patient's weight
Δt_i	duration of the i^{th} frame
λ	decay constant for ^{18}F

1 Introduction

1.1 Lung cancer

Lung cancer is the most common cancer in the world, representing 19% of all cancer deaths in 2011 (Jemal *et al.*, 2011). Around 44,000 people in the UK were diagnosed with lung cancer in 2012, and in the same year there were over 35,000 deaths due to lung cancer, reflecting the poor prognosis associated with the disease. It is the most common cause of cancer death, accounting for more than 1 in 5 deaths. The UK ten-year survival rate in 2010 was 5%, a level that has remained relatively constant for 40 years. (Cancer Research UK, 2016)

The majority of lung cancers (80%) are classified as non-small cell lung cancer (NSCLC) (Devesa *et al.*, 2005) of which there are three main types: adenocarcinoma, squamous cell cancer, and large cell carcinoma. The most common histological type of NSCLC is adenocarcinoma followed by squamous cell carcinoma (Akin *et al.*, 2012). Adenocarcinoma develops from mucus making cells in the lining of the airways whereas squamous cell cancer develops from the surface covering cells in the airways.

Lung cancers are categorised as Stage I-IV depending on the size of the tumour, involvement of nodes and presence or absence of distant metastases; staging details are described elsewhere (Goldstraw *et al.*, 2007). In summary, in Stage I disease tumours are small (up to 5 cm) and localised in one area of the lung. In Stages II-III tumours are larger (up to 7 cm) and patients may have cancer cells in the lymph nodes or surrounding tissue. Stage IV relates to patients whose lung cancer has metastasised. In the UK the majority (47%) of patients are diagnosed with Stage IV lung cancer (National Cancer Intelligence Network, 2015).

The Stage of the cancer determines the treatment received by the patient. In Oxford, Stages I-II are generally treated by surgical resection; medically inoperable patients are given radical radiotherapy (RT) of 66 Gy (33 fractions) or, for small peripheral lung tumours, stereotactic ablative radiotherapy (SABR). Higher-risk patients are administered adjuvant chemotherapy. If the tumour is not completely surgically excised then patients are also treated with radical RT of 60 Gy (30 fractions) for microscopically incomplete resection or 66 Gy (33 fractions) for macroscopically incomplete resection. Stage III treatments vary depending on the patient's fitness: the optimal treatment is concurrent chemo-RT (with a 'platinum doublet') but is associated with potentially significant side effects. Sequential chemo-RT treatment, or even RT alone may be delivered to less fit patients. Stage IV patients receive palliative thoracic RT in order to alleviate their symptoms.

The overall prognosis associated with NSCLC remains very poor despite improvements associated with combined modality treatment and advances in the physics and engineering of RT, which enable the accurate delivery of high doses of radiation to precisely sculpted treatment volumes. One possible reason for the poor prognosis is the hypoxic nature of the tumours (Horsman *et al.*, 2012).

1.2 Hypoxia

Hypoxia occurs when there is an imbalance between oxygen consumption and oxygen delivery (Span and Bussink, 2015). In simplified terms, hypoxia occurs in tumours which outgrow their blood supply, causing local oxygen concentrations within some tumour regions to fall well below that of normal tissue. Hypoxia can be chronic or acute. Chronic hypoxia was first reported by Thomlinson and Gray in 1955 and is caused by some tumour cells lying too far from blood vessels to receive a normal level of oxygen, a state termed diffusion-limited hypoxia. Oxygen diffuses approximately 100-180 μm from capillaries before being fully metabolised, and so cells located further than this distance from their nearest capillaries are likely to be hypoxic (Powis and Kirkpatrick, 2004). Acute hypoxia was first reported by Brown in 1979, and was attributed to perfusion fluctuations in the tumour vasculature and termed transient hypoxia.

Thomlinson and Gray (1955) discovered that tumour cells *in vitro* irradiated under hypoxic conditions were much more radioresistant than those irradiated under normoxic conditions. In normoxic tissues the presence of oxygen increases deoxyribonucleic acid (DNA) damage via the formation of oxygen free radicals after RT. In hypoxic tissues with low oxygen concentrations (partial pressures (pO_2) less than 10-15 mm Hg, Höckel and Vaupel, 2001) the radiation dose-levels required for a fixed level of cell kill are much greater than those required in normoxic cells (Figure 1.1). The relative reduction in dose required to achieve an isoeffective cell-kill level in the presence of oxygen is termed the oxygen enhancement ratio (OER) and can reach levels approaching a factor of three.

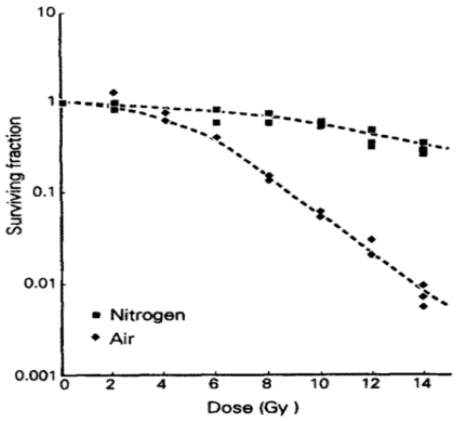


Figure 1.1. Cell survival levels following irradiation in hypoxic (nitrogen) and normoxic (air) conditions (Prekeges *et al.*, 1991).

In addition to being less responsive to RT treatments, hypoxia promotes a more aggressive phenotype with tumours shown to metastasise more readily and be more resistant to chemotherapy (Walsh *et al.*, 2014). It is thus a priority to identify tumours in which hypoxia exists (using imaging, circulating biomarkers, or genomics techniques) and to develop strategies to reduce it.

1.2.1 Imaging hypoxia

Hypoxia within tumours can be assessed by several methods: directly measuring oxygen levels, evaluating the expression of endogenous markers as a response to hypoxia, and measuring physiological processes involving oxygen molecules (Walsh *et al.*, 2014).

Direct measurement of pO_2 is possible using Eppendorf needles: a probe is invasively inserted into the region of interest and pO_2 measurements are taken along measurement tracks (Kallinowski *et al.*, 1990; Vaupel *et al.*, 1991). Studies have shown a link between hypoxia as measured via the needles and treatment outcome, but the needles can only be used for superficial tumours and are invasive for patients (Höckel *et al.*, 1993; Nordsmark *et al.*, 1996). Eppendorf needle pO_2 measurements were performed on a number of lung tumours in 2006 intraoperatively during surgical resection, revealing that tumour hypoxia

(as measured by Eppendorf needle and endogenous markers) correlated with poor prognosis (Le *et al.*, 2006). Whilst Eppendorf needle measurements are considered the 'gold standard' for hypoxia measurement, their invasive nature means that other imaging methods are generally preferred.

Many studies have been performed exploring links between endogenous markers of hypoxia and overall survival: in particular, expression levels of hypoxia-regulated proteins such as hypoxia inducible factor-1 (HIF1) and some of its targets such as glucose transporter 1 (GLUT1) and carbonic anhydrase-IX (CAIX) have been measured, with several studies finding significant links between biomarker expression and overall survival (Graves *et al.*, 2010).

An alternative approach is to gauge hypoxia using imaging techniques such as Positron Emission Tomography (PET, described in section 1.3) and Magnetic Resonance Imaging (MRI). A technique called blood-oxygen-level dependent (BOLD) MRI indirectly measures hypoxia by detecting the deoxyhaemoglobin concentration. The presence of deoxyhaemoglobin is associated with an increase in the MR relaxation time of water in blood and surrounding tissue. For PET imaging, pharmaceutical hypoxia tracers are labelled with radioisotopes, typically Fluorine (^{18}F) or Copper (^{64}Cu), as described in the next section.

1.2.2 PET hypoxia tracers

Several of the tracers developed for PET imaging of hypoxia are based on nitroimidazoles. These compounds were first introduced in 1979 as hypoxia markers (Chapman, 1979): in normoxic cells they are reduced and then re-oxidised, but in hypoxic cells further reduction of the nitro radical anion occurs, causing the compound to become irreversibly trapped within the cell. This reduction begins to occur at pO_2 levels lower than 10 mg Hg, and the degree of reduction increases as pO_2 values decrease (Gross *et al.*, 1995).

^{18}F-Fluoromisonidazole (FMISO) is the most extensively studied of the various PET tracers used to detect hypoxia. FMISO was developed in 1989 (Rasey *et al.*, 1989), and the earliest published experience of its clinical use was in 1992 (Koh *et al.*, 1992). FMISO is a lipophilic molecule, and with a partition fraction of 0.4 it diffuses readily across cell membranes; unlike some other PET tracers, movement of FMISO is driven entirely by passive diffusion. As with other nitroimidazoles, FMISO is reversibly reduced and re-oxidised in normoxic cells, and subsequently diffuses out of them, whereas in hypoxic cells it is further reduced and bound covalently (and irreversibly) to intracellular molecules at rates inversely proportional to intracellular oxygen concentrations (Figure 1.2). The process of passive diffusion leads to slow tracer accumulation in hypoxic tissue and slow clearance from normal tissue, generating a low target-to-background contrast and requiring imaging at late time-points (Prekeges *et al.*, 1991). Previous FMISO PET studies carrying out static imaging of lung cancer have scheduled scans at more than 2 hours post-injection (Koh *et al.*, 1995; Gagel *et al.*, 2006; Thureau *et al.*, 2013; Vera *et al.*, 2011). A good correlation has been reported between late FMISO uptake and Eppendorf needle pO_2 measurements (Gagel *et al.*, 2004); the degree of FMISO uptake increases sharply as pO_2 levels fall to values at which the OER decreases, thus making FMISO PET a potentially useful tool for imaging levels of hypoxia likely to impact on RT (Figure 1.2(B)).

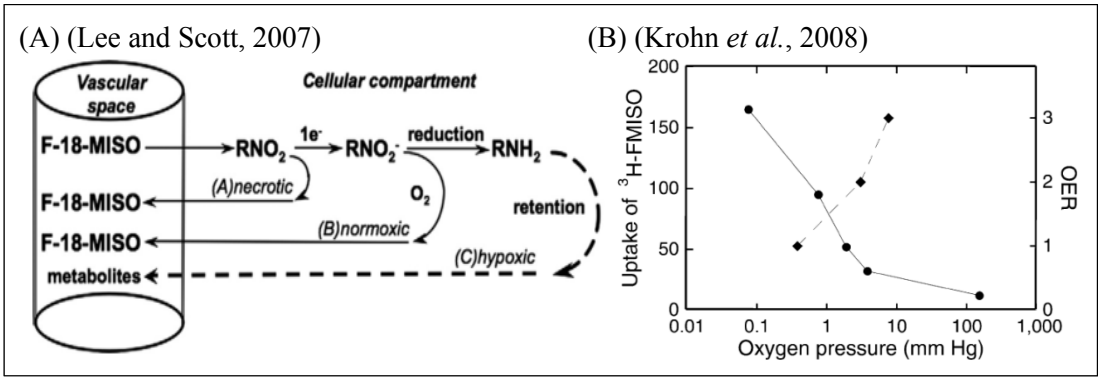


Figure 1.2. (A) FMISO behaviour in different cell types and (B) FMISO retention (solid line) in relation to OER (dotted line) and oxygen pressure. Figure 1.2B was originally published in JNM. Krohn *et al*. *Molecular Imaging of Hypoxia* 2008;49:129S-148S © by the Society of Nuclear Medicine and Molecular Imaging, Inc.

Relatively little work has been carried out to determine the reproducibility of derived parameters from FMISO scans. Recently a study in Japan on patients with head and neck cancer performed two FMISO scans 48 hours apart, imaging at 4 hours post-injection: the differences between mean hypothesised hypoxic volumes (in this case, tumour volumes having a tumour-to-blood ratio greater than 1.5) measured on the two scans had a one s.d. spread of ±9.9% (Okamoto et al., 2013). Another study showed geographically stable FMISO hypoxic subvolumes at 5 weeks into RT for patients with hypoxia on images 2 weeks into RT (Bittner *et al.*, 2013). Poorer reproducibility had been found in earlier studies of head and neck cancer patients which used an older generation PET scanner and imaged FMISO uptake at an earlier time point of 2.5 hours post-injection (Nehmeh *et al.*, 2008; Lin *et al.*, 2008). A new trial soon to open in Oxford involves FMISO PET imaging of NSCLC patients, and will include a test-retest reproducibility assessment.

In summary, the main limitations of the FMISO radiotracer are its slow pharmacokinetic profile and modest hypoxic-to-normal tissue ratios. However, at the time of this work FMISO was the only hypoxia tracer that could practically be transported to the site.

Other nitroimidazole-based hypoxia tracers, such as fluoroazomycinarabinofuranoside (FAZA), fluoroerythronitroimidazole (FETNIM) and flortanidazole (HX4), have been

developed with the aim of improving on the pharmacokinetics profile of FMISO (Carlin and Humm, 2012; Horsman *et al.*, 2012; Fleming *et al.*, 2014).

The pharmacokinetics of FAZA are in fact more favourable than those of FMISO, as FAZA is more rapidly cleared from the blood (Piert *et al.*, 2005). Consequently FAZA has been used in several studies to image hypothesised hypoxic volumes within lung tumours (Postema *et al.*, 2009; Trinkaus *et al.*, 2013; Bollineni *et al.*, 2013). One study of lung cancer patients imaged using the hydrophilic radiotracer FETNIM (Yang *et al.*, 1995) reported a strong correlation between sizes of hypothesised hypoxic volumes and overall survival (Hu *et al.*, 2013). An imaging study compared uptake of the HX4 tracer with FMISO and ^{18}F-fluorodeoxyglucose (FDG) uptake in 12 patients, and found that FMISO and HX4 images correlated positively, indicating promise for the use of HX4 (Chen *et al.*, 2012). The uptake pattern of HX4 between 2 hours and 4 hours post-injection has been found to be stable (Zegers *et al.*, 2013). Simulation work comparing FMISO, FAZA, and HX4 suggests that whilst HX4 has the highest contrast ratio, FMISO uptake is more reproducible (Wack *et al.*, 2015).

Another family of PET hypoxia tracers is based on diacetyl-bis analogues. Copper(II)-diacetyl-bis(N4-methylthiosemicarbazone) (Cu-ATSM) can be labelled with either Cu-60 (24 minute half-life) or Cu-64 (13 hour half-life). The longer half-life of Cu-64 means that Cu-ATSM labelled with this isotope is in many ways the more practically useful radiotracer, allowing late scanning to be undertaken and permitting the tracer to be used at PET-CT facilities not possessing in-house cyclotrons. Cu-ATSM has been utilised to image lung cancer patients (Lohith *et al.*, 2009), although recent work suggests that some Cu uptake is due to the behaviour of ionic Cu formed from hepatic metabolism rather than just hypoxia (Hueting *et al.*, 2014).

1.3 PET imaging

PET cameras indirectly image positron-emitting radionuclides by detecting back-to-back pairs of 511 keV gamma ray photons created by electron-positron annihilation (Cherry *et al.*, 2012). The positrons are generated by the β+ decay of radioisotopes such as ^{18}F, which are attached to tracers whose pharmacokinetics imitate at least part of the pathway of a physiological process of interest. For example, a common tracer used within oncology is FDG which, being a glucose analogue, is taken up at sites of high glycolytic activity.

PET systems detect both of the back-to-back gamma-ray photons, and the image reconstruction process assumes that the positron annihilation event generating them occurred somewhere along the line of response (LOR) between the two points at which the pair of photons are detected. Modern PET scanners operate in 3D time of flight (ToF) mode and record the time interval between detection of the two gamma ray photons, allowing the estimated location of the annihilation event to be narrowed to an interval along the LOR, unlike detectors operating in non-ToF mode, which cannot localise the event to any LOR sub-length. As a result, signal-to-noise ratios are higher in images obtained from scanners using ToF mode.

In 2D PET scanners lead septa separate neighbouring detector rings, so that only back-to-back photon pairs travelling in or close to each ring plane are detected. In contrast, more modern 3D systems detect photon pairs irrespective of their plane of travel provided that both photons are incident somewhere on the detector system.

Typically a CT image is generated using a 40-140 keV X-ray beam, and is hard-registered to a PET image. This can be transformed to represent 511 keV gamma ray photons to produce a PET attenuation map. This map is then used to correct for the attenuation that gamma ray photons undergo as they emerge from different parts of the body.

The coincident pairs of gamma photons detected by PET scanners during imaging are recorded and the data used to reconstruct an image of the distribution of the radiotracer within the body. This is most commonly achieved using the ordered subset expectation maximization (OSEM) iterative image reconstruction algorithm (Hudson and Larkin, 1994); a ToF version of this algorithm is implemented as VPFX on GE Healthcare PET scanners. Prior to the widespread use of OSEM an analytical algorithm called filtered back projection (FBP) was used. The shift towards OSEM was prompted by its improved signal-to-noise ratio compared to FBP (Adams *et al.*, 2010; Kinahan and Fletcher, 2010). Another type of iterative PET reconstruction algorithm (Bayesian penalised likelihood, BPL) has recently been released for clinical use ('Q.Clear', GE Healthcare, Milwaukee, USA). Studies on this new algorithm have been published showing an improvement in BPL-reconstructed images compared to OSEM (Teoh*/McGowan* *et al*. 2015, Teoh *et al*. 2015, Parvizi *et al*. 2015); Teoh*/McGowan* *et al*. (2015) is provided in the Appendix to the thesis.

A common way of reporting PET uptake values is via the standardised uptake value (SUV), defined as

$$SUV = \frac{c}{A/m} \tag{1.1}$$

where c is the radioactivity concentration, A the injected activity, and m the patient's weight.

The typical spatial resolution of a PET scanner is 5 mm (Bettinardi *et al.*, 2011) and consequently PET image voxels have widths of this length or a little smaller. This length-scale is much greater than that of typical hypoxic regions, and so each PET voxel is likely to consist of an average of both normoxic and hypoxic tumour regions, as is the case for the histopathology section shown in Figure 1.3. This leads to relatively low FMISO PET

SUV values compared to FDG, which is more uniformly taken up within individual voxels (Carlin and Humm, 2012).

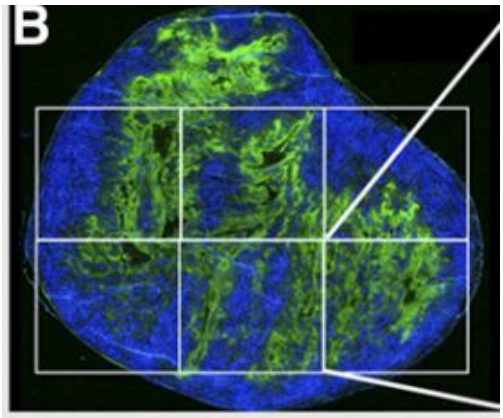


Figure 1.3 – Histology tumour section showing hypoxic regions (pimonidazole, green) and vascular perfusion (Hoechst 33342, blue) with white grid indicating example 5 mm^2 PET voxels. Originally published in JNM. Carlin and Humm *PET of Hypoxia: Current and Future Perspectives* 2012;53:1171-1174 © by the Society of Nuclear Medicine and Molecular Imaging, Inc.

1.3.1 Dynamic PET

When a PET acquisition is performed in 'list' mode it is possible to retrospectively reprocess scan data into distinct time windows. If scanning has been carried out over a lengthy timespan, the result is known as a dynamic PET (dPET) scan and consists of a 'cine' image showing the evolution of tracer uptake within a 3D volume over time. Time-courses of imaged uptake within single voxels or larger volumes (referred to as time-activity curves or TACs) can be easily extracted from dPET scans, and analysed with the aim of generating quantitative information about the mechanisms of tracer uptake. Different analysis methods can be used: most commonly compartment models have been fitted to the TACs, but approaches based on spectral analysis, graphical methods (essentially simplified compartment models, called Patlak plot and Logan plot for irreversible and reversible uptake respectively), and basis function techniques have also been attempted (Turkheimer *et al.*, 1994; Patlak and Blasberg, 1985; Logan *et al.*, 1990; Gunn *et al.*, 1997).

Compartment models use fitted rate-constants to characterise how a tracer moves from one compartment to another. Most rate-constants have units of min^{-1}, but the rate-constant K_1 which describes tracer movement from the vasculature to the first modelled compartment is capitalised to indicate its different units of mL blood per min per gram of tissue. Differential equations are used to describe tracer movement between compartments: for a simple two-tissue compartment model (shown in Figure 1.4) the equations describing the system are:

$$\frac{dC_1}{dt} = K_1 C_B(t) - k_2 C_1(t) - k_3 C_1(t) + k_4 C_2(t) \quad (1.2)$$

$$\frac{dC_2}{dt} = k_3 C_1(t) - k_4 C_2(t) \quad (1.3)$$

$$C_{PET}(t) = (1 - v_B)\big(C_1(t) + C_2(t)\big) + v_B C_B(t) \quad (1.4)$$

where k_2-k_4 are the other rate-constants, C_B the activity concentration of tracer in the blood (input function), C_1 the activity concentration of tracer in compartment 1 (free intracellular tracer, for example), C_2 the tracer activity concentration in compartment 2 (bound intraceullular tracer, for example), C_{PET} the total measured activity concentration, and v_B the fractional blood volume. In this thesis an x compartment model refers to an x tissue-compartment model.

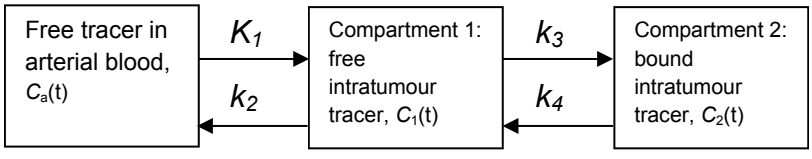


Figure 1.4 – Two-compartment model (Sokoloff *et al.*, 1977).

When FMISO scans are acquired in dynamic mode they are most commonly analysed using the two-compartment model shown in Figure 1.4 (Bruehlmeier *et al*., 2004; Wang *et al*., 2009). This two-compartment model has been adapted by redefining using weighting factors for diffusive and accumulative compartments (Thorwarth *et al*., 2005) and also by including parameters to model the spatial diffusion of FMISO (Kelly and Brady, 2006). A compartment model has also been developed consisting of ten parameters modelling the full chemical details of FMISO reduction (Casciari *et al.*, 1995).

Whilst it is straightforward to analyse late static images to assess regions of hypoxia there is a possibility that low uptake levels could be the result of two different mechanisms: either from decreased binding within cells indicating higher pO_2 levels, or from slower diffusion into the hypoxic cells owing to a lower local level of perfusion (Wang *et al.*, 2009). Kinetic analysis aims to distinguish between these two situations by analysing overall uptake time-courses as opposed to single late time-point images. In previous work a correlation was found between parameter fits obtained from FMISO kinetic analysis and radiation treatment outcomes (Thorwarth *et al.*, 2005).

Input functions can be obtained either from direct measurement of serial blood samples taken directly from a vessel (often a radial artery) or from the PET images themselves. Whilst direct arterial measurement is the gold standard, it is very invasive for patients and has a risk of adverse effects (Hall, 1971; Machleder *et al.*, 1972). An alternative used widely since 1989 is to determine the input function from a region of blood visible on the PET images (Gambhir *et al.*, 1989). Good agreement has been shown between input functions obtained from arterial sampling and image-derived input functions generated from volumes outlined within the ascending or descending aorta (van der Weerdt *et al.*, 2001).

FBP algorithms have historically been used to reconstruct images for dynamic PET analysis, but increasingly OSEM algorithms are now being used (Boellaard *et al.*, 2001; Reinders *et al.*, 2002; Mesina *et al.*, 2003; Lubberink *et al.*, 2004; Morimoto *et al.*, 2006; Søndergaard *et al.*, 2007). Images reconstructed from short time-frames using FBP suffer from very high noise, whilst OSEM has a non-negativity constraint that can lead to a positive bias in quantitative uptake values. In studies comparing FBP and OSEM reconstruction algorithms for dynamic PET, very similar rate-constants were obtained from kinetic analyses of both types of image, except for volumes-of-interest (VOIs) enclosed within a much hotter region with an activity concentration 5 to 10 times that of the VOI (Boellaard *et al.*, 2001). For lung tumour work this situation is unlikely to occur, as high activity concentration tumours are surrounded by lung tissue containing large volumes of air, and throughout this work image-derived input functions have been measured in a region of the descending aorta so as to be further from hot regions. OSEM reconstructions (2 iterations, 24 subsets, 6.4 mm Gaussian filter) have been used throughout the work described in this thesis, owing to the reduced image noise it produces. Further work is currently underway to compare FBP, OSEM, and BPL reconstructions for PET kinetic analysis.

1.4 CT imaging

Computed tomography (CT) is carried out using an x-ray source collimated to a fan-beam of X-rays. The mAs (tube current multiplied by exposure time) of a beam is proportional to the x-ray fluence produced. Both the source and detector rotate around the patient and FBP reconstruction can be used to reconstruct an image of the object that would cause the attenuation measured by the fan beam as it rotates around the patient. Thus by reconstructing the 2D distribution of attenuations in each slice, a 3D distribution of

attenuation in the patient can be produced. In order to reduce patient dose, mA modulation is used: a certain noise-level is selected, and the mA is modulated to ensure that sufficient photons traverse all the way through the patient at all beam angles to maintain this consistent noise level.

CT scanners can be used in three main modes: axial, helical, and cine. In axial mode the imaging system rotates once, generating the data for a single-slice image, then the patient is translated perpendicular to the plane of rotation of the scanner and another acquisition occurs; the process is repeated, generating a stack of axial slices until the slices span the whole volume required to be imaged. In helical mode the table moves continuously whilst the beam remains on and the imaging system rotates, creating a helical motion around the patient with a pitch defined as the ratio of table translation-per-imager rotation to the width of the radiation beam. Both helical and axial CT images essentially represent a snapshot in time, whereas in cine mode a movie is created by performing axial acquisitions repeatedly over time at the same table position.

During perfusion CT (pCT) cine imaging a bolus of CT contrast agent is injected into a patient and its passage through the patient is recorded, creating a dynamic image series similar to dPET, though on a shorter time-scale. Data from pCT images has been analysed by fitting a number of models to time-courses of imaged contrast agent concentrations within patients, generating fitted parameter values which in principle correspond to the flow of blood into a region, and the volume of blood in the region (Prezzi *et al.*, 2015).

1.5 PI3K Inhibitors: Buparlisib

Numerous strategies have been investigated to modify tumour hypoxia in order to improve RT outcomes. In Oxford, lab work has shown that altering oncogenic signalling can cause a reduction in tumour hypoxia within cells.

The phosphatidylinositol 3-kinase (PI3K) / protein kinase B (Akt) / mammalian target of rapamycin (mTOR) pathway (shown in Figure 1.5) controls tumour cell proliferation, growth, and survival after DNA damage, and is frequently dysregulated in human cancers (Samuels *et al.*, 2004; Yuan and Cantley, 2008; Maity and Bernhard, 2010). Activation of this pathway is common (50-83%) in NSCLC (Lee *et al.*, 2002; Gupta *et al.*, 2004; Graves *et al.*, 2010) and many other cancers, and causes an increase in tumour proliferation and resistance to cell death in tumours. In NSCLC activation can occur via several mechanisms such as mutation of the epidermal growth factor receptor (EGFR) gene (~8% NSCLC), mutation of the KRas oncogene (~20-25% NSCLC), and PI3K mutation (~1-2% NSCLC) (McKenna *et al.*, 2003; Maity and Bernhard, 2010).

Cell-death caused by ionising radiation is most often the result of radiation-induced DNA double-strand breaks. Studies have shown that activated Akt accelerates the repair of double strand breaks and therefore improves post-irradiation tumour cell survival (Toulany and Rodemann, 2013). Activation of the EGFR-Ras-PI3K pathway can cause cells to become intrinsically radioresistant (Kim *et al.*, 2005; Shaw and Cantley, 2006; Liu *et al.*, 2009; Akinleye *et al.*, 2013). Therefore it is of interest to investigate PI3K inhibitors, particularly when their use is combined with RT. In addition to altering intrinsic radiosensitivity, there is emerging evidence that inhibition of this pathway reduces tumour hypoxia, thereby reducing extrinsic radioresistance (Qayum *et al.*, 2009).

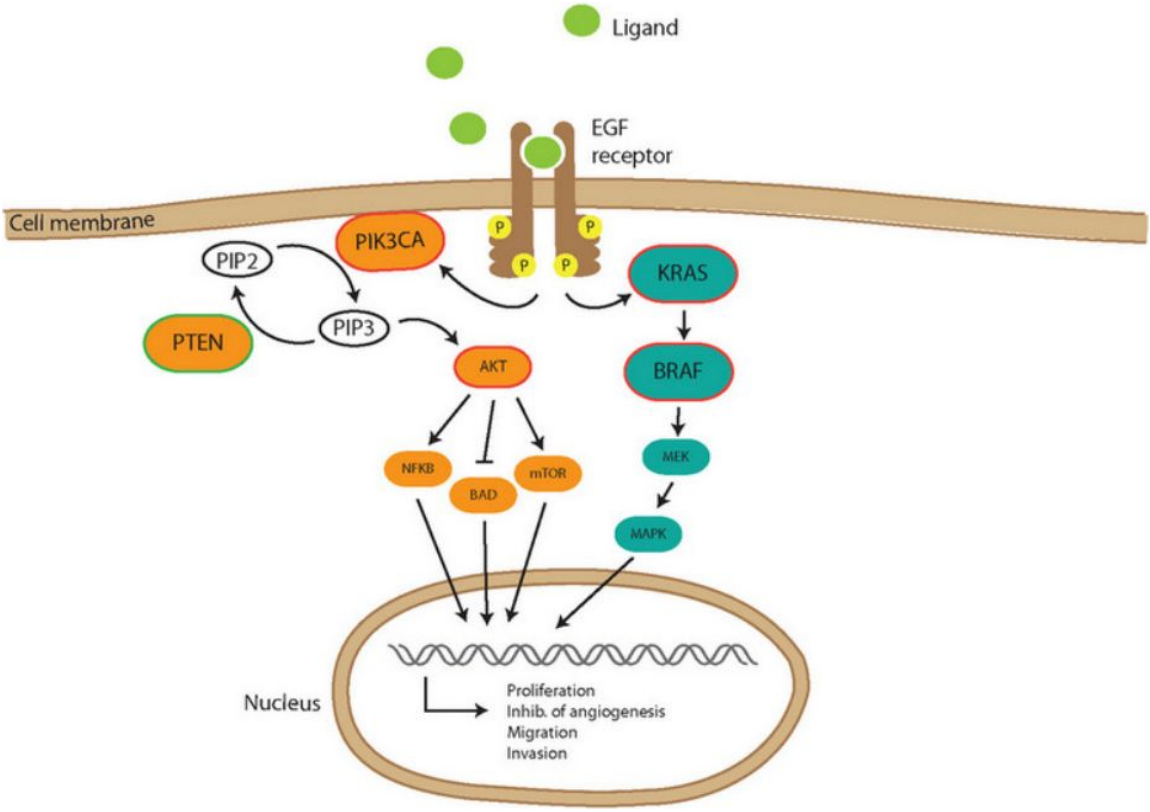


Figure 1.5 – The PI3K pathway: in addition to activation of KRAS, EGFR activates the oncogene PIK3CA which activates AKT and mTOR leading to proliferation and cell growth (Berg and Soreide, 2012).

Tumours typically have grossly abnormal blood vessels, leading to poor tumour perfusion and hypoxia. However after seven daily doses of Buparlisib (a specific PI3K inhibitor formerly called BKM120) in xenograft models tumour perfusion was increased (Figure 1.6), tumour vasculature was normalised, and tumour hypoxia was reduced (Figure 1.7) (Fokas *et al.*, 2012). Buparlisib induces 'vascular normalisation', causing vessels to become longer, wider and less tortuous. A significant increase in tumour regrowth delay resulted when animals were treated with radiation and drug (BEZ235, another PI3K inhibitor) compared to either treatment delivered alone (Figure 1.8), suggesting that it might be beneficial to combine PI3K inhibitors with RT. BEZ235 has off-target effects including on DNA-PK (protein kinase that initiates signalling in response to DNA damage) and as this could exacerbate RT induced side effects only Buparlisib was considered for the trial.

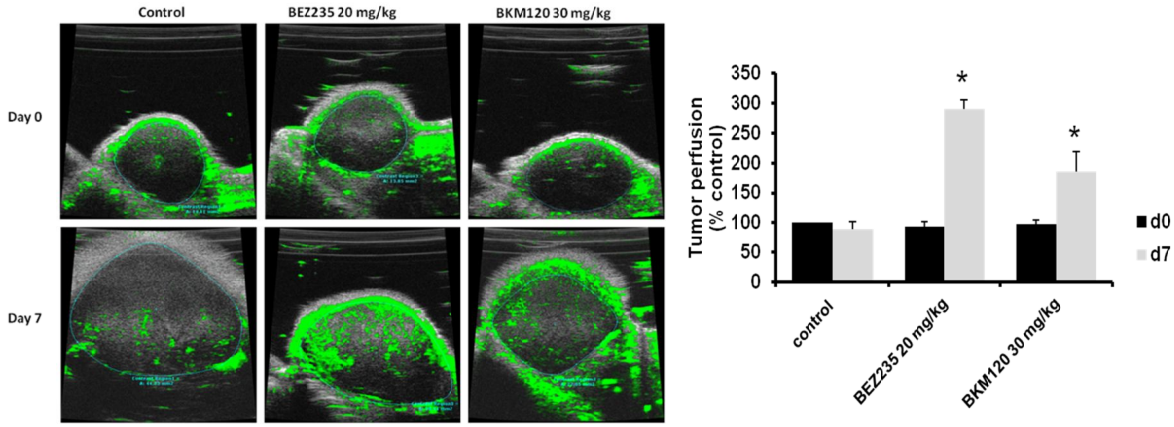


Figure 1.6 – Perfusion changes before and after seven days of BKM120 (now Buparlisib) and BEZ235 (Fokas *et al.*, 2012). A significant increase in perfusion (measured by microbubble ultrasound) occurs after both BKM120 and BEZ235, but with no significant change in the control group.

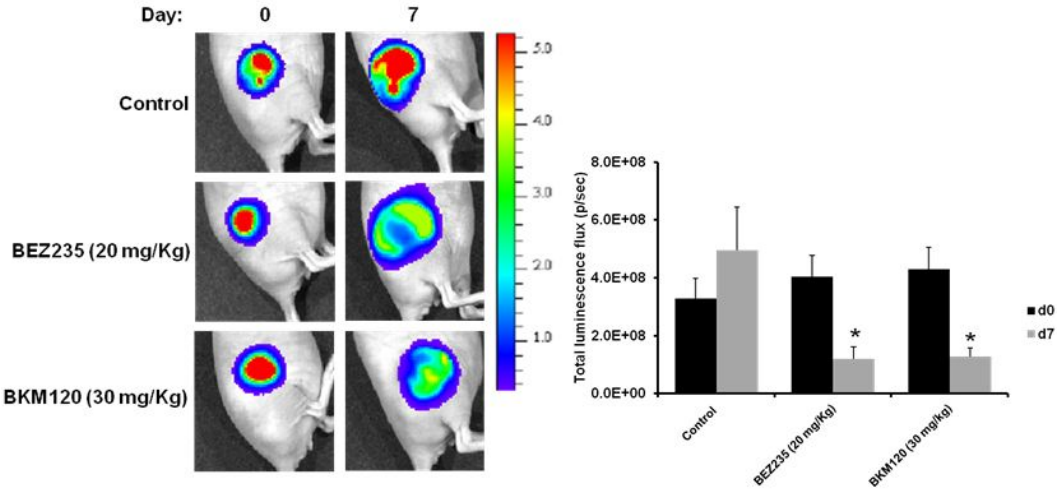


Figure 1.7 – Hypoxia changes before and after seven days of BKM120 (now Buparlisib) and BEZ235 (Fokas *et al.*, 2012). A significant reduction in hypoxia (measured by bioluminescent imaging) occurs after both BKM120 and BEZ235, with no significant change in the control group.

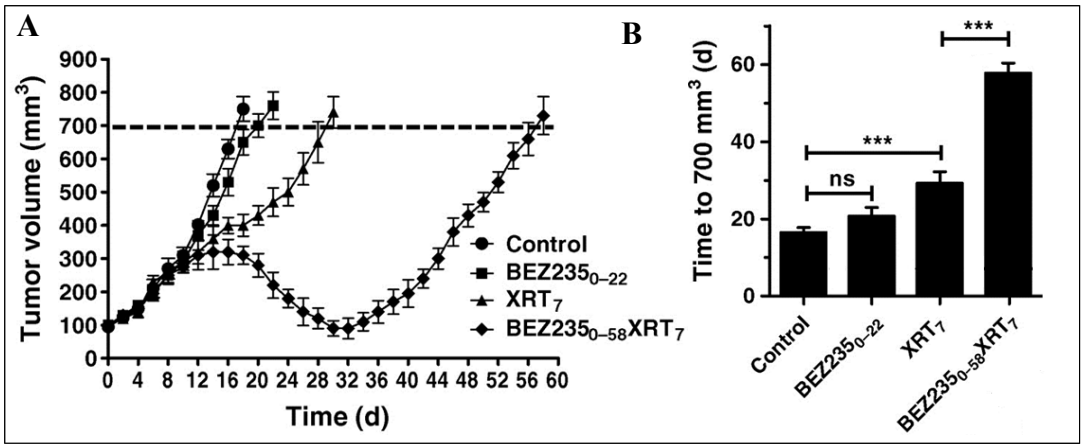


Figure 1.8 – Tumour volume changes following BEZ235 and RT (Fokas *et al.*, 2012). (A) The evolution of tumour volume over time is plotted for tumours of approximately 100 mm^3 initial volume, treated according to one of the following schedules: control, no treatment (circles); daily doses of BEZ235 (20 mg/kg) on days 0-22 (squares); 6 Gy RT on day 7 (triangles); daily doses of BEZ235 on days 0-58 with 6 Gy at day 7 (diamonds). (B) Mean times taken for tumours in each treatment schedule to reach 700 mm^3.

If Buparlisib had a similar effect in the human tumour microenvironment, the effectiveness of RT treatments might well be substantially increased by its administration. A Phase I study underway in Oxford is assessing the effect of Buparlisib in advanced stage NSCLC patients, using PET and pCT imaging. Buparlisib has been used as a single agent in several clinical trials that have administered daily doses of 100 mg (for example BELLE-3, a Phase III trial in breast cancer), following the result of a prior single agent study which determined this to be the maximum tolerated dose (MTD) of Buparlisib (Bendell *et al.*, 2012). The Oxford trial is the first to administer Buparlisib in combination with RT in NSCLC patients, potentially improving the efficacy of the RT treatment.

This thesis analyses static and dynamic PET and perfusion CT imaging data collected during the Oxford trial, according to the following framework:

- Chapter 2 – movement correction systems for PET scanners
- Chapter 3 – description of the trial structure, an investigation of hypothesised hypoxic volumes determined from late time-point static FMISO data, and the impact of Buparlisib on these volumes
- Chapter 4 – exploration of the compartment model that best describes whole tumour FMISO TACs and whose fitted parameter values most accurately match the kinetics of FMISO uptake, and an exploration of the impact of Buparlisib on the fitted model rate-constants
- Chapter 5 – exploration of the optimum compartment model for analysing FMISO kinetics on a voxel-by-voxel basis, an exploration of the fitted model rate-constants, the impact of Buparlisib on these derived rate-constants, and a description of the changes in derived rate-constants depending on the distance from the edge of the tumour
- Chapter 6 – analysis of the pCT data, and evaluation of the impact of Buparlisib on the derived perfusion parameters
- Chapter 7 – summary of the results obtained, and conclusions drawn.

2 Movement

2.1 Abstract

A study is presented which evaluates different PET/CT motion correction systems (uncorrected, gated, QFreeze). A phantom containing three active volumes was imaged using a GE 690 PET/CT scanner. Three metrics were used to analyse the data: the maximum standardised uptake values within an active volume (SUV_{max}), concentration volume histograms (CVHs), and receiver operating characteristic (ROC) curves. CVHs describe volumes with uptakes greater than a range of fractions of SUV_{max}, and are analogous to dose volume histograms. ROC curves describe the sensitivities and specificities with which 'boost voxels', having imaged uptakes greater than 50% of SUV_{max} in the absence of movement, are identified in scans of the moving phantom.

Gated and QFreeze images substantially reduce the distortion introduced into PET images by periodic movement, eliminating around 60% of the error on SUV_{max} and 80% of the loss in the area under the curve (AUC) of ROC curves for a 0.5 mL active volume. Boost voxels could be identified on uncorrected scans with a sensitivity of 95% and false positive rate of 35%, which is reduced to 10% using QFreeze or gated images. Overall QFreeze performed slightly better than gating at clinically typical noise levels, having a significantly higher AUC for the ROC curve of a 20 mL active volume.

The impact of regular respiratory motion on PET images was substantially reduced by both gating and QFreeze. At clinically typical noise levels QFreeze provided small, but significant, advantages over gating.

2.2 Introduction

Radiotherapy (RT) dose-painting is an emerging technique that delivers escalated radiation doses to specific tumour subvolumes selected on the basis of functional or molecular imaging (Ling *et al.*, 2000). One approach uses PET images of the ^{18}F-fluorodeoxyglucose (^{18}F-FDG) radiotracer (Bentzen, 2005; Søvik *et al.*, 2009) and boosts subvolumes in which radiotracer uptake exceeds a given percentage (typically 40-50%) of the maximum uptake within the tumour (Mah *et al.*, 2002; Erdi *et al.*, 2002; Bradley *et al.*, 2004; Deniaud-Alexandre *et al.*, 2005; van Elmpt *et al.*, 2012). Levels of tracer uptake in PET scans are commonly represented as SUV (Zasadny and Wahl, 1993). The maximum uptake within a tumour is denoted as SUV_{max}.

Over the respiratory cycle lung tumours typically move a distance of around 0.5 cm, although some tumours can move by more than 2 cm (Seppenwoolde *et al.*, 2002). This respiratory motion blurs PET images, potentially distorting both SUV_{max} values and the shape of tumour subvolumes that have imaged uptake levels in excess of a given fraction of SUV_{max}. Several methods have been developed to minimise or correct for movement during PET imaging, and in this work we evaluate some of them using phantom measurements analysed via metrics specific to dose-painting. In particular, we determine their impact on SUV_{max} (which, being a point measure, is susceptible to both blur and image noise), and on the sizes and shapes of volumes with imaged uptakes in excess of different fractions of SUV_{max}.

Motion correction systems typically record a trace of a tumour movement surrogate during imaging, and use it either to gate PET images, retaining data collected during one particular phase interval of the breathing cycle, or to construct 4D-PET images comprising a sequence of 3D images collected in each breathing phase. The Varian Real-time Position

Management system (RPM, Varian Medical Systems, Palo Alto, CA) uses a block with six infrared (IR) markers, placed on a patient's chest and imaged by an IR camera, to generate a trace. Alternatives include a belt placed around the chest which detects pressure changes over the breathing cycle (Anzai Medical Systems, Tokyo, Japan), and systems that detect air flow or changes in air temperature during breathing. The same systems can be used to guide the collection of 4D-CT images. PET data collected during particular phase intervals of the respiratory cycle should, in principle, be attenuation-corrected using 'phase-matched' CT images collected during the same intervals; however, more approximate corrections can be made using a 3D-CT image collected during the course of quiet breathing.

QFreeze is a motion correction algorithm developed by General Electric (GE, Milwaukee, USA). 4D-CT and 4D-PET data are collected using the Varian RPM system, and the PET data is attenuation-corrected using the phase-matched 4D-CT. PET images of the different phase intervals are then deformably co-registered to a reference phase (chosen by the user) and added together to create a single PET image, a process that has been summarised as 'reconstruct, register, average' (Wollenweber *et al.*, 2012). Owing to that fact that the QFreeze PET images comprise data collected across all phase intervals, they are in principle less noisy than images constructed from data obtained during a single phase.

The impact of the QFreeze algorithm on SUV_{max} values has previously been evaluated (Wollenweber *et al.*, 2012) using a phantom containing moving spheres of 17 mm internal diameter filled with the germanium radioisotope ^{68}Ge. In this work we use a phantom containing ^{18}F-FDG-bearing volumes with diameters between 10 and 30 mm. Alongside SUV_{max} measurements, we determine the sensitivities and specificities with which voxels whose imaged uptakes are 50% or more of SUV_{max} in PET scans of the static phantom can be identified in scans of the moving phantom, either making no correction for movement, or using QFreeze, or a gated PET image.

2.3 Materials and methods

2.3.1 Experimental methods

Three syringes (BD Medical, Franklin Lakes, USA) were attached to a semi-cylindrical cedar insert located within a cylindrical cut-out in a perspex QUASAR phantom (Modus Medical, London, Canada). The syringes had diameters of 0.9, 2.3, 2.8 cm, and were filled with 0.5, 10 and 20 mL of uniform ^{18}F-FDG activity (350 kBq/mL). These active volumes are denoted here as small, medium and large (Table 2.1) and are smaller than the capacities of the syringes, whose remaining unfilled volumes contained air (Figure 2.1A). The rest of the phantom contained no activity. Top-down and axial schematics of the phantom geometry are shown in Figures 2.1B and 2.1C. The phantom was programmed to move sinusoidally along a horizontal (z) direction for a 3-second period with a peak-to-peak amplitude of 2 cm, which is close to the maximum movement expected for patients (Seppenwoolde *et al.*, 2002), thereby providing a demanding test of the motion correction software. An auxiliary plate was programmed to move sinusoidally in the vertical (y) direction with a peak-to-peak amplitude of 1 cm, and an IR marker box was placed on the plate.

Syringe	^{18}F volume in syringe	Syringe maximum volume	Active length shown as dotted line in Figure 2.1A	Active diameter shown as dotted line in Figure 2.1A
Small	0.5 mL	2 mL	0.8 cm	0.9 cm
Medium	10.0 mL	40 mL	2.5 cm	2.3 cm
Large	20.0 mL	60 mL	3.3 cm	2.8 cm

Table 2.1 – Details of active volumes used in the experiment.

The phantom was scanned using a GE Discovery 690 PET/CT camera (GE, Milwaukee, USA) with the z-direction aligned with the scanner axis, following a PET/CT protocol comprising a scout scan for positioning, a standard 3D-CT scan, a 4D-CT scan (when phantom movement was enabled), and a 4-minute PET scan in one bed position. The Varian RPM system was used to record the respiratory motion trace. Imaging was carried out three times with phantom movement turned off, and a further three times with the phantom moving, collecting a total of six complete datasets.

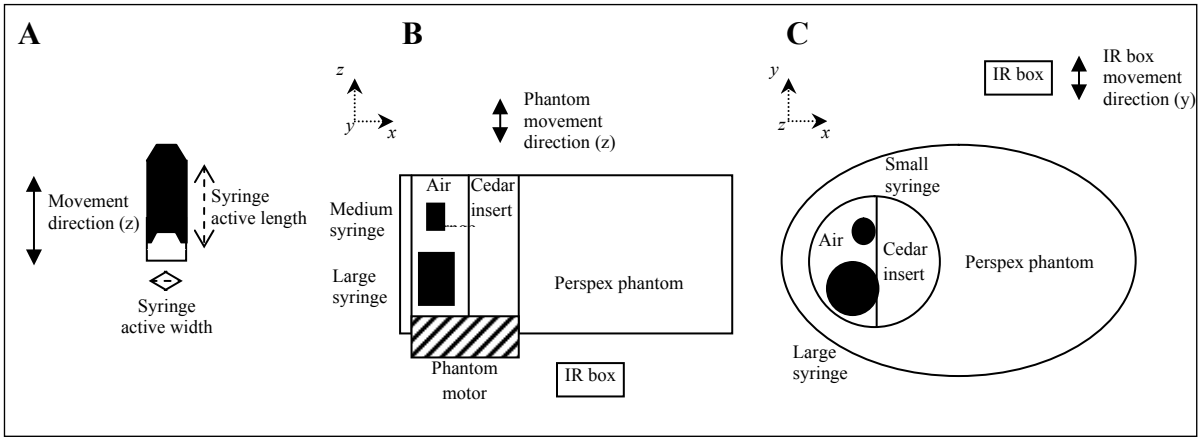


Figure 2.1 – The phantom geometry: (A) a schematic filled syringe, the solid black region representing an active (^{18}F) volume whose dimensions are detailed in Table 2.1; (B) a top-down view of the phantom geometry; (C) axial view.

PET data collected for the static phantom were attenuation-corrected using a static 3D-CT scan, and then reconstructed to create images denoted here as ‘static’. Data collected for the moving phantom were initially attenuation-corrected using a standard 3D-CT scan of the moving phantom and reconstructed without correcting for movement, creating images denoted as ‘uncorr’. Image reconstruction was carried out using a ToF OSEM algorithm (2 iterations, 24 subsets, 6.4 mm Gaussian filter, 2.7×2.7×3.3 mm voxel).

Using the GE MotionMatch software, PET list-mode data collected for the moving phantom was also retrospectively split into bins comprising data collected during five equal phase intervals of the breathing cycle, identified from the respiratory trace. These intervals are termed A-E and summarised in Table 2.2: each covers 20% of the cycle, phase A starting at 0% which corresponds to the end of inhalation. The midpoints of the PET phase ranges were used as the target phases when collecting 4D-CT images.

Phase Name	CT target phase %	PET phase range %
A	10	0-20
B	30	20-40
C	50	40-60
D	70	60-80
E	90	80-100

Table 2.2 – Details of the five phase bins.

4D-PET images were generated by separately reconstructing the PET data of each phase bin. Two 4D-PET images were created from each data-set. The first, here termed ‘gated single CT attenuation correction’ or Gate-SC, was reconstructed from data corrected for attenuation using a single conventional 3D-CT image. The second, termed ‘gated phase-matched attenuation correction’ or Gate-PM, was reconstructed from data that had been attenuation-corrected phase-specifically using a 4D-CT.

Both 4D-PET images consist of five individual 3D images corresponding to phases A-E of the respiratory cycle. From the Gate-PM and 4D-CT data, the QFreeze system generates a single motion-corrected 3D-PET QFreeze image by registering the different phases of the Gate-PM scan to one reference phase. Some details of this algorithm are given in the published white paper (Ganin and Wollenweber, 2012) but the exact details are commercially confidential. As part of this study each of the five phases was sequentially selected as the reference, thus generating five QFreeze images from each scan of the moving phantom. The various images generated in the study are summarised in Table 2.3.

In Oxford, patients are typically injected with 4 MBq/kg of ^{18}F-FDG and imaged for 4 minutes per bed position at 90 minutes post-injection. Accounting for physical decay, a patient's average specific activity at imaging is 4×0.566 or 2.27 MBq/kg. Consequently an SUV of 2 corresponds to an activity concentration of 4.54 kBq/mL, assuming a unit mass density. The 350 kBq/mL activity concentration within the active volumes imaged in this work is higher by a factor of 77, and from Poisson statistics the signal-to-noise ratio in these volumes, imaged for 4 minutes, will be approximately a factor of 9 greater than that for lesions with an SUV of 2, and these images are therefore termed 'low noise'. To represent scans with clinically typical noise images were reconstructed from data taken from 3-second intervals of one of the 4-minute scans, thus reducing the counts to approximately the number obtained in 4 minutes from a lesion with an SUV of 2. We have taken data from three such intervals of 3 seconds, to create repeat 'clinically typical noise' scans.

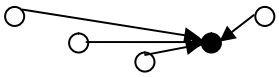
Image description	Schematic illustration	Number of images generated	CT scan used for attenuation correction
Uncorrected movement (Uncorr)		1	3D-CT
Gated, single CT AC (Gate-SC)		5, 1 for each phase	3D-CT
Gated, phase-matched AC (Gate-PM)		5, 1 for each phase	4D-CT
QFreeze		Usually 1, for a single reference phase, but 5 here	4D-CT
Static		1	3D-CT

Table 2.3 – A comparison of image types generated in the study. Three repeat scans have been obtained for each type of image, both for low and clinically typical noise conditions. AC indicates attenuation-corrected.

2.3.2 Analysis methods

The image data was analysed using three quantitative metrics: SUV_{max}, concentration volume histograms (CVHs), and receiver operating characteristic (ROC) curves (Lusted, 1971).

Using the Hermes Hybrid Viewer (Hermes Medical Solutions AB, Sweden), SUV_{max} values were determined for each active volume in images of the moving and static phantom. Values obtained from scans of the moving phantom were then calculated as percentages of the SUV_{max} values obtained for the same active volumes from the static

scan. Although the true activity concentration within the active volumes is known, it is preferable to normalise to static scan SUV_{max} values because the partial volume effect causes the imaged activity concentrations of small active volumes to lie below the true concentration, even in the absence of phantom movement (Soret *et al.*, 2007).

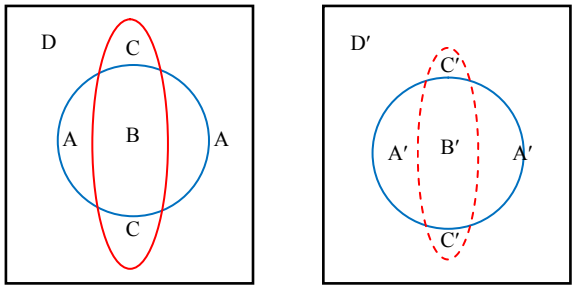
CVHs are defined similarly to the previously described activity volume histograms (AVHs) (Thorndyke *et al.*, 2008). They are analogous to the dose volume histograms (DVHs) commonly used in RT planning, but display relative imaged activity concentrations (SUV/SUV_{max}) rather than doses along the x-axis. The y-axis of cumulative CVHs shows volumes with activity concentrations greater than the fractions of SUV_{max} displayed on the x-axis. In current RT dose-painting protocols, tumour subvolumes selected for boosting typically have uptakes in excess of 40-50% of SUV_{max} (Mah *et al.*, 2002; Erdi *et al.*, 2002; Bradley *et al.*, 2004; Deniaud-Alexandre *et al.*, 2005), and so CVHs above a lower threshold of 40% of SUV_{max} were plotted, allowing absolute volumes with uptakes in the range 40-100% of SUV_{max} to be compared between different images.

In order to compare the shapes and locations of these regions, as well as their volumes, we have used the ROC methodology (Lusted, 1971). Specifically, graphs were plotted of the sensitivities and specificities for volumes comprised of voxels having uptakes greater than 50% of SUV_{max} in a static phantom image, which can be identified in other images by selecting voxels with imaged uptakes greater than a given fraction (varied from 20-90% in increments of 5%) of SUV_{max} in those images. Here sensitivity is defined as the number of voxels correctly identified in the moving scan, expressed as a fraction of the number of voxels with uptakes greater than 50% of SUV_{max} in the static phantom scan (the overlapping volume B of Figure 2.2 divided by volumes A and B). Additionally, (1-specificity), the false positive rate, is defined as the number of voxels incorrectly identified from the moving scan, expressed as a fraction of the total number of voxels with uptakes

less than 50% and greater than 5% of SUV_{max} in the static scan (volume C divided by volumes C and D). The lower 5% limit on volume D has been chosen in order to limit the voxels being considered to those within a reasonable distance of the active volume. In the context of RT dose-painting, sensitivity is the fraction of the voxels that would be chosen for boosting on the basis of an image unaffected by movement, which are actually boosted on the basis of a scan in which respiratory movement occurs. The measure (1-specificity) is the number of voxels incorrectly selected for boosting using a scan of a moving patient, expressed as a fraction of the total number of voxels that would not be boosted on the basis of an image free of movement effects.

We carried out ROC comparisons of QFreeze, Gate-PM, and uncorrected images versus a static image. For each non-standard image type we obtained sensitivities and specificities from three repeated scans, to estimate both average values and their degree of variability. In order to determine which voxels (regions of space) in the non-standard images correspond to voxels in the static image, QFreeze images were rigidly registered to the static scan; registration of the gated and uncorrected images with the static scan then followed automatically. ROC analyses of the three repeated static scans versus each other were also carried out, to assess the degree to which image noise causes the ROC curves for static images to depart from the ideal of unit sensitivity and specificity. The AUC (Hanley *et al.*, 1982) was used as a metric of ROC quality: for an ideal image in which boost voxels could be identified perfectly, the AUC would be 1; whereas for an image containing no information, in which the locations of boost voxels were randomly guessed, the AUC would be 0.5.

The blue lines represent the region in which imaged uptake is greater than 50% of SUV_{max} in a scan unaffected by movement.



The red lines represent regions in which voxels have an imaged uptake greater than some fraction of SUV_{max} in a scan affected by respiratory motion, the fraction being higher for the dashed than for the continuous line.

Figure 2.2 – Schematic of the ROC analysis.

2.3.3 Statistical methods

Averages, variances and sample standard deviations of SUV_{max} values, CVH volumes, sensitivities, specificities and AUCs of ROC curves were calculated from individual values obtained from three repeat images, collected under both clinically typical and low noise conditions. Figures plotted in the Results (section 2.4) show average values together with error bars representing scan-to-scan variation at the one standard deviation level. A two-tailed paired t-test was used to determine the significance of differences in values between QFreeze, gated, uncorrected and static scans. Differences in variance were assessed using the F-test.

2.4 Results

Figure 2.3 shows the difference between uncorrected and QFreeze images of the moving phantom at low and clinically typical noise levels.

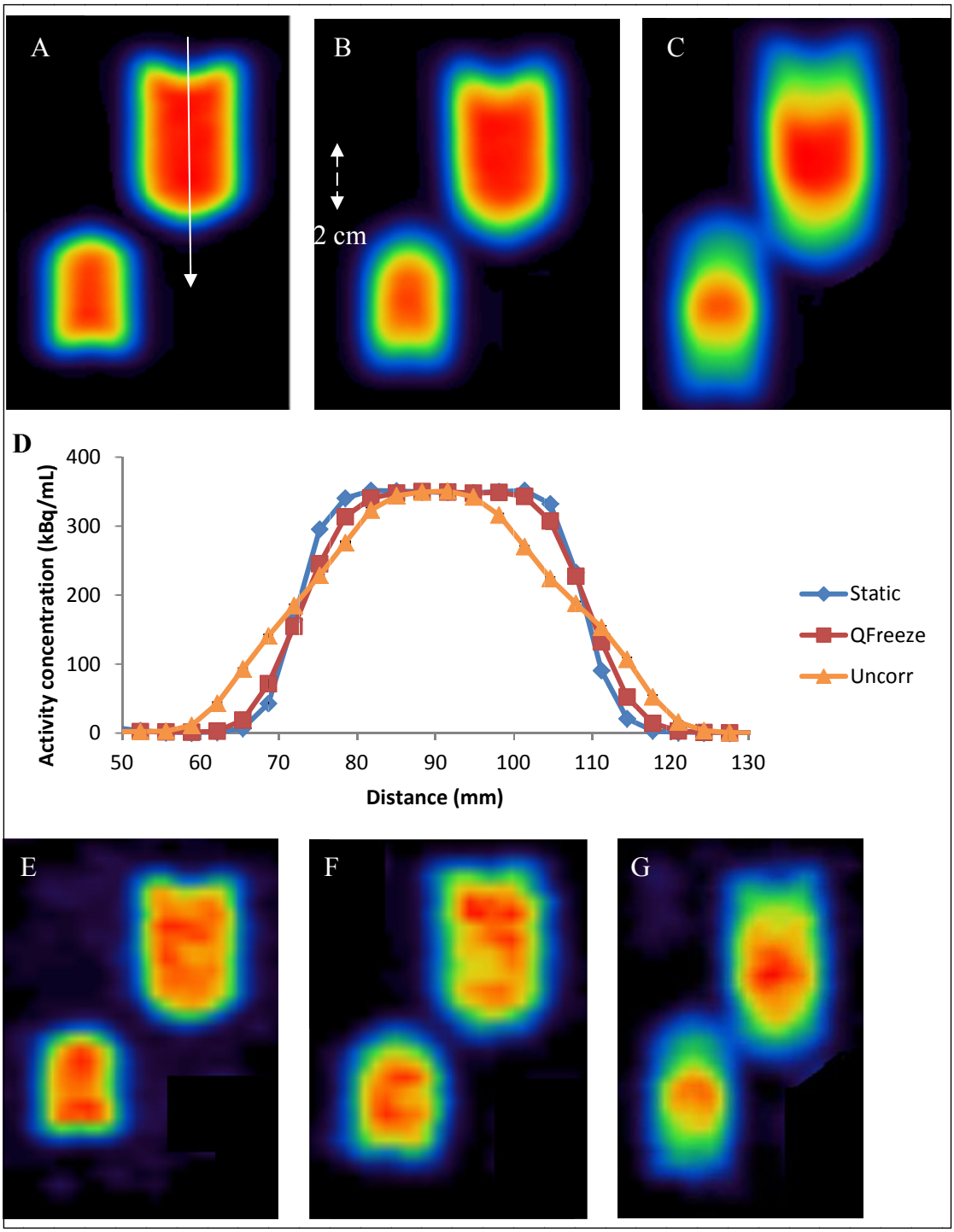


Figure 2.3 – Cross-sections through (A) static, (B) QFreeze, and (C) uncorr images of the moving phantom (2 cm peak-to-peak movement, low noise), (D) a line profile along the axis of the large active volume (white arrow in (A)) and cross-sections through (E) static, (F) QFreeze, and (G) uncorr images of the moving phantom (2 cm peak-to-peak movement, clinically typical noise).

Cross-sections from the QFreeze and uncorrected scans are plotted alongside a cross-section from a scan of the static phantom, together with line profiles along the axis of the large active volume indicated by the white arrow in Figure 2.3A. The QFreeze profile is more blurred than the static profile but much sharper than the profile of the uncorrected scan. The Gate-PM and Gate-SC images are similar to the QFreeze image and are not shown in the figure.

2.4.1 Maximum imaged uptake levels (SUV_{max})

The SUV_{max} values of the three active volumes are shown in Figures 2.4-2.7 for the different image types, normalised to the average SUV_{max} values of the respective active volumes in the static images. Although the activity concentration within each active volume was the same, the imaged SUV_{max} of the small active volume in the static scan was only 46% of the SUV_{max} values of the medium and large active volumes due to the partial volume effect (Soret *et al.*, 2007), and as this difference is unrelated to movement we have normalised it out.

Figure 2.4A shows normalised SUV_{max} values at low noise levels in static, QFreeze (reference phase D), gated (phase D) and uncorrected images. Figure 2.4B shows equivalent data for scans with clinically typical noise levels. Figure 2.5 shows the variation by phase of the normalised SUV_{max} values of the QFreeze and gated scans. Tables 2.4 and 2.5 list the significances of differences between the SUV_{max} values of the various image types at low and clinically typical noise levels.

SUV_{max} values in uncorrected scans are lower than those in static scans, as expected, to a degree that reaches significance in our dataset for the small and medium active volumes imaged at low noise levels, and for the small active volume at clinically typical noise levels ($p<0.05$). SUV_{max} values in Gate-SC, Gate-PM and QFreeze images lie closer to the

static image values, with differences only being obvious for the small active volume (Figure 2.4) and only reaching significance at low noise levels in this dataset.

Differences between SUV_{max} values in QFreeze images and those in Gate-PM and Gate-SC images are insignificant at low noise levels. However, at clinically typical noise levels SUV_{max} values in the Gate-PM and Gate-SC images exceed those in QFreeze images, significantly so for the medium and large active volumes. This occurs because QFreeze images comprise data from all five respiratory phases, whereas Gate-SC and Gate-PM images contain data from only one phase and are thus noisier; since SUV_{max} is a point-based measure reflecting the highest imaged uptake within a region, its value rises as the noise level increases. For the medium and large volumes, SUV_{max} values in the QFreeze images are very similar to those in static images, and consequently SUV_{max} values in the Gate-SC and Gate-PM images overestimate those in the static images. For the small active volume, however, QFreeze SUV_{max} values are lower than those in static images (significantly so at low noise levels) because of the residual movement within each respiratory phase; small active volume SUV_{max} values in Gate-SC and Gate-PM images are closer to those in static images than QFreeze SUV_{max} values.

For the Gate-PM and QFreeze images the variances of SUV_{max} values were calculated across all five phases of the three repeated scans. As expected, the variances of Gate-PM SUV_{max} values are significantly greater than those of QFreeze SUV_{max} values at clinically typical noise levels for the small and medium active volumes ($p = 0.02$ and 2×10^{-6} respectively). However, no significant differences in variance were seen for the large active volume at clinically typical noise levels, or for any volume at low noise levels.

In summary, SUV_{max} values in gated and QFreeze images are closer to those in static images than SUV_{max} values in uncorrected images of the moving phantom. However,

SUV_{max} values of small active volumes in QFreeze images still lie below those in static images, due to residual movement within each respiratory phase, although for larger volumes QFreeze SUV_{max} values are very close to static values. At clinically typical noise levels, SUV_{max} values in gated images have higher variances and mean values than QFreeze SUV_{max} values. Consequently, for larger active volumes the SUV_{max} values in Gate-SC and Gate-PM images overestimate SUV_{max} values in static images, but for smaller active volumes they provide better estimates than QFreeze values.

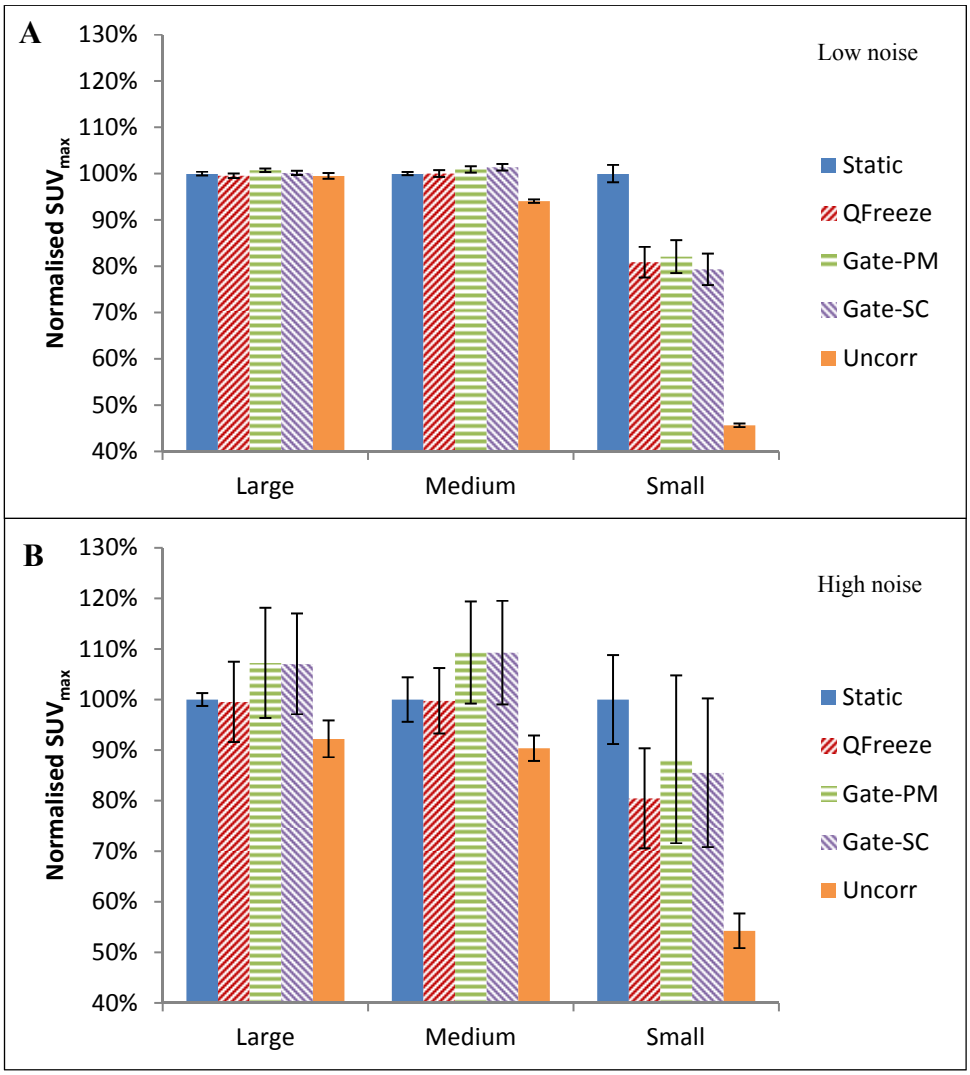


Figure 2.4 – Normalised SUV_{max} values of the large, medium and small active volumes in static and uncorrected images, phase D of gated images, and QFreeze images generated using phase D as the reference, at (A) low, and (B) clinically typical noise levels. The phantom oscillated with a peak-to-peak amplitude of 2 cm during the non-static scans.

Syringe size	Scan type	QFreeze	Gate-PM	Gate-SC	Uncorr
Large		-	0.02	-	-
Medium	Static	-	-	-	0.001
Small		0.02	0.02	0.01	0.002
Large			-	-	-
Medium	QFreeze		-	-	0.001
Small			-	-	0.004
Large				-	-
Medium	Gate-PM			0.003	0.004
Small				0.001	0.003
Large					-
Medium	Gate-SC				0.004
Small					0.003

Table 2.4 – Significances of differences between SUV_{max} values of the active volumes in the various image types at low noise levels (phase D of gated images, QFreeze reference phase D). Dashes indicate $p>0.05$.

Syringe size	Scan type	QFreeze	Gate-PM	Gate-SC	Uncorr
Large		-	-	-	-
Medium	Static	-	-	-	-
Small		-	-	-	0.007
Large			0.009	0.004	-
Medium	QFreeze		-	0.05	-
Small			-	-	0.02
Large				-	-
Medium	Gate-PM			-	-
Small				-	0.05
Large					-
Medium	Gate-SC				-
Small					-

Table 2.5 – Significances of differences between SUV_{max} values of the active volumes in the various image types at clinically typical noise levels (phase D of gated images, QFreeze reference phase D). Dashes indicate $p>0.05$.

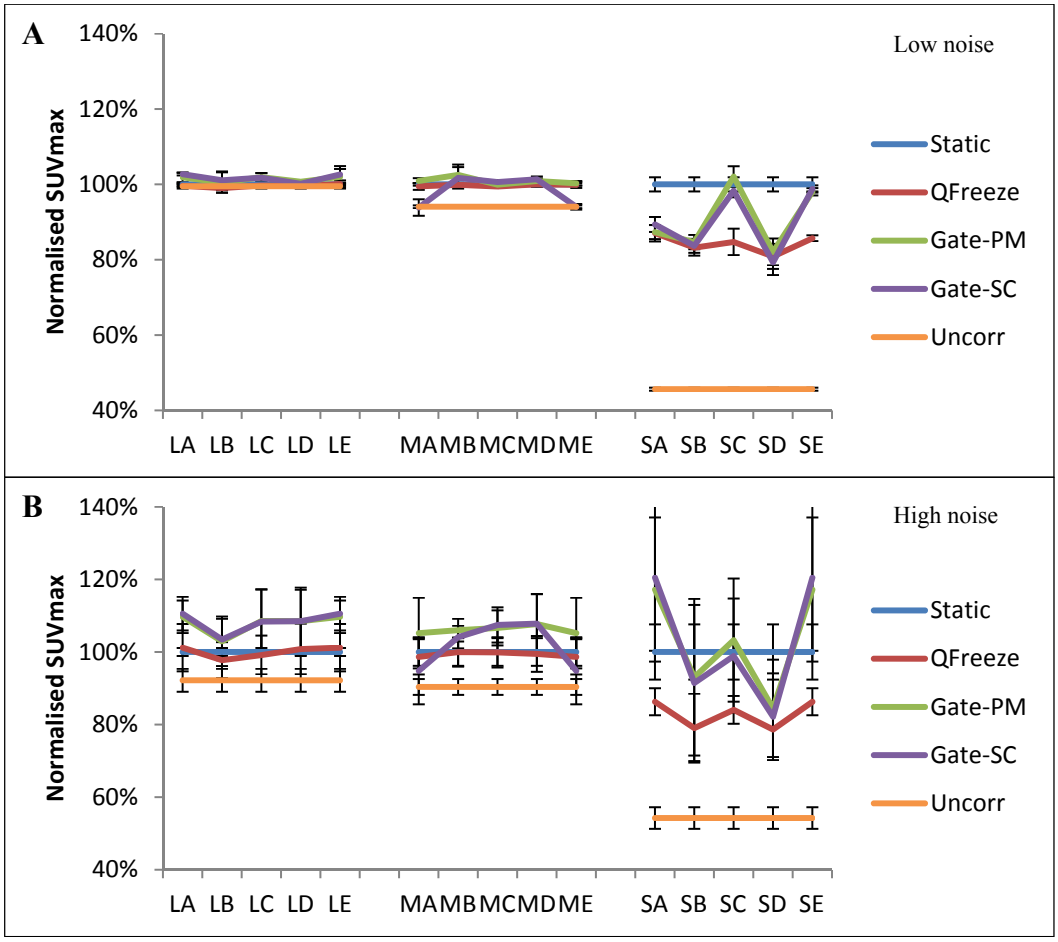


Figure 2.5 – Normalised SUV_{max} values of the large (L), medium (M) and small (S) active volumes in static and uncorrected images, phases (A-E) of gated images, and in QFreeze images generated using each phase as the reference, at (A) low and (B) clinically typical noise levels.

2.4.2 Concentration volume histograms

CVHs are shown in Figure 2.6 for the large active volume in the different image types at low and clinically typical noise levels. Figure 2.7 shows volumes with uptakes greater than 50% of SUV_{max} in the different images (denoted as $V_{50\%}$, and expressed as percentages of the $V_{50\%}$ volumes in static scans) for the small, medium and large active volumes at low and clinically typical noise levels.

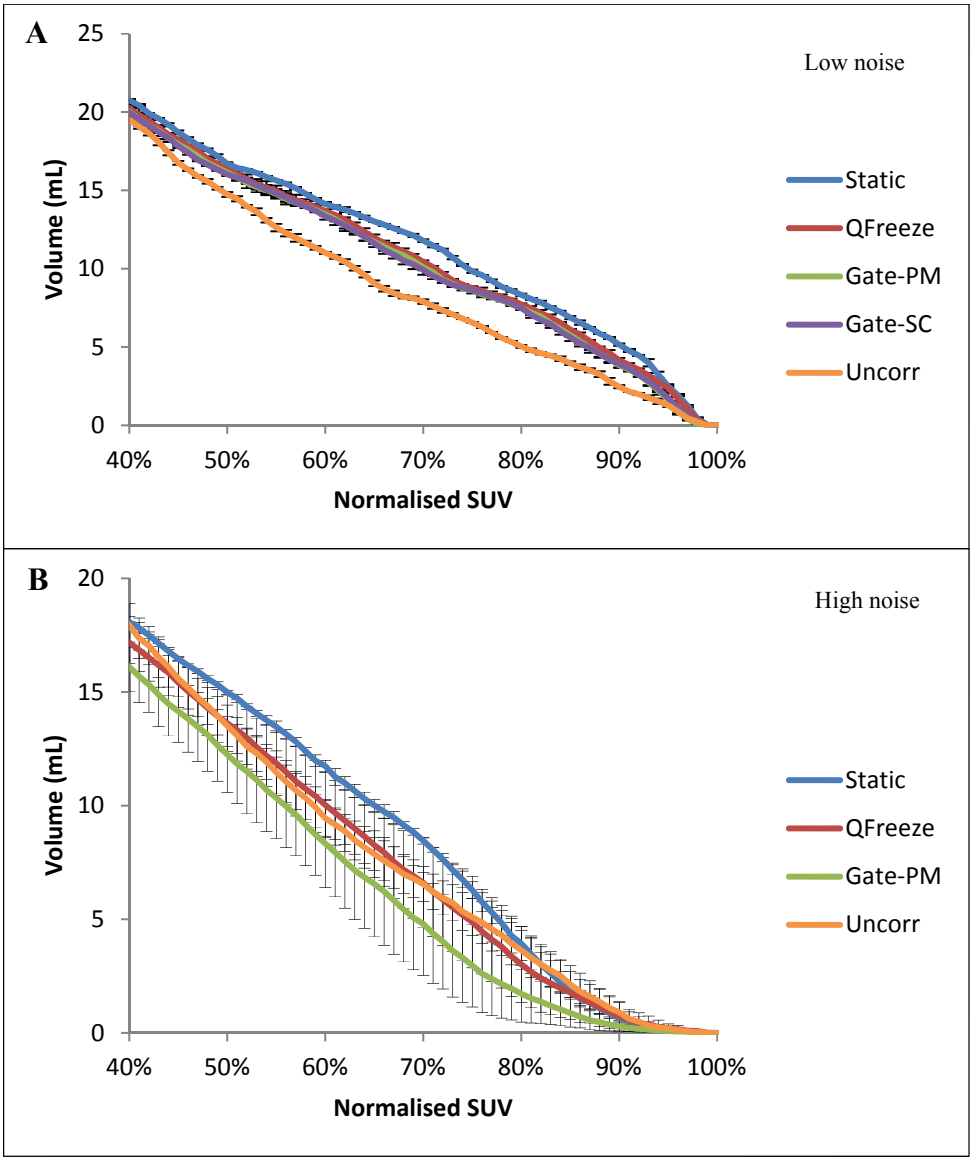


Figure 2.6 – CVHs of the large active volume at (A) low, and (B) clinically typical noise levels (phase D of gated images, QFreeze reference phase D). For clarity the Gate-SC CVH is not shown in B.

Several patterns are apparent in the low noise image data. Firstly, for the small active volume, $V_{50\%}$ is much greater in uncorrected images than in the other images ($p<0.02$); $V_{50\%}$ values in the QFreeze and gated images are slightly, but insignificantly, larger than in the static images. Secondly, this trend is reversed for the medium and large active volumes, with $V_{50\%}$ being significantly smaller in uncorrected scans than in static images ($p<0.02$). Lastly, there are no significant or substantial differences between $V_{50\%}$ values in the gated and QFreeze images.

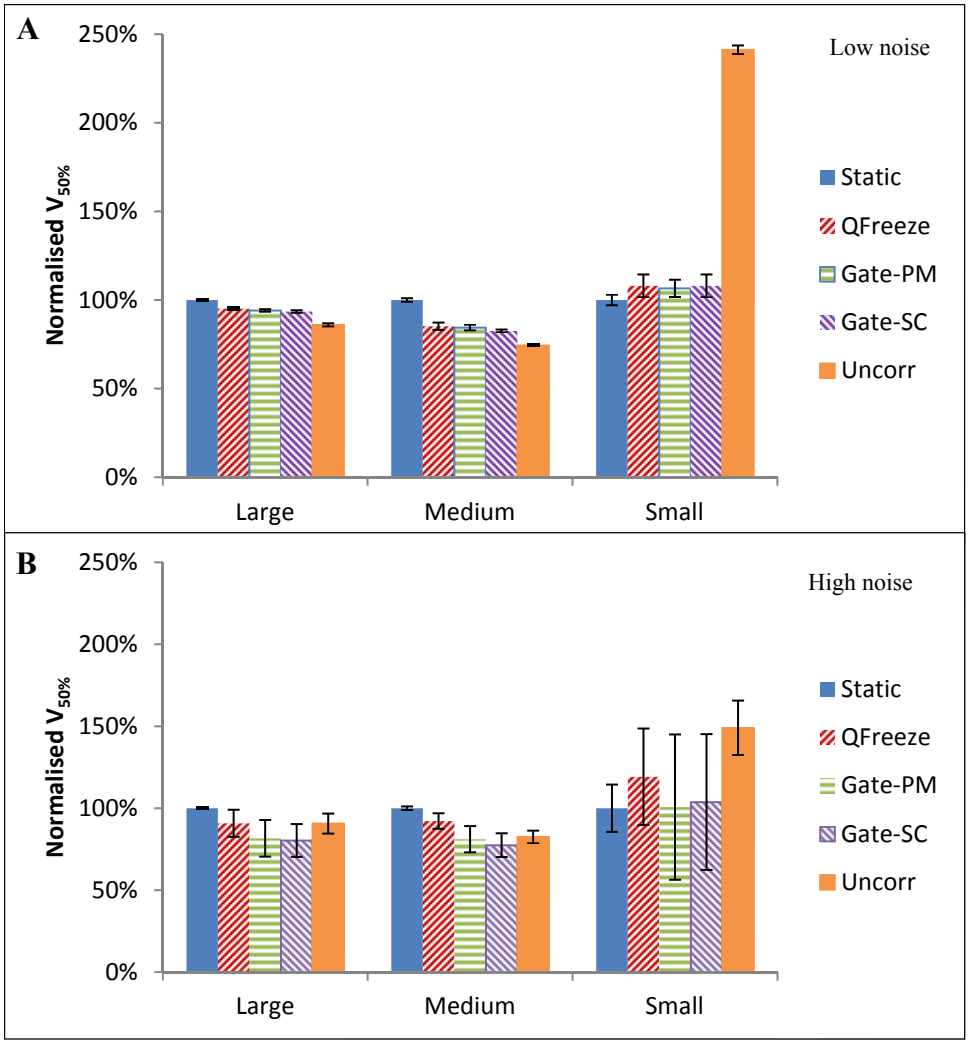


Figure 2.7 – Volumes with imaged uptakes greater than 50% of SUV_{max} in static and uncorrected images, QFreeze reference phase D, and phase D of gated images, plotted as percentages of the volumes with > 50% SUV_{max} uptake in static images for the small, medium and large active volumes, at (A) low, and (B) clinically typical noise levels.

The first pattern occurs because SUV_{max} values are much lower in uncorrected images of the small active volume than in other images; consequently much larger volumes have imaged uptakes greater than 50% of SUV_{max} in uncorrected images. Similarly, SUV_{max} values in the gated and QFreeze images are slightly lower than those in static images, due to residual movement in each image gate, and so $V_{50\%}$ values in gated and QFreeze images are slightly higher than those in static images. The second trend is unrelated to differences between SUV_{max} values in the various images (these differences being very small for the medium and large active volumes) and instead reflects more subtle differences in volumes with a particular absolute level of imaged uptake. Finally, the lack of difference between $V_{50\%}$ values in QFreeze and gated images reflects the low noise levels in all of the images.

At clinically typical noise levels the first two patterns are also observed, but the third is not: $V_{50\%}$ values in gated images are slightly smaller than those in QFreeze images ($p=0.02$ significance for the medium active volume, $p = 0.08$ and 0.09 borderline significance for the small and large active volumes) because at these noise levels SUV_{max} values in gated images are higher than those in QFreeze images. Consequently, at clinically typical noise levels QFreeze provides better estimates of $V_{50\%}$ for the large and medium active volumes than the gated images, but less accurate estimates for the small active volume.

2.4.3 ROC curves

ROC curves were plotted of the sensitivities and specificities for volumes comprised of voxels having uptakes greater than 50% of SUV_{max} in a static phantom image, which can be identified in other images by selecting voxels with imaged uptakes greater than a given fraction (varied from 20-90% in increments of 5%) of SUV_{max} in those images (section 2.3.2, Figure 2.2). ROC curves obtained for the large active volume from images with low

and clinically typical noise are shown in Figure 2.8. Curves obtained for the small active volume are shown in Figure 2.9. The plots show sensitivities and specificities with which boost voxels are identified in QFreeze, Gate-PM and uncorrected images of the moving phantom, and in another scan of the static phantom. ROC curves obtained from Gate-SC images are very similar to those obtained from Gate-PM images, and have therefore not been plotted.

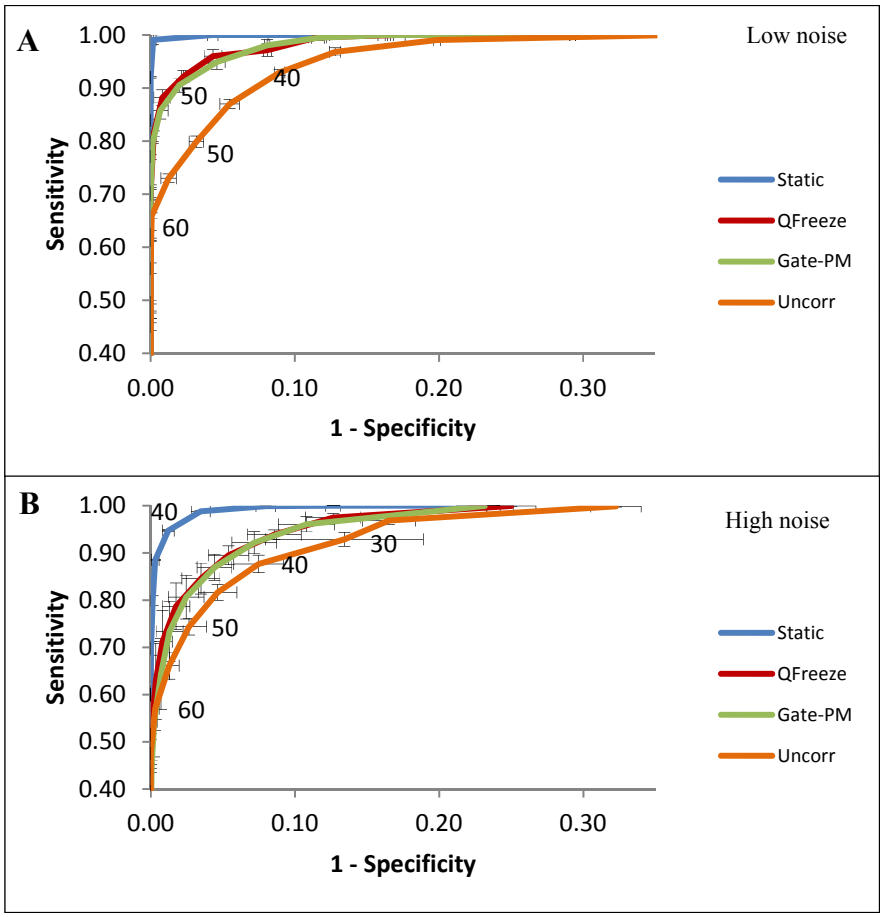


Figure 2.8 – ROC curves for the large active volume in static, QFreeze (reference phase D), Gate-PM (phase D) and uncorrected images at (A) low, and (B) clinically typical noise levels. The numbers by the curves show the percentages of SUV_{max} associated with the nearest points on the curves.

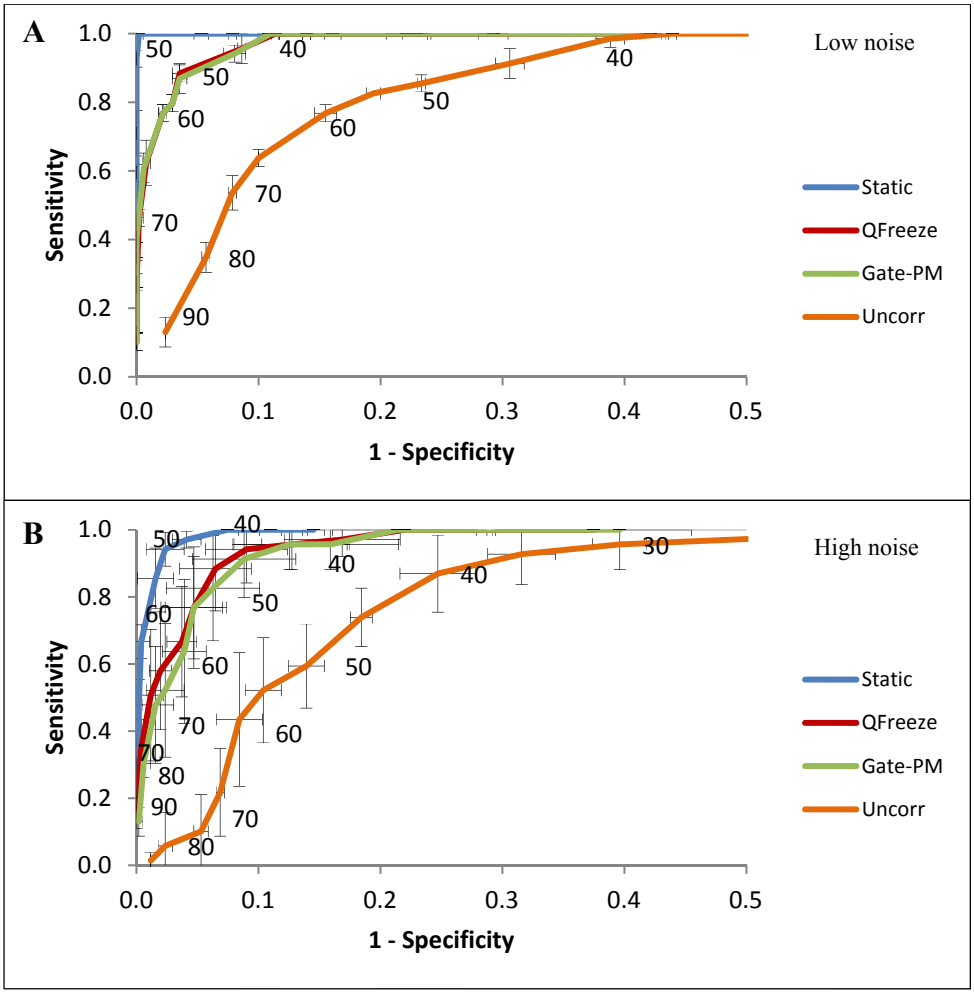


Figure 2.9 – ROC curves for the small active volume in static, QFreeze (reference phase D), Gate-PM (phase D) and uncorrected images at (A) low, and (B) clinically typical noise levels.

In all the graphs, the ROC curves obtained from uncorrected images lie below those obtained from QFreeze and gated images, which in turn lie below those obtained from static images. QFreeze and gated ROC curves are similar, with the gated curves lying just below those obtained from QFreeze images at clinically typical noise levels.

Mean AUCs and standard errors on the means are listed in Table 2.6 and Table 2.7 for the different image types at low and clinically typical noise levels respectively. Significances of AUC differences are shown in Table 2.8 and Table 2.9. AUCs of ROC curves obtained from uncorrected images are significantly smaller than those obtained from static images, by around 2.5% and 12% for the large and small active volumes. For QFreeze and static images, AUCs are also significantly smaller than those obtained for static images, but only by around 1% and 2% for the large and small active volumes. AUCs of ROC curves obtained from QFreeze and Gate-PM images are similar, although for the large active volume at clinically typical noise levels the QFreeze AUC is slightly (0.2%) but significantly (p=0.005) higher.

Syringe size	Static	QFreeze	Gate-PM	Uncorr
Large	1.000(0.000)	0.993(0.002)	0.993(0.003)	0.980(0.002)
Small	1.000(0.000)	0.985(0.002)	0.984(0.002)	0.887(0.009)

Table 2.6 – Mean AUCs for ROC curves obtained from the various image types at low noise levels. Values in brackets are standard deviations on the means.

Syringe size	Static	QFreeze	Gate-PM	Uncorr
Large	0.998(0.001)	0.984(0.003)	0.982(0.003)	0.973(0.003)
Small	0.993(0.007)	0.974(0.005)	0.969(0.005)	0.856(0.049)

Table 2.7 – Mean AUCs for ROC curves obtained at clinically typical noise levels.

Scan type	QFreeze	Gate-PM	Uncorr	Syringe size
Static	0.009	0.01	0.002	Large
	0.003	0.003	0.001	Small
QFreeze		-	4×10^{-4}	Large
		-	0.002	Small
Gate-PM			0.01	Large
			0.002	Small

Table 2.8 – Significances of AUC differences between the various image types at low noise levels. Dashes indicate p>0.05.

Scan type	QFreeze	Gate-PM	Uncorr	Syringe size
Static	0.005	0.004	0.002	Large
	0.004	0.02	0.02	Small
QFreeze		0.005	0.02	Large
		-	0.03	Small
Gate-PM			0.02	Large
			0.03	Small

Table 2.9 – Significances of AUC differences between the various image types at clinically typical noise levels.

In summary, for the large and small active volumes around 50% and 80% respectively of the AUC of ROC curves that is lost due to uncorrected movement can be recovered using QFreeze or gated images. These two image types produce similar results, with QFreeze performing slightly better in the case of the small active volume at clinically typical noise levels.

2.5 Discussion

Much of the distortion introduced into PET scans by respiratory motion is eliminated in QFreeze and gated images. For the small (0.5 mL) active volume studied in this work, around 60% of the error on SUV_{max} and 80% of the loss in AUC of ROC curves was eliminated, irrespective of noise levels. Using uncorrected scans, boost voxels could be identified with a sensitivity of 95% and a false positive rate of 35%, whereas using QFreeze or gated scans the false positive rate was reduced to 10% for the same sensitivity. At low noise levels, roughly 95% of the error on $V_{50\%}$ for the small active volume was eliminated using these images.

QFreeze and gating perform very similarly at low noise levels, however at clinically typical noise levels the greater degree of noise in the gated images leads to certain differences: for the small active volume, SUV_{max} values in static scans are more accurately matched by those in gated rather than QFreeze images; whereas for the medium and large (10 and 20 mL) active volumes, the static SUV_{max} values are more closely matched by QFreeze than by gated SUV_{max} values. Likewise, for the small active volume, $V_{50\%}$ values in gated images are closer to the static image $V_{50\%}$ values than those in QFreeze images, but for the medium and large active volumes QFreeze performs better. Overall, QFreeze represents a marginal improvement at clinically typical noise levels, having a slightly but significantly higher AUC for the ROC curve of the large active volume. Notice that in this study scans were split into five phase intervals: had more phases been used, the gated images would have been noisier and differences between QFreeze and gated images larger. Fives gates were used as it is the number used clinically in Oxford; other sites have also reported on work using five gates (Nehmeh *et al.*, 2003).

The experiments described here are a simplification of the clinical situation, in which respiration is not entirely regular, baseline shifts occur, and tumours and surrounding tissues deform rather than translating rigidly, all of which present greater challenges for systems aiming to mitigate the impact of movement on imaging. Thus our results likely provide upper estimates of the gains achievable using QFreeze and gated images. Unlike patients, the Perspex phantom we imaged contained no background activity (in common with previous work (Wollenweber *et al.*, 2012)); however average activity concentrations throughout patients' bodies are low in comparison with typical uptake levels reported for non-small cell lung cancers (average SUV being 1 by definition compared to 15 for tumour SUV_{max} (Ozgül *et al.*, 2013)), and the phantom did contain multiple active volumes.

Future work is planned to address the limitations in this study. A body phantom is currently being built (based on a human MRI scan) that will allow different sizes of lesions to be placed in it, surrounded by background activity. This body phantom, when made, will be placed on a new movement platform allowing a range of more realistic patient respiratory movements.

2.6 Conclusion

In conclusion, the impact of regular respiratory motion on PET images can be substantially reduced using gating or the QFreeze system, potentially improving the specificity with which tumour subvolumes can be identified for boosting in RT dose-painting. There was relatively little difference between QFreeze and gating seen in this study. However, QFreeze provided very small, but statistically significant, advantages over gating at clinically typical noise levels.

3 Buparlisib Trial

3.1 Abstract

A clinical trial is being run to investigate the impact of a potentially hypoxia-modifying drug, Buparlisib, on perfusion and hypoxia in NSCLC patients. The trial design is a 3+3 dose escalation study exploring the effect of daily doses of 50 mg, 80 mg, and 100 mg Buparlisib. The imaging protocol of the trial is described and characteristics of the nine patients recruited to date are detailed. Some representative FMISO images are presented.

The results of an analysis of tumour volumes with tumour-to-blood ratios (TBR) greater than 1.4 (considered to indicate hypoxia) are described: specifically, changes in these hypothesised hypoxic volumes before and after one week of Buparlisib have been determined. The hypothesised hypoxic volumes generally decrease following daily administration of more than 50 mg of Buparlisib. A significant difference ($p=0.02$) was seen in this measure between Cohort 1 (50 mg daily) and Cohorts 2 and 3 (80 and 100 mg daily respectively).

3.2 Introduction

3.2.1 Trial background

In a single-centre trial underway in Oxford (NCT02128724, 'CR-UK Phase I Study of BKM120 in Patients with Non-small Cell Lung Cancer Receiving Thoracic Radiotherapy'), the impact of Buparlisib (Novartis, Basel, Switzerland) on perfusion and hypoxia levels in NSCLC is being explored. The primary objective of the trial is to determine the MTD of Buparlisib in NSCLC patients when given concurrently with palliative radiotherapy. Secondary objectives are to measure changes in PET-imaged FMISO uptake, indicating changes in tumour hypoxia, and changes in tumour perfusion parameters obtained from analysis of perfusion CT images. The trial is a single-centre open-label 3+3 cohort study of dose escalation; FMISO PET and perfusion CT images of patients are obtained at baseline and seven days after administration of Buparlisib without any other intervention.

3.2.2 FMISO scans

FMISO is transported between the vasculature and tumour cells by passive diffusion, which is a slow process. FMISO PET imaging is carried out several hours post-injection, to allow influx of the radiotracer to hypoxic cells and washout from normoxic cells (Prekeges *et al.*, 1991). Dynamic PET imaging of FMISO uptake at multiple time points has also been investigated alongside static imaging at a late time point. In this trial both dynamic and static FMISO scan data is collected and analysed in order to maximise the information gathered.

Late static FMISO images are commonly analysed using the measures of TBR or tumour to muscle ratio (TMR) as a means of classifying individual voxels as either hypoxic or not, rather than the SUV measures frequently used in PET (Koh *et al.*, 1995; Gagel *et al.*, 2006;

Thureau *et al.*, 2013; Vera *et al.*, 2011). Various researchers have used a range of thresholds signifying supposedly hypoxic voxels, ranging from 1.2-2 (Swanson et al., 2009; IND, 2013; Zips et al., 2012; Okamoto et al., 2013). Most commonly, voxels with a TBR greater than 1.4 are considered to be hypoxic, and this is the threshold used throughout this work, in which volumes with TBR ≥ 1.4 (often termed 'hypoxic volumes') are denoted as '$V_{TBR1.4}$'. This work also uses a TBR colour scale such that red regions depict a TBR greater than 1.4 and no visible PET depicts a TBR lower than 1.

3.3 Materials and methods

3.3.1 Trial details

The trial has been set up as shown in Figure 3.1, with patients being imaged pre-Buparlisib and after seven days of Buparlisib. The amount of drug varies according to cohort: for Cohort 1, a dose of 50mg Buparlisib was administered once daily, patients in Cohort 2 were given 80 mg, and those in Cohort 3 are given 100mg. The imaging protocol for FMISO PET/CT is described in detail below, while the protocol for pCT is described in Chapter 6.

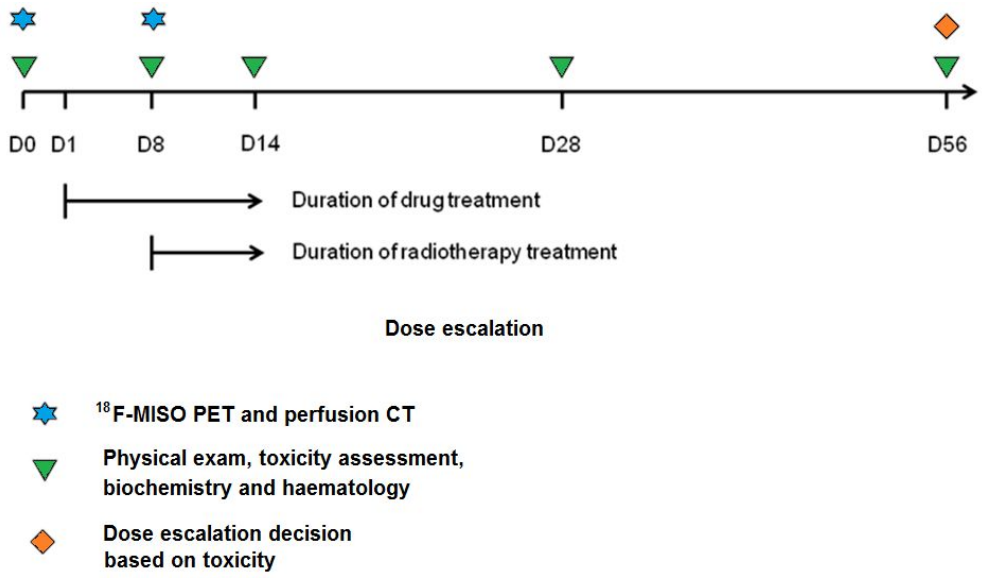


Figure 3.1 – Trial schematic (Higgins *et al.*, 2014).

3.3.2 Trial imaging protocol

The FMISO PET/CT imaging protocol consists of three scans in separate time-frames: 0-45 mins, then a ten-minute scan at 2 hours post-injection, and a further ten-minute scan at 4 hours post-injection (immediately followed by the pCT scan), as shown schematically in Figure 3.2. Patients are imaged in a supine position with their arms by their side in a single bed position using a GE Discovery 690 PET/CT scanner (GE Healthcare, Milwaukee, USA). They are injected with 370 MBq ^{18}F-FMISO 30 seconds after the start of the first 45-minute scan. The four hour image ensures good contrast between hypoxic tissue and background. A separate CT scan is performed for localisation and PET attenuation correction prior to collection of PET data in each scan.

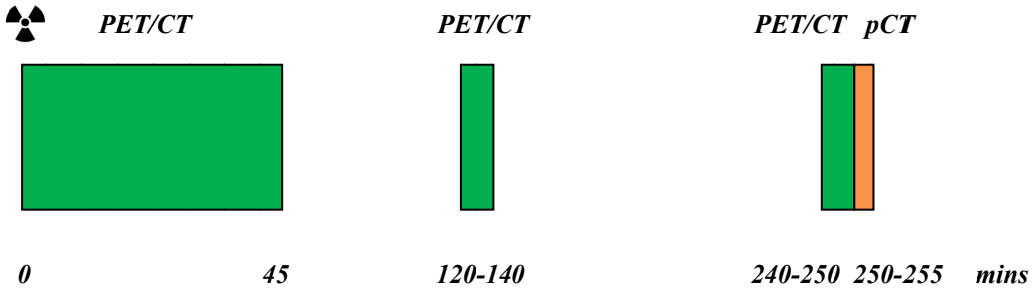


Figure 3.2 – Trial imaging protocol schematic.

3.3.3 Analysis of late FMISO images

Analysis of the four hour FMISO images has been performed using the Hermes Hybrid Viewer software (Hermes Medical Solutions AB, Sweden). Tumour outlining has been carried out with the assistance of a PET consultant radiologist, who contoured the tumour regions on each CT image obtained from the dynamic, two hour, and four hour PET/CT scans. The outlined volumes were transferred to the associated PET images (the PET and CT images being hard-registered via the PET/CT imaging system). If any metastases or involved nodes were identified within the PET field of view in addition to the tumour, they were also outlined by the radiologist. Where appropriate, prior patient imaging was used to guide tumour outlining. A blood volume was also drawn in the centre of the descending aorta, on at least five PET slices. For each volume outlined, SUV_{mean} values were generated using the Hermes Hybrid Viewer software (Hermes Medical Solutions AB, Sweden). TBR_{mean} values were then calculated by dividing these SUV_{mean} values by the blood SUV_{mean} obtained for the outlined descending aorta.

To calculate the $V_{TBR1.4}$ volumes, voxel-by-voxel SUVs were exported for the outlined volumes and divided by the associated blood SUV_{mean}, after which volumes of voxels with TBRs greater than 1.4 were summed. For each patient the total $V_{TBR1.4}$ denotes the sum of $V_{TBR1.4}$s obtained for the outlined tumour, metastases, and involved nodes within the PET field of view. Fractional pre- to post-Buparlisib changes in $V_{TBR1.4}$ were calculated by dividing the post-Buparlisib total $V_{TBR1.4}$ value by the corresponding pre-Buparlisib $V_{TBR1.4}$. Analysis of the dynamic PET data is considered in Chapters 4 and 5.

3.3.4 4D-PET/CT

4D-PET/CT imaging was attempted for all trial patients using the Varian RPM system described in Chapter 2. However, technical problems with 4D-PET/CT acquisition were encountered for the majority of patients scanned. Initially the IR camera had been mounted on the wall and the breathing trace could not be detected due to the extended distance between the IR camera and patient. Even when the IR camera was moved to the couch, the RPM system was still unable to detect a suitable trace from the patient for the entire procedure. This was principally because the trial patients were positioned feet-first in the scanner (to allow dynamic injection), forcing the marker box to be placed at the superior border of the patient's chest where there is less respiratory movement than at the normal inferior chest position. The limited movement of the marker box in this compromised position was usually insufficient for the gating system to consistently detect a breathing trace.

For patients where at least one 4D-CT in the FMISO imaging procedure was successful (8 out of 9), the maximum observed movement was 5 mm in the cranial-caudal direction for a tumour length of 60 mm in the same direction; this is much less than the worst case of 20 mm movement studied in Chapter 2. Average movements for tumours of length 30-110 mm (median 75 mm) were 3 mm in the cranial-caudal direction with an average fractional movement of 6%. Generally the tumours investigated were situated in the upper lobe which has been shown to experience less breathing movement than tumours in the lower lobe (Seppenwoolde *et al.*, 2002). For consistency this work will analyse all of the patients considering only the 3D-PET/CT scans.

3.4 Results

3.4.1 Patients scanned

To date, analysable FMISO PET data has been collected pre- and post-Buparlisib for nine patients whose basic clinical details are shown in Table 3.1.

Patient	Cohort	Sex	PS	Stage	Histological Type	Tumour Volume (mL)
1	1	F	1	IV	adeno	13
2	1	M	1	IIIa	squam	510
3	1	M	1	IV	adeno	105
4	2	M	1	IV	squam	29
5	2	F	1	IV	adeno	135
6	2	F	1	IV	squam	180
7	3	F	1	IV	adeno	44
8	3	M	0	IV	squam	188
9	3	M	1	IV	squam	40

Table 3.1 – Details of Buparlisib trial patients; 'PS' represents performance status, 'adeno' adenocarcinoma, and 'squam' squamous cell.

The first three patients in Cohort 3 have completed follow-up with no dose-limiting toxicities reported, and thus the MTD of Buparlisib in combination with palliative thoracic radiotherapy has been declared as 100 mg a day. The trial was designed not to investigate daily doses higher than 100 mg as a result of a prior single agent Buparlisib MTD study (Bendell *et al.*, 2012). Currently Cohort 3 is being expanded, requiring six additional patients to be recruited, each also receiving 100 mg of Buparlisib daily. These additional patients will provide supplementary data regarding changes in hypoxia and perfusion post-Buparlisib. However, in this work data will only be reported and analysed for the nine patients who have completed the trial thus far.

3.4.2 Example FMISO PET images

This section shows images illustrating the typical time-course of FMISO uptake. Additionally, late time-point images are presented for two patients, showing post-Buparlisib change and no change respectively. In the next section the late time-point images of all nine patients are analysed quantitatively.

FMISO PET images from one patient, reconstructed from data collected over 10-minute time-frames beginning at 35 minutes and at two and four hours post-injection, are shown in Figure 3.3. FMISO gradually diffuses away in normoxic regions (e.g. muscle) and accumulates in hypoxic regions. FMISO does not accumulate in necrotic tissue, as can be seen in Figure 3.4.

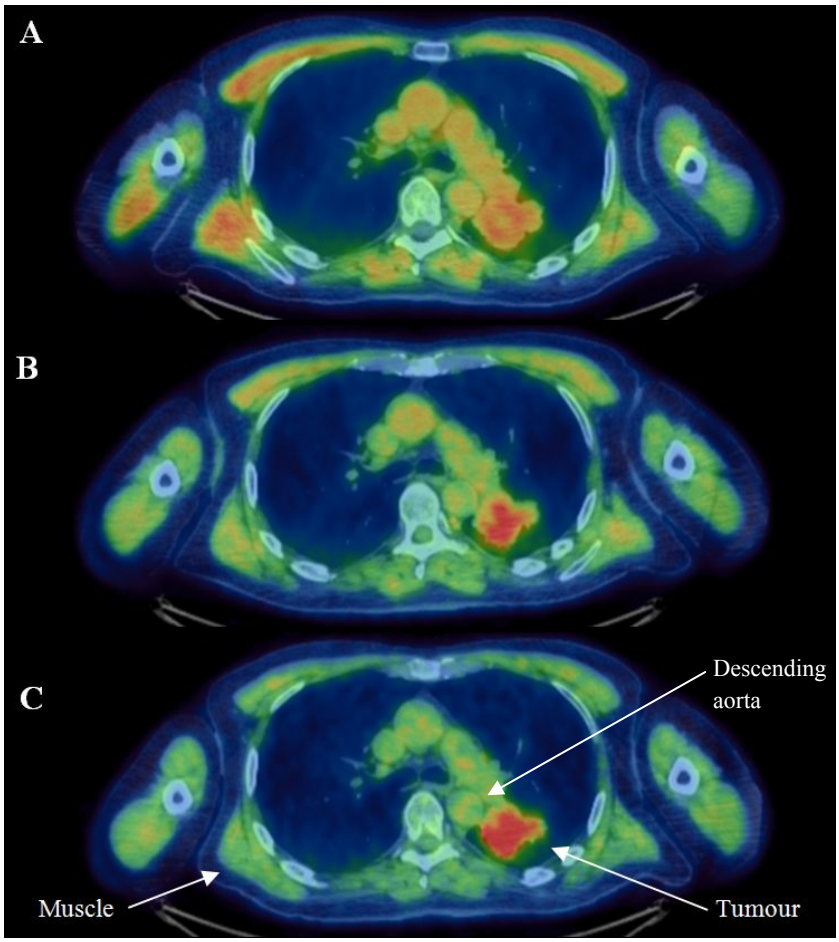


Figure 3.3 – Fused FMISO images for Patient 3_{preC1} at (A) 35 mins, (B) two hours, and (C) four hours post-injection. All images are of 10 minutes duration and use an SUV scale of 0-3, such that red corresponds to SUV values greater than 3.

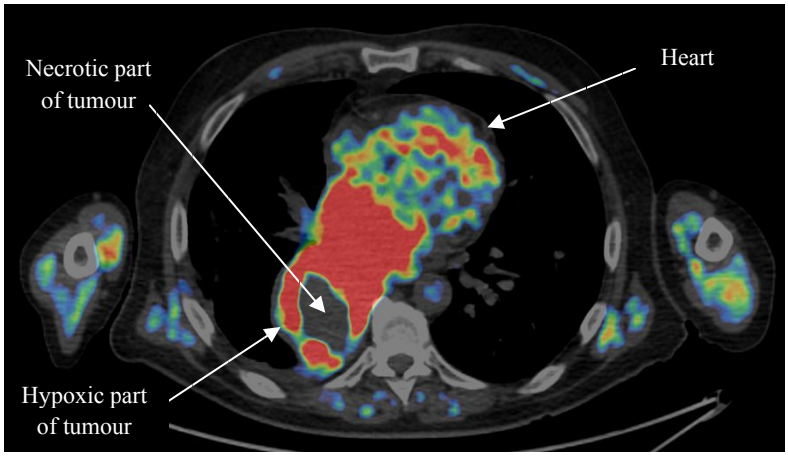


Figure 3.4 – Example four hour FMISO image (from Patient 2_{preC1}) showing no uptake in necrotic part of the tumour (shown to be necrotic in prior CT imaging), using a TBR colour scale 1-1.4.

A single coronal slice is presented in Figure 3.5, showing several involved hypoxic nodes and a tumour. The FMISO scan for this particular patient detected a scapular metastasis that was not detected on routine CT imaging. This was subsequently confirmed by an MRI image (McGowan *et al.*, 2016) and is discussed further in the Appendix.

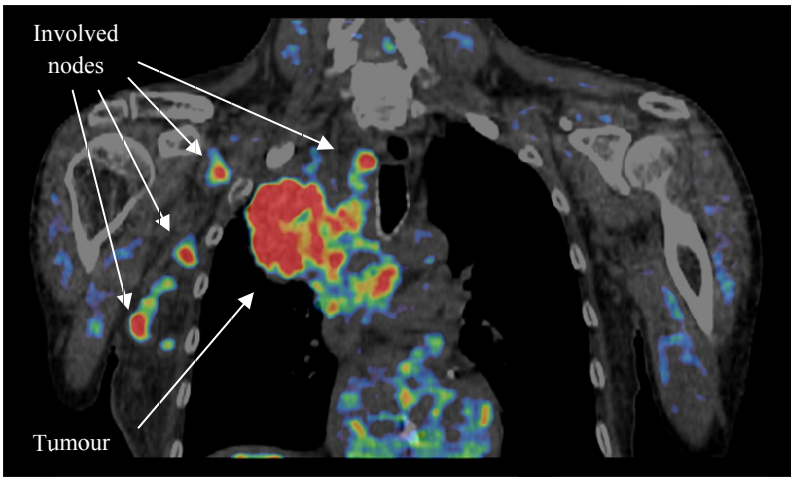


Figure 3.5 – Example four hour FMISO image (from Patient 8_{preC3}) showing several involved nodes and tumour, using a TBR colour scale 1-1.4.

Axial slices taken from pre- and post-Buparlisib FMISO PET images collected at four hours post-injection are shown for one patient showing a change in FMISO uptake (Figure 3.6) and one with no change (Figure 3.7).

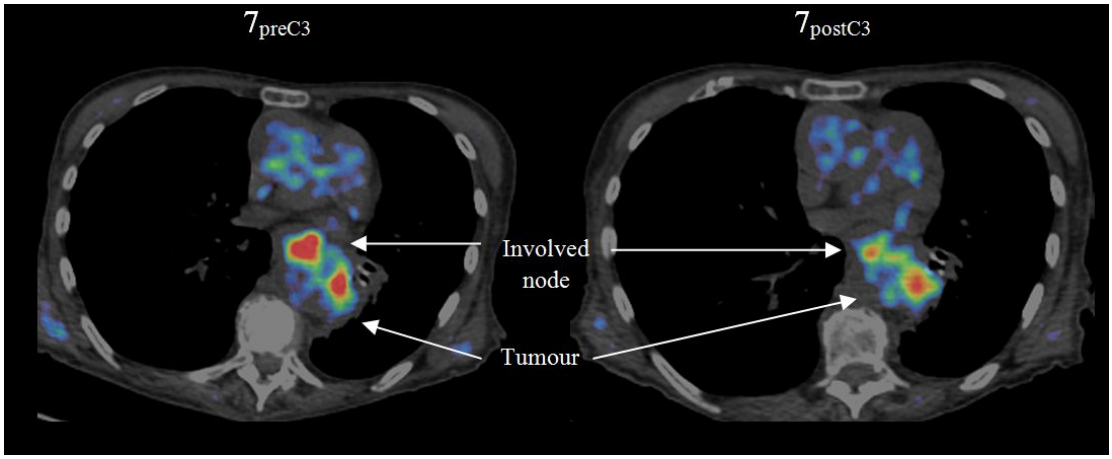


Figure 3.6 – Four hour images showing a change in FMISO uptake from pre-Buparlisib (left) to post-Buparlisib (right), using a TBR colour scale 1-1.4.

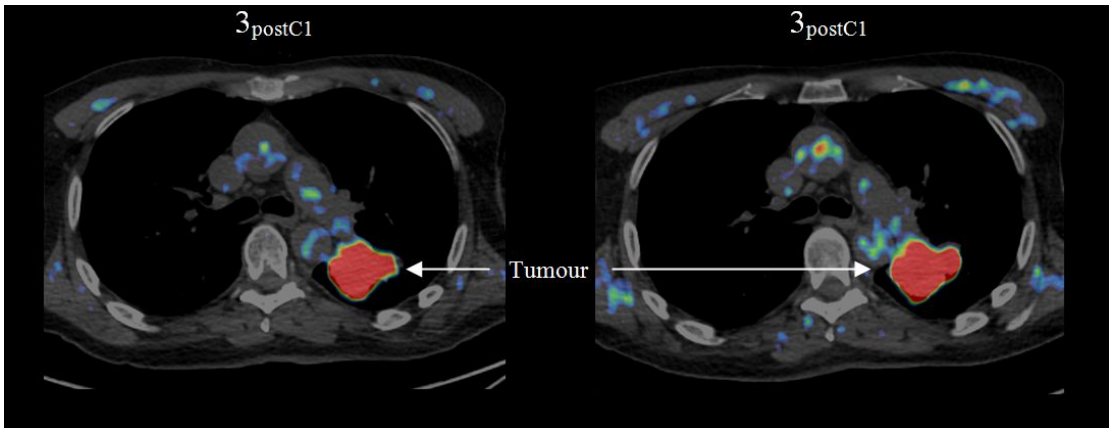


Figure 3.7 – Four hour images showing no change in FMISO uptake from pre-Buparlisib (left) to post-Buparlisib (right), using a TBR colour scale 1-1.4.

3.4.3 Quantitative FMISO data

Table 3.2 lists the metrics of FMISO uptake at four hours post-injection for all outlined volumes (tumours, metastases and involved nodes) of the nine patients imaged so far. Pre-Buparlisib $V_{TBR1.4}$ volumes are plotted against tumour, metastasis, or involved node volume in Figure 3.8, along with fractional $V_{TBR1.4}$ volumes ($V_{TBR1.4}$ divided by volume). The fractional $V_{TBR1.4}$ volumes lie at approximately 40% for volumes greater than 100 mL, showing that size is not a good predictor of hypoxic volume, in common with other work.

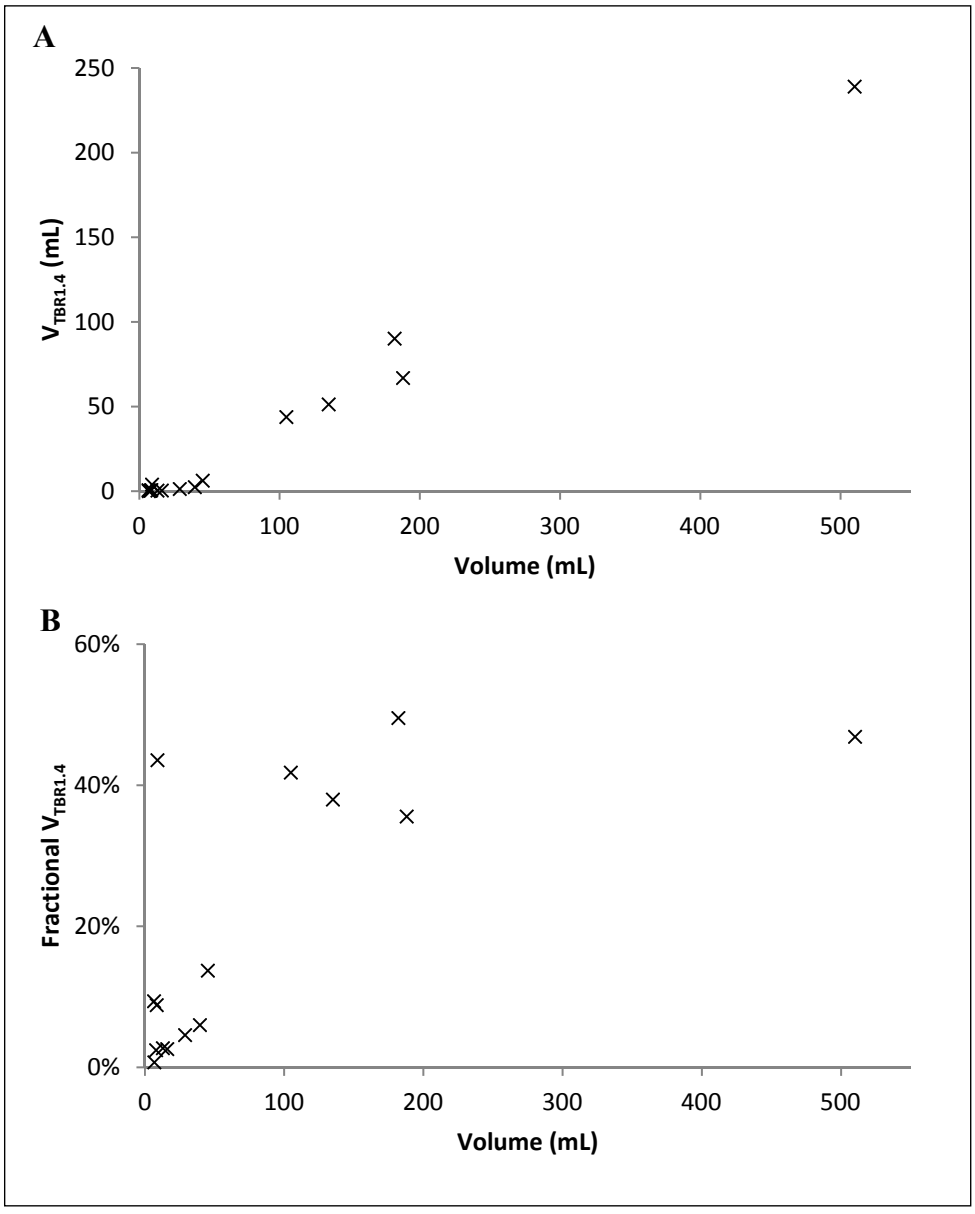


Figure 3.8 – Graph showing the relationship between tumour, metastasis, or involved node volume and $V_{TBR1.4}$ for all TACs pre-Buparlisib.

For smaller volumes the fractional $V_{TBR1.4}$ volumes are lower (<20%) in all but one case. However, for small volumes the partial volume effect is likely to reduce the detected uptake and so will alter the derived parameters (Cherry *et al.*, 2012).

	Patient	SUV_{mean}	TBR_{mean}	Volume (mL)	$V_{TBR1.4}$ (mL)	$\Delta V_{TBR1.4}$ post-Buparlisib
Cohort 1 (50 mg Buparlisib)	1_{preC1}	1.1	0.9	13	0.4	+14%
	1_{postC1}	1.2	1.0	13	0.4	
	2_{preC1}	1.5	1.4	510	239	-3%
	2_{postC1}	1.9	1.5	490	233	
	3_{preC1}	2.1	1.4	105	44	+7%
	3_{postC1}	2.0	1.4	98	46	
	$3M_{preC1}$	1.6	1.0	6.5	0.6	
	$3M_{postC1}$	1.5	1.1	6.7	1.1	
Cohort 2 (80 mg Buparlisib)	4_{preC2}	1.3	0.9	29	1.3	-17%
	4_{postC2}	1.4	0.9	27	1.1	
	5_{preC2}	2.0	1.2	135	51	-18%
	5_{postC2}	1.8	1.1	141	42	
	6_{preC2}	1.6	1.4	180	90	-26%
	6_{postC2}	1.5	1.4	175	70	
	$6N_{preC2}$	1.6	1.4	9.0	3.9	
	$6N_{postC2}$	1.5	1.3	9.2	3.2	
Cohort 3 (100 mg Buparlisib)	7_{preC3}	1.5	1.2	44	6.2	-49%
	7_{postC3}	1.4	1.1	37	3.4	
	$7N_{preC3}$	1.5	1.2	8.5	0.8	
	$7N_{postC3}$	1.4	1.1	6.0	0.1	
	8_{preC3}	1.7	1.3	188	67	-21%
	8_{postC3}	1.5	1.2	182	53	
	$8N1_{preC3}$	1.3	1.0	6.8	0.1	
	$8N1_{postC3}$	1.3	1.1	7.0	0.3	
	$8N2_{preC3}$	1.4	1.1	16	0.4	
	$8N2_{postC3}$	1.3	1.0	15	0.2	
	$8N3_{preC3}$	1.3	1.0	8.0	0.2	
	$8N3_{postC3}$	1.2	1.0	7.0	0.05	
	9_{preC3}	1.6	1.1	40	2.4	-37%
	9_{postC3}	1.6	1.0	37	1.5	

Table 3.2 – Details of trial patients with $\Delta V_{TBR1.4}$ the percentage change from all outlined volumes from pre- to post-Buparlisib; 'M' represents metastases and 'N' node.

Pre- and post-Buparlisib TBR_{mean} values are plotted in Figure 3.9. For Cohort 1 patients the post-Buparlisib values are higher, but for the majority of Cohort 2 and 3 patients they are lower, suggesting that tumour hypoxia may be reduced only in patients receiving daily doses greater than 50 mg Buparlisib. Using a two-tailed paired t-test, differences between pre- and post-Buparlisib TBR_{mean} values are not significant for any individual cohort. However, when Cohorts 2 and 3 are considered together the decrease in TBR_{mean} is significant (p=0.02).

SUV_{mean} data is plotted in Figure 4.3 of Chapter 4, in order to study how it correlates with kinetic parameters of FMISO uptake in comparison to TBR_{mean}. However, in this chapter SUV_{mean} is not plotted because TBR is generally considered to be a more useful measure for FMISO PET (Koh *et al.*, 1992; Fleming *et al.*, 2014).

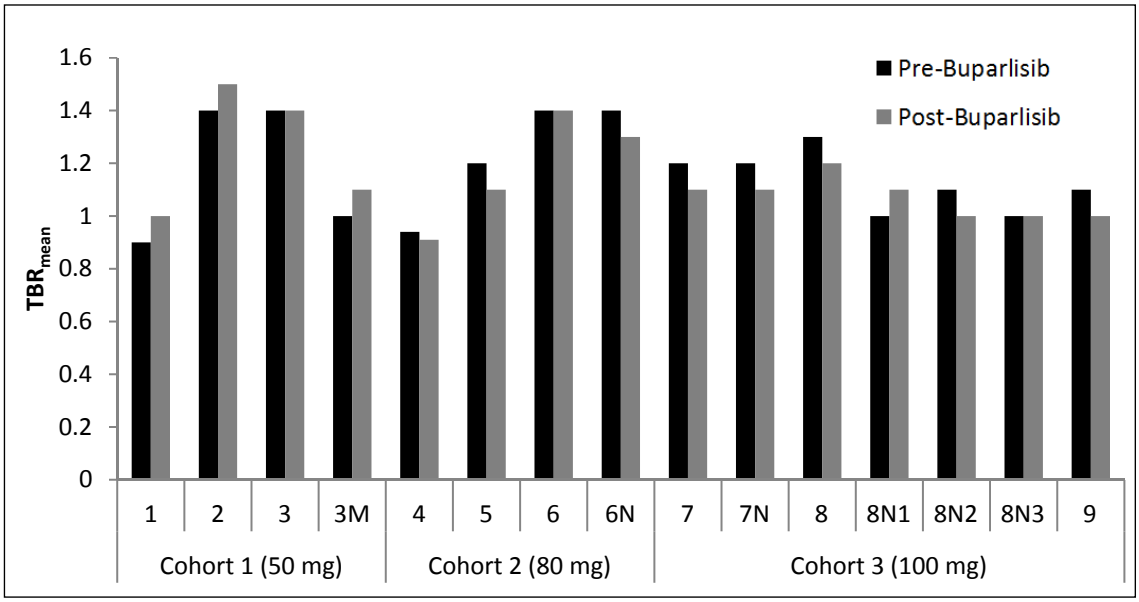


Figure 3.9 – Graphs showing TBR_{mean} values of nine patients pre- and post-Buparlisib; 'M' represents metastases and 'N' node.

Pre- to post-Buparlisib changes in $V_{TBR1.4}$ are shown in Figure 3.10. Percentage rather than absolute changes are graphed as there were large variations in volumes between patients. Data for the outlined tumour, metastases, and nodal volumes summed for each patient is shown in Figure 3.10A, and for tumour volumes alone in Figure 3.10B. Both figures show very similar results. In subsequent chapters analysis is carried out only for FMISO uptake by primary tumours. Data for non-primary tumour volumes has not been included due to the low number (one) of such volumes in Cohorts 1 and 2.

For the three Cohort 1 patients, $V_{TBR1.4}$ increases or only slightly decreases after one week of Buparlisib, whereas for all six Cohort 2 and 3 patients $V_{TBR1.4}$ falls. This mirrors the pattern seen for TBR_{mean} changes and again suggests that a 50 mg daily dose of Buparlisib for a week is insufficient to reduce hypoxia within tumours. Using the two-tailed t-test there is a significant difference between $V_{TBR1.4}$ changes seen in Cohorts 1 and 2 ($p=0.02$), and Cohorts 1 and 3 ($p=0.02$), but with no significant difference between $V_{TBR1.4}$ changes seen in Cohorts 2 and 3 ($p=0.2$).

Figure 3.11 further summarises the data shown in Figure 3.10A, this time plotting the mean $V_{TBR1.4}$ and one standard deviation for each cohort (each individual patient's $V_{TBR1.4}$ includes tumour, metastasis, and nodal volumes). There is a significant ($p=0.005$) positive correlation ($r=0.9$) between the $V_{TBR1.4}$ decrease achieved and the daily dose of drug used. The gradient of the trendline has a 95% confidence interval of 0.5-1%/mg showing a positive relationship between amount of drug and $V_{TBR1.4}$ decrease. Further validation of this finding will be sought in the results from six additional patients currently being recruited to the expanded Cohort 3 (each receiving 100 mg Buparlisib daily).

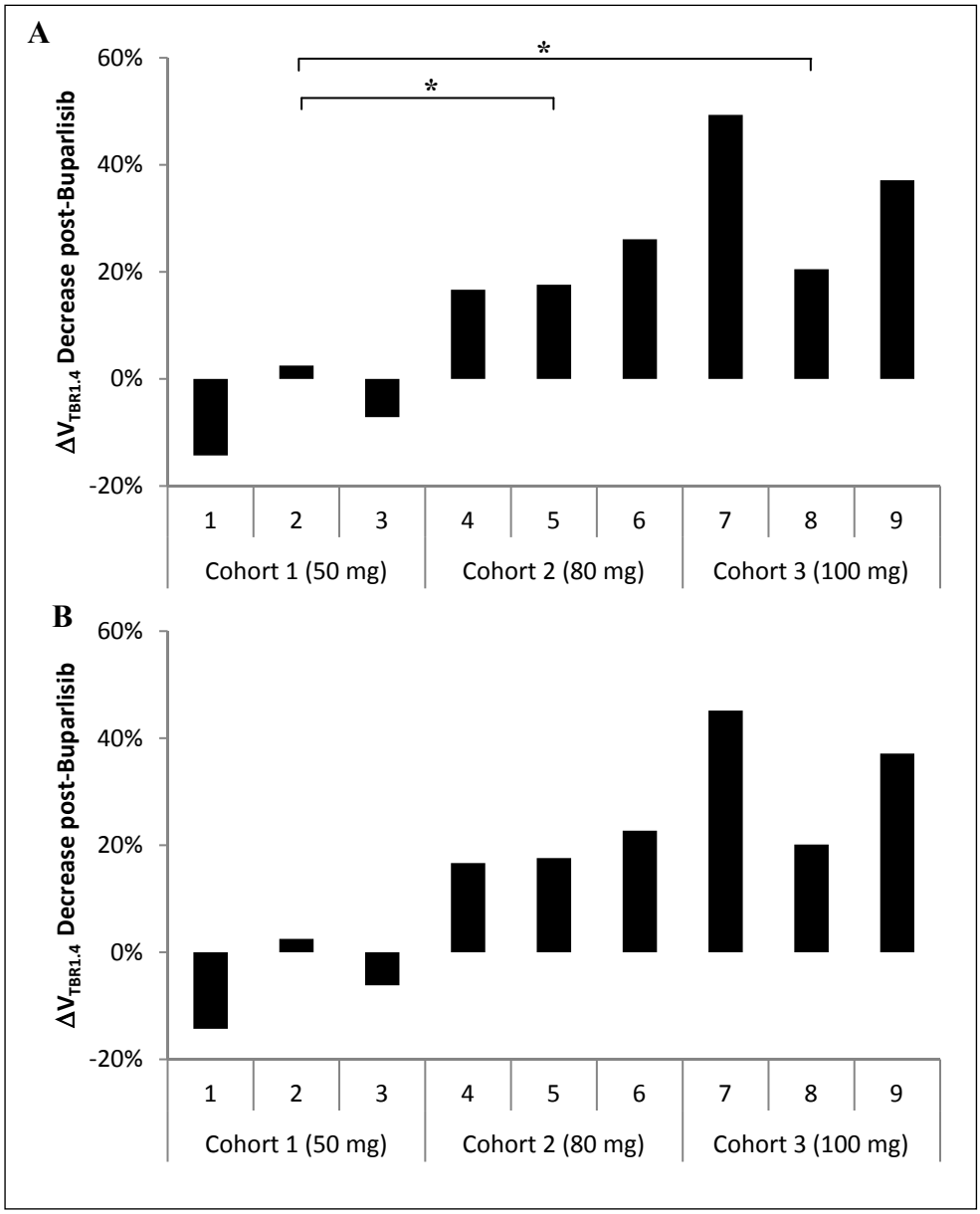


Figure 3.10 – Summary of the decrease in post-Buparlisib $V_{TBR1.4}$ for each patient (original values shown in Table 3.2) as a percentage of the pre-Buparlisib $V_{TBR1.4}$ for (A) all outlined volumes (tumour, metastases, involved nodes) and (B) all outlined tumour volumes. There is a significant difference between Cohorts 1 and 2 (p=0.02) and Cohorts 1 and 3 (p=0.02).

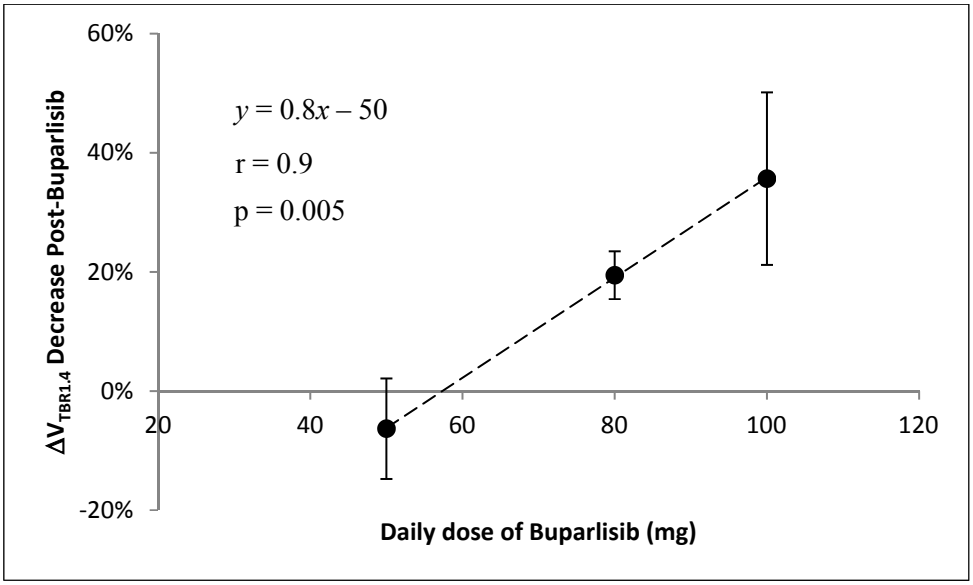


Figure 3.11 – Mean percentage $V_{TBR1.4}$ decrease post-Buparlisib for each cohort with error bars showing one standard deviation patient-to-patient within the cohort (n=3). A linear fit gives a significantly (p=0.005) positive correlation (r=0.9) with a gradient of 0.8%/mg. The 95% confidence interval for the gradient is 0.5-1%/mg.

3.5 Conclusion

An FMISO imaging protocol designed to investigate changes in perfusion and hypoxia as part of the Buparlisib trial has been successfully implemented.

The volume of voxels with a TBR greater than 1.4 (the hypothesised hypoxic volume, $V_{TBR1.4}$) decreases after daily administration of Buparlisib for doses exceeding 50 mg. A significant difference is seen in $V_{TBR1.4}$ between Cohorts 1 and 2 (50 and 80 mg per day, p=0.02) and between Cohorts 1 and 3 (50 and 100 mg per day p=0.02). A significant positive correlation exists between daily dose of Buparlisib and decrease in $V_{TBR1.4}$ (r=0.9, p=0.005).

In Chapters 4 and 5 of this thesis the dynamic FMISO PET scans are analysed in their entirety, exploring time-courses of FMISO uptake and any changes in the uptake kinetics before and after administration of Buparlisib. Later, in Chapter 6 perfusion CT data is analysed, exploring how tumour perfusion changes after Buparlisib.

4 Whole Tumour PET Kinetics

4.1 Abstract

A study is presented of the relative abilities of different kinetic models to describe tumour FMISO uptake TACs obtained from dynamic PET images of patients recruited to the Buparlisib trial. Models were fitted to 30 TACs and the appropriateness of the fits compared using the Wald-Wolfowitz runs test, Akaike and Bayesian information criteria (AIC/BIC), and leave-one-out cross-validation. Statistical simulation procedures were used to determine the accuracy and precision with which the uptake kinetics were estimated by fitted model parameters.

An irreversible three-compartment model, 3C5K, provided the best overall description of FMISO uptake time-courses according to AIC, BIC and cross-validation scores totalled across all TACs. Fits of the 3C5K model yielded the most accurate and precise estimates of FMISO uptake kinetics (10% precision for tumour fractional blood volumes and FMISO binding rate-constants). Whole tumour changes in perfusion and hypoxia before and after administration of Buparlisib have been characterised using fits of this model, and the results compared to the patients' TBR and SUV (given in Table 3.2).

4.2 Introduction

The radiotracer ^{18}F-fluoromisonidazole (FMISO) diffuses passively into cells, where it is reduced and, in hypoxic environments, is irreversibly bound. Thus PET imaging of FMISO uptake can be used to identify hypoxic tumour subvolumes (Rasey *et al.*, 1989; Casciari *et al.*, 1995; Eschmann and Paulsen, 2005). For locally-advanced stage non-small cell lung cancer (NSCLC) the current poor survival rates following chemoradiotherapy (CRT) might be improved by selectively boosting radiation doses delivered to these tumour subvolumes (Rami-Porta *et al.*, 2007; Eschmann and Paulsen, 2005; Zegers *et al.*, 2013; Bollineni *et al.*, 2013; Ling *et al.*, 2000).

To achieve this most effectively requires knowledge of the degree of hypoxia, which can be estimated either from uptake levels in single FMISO images collected 2-4 hours after tracer injection (Koh *et al.*, 1992), or by analysing the kinetics of FMISO uptake in dynamic sequences of PET images (dPET) in order to determine the rate-constant of FMISO intracellular binding. The cellular oxygen concentration is more directly associated with the estimation of the FMISO intracellular binding rate-constant than FMISO uptake at a late time point (Casciari *et al*., 1995). Previous kinetic analysis of FMISO uptake in head and neck cancers has generated indices that correlate with RT outcomes (Thorwarth *et al*., 2005b). In Chapter 3, data from fixed time points was considered, whilst here and in Chapter 5, PET kinetic data is analysed.

Several methods have been used to analyse dPET data, the most common being compartment modelling (Muzi *et al.*, 2012; PMOD, 2012). Time-courses of tumour tracer uptake are often described using a model with two compartments representing the free intratumor and bound tracer, which, together with blood-borne tracer, make up the total tumour tracer uptake (Li *et al.*, 2014; Sokoloff *et al.*, 1977). Figure 4.1 shows a schematic

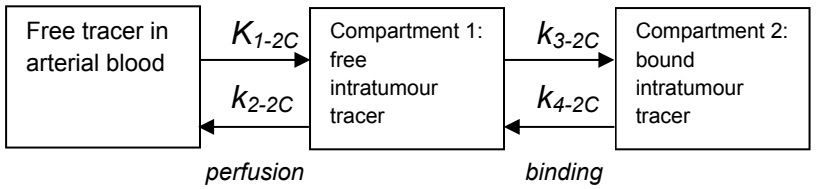
illustration of this model alongside an alternative and less frequently used three-compartment model whose additional compartment describes the tumour interstitium lying between the vasculature and cells (Bertoldo *et al.*, 2001).

In this thesis *x*C*y*K denotes a model comprising a linear chain of *x* tissue compartments (excluding blood-borne tracer) and *y* rate-constants, and so the two- and three-compartment models with reversible flow between each compartment are named 2C4K and 3C6K. Rate-constants are numbered sequentially in pairs, K_1 and k_2 describing reversible flow between the blood and first compartment, for example. For the three-compartment model the K_1 rate-constant is normalised to ensure that the model is uniquely identifiable (Bertoldo *et al.*, 2001). In order to associate rate-constants with particular models the subscript *x*C is added to their names, except for rate-constants of the 3C model.

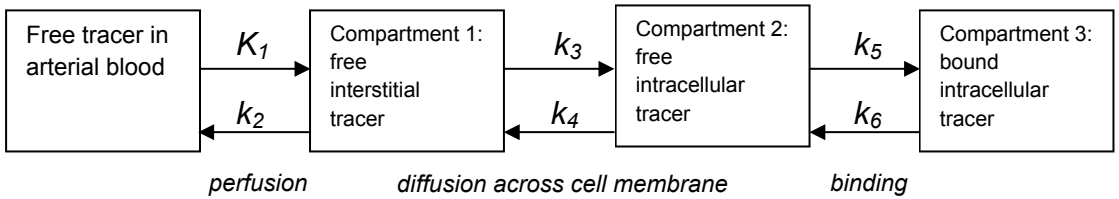
Alongside the 2C4K and 3C6K models it is also useful to explore the simpler two- and three-compartment models 2C3K and 3C5K, in which the rate-constant describing flow from the bound to unbound tracer compartments is set to zero, since binding of reduced FMISO is considered to be irreversible. More complex 4C7K and 4C8K four-compartment models are also illustrated in Figure 4.1, with the additional compartment representing reduced but unbound intracellular FMISO tracer.

In addition to the models already described, which comprise linear chains of compartments, there is also a branching compartment model developed by Casciari *et al.* (1995) that more fully reflects the pathways taken through tumours by FMISO. This consists of ten fitting parameters and so is unlikely to be used in kinetic modelling; it has been simplified into a branching model with three compartments, five effective rate-constants and six fittable parameters, illustrated schematically in Figure 4.2 and denoted as 3C5Kb. For clarity this model will be considered in the discussion (section 4.5) with the results (section 4.4) focusing on comparing the linear chain compartment models.

A. Two-compartment model 2C4K (2C3K for irreversible binding when $k_{4\text{-}2cpt} = 0$)



B. Three-compartment model 3C6K (3C5K for irreversible binding when $k_6 = 0$)



C. Four-compartment model 4C8K (4C7K for irreversible binding when $k_8 = 0$)

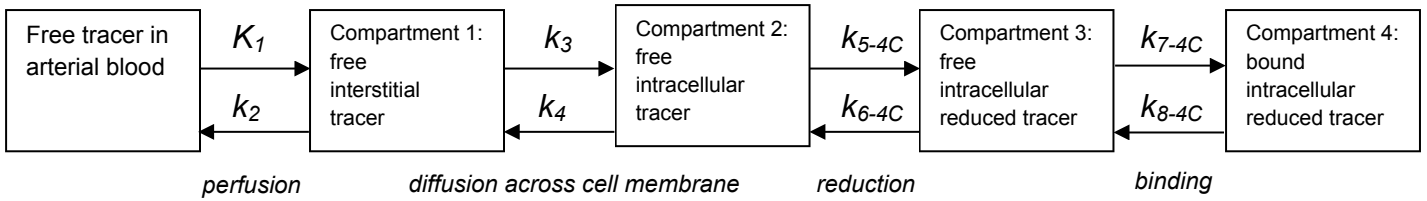


Figure 4.1 – (A) Two-, (B) three-, and (C) four-compartment models. Flow rates from one compartment to another are defined by rate-constants (k values), tracer concentrations in the various compartments, and compartment volumes.

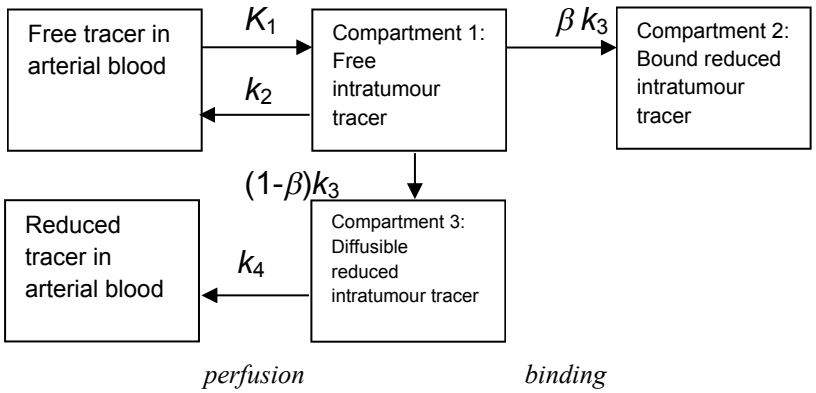


Figure 4.2 – A simplified compartment model, 3C5Kb, based on the FMISO transport schema of Casciari *et al.* 1995. The model effectively has five rate-constants, including βk_3 and $(1-\beta)k_3$ where β is the proportion of reduced FMISO that is bound and retained within the tumour.

4.3 Materials and methods

4.3.1 FMISO scan data

Analysis was undertaken of 30 FMISO TACs, obtained for 15 volumes of interest (VOI) (9 primary tumours, 5 involved nodes, and 1 metastasis) in 9 patients who were each imaged twice, before and after a week of Buparlisib (Table 4.1).

Patient	Cohort	VOIs and TAC reference numbers in images taken pre/post drug administration
1	1	primary (1/2)
2	1	primary (3/4)
3	1	primary (5/6), metastasis (7/8)
4	2	primary (9/10)
5	2	primary (11/12)
6	2	primary (13/14), node (15/16)
7	3	primary (17/18), node (19/20)
8	3	primary (21/22), node 1 (23/24), node 2 (25/26), node 3 (27/28)
9	3	primary (29/30)

Table 4.1 – Details of the 30 TACs analysed for the AIC/BIC/statistical simulation work.

The PET data was reconstructed using a ToF OSEM algorithm (2 iterations, 24 subsets, 6.4 mm Gaussian filter, VPFX, GE Healthcare, Milwaukee, USA). The first 45 minutes of data were binned into two time sequences, S1 (1×30s, 12×5s, 6×10s, 6×30s, 10×60s, 6×300s) and S2 (1×30s, 60×1s, 12×10s, 4×30, 10×60s, 6×300s), and the sequences of images were reconstructed on a matrix of 5.5×5.5×3.3 mm^3 voxels. Data collected during the two ten-minute intervals (at two hours and four hours post-injection) were processed as single frames and reconstructed on a 2.7×2.7×3.3 mm^3 matrix.

Primary tumours, involved nodes and metastases were outlined on each of the images (dynamic, two hour, four hour) by an experienced PET radiologist, and blood volumes were defined within the central part of the descending aorta on five or more consecutive PET axial slices (van der Weerdt *et al.*, 2001). TACs representing time-courses of mean tracer activity concentrations within each of the tumour VOIs and blood volumes were obtained from image sequences S1 and S2 respectively. Activity data from the two later ten-minute frames were appended to the TACs to extend them to four hours post-injection.

4.3.2. Kinetic analysis and model fitting

Kinetic analysis was carried out using the PMOD system (version 3.6, PMOD Technologies, Zurich) and in-house MATLAB code (version R2014a, MathWorks, Natick, Massachusetts). A standard eight-parameter exponential model (Feng *et al.*, 1993; Liu *et al.*, 2014) was fitted to decay-corrected blood TACs, and the model fits were subsequently used as input functions describing tracer flow into tumour VOIs. Model parameters were optimised using the Marquard-Levenberg algorithm to achieve the best match between the measured and modelled blood data as gauged by the weighted sum of squares:

$$SS = \sum_{i=1}^{T} w_i \left(C_{PET}(t_i) - C_{model}(t_i)\right)^2 \quad (4.1)$$

in which $C_{PET}(t_i)$ and $C_{model}(t_i)$ are the imaged and modelled mean activity concentrations at time t_i (the midpoint of the ith of T time-points) and w_i is the relative weighting factor for the time-point, given by

$$w_i = \Delta t_i \exp(-\lambda t_i) \,/\, C_{PET}(t_i) \tag{4.2}$$

where Δt_i is the duration of the i^{th} frame and λ is the decay constant for ^{18}F (Bertoldo *et al.*, 2001; Liu *et al.*, 2014).

Mathematical representations of the two-, three- and four-compartment models of tumour tracer uptake were fitted to the decay-corrected FMISO uptake TACs of the 30 VOIs listed in Table 4.1, using the blood TAC fits as input functions (Sokoloff *et al.*, 1977; Bertoldo *et al.*, 2001). Again, fitted model parameters were optimised to achieve the best match between imaged and modelled TACs according to the weighted sum-of-squares (equation 4.1) with an additional fitting parameter, v_B, representing the fractional tumour volume occupied by the blood.

4.3.3. Assessment of model descriptions of the data

The non-parametric Wald-Wolfowitz runs test was used to determine whether imaged FMISO tumour uptake TACs were adequately described by compartment model fits (Wald and Wolfowitz, 1940; Cobelli *et al.*, 2001). To further assess the relative abilities of the different models to describe the data, we used the Akaike information criterion (AIC) (Akaike, 1974) corrected for small sample size (Burnham, 2004) and the Bayesian information criterion (BIC) (Burnham, 2004). Both were calculated by adding a scaled form of the sum-of-squares of equation 4.1 to terms that increase with the number of model parameters, N, with the best model having the lowest calculated AIC or BIC score. Specifically, the sum of squares was normalised by a scale factor SF so that

$$SS_{scale} = \frac{SS}{SF} = \sum_{i=1}^{T} w_i \frac{\left(C_{PET}(t_i) - C_{model}(t_i)\right)^2}{SF} = \sum_{i=1}^{T} \frac{\left(C_{PET}(t_i) - C_{model}(t_i)\right)^2}{\sigma^2_{est\ C_{PET}(t_i)}} \quad (4.3)$$

In equation 4.3 $\sigma^2_{est\ C_{PET}(t_i)}$ is an estimate of the statistical variance of the imaged mean activity concentration of the VOI, given by

$$\sigma^2_{est\ C_{PET}(t_i)} = \left(SF\ C_{PET}(t_i)\ \exp(\lambda t_i)\right) / \Delta t_i \quad (4.4)$$

The number of radioisotope decays occurring within the VOI is Poisson-distributed, and if all these decays were detected during PET imaging and accurately attributed to the VOI, the scale factor would be the reciprocal of the VOI volume. However *SF* takes a much larger value in practice, due to both noise-propagation during image reconstruction and because many decays go undetected (Bertoldo *et al.*, 2001; Cobelli *et al.*, 2001; Liu *et al.*, 2014). For good fits to TAC data and accurate estimates of the variance of imaged activity concentrations, SS_{scale} has a chi-square distribution with (*T*-*N*) degrees of freedom. For each TAC the *SF* was estimated as the value of *SS*/(*T*-*N*) obtained from the fit of the model with the lowest number of parameters of any model that passed the runs test for that TAC, and was used to estimate the information criteria scores for these TACs.

Model performance was also compared using leave-one-out cross-validation (Picard and Cook, 1984), a method that does not depend on estimated scale factors. Following this approach a model is fitted to a complete dataset minus one point, and the difference between the model fit and the omitted data-point is calculated. This process is repeated sequentially *T* times, leaving out a different data-point each time. Finally all the calculated errors are squared and summed to form the mean square error of prediction (MSEP), with the model with the lowest MSEP being considered best.

4.3.4. Accuracy and precision of fitted rate-constants

The primary purpose of kinetic analysis is to estimate the rate-constants of the various steps in the tracer uptake process. However, the model providing the best description of tumour TACs may not necessarily be the one whose fitted parameter values also lie closest to the true rate-constants, and so a statistical simulation procedure was used to assess which model produces the most accurate and precise rate-constant estimates.

Fits of the 3C5K model to the 30 measured TACs were used to create noise-free 'ground-truth' TACs binned into frames of the same lengths as the original measurements, and parameter values from these fits were taken as ground-truth rate-constants. This model was chosen as it had the lowest AIC/BIC/MSEP of the models investigated (Table 4.2). For each of the 30 ground-truth TACs, 1000 noisy TACs were simulated by adding normally distributed random variables, with variances given by equation 4.4 to the activity concentrations of the individual time-frames (Liu *et al.*, 2014; Bertoldo *et al.*, 2001; Wang *et al.*, 2009; Guo *et al.*, 2009; Muzi *et al.*, 2005). The simulated TACs were then fitted using the 2C3K, 2C4K, 3C5K and 3C6K models.

The simulated noise introduces random uncertainty and systematic error (bias) into fitted parameter values, adding to any bias that might result from mismatches between the models fitted to the simulated TACs and the 3C5K model used to generate them. For each ground-truth TAC the bias in a fitted model parameter was calculated as the difference between the mean of the parameter values obtained from fits to the 1000 simulated noisy TACs and the value of the related ground-truth parameter, and the variance was obtained from the spread around the mean parameter value. Individual biases obtained for the 30 ground-truth TACs were combined to calculate the mean bias (*MB*) and variance of bias values (σ_B^2). The mean variance (σ_P^2) was calculated for each parameter as the average of

the parameter variances obtained for the 30 ground-truth TACs. For any particular TAC, the difference between the mean bias and the unknown individual TAC bias combines to create a total one standard deviation (1 s.d.) uncertainty on a fitted parameter given by

$$\sigma_T = (\sigma_B^2 + \sigma_P^2)^{1/2} \quad (4.5)$$

This total uncertainty is of particular interest since it cannot be eliminated, whereas *MB* represents a constant offset on fitted values, which would effectively cancel out if parameter values were interpreted in the light of research into associations between outcomes and previous parametric images with the same mean bias (Thorwarth *et al.*, 2005; Eschmann and Paulsen, 2005).

MB, σ_B, σ_P and σ_T were calculated for the composite flux constants $k_{\text{flux-2C}}$ and k_{flux} of the fitted two- and three-compartment models (equations 4.6 and 4.7, Bertoldo *et al.*, 2001; Liu *et al.*, 2014).

$$k_{flux-2C} = K_{1-2C}k_{3-2C}/(k_{2-2C} + k_{3-2C}) \quad (4.6)$$

$$k_{flux} = K_1k_3k_5/(k_2k_4 + k_2k_5 + k_3k_5) \quad (4.7)$$

Similarly, *MB*, σ_B, σ_P and σ_T values were also calculated for all individual rate-constants of the fitted models for which directly related rate-constants exist within the ground-truth 3C5K model. Some rate-constants of fitted models are not directly related to any single 3C5K model parameter: processes described by the two-compartment $K_{1\text{-}2C}$ parameter, for example, are split between rate-constants K_1 and k_3 in three-compartment models. For such rate-constants only σ_P values were calculated.

Finally the entire process was repeated, this time generating noise-free TACs from fits of the 3C6K model to the measured TACs, rather than from 3C5K fits, and viewing the ground-truth uptake process as being represented by fitted 3C6K parameter values.

4.4. Results

Numbers of tumour TACs for which the fits of each model pass the runs test are listed in Table 4.2. Total AIC, BIC and MSEP scores summed over all TACs are also shown in the table for the different models, together with the numbers of TACs for which the models achieve the lowest AIC, BIC or MSEP scores. Individual runs test results and AIC, BIC and MSEP scores are presented TAC-by-TAC in Tables 4.3 and 4.4.

The three- and four-compartment models describe the tumour TACs better than two-compartment models: fits of the 4C7K, 3C5K, 2C4K and 2C3K models pass the runs test for 100%, 83%, 20% and 0% of the TACs respectively, and summed AIC, BIC and MSEP scores are much lower for the three- and four-compartment than for two-compartment models, indicating that the simpler two-compartment models lack the flexibility to fully describe the data (Raftery, 1995). Fits of the 3C5K model to TACs 11 and 24 are plotted in Figure 4.3 alongside 2C3K and 2C4K model fits, which visually describe the data less well.

Total AIC, BIC and MSEP scores summed over all tumour TACs are lowest for the 3C5K model. Likewise, most (21-24) individual TAC scores are lowest for the 3C5K model; the generally higher individual scores obtained for the 3C6K, 4C7K, and 4C8K models show that their additional complexity is usually unnecessary to describe the TAC data. Interestingly, the more highly-parameterised 4C7K model whose final stage is irreversible achieves the lowest scores for larger numbers of TACs than the reversible 3C6K model, implying that the inclusion of a reversible final step is not useful when describing the kinetics of the FMISO uptake process.

Model	2C3K	2C4K	3C5K	3C6K	4C7K	4C8K
Runs test passes from fits to all 30 TACs						
Runs	0	6	25	26	30	30
Information criteria and cross-validation scores summed for all TACs						
AIC	13970	8600	1544	1554	1557	1644
BIC	14174	8836	1813	1840	1850	1945
MSEP	219	131	24.6	26.9	24.8	27.9
Numbers of TACs for which each model has the lowest scores						
AIC	0	0	21	3	6	0
BIC	0	0	22	3	5	0
MSEP	0	0	24	3	3	0

Table 4.2 – Summary of runs test results and AIC, BIC and MSEP scores for model fits to all 30 TACs. The lowest AIC, BIC and MSEP scores have been underlined, indicating the best model according to that measure.

TAC	2C3K	2C4K	3C5K	3C6K	4C7K	4C8K
1		✓	✓	✓	✓	✓
2			✓	✓	✓	✓
3					✓	✓
4					✓	✓
5			✓	✓	✓	✓
6			✓	✓	✓	✓
7			✓	✓	✓	✓
8			✓	✓	✓	✓
9		✓	✓	✓	✓	✓
10				✓	✓	✓
11			✓	✓	✓	✓
12			✓	✓	✓	✓
13			✓	✓	✓	✓
14			✓	✓	✓	✓
15			✓	✓	✓	✓
16			✓	✓	✓	✓
17					✓	✓
18					✓	✓
19		✓	✓	✓	✓	✓
20			✓	✓	✓	✓
21			✓	✓	✓	✓
22			✓	✓	✓	✓
23			✓	✓	✓	✓
24			✓	✓	✓	✓
25			✓	✓	✓	✓
26			✓	✓	✓	✓
27		✓	✓	✓	✓	✓
28		✓	✓	✓	✓	✓
29		✓	✓	✓	✓	✓
30			✓	✓	✓	✓

Table 4.3 – TAC-by-TAC results of the Wald-Wolfowitz runs test for fits of the different models (ticks indicate runs test passes).

Models	**2C3K**			**2C4K**			**3C5K**			**3C6K**			**4C7K**			**4C8K**		
TAC	*AIC*	*BIC*	*MSEP*	*AIC*	*BIC*	*MSEP*	*AIC*	*BIC*	*MSEP*	*AIC*	*BIC*	*MSEP*	*AIC*	*BIC*	*MSEP*	*AIC*	*BIC*	*MSEP*
1	170	170	7.4	49	57	2.4	34	43	0.52	37	47	0.52	41	50	0.54	44	54	0.52
2	540	550	8.2	310	320	1.6	51	60	0.70	55	64	0.70	58	68	0.71	62	71	0.70
3	580	590	17	260	270	5.0	78	87	1.8	81	90	3.1	56	65	1.5	59	69	3.1
4	260	270	14	88	96	4.6	62	70	4.2	65	74	4.1	56	65	4.1	59	69	4.1
5	720	730	4.4	490	500	4.7	52	61	0.36	49	59	0.39	52	62	0.37	56	66	0.37
6	800	800	2.5	380	390	1.7	52	61	0.15	54	63	0.15	57	67	0.15	61	71	0.15
7	230	230	4.3	150	160	4.8	51	60	1.1	55	64	1.8	58	67	1.3	62	71	1.5
8	350	350	4.4	130	140	2.5	52	61	0.49	52	62	0.63	59	69	0.57	59	69	0.59
9	140	150	23	48	55	2.2	29	37	1.1	32	41	1.9	33	42	1.6	36	45	1.8
10	590	590	6.2	200	210	2.3	100	120	1.7	54	64	0.46	58	68	1.3	56	66	1.3
11	860	870	19	700	710	18	52	61	0.65	55	65	0.78	59	69	0.85	63	73	0.77
12	400	410	2.6	260	270	4.8	52	61	0.51	37	46	0.51	41	51	0.51	44	54	0.54
13	1100	1100	11	650	660	4.7	51	60	0.29	52	62	0.29	58	67	0.34	59	69	0.31
14	1500	1500	6.4	970	980	5.3	51	60	0.14	54	63	0.17	50	60	0.17	54	64	0.56
15	250	260	21	190	200	9.6	51	59	1.3	54	63	1.3	57	66	1.3	61	70	1.3
16	510	520	15	350	360	10	51	60	0.60	54	63	1.0	45	55	0.4	49	58	1.1
17	1900	1900	6.1	1000	1100	4.0	80	89	0.18	83	93	0.21	57	67	0.15	60	70	0.18
18	490	490	5.8	410	410	5.7	75	84	0.73	72	82	0.70	57	67	0.75	58	68	0.72
19	82	89	4.6	50	58	3.9	24	33	0.48	28	37	0.48	31	41	0.48	34	45	0.48
20	120	120	3.7	100	110	2.7	52	61	0.82	53	62	0.85	71	81	0.84	73	84	0.90
21	330	330	1.7	300	310	3.1	51	60	0.74	54	64	0.74	58	67	0.74	61	71	0.74
22	370	380	3.8	260	270	3.8	52	61	0.33	52	62	0.34	55	65	0.34	58	69	0.34
23	130	140	1.6	130	140	1.8	52	61	0.61	61	71	0.61	59	69	0.61	62	72	0.61
24	510	520	2.9	460	470	4.9	52	61	0.31	55	65	0.31	59	69	0.31	62	72	0.31
25	220	230	2.5	200	210	2.4	52	61	0.59	55	65	0.59	59	69	0.59	62	72	0.59
26	420	430	3.8	210	220	4.3	51	60	0.57	54	65	0.57	58	69	0.57	61	72	0.57
27	67	74	2.6	50	58	2.7	31	40	0.89	38	47	0.89	37	47	0.90	41	51	0.92
28	160	170	5.1	50	58	3.2	42	51	1.4	46	55	1.4	49	59	1.4	52	63	1.4
29	106	113	3.9	50	58	1.5	31	40	0.60	34	44	0.69	37	47	0.66	41	51	0.63
30	84	91	4.5	50	58	2.3	26	34	0.65	29	38	0.68	32	42	0.66	35	46	0.66

Table 4.4 – Individual AIC, BIC and MSEP scores listed by each TAC for the various models. The lowest AIC, BIC and MSEP scores have been underlined for each TAC, indicating the best model according to that measure.

Table 4.5 displays the results of the statistical simulation analysis of parameter accuracy and precision for ground-truths represented by fits of either the irreversible 3C5K or reversible 3C6K three-compartment model to measured TACs. Mean biases and total uncertainties of fitted 2C4K model parameters are very large, irrespective of the ground-truth model used. Mean biases and uncertainties of 2C3K model parameters are not as large as for 2C4K, again demonstrating the superiority of irreversible models for modelling FMISO uptake; however 2C3K model biases and uncertainties are still

somewhat larger than those calculated for three-compartment models. Mean biases and total uncertainties on the parameters v_B, K_1, k_2, k_3 and k_4 are very similar for 3C5K and 3C6K model fits, regardless of the ground-truth model used. For the k_5 and k_{flux} parameters, mean biases and total uncertainties are notably lower for the 3C5K model than for 3C6K when the ground-truth is considered to be represented by fits of the irreversible 3C5K model to measured TACs; however when 3C6K model fits are instead viewed as ground-truth, biases and uncertainties in k_5 and k_{flux} are similar for the two models. Overall, the 3C5K model provides the most accurate and precise estimates of FMISO uptake kinetics.

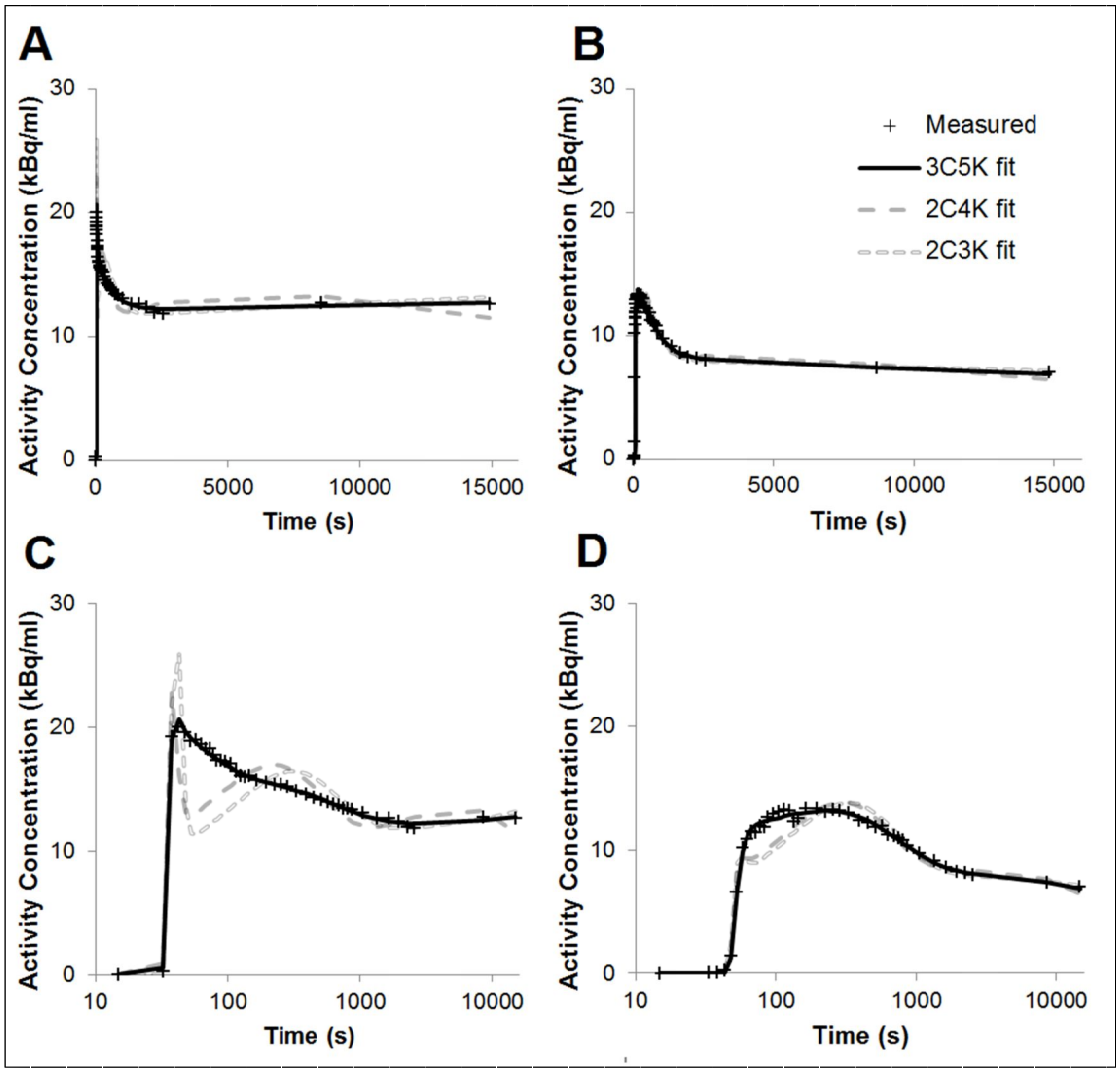


Figure 4.3 – Fits of the 2C3K, 2C4K and 3C5K models to TACs 11 (plots A and C) and 24 (B and D). Time post-injection is plotted on linear (A/B) and logarithmic (C/D) scales.

Ground truth 3C5K model								
Model fitted	*Fitted model parameters*							
2C3K	$v_{B\text{-}2C}$	$K_{1\text{-}2C}$	$k_{2\text{-}2C}$			$k_{3\text{-}2C}$		$k_{flux\text{-}2C}$
MB (%)	47	-	-			-40		25
σ_B (%)	24	-	-			23		11
σ_P (%)	14	30	36			12		13
σ_T (%)	27	-	-			26		17
2C4K	$v_{B\text{-}2C}$	$K_{1\text{-}2C}$	$k_{2\text{-}2C}$			$k_{3\text{-}2C}$	$k_{4\text{-}2C}$	$k_{flux\text{-}2C}$
MB (%)	25	-	-			565	-	1014
σ_B (%)	20	-	-			700	-	1049
σ_P (%)	9	7	13			73	45	47
σ_T (%)	22	-	-			704	-	1050
3C5K	v_B	K_1	k_2	k_3	k_4	k_5		k_{flux}
MB (%)	0	-1	0	1	0	0		0
σ_B (%)	4	1	2	4	3	1		0
σ_P (%)	10	5	10	20	11	10		6
σ_T (%)	10	5	10	20	12	10		6
3C6K	v_B	K_1	k_2	k_3	k_4	k_5	k_6	k_{flux}
MB (%)	0	0	1	5	3	13	-	13
σ_B (%)	4	1	2	7	5	17	-	16
σ_P (%)	10	5	9	22	13	34	384	31
σ_T (%)	11	5	9	23	14	38	-	35
Ground truth 3C6K model								
Model fitted	*Fitted model parameters*							
2C3K	$v_{B\text{-}2C}$	$K_{1\text{-}2C}$	$k_{2\text{-}2C}$			$k_{3\text{-}2C}$		$k_{flux\text{-}2C}$
MB (%)	48	-	-			-51		4
σ_B (%)	24	-	-			49		33
σ_P (%)	13	36	42			11		10
σ_T (%)	27	-	-			50		35
2C4K	$v_{B\text{-}2C}$	$K_{1\text{-}2C}$	$k_{2\text{-}2C}$			$k_{3\text{-}2C}$	$k_{4\text{-}2C}$	$k_{flux\text{-}2C}$
MB (%)	25	-	-			510	3173	914
σ_B (%)	20	-	-			748	3768	1122
σ_P (%)	11	16	36			65	44	40
σ_T (%)	23	-	-			751	3768	1123
3C5K	v_B	K_1	k_2	k_3	k_4	k_5		k_{flux}
MB (%)	0	-1	-2	-5	-5	-18		-11
σ_B (%)	3	1	2	7	5	21		18
σ_P (%)	10	5	10	21	11	10		8
σ_T (%)	10	5	10	22	12	23		19
3C6K	v_B	K_1	k_2	k_3	k_4	k_5	k_6	k_{flux}
MB (%)	-1	0	1	4	3	8	86	6
σ_B (%)	3	1	2	5	4	9	131	8
σ_P (%)	10	5	9	21	13	27	164	22
σ_T (%)	10	5	9	21	13	29	210	24

Table 4.5 – Estimates of the accuracy and precision of parameter values obtained from fits of the 2C3K, 2C4K, 3C5K, 3C6K models to TAC data simulated using ground-truth 3C5K or 3C6K models. Values of *MB*, σ_B, σ_P and σ_T are shown for fitted parameters as percentages of the mean values of directly related ground-truth parameters of the 3C5K or 3C6K models. When no directly related parameter exists, σ_P values alone are shown as percentages of the mean fitted parameter values.

Tumour TBR_{mean} and SUV_{mean} obtained from four hour post-injection PET images (Table 3.2) are plotted in Figure 4.4 against the k_{flux} and k_5 rate-constants of 3C5K model fits to the 30 measured TACs. TBR_{mean} correlates slightly more strongly with k_{flux} than SUV_{mean} (Pearson r coefficient of 0.83 versus 0.74). Correlations at two hours are weaker, both for k_{flux} (r = 0.73 at two hours versus 0.83 at four hours for TBR_{mean}) and for k_5 (0.54 at two hours versus 0.71 at four hours for TBR_{mean}).

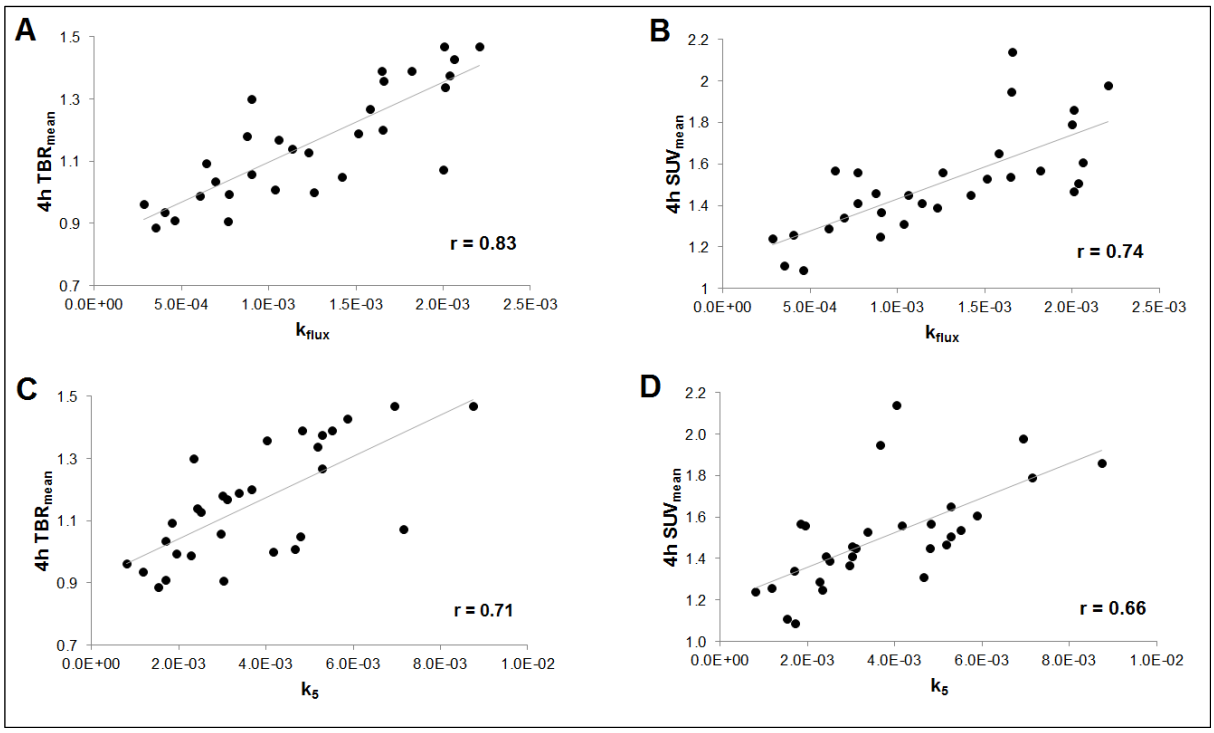


Figure 4.4 – Tumour to blood ratios (TBR_{mean}) and tumour mean standardised uptake values (SUV_{mean}) at four hour post-injection plotted against k_{flux} and k_5 rate-constants of 3C5K model fits to the 30 measured TACs.

The kinetic parameters of patients pre- and post-Buparlisib are shown for the 3C5K model in Figure 4.5. Only tumour data has been considered for this set of graphs, in order to match the studies of voxel kinetics (Chapter 5) and pCT (Chapter 6). Figure 4.5 shows that independent of cohort there are significant increases in v_B (p=0.02) and decreases in K_1 (p=0.001) using a paired t-test. The rate-constant associated with hypoxia, k_5, increases after Buparlisib in Patients 1-5 and then decreases for Patients 6-8 with no change for Patient 9 (not significant over the nine patients). A paired t-test was chosen to assess Buparlisib response as the error on the parameters is considerably less than the patient-to-patient parameter variability.

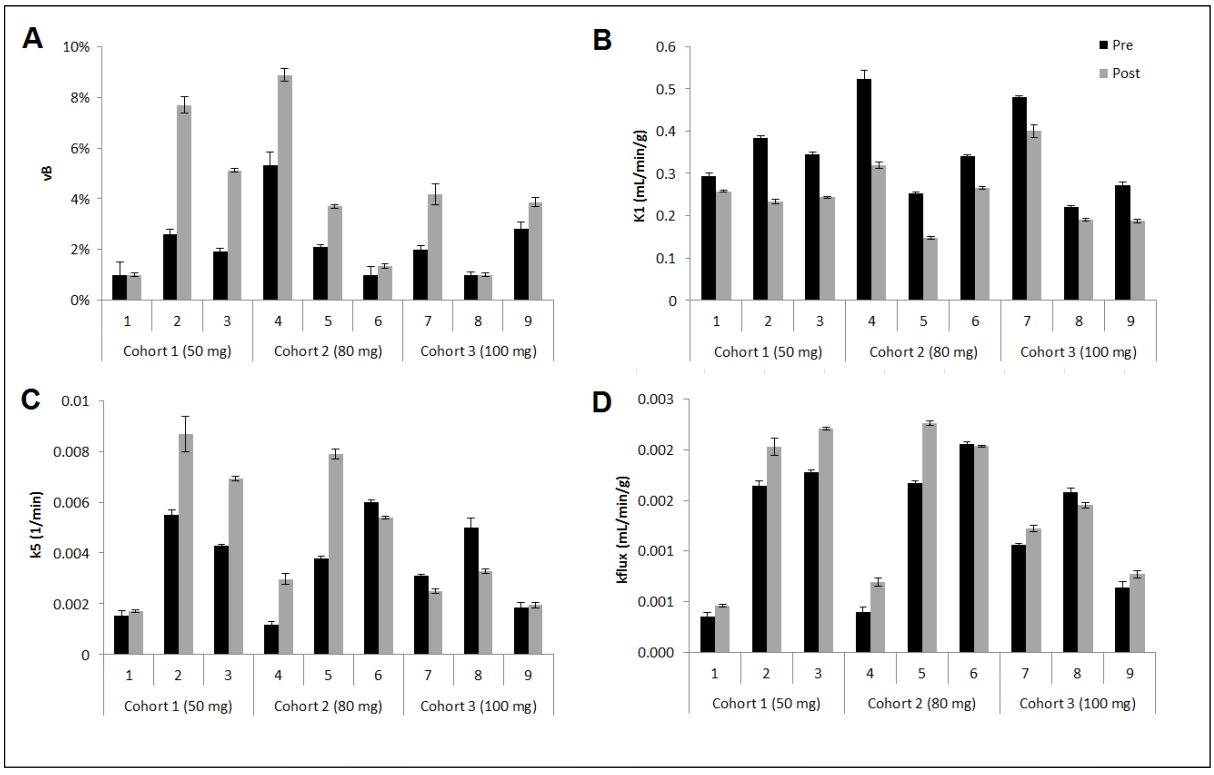


Figure 4.5 – Graphs showing values of 3C5K kinetic parameters for whole tumours of nine patients pre- and post-Buparlisib for (A) v_B, (B) K_1, (C) k_5, and (D) k_{flux}.

4.5 Discussion

TACs describing FMISO uptake by advanced stage NSCLC are described better by three- than by two-compartment models, mirroring an earlier finding for ^{18}F-fluorothymidine (FLT) uptake in head and neck tumours (Liu *et al.*, 2014). According to information criteria and cross-validation scores, the irreversible three-compartment model, 3C5K, describes FMISO uptake better than the reversible 3C6K model, reflecting the irreversible nature of FMISO intracellular binding. More highly parameterised four-compartment models also perform less well than 3C5K. Statistical simulations show that when the ground-truth is represented by either 3C5K or 3C6K models, fits of 3C5K provide more precise estimates of kinetic rate-constants than fits of the other two- and three-compartment models tested.

For 6 of the 30 TACs studied, however, fits of the 4C7K model achieve lower AIC, BIC or MSEP scores than the three-compartment model fits (Table 4.4), raising the question of whether the more complex four-compartment model should be used to analyse the kinetics of these TACs. We have therefore carried out a statistical simulation sub-study for these six TACs, this time viewing 3C5K or 4C7K fits to the TACs as ground-truth, adding noise to them, and calculating the accuracy and precision of parameter values of model fits to the simulated noisy data. Values of the k_5 FMISO binding rate-constant obtained from 3C5K model fits were compared to a composite effective binding rate-constant k_{eb} defined for the 4C7K model as the combination ($k_{5\text{-}4C}\, k_{7\text{-}4C} / (k_{6\text{-}4C} + k_{7\text{-}4C})$) of fitted 4C7K parameters. The equivalence between k_5 and k_{eb} follows if the $k_{5\text{-}4C}$ and $k_{6\text{-}4C}$ rate-constants of FMISO reduction and re-oxygenation are large compared to FMISO binding ($k_{7\text{-}4C}$) (Casciari *et al.*, 1995). Results of the sub-study are summarised in Table 4.6: even for these six TACs, fits

of the 3C5K model provide more precise estimates of the effective FMISO binding rate-constant than the 4C7K fits, regardless of which model is viewed as the ground-truth.

Ground truth 3C5K model						
Model fitted	*Fitted model parameters*					
3C5K	v_B	K_1	k_2	k_3	k_4	k_5
MB (%)	0	0	0	0	0	0
σ_B (%)	1	0	1	2	0	1
σ_P (%)	9	4	7	11	6	8
σ_T (%)	9	4	7	11	6	8
4C7K	v_B	K_1	k_2	k_3	k_4	k_{eb}
MB (%)	-7	9	68	1044	1226	103
σ_B (%)	8	9	76	1141	860	96
σ_P (%)	10	5	28	76	48	51
σ_T (%)	13	10	81	1143	861	109
Ground truth 4C7K model						
Model fitted	*Fitted model parameters*					
3C5K	v_B	K_1	k_2	k_3	k_4	k_5
MB (%)	13	-12	-29	-48	-60	-19
σ_B (%)	19	15	39	66	54	33
σ_P (%)	11	4	5	2	1	7
σ_T (%)	22	15	39	66	54	33
4C7K	v_B	K_1	k_2	k_3	k_4	k_{eb}
MB (%)	-7	8	30	42	37	17
σ_B (%)	17	16	55	71	46	30
σ_P (%)	14	12	48	59	39	62
σ_T (%)	22	20	73	93	60	68

Table 4.6 – Summary of a statistical simulation sub-study carried out for six TACs that were better described by the 4C7K model than by three-compartment model fits. Values of *MB*, σ_B, σ_P and σ_T are tabulated for 3C5K and 4C7K models fitted to noisy TACs simulated from underlying ground-truth 3C5K and 4C7K model data fits. The 4C7K effective rate-constant for FMISO binding, $k_{eb} = (k_{5\text{-}4C}\, k_{7\text{-}4C} / (k_{6\text{-}4C} + k_{7\text{-}4C}))$.

Alongside the models already tested, which comprise linear chains of compartments, the branching compartment model developed by Casciari *et al.* (1995) that more fully reflects the pathways taken through tumours by FMISO has been explored. Greater information criteria and MSEP scores are obtained when fitting the full ten parameter model of Casciari than when fitting 3C5K, a finding that follows the overall trend seen in Table 4.2 whereby scores increase as model parameter numbers rise from six (for the 3C5K model)

to nine (for the 4C8K model). A simplification of the full Casciari model as a branching model with three compartments, five effective rate-constants and six fittable parameters is shown in Figure 4.2 and denoted as 3C5Kb. Transformation analysis (Cobelli and DiStefano, 1980) shows, however, that the 3C5Kb and conventional 3C5K linear chain models are directly equivalent, a result that is confirmed by the matching AIC, BIC and MSEP scores obtained for fits of the two models.

We have also explored the impact of constraining the rate-constants of the standard 3C5K model. Specifically, FMISO transport between the vasculature and tumour interstitium is known to be passive, as is transport between the interstitium and cells (Grunbaum *et al.*, 1987), properties which in principle allow the 3C5K transport rate-constants K_1, k_2, k_3, k_4 and k_5 to be reduced to just four fittable parameters. In practice, though, greater information criteria and MSEP scores are obtained when this constraint is applied than when the 3C5K model is left unconstrained.

The kinetic analysis described has been performed using image-derived input functions (IDIFs) calculated from mean tracer activity concentrations within volumes drawn in the descending aorta. The gold-standard method for determining input functions is direct arterial line sampling; we chose to use IDIFs, however, for patient comfort and safety and because good agreement has been demonstrated between directly sampled input functions and IDIFs obtained from the descending aorta (van der Weerdt *et al.*, 2001).

TBR_{mean} and SUV_{mean} measures of tracer uptake obtained from four hour scans correlate more closely with the FMISO flux-constant k_{flux} and intracellular binding rate-constant k_5 than uptake measures obtained from two hour scans. Mechanistically, uptake is expected to correlate with k_{flux} and k_5 more strongly at later time-points, k_{flux} being interpretable as proportional to SUV_{mean} at infinite time. The data shows that this expected trend is not overcome by increases in statistical noise at late time-points due to tracer decay, and therefore patients will continue to be imaged at the four hour time-point. Stronger correlations were obtained for TBR_{mean} than for tumour SUV_{mean}, agreeing with previous work using TBR-based measures to represent the amount of tissue hypoxia (Koh *et al.*, 1992; Fleming *et al.*, 2014).

Overall, then, for the NSCLC tumours studied here the 3C5K model provides better descriptions of FMISO uptake TACs, and more accurate and precise estimates of FMISO uptake kinetics, than other two-, three- and four-compartment models. Using this model to assess tumour response to Buparlisib, it can be seen in Figure 4.5 that decreases in the K_1 rate-constant and increases in the modelled fractional blood volume v_B occur in significant numbers of patients after administration of Buparlisib (p=0.001 for K_1, p=0.02 for v_B). Thus the rate-constant of tracer flow (perfusion) from the blood to the tumour interstitium is lower after Buparlisib administration, and the modelled blood volume within the tumour is higher. These results concord with some findings of pre-clinical studies with Buparlisib (described in Chapter 1) in which blood vessels underwent vascular normalisation. Independent perfusion data tested using pCT is described in Chapter 6.

4.6 Conclusion

According to information criteria and cross-validation measures, a 3C5K model of irreversible tumour FMISO uptake comprising a linear chain of three compartments describes time-courses of FMISO uptake by advanced stage NSCLC better than the other two-, three- and four-compartment models investigated. Fits of this model also provide the most accurate and precise estimates of FMISO uptake kinetics, with statistical simulation studies indicating a precision of 10% for fitted values of the tumour fractional blood volume and intracellular FMISO binding rate-constant.

Using this model to assess whole tumour changes in response to Buparlisib it was seen that, independent of cohort, K_1 significantly decreased and v_B significantly increased, a finding which is consistent with vascular normalisation seen in pre-clinical studies.

The k_5 rate constant (linked to hypoxia) increased for Patients 1-5 and decreased for Patients 6-8 post-Buparlisib, differing from the $V_{TBR1.4}$ results in Chapter 3 which decreased for all patients in Cohorts 2 and 3 (Patients 4-9). The change in whole tumour k_5 values are not significant ($p>0.05$) across all nine patients, perhaps indicating that these indices show a more diverse and cohort-dependent response to Buparlisib. This will be examined in Chapter 5 where the kinetic analysis is extended to consider voxel PET kinetics and the optimum model for this case, where noise levels are higher, is identified.

5 Voxel-by-voxel analysis of FMISO uptake kinetics

This chapter on voxel-by-voxel kinetic analysis is divided into three sections:

A. model selection

B. kinetic parameter values, clustering, and the overall effect of Buparlisib

C. variation of kinetic parameter values with distance from the tumour edge, and the effect of Buparlisib at different distances from the tumour edge

A) A study is presented of the relative abilities of different kinetic models to describe voxel FMISO uptake TACs. Models were fitted to 30 TACs and the appropriateness of the fits compared using the Wald-Wolfowitz runs test and AIC/BIC measures. Statistical simulation procedures were used to determine the accuracy and precision with which the uptake kinetics were estimated by fitted model parameters. An irreversible two-compartment model, 2C3K, was found to provide the best overall model description.

B) A clustering method is presented to distil the information from the voxel kinetic analysis into the minimum number of clusters that adequately describe the data. This highlighted a limitation of the TBR metric for voxels with low $K_{1\text{-}2C}$ and high $k_{3\text{-}2C}$. Analysis using clustered data and weighted means showed a significant increase in $v_{B\text{-}2C}$ and significant decrease in $K_{1\text{-}2C}$ for all cohorts post-Buparlisib, with a cohort-dependent change in $k_{3\text{-}2C}$.

C) Categorising parameter values according to their distance from the tumour edge showed $v_{B\text{-}2C}$ and $K_{1\text{-}2C}$ increasing significantly towards the tumour edge and $k_{3\text{-}2C}$ decreasing significantly towards the tumour edge. Post-Buparlisib significant changes have been shown in the weighted means of $v_{B\text{-}2C}$ and of $K_{1\text{-}2C}$ (throughout the tumour volume). The response of $k_{3\text{-}2C}$ to Buparlisib is distance dependent: weighted mean values decrease in edge and outer distance categories and increase in inner and central categories.

5.A Model selection

5.A.1 Introduction

In Chapter 4 the 3C5K irreversible three-compartment model was found to be optimal for analysing the kinetics of FMISO uptake by whole tumours, whose TACs contain relatively little noise as the signal is averaged across their entire volume. Noise levels on the TACs of individual voxels are much higher, and so a simpler model might provide a more stable description of the kinetic rate-constants. Models similar to the 2C3K irreversible two-compartment model described in Chapter 4 have previously been used to perform voxel-based kinetic analyses of FMISO PET uptake data, but these earlier studies did not investigate other models (Bruehlmeier *et al*., 2004; Thorwarth *et al*., 2005; Wang *et al*., 2009) whereas here the performance of a range of different compartment models is compared.

5.A.2 Materials and methods

The PET/CT images collected at two and four hours post-injection (10-minute PET scan each) were rigidly registered to the dynamic PET/CT image (45-minute PET scan commencing upon tracer injection) using an automatic registration tool in the Hermes Hybrid Viewer (Hermes Medical Solutions AB, Stockholm) and followed by manual adjustment when required. The activity concentrations and coordinates of the tumour PET voxel values outlined on the registered two and four hour post-injection images were exported and (using MATLAB code, version R2014a, MathWorks, Natick, Massachusetts) the activity concentrations were interpolated onto the coordinates of voxels within the tumour in the dynamic scan. The interpolated PET values were then appended to the voxel TACs obtained from the dynamic scan.

FMISO uptake kinetics have been analysed for a total of 24,662 voxels, comprising 9 primary tumours in 9 patients each imaged twice, as detailed in Table 5.1. The kinetic analysis uses the methods described in the whole tumour kinetic modelling work (4.3.2). For each voxel the compartment model fitting process was initiated from 100 sets of random starting values, to attempt to reach global rather than local best fits. Runs tests and AIC and BIC scoring were also performed using the methods described in Chapter 4 (4.3.3). However, for practical reasons SF values were calculated (4.3.3) for three randomly selected voxels within each tumour volume and the average SF applied to all of the voxels in that volume, as opposed to calculating an individual SF for every voxel. Also, rather than calculating AIC and BIC scores for all 24,662 voxels, they were instead calculated for a subset of 30 voxels comprising 3 voxels randomly selected from each of 10 outlined tumour volumes in 5 patients, each imaged twice.

	Patient	Voxels	Volume (mL)
Cohort 1 (50 mg Buparlisib)	1_{preC1}	140	13
	1_{postC1}	135	13
	2_{preC1}	5116	510
	2_{postC1}	5023	490
	3_{preC1}	1065	105
	3_{postC1}	1102	98
Cohort 2 (80 mg Buparlisib)	4_{preC2}	261	29
	4_{postC2}	268	27
	5_{preC2}	1332	135
	5_{postC2}	1370	141
	6_{preC2}	1918	180
	6_{postC2}	1803	175
Cohort 3 (100 mg Buparlisib)	7_{preC3}	444	44
	7_{postC3}	344	37
	8_{preC3}	1973	188
	8_{postC3}	1858	182
	9_{preC3}	259	40
	9_{postC3}	251	37

Table 5.1 – Details of the outlined tumour volumes and numbers of voxels they contain.

A numerical simulation study similar to that described in Chapter 4 was carried out to determine the accuracy and precision of the fitted rate-constants derived from different model fits to the voxel TACs. The same methodology was followed for this voxel-by-voxel analysis as for the whole tumour case, described in 4.3.4, although rather than using fits to the TACs measured for all 24,662 voxels to provide the ground-truth, fits to the TACs of just the 30 voxels selected for the AIC/BIC analysis were used instead.

Initially 3C5K model fits to these 30 TACs were used to create noise-free 'ground-truth' TACs, and parameter values from these fits were taken as ground-truth rate-constants. For each of the 30 ground-truth TACs, 1000 noisy TACs were simulated by adding normally distributed voxel-level noise (described in 4.3.4). The simulated TACs were then fitted using the 2C3K and 3C5K models, and these simulated data fits were used to calculate the MB, σ_B, σ_P, σ_T for each parameter (method described in 4.3.4). This process was then repeated, now using 2C3K model fits to the 30 measured TACs to represent the ground-truths. Only irreversible models (2C3K and 3C5K) were considered for the voxel statistical simulation work, as prompted by the findings of the whole tumour kinetic modelling study.

5.A.3 Results

Runs test results for all voxels are shown in Table 5.2, which lists the percentages of voxels within each volume that pass the runs test at significance levels p=0.05 and p=0.01. Table 5.2 shows that at both significance levels a greater proportion of voxels pass the runs test for the 3C5K model than for 2C3K (93% compared to 72% for p=0.05, with 93% being very close to the 95% value expected by chance for p=0.05 and a well-fitting model).

Models	2C3K	3C5K	2C3K	3C5K
Patient/Scan	p = 0.05		p = 0.01	
1_{preC1}	73%	99%	86%	99%
1_{postC1}	70%	96%	88%	100%
2_{preC1}	76%	93%	90%	98%
2_{postC1}	71%	90%	87%	97%
3_{preC1}	56%	89%	76%	97%
3_{postC1}	70%	95%	84%	99%
4_{preC2}	79%	93%	92%	98%
4_{postC2}	72%	97%	87%	99%
5_{preC2}	46%	95%	65%	99%
5_{postC2}	84%	93%	94%	98%
6_{preC2}	63%	95%	78%	99%
6_{postC2}	57%	94%	74%	99%
7_{preC3}	61%	95%	77%	99%
7_{postC3}	61%	92%	80%	97%
8_{preC3}	90%	94%	97%	99%
8_{postC3}	88%	95%	96%	99%
9_{preC3}	92%	94%	98%	98%
9_{postC3}	92%	94%	96%	98%
Total	72%	93%	86%	98%

Table 5.2 – Proportion of voxels in each volume for which fits of the 2C3K and 3C5K kinetic models pass the runs test at significance levels of p=0.05 and p=0.01.

Table 5.3 shows the number of runs test passes and the AIC and BIC results for the TACs of the subgroup of 30 voxels investigated (shown in Figure 5.3). For this subgroup of voxels the 3C5K model had 100% runs test passes and achieved the lowest overall AIC and BIC scores of the four models tested, as well as having the lowest individual AIC and BIC scores for the largest number of voxel TACs. The same result was seen for whole tumours, although the difference in AIC and BIC values between the 2C and 3C models was much larger for whole tumours than for voxels (Tables 4.2 and 5.3). Thus the more complex 3C5K model is able to describe the shape of individual voxel TACs better than the simple 2C3K model. However, the primary aim of this work is to explore the potential of fitted model parameters to gauge the degree of hypoxia, and for this purpose the choice of model should be guided more by the accuracy and precision of the fitted model parameters than by the model's ability to describe TAC shapes.

Model	2C3K	2C4K	3C5K	3C6K
Runs test passes from fits to all 30 TACs				
Runs	22	27	30	30
Information criteria summed for all statistical simulation TACs				
AIC	1793	1500	<u>1424</u>	1508
BIC	2002	1738	<u>1687</u>	1790
Numbers of TACs for which each model has the lowest scores				
AIC	4	8	18	0
BIC	8	7	15	0

Table 5.3 – Summary of runs test results and AIC and BIC scores for model fits to all 30 TACs used in the statistical simulation analysis. The lowest AIC and BIC scores have been underlined, indicating the best model according to that measure.

Table 5.4 displays the results of the statistical simulation analysis for ground-truths represented by fits of either the 2C3K or 3C5K models to measured TACs. For 2C3K ground-truth TACs the 2C3K model fits have lower MB, σ_B and σ_P values than the 3C5K model fits, and so unsurprisingly 2C3K can be considered to be more accurate and precise than 3C5K for this case.

For 3C5K ground-truth TACs, the 2C3K model fits have lower σ_P values than the 3C5K model fits, suggesting that 2C3K is more precise. In contrast, 2C3K model fits have higher MB and σ_B values than the 3C5K model fits, suggesting that 2C3K is not as accurate in this respect. However, for $k_{3\text{-}2C}$ the total uncertainty (combining the random uncertainty σ_P in quadrature with voxel-to-voxel variation in bias σ_B) is lower for 2C3K than 3C5K model fits, while for k_{flux} the total uncertainty is the same for 2C3K as for 3C5K model fits. Since the mean bias could in theory be eliminated by appropriate normalisation using previous imaging and histopathology data, the 2C3K model is optimal as it provides results that are more precise overall, with total uncertainties of 21% and 37% for $K_{1\text{-}2C}$ and $k_{3\text{-}2C}$ respectively.

Ground-truth 2C3K model							
Model fitted	*Fitted model parameters*						
2C3K	$v_{B\text{-}2C}$	$K_{1\text{-}2C}$	$k_{2\text{-}2C}$			$k_{3\text{-}2C}$	$k_{flux\text{-}2C}$
MB (%)	-12	-4	-4			-4	-4
σ_B (%)	13	7	7			6	5
σ_P (%)	36	20	28			37	29
σ_T (%)	39	21	29			37	30
3C5K	v_B	K_1	k_2	k_3	k_4	k_5	k_{flux}
MB (%)	-22	-	-	-	-	398	-5
σ_B (%)	39	-	-	-	-	591	11
σ_P (%)	43	199	155	61	64	89	32
σ_T (%)	58	-	-	-	-	597	34
Ground-truth 3C5K model							
Model fitted	*Fitted model parameters*						
2C3K	$v_{B\text{-}2C}$	$K_{1\text{-}2C}$	$k_{2\text{-}2C}$			$k_{3\text{-}2C}$	$k_{flux\text{-}2C}$
MB (%)	53	-	-			-42	24
σ_B (%)	85	-	-			56	37
σ_P (%)	44	32	42			33	31
σ_T (%)	96	-	-			65	47
3C5K	v_B	K_1	k_2	k_3	k_4	k_5	k_{flux}
MB (%)	-5	25	104	90	67	27	-1
σ_B (%)	15	174	535	208	104	58	13
σ_P (%)	56	188	227	119	125	142	44
σ_T (%)	58	257	581	240	162	153	47

Table 5.4 – Estimates of the accuracy and precision of parameter values obtained from fits of the 2C3K and 3C5K models to TAC data simulated using ground-truth 2C3K or 3C5K models. Values of *MB*, σ_B, σ_P and σ_T are shown for fitted parameters as percentages of the mean values of directly related ground-truth parameters of the 2C3K or 3C5K models. When no directly-related parameter exists, σ_P values alone are shown as percentages of the mean fitted parameter values.

5.A.4 Conclusion

Total uncertainties (combining statistical uncertainties with patient-to-patient bias variation) on parameter values obtained from fits of the 2C3K compartment model to individual voxel TACs are lower than the total uncertainties on parameter values obtained from 3C5K model fits. Therefore throughout the rest of this work the 2C3K model is used to analyse the kinetics of FMISO uptake in individual voxels.

5.B Kinetic parameter values, clustering, and effect of Buparlisib

5.B.1 Introduction

24,662 individual voxel TACs were acquired as part of the Buparlisib trial. In Chapter 4 the derived rate-constant values obtained from fits to the 30 TACs of the whole tumour volumes were examined individually. This approach is not as practical for some aspects of analysing fits to the many individual voxel TACs, and therefore a clustering method is applied to the rate-constant values obtained for the individual voxels, to assist with data interpretation and visualisation. Previously Thorwarth *et al.* (2005b) informally grouped data based on regions drawn around a graph of voxel perfusion and hypoxia measures obtained from kinetic analysis:

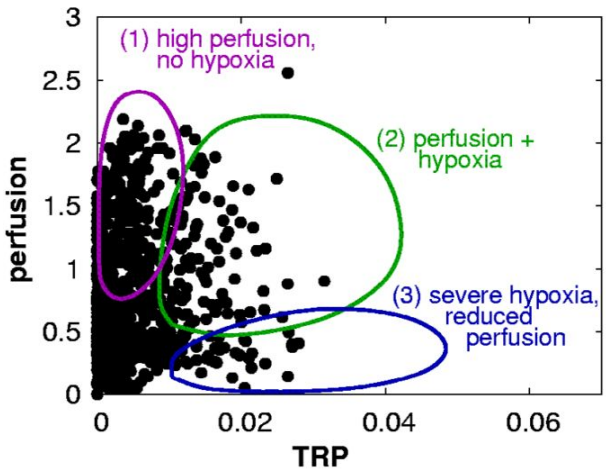


Figure 5.1 – Example clustering (Thorwarth *et al.*, 2005b).

The clustering technique used in this work is based on a statistical method developed by Liu *et al.* (2014) which accounts for the uncertainties on fitted parameter values of each voxel, and, using an expectation-maximization–Bayesian information criterion approach, determines the number of clusters that best describe the underlying data.

5.B.2 Materials and methods

Voxel TACs were generated and kinetic analysis performed on them as detailed in 5.A.2, using the 2C3K model. When appropriate, the derived parameter values were clustered (Table 5.4) using in-house MATLAB code (version R2014a, MathWorks, Natick, Massachusetts) according to the method of Liu *et al.* (2014), extended to factor patient-to-patient bias variation into estimates of total uncertainties on fitted parameter values, as detailed below:

1. A number of discrete cluster levels K is assumed.
2. Parameter values x_i of each voxel i (1, ..., N) are initially grouped into K discrete levels of value $X_{cluster\text{-}j}$ (j = 1, ..., K) using a weighted k-means clustering algorithm. The weights associated with each point, w_i, are defined as

$$w_i = (\sigma_B{}^2 + \sigma_i{}^2)^{-1} \quad (5.1)$$

   where σ_B is the variability (one s.d.) of patient-specific bias for that parameter (from Table 5.4) and σ_i is the statistical uncertainty (one s.d.) on the fitted parameter value.
3. Matrix Z_{ij} is generated where a value of 1 is assigned if voxel i is assigned to cluster level j and 0 assigned for other matrix elements.
4. An iterative expectation-maximization (EM) algorithm is used to refine the initial clustering, adjusting $X_{cluster}$ and Z_{ij} values to maximise the mixture-likelihood.
5. After convergence, the *BIC* is calculated as BIC(K) = -2 ln L(K) + K ln$N$.
6. Steps 1 to 5 are repeated for a range of K levels (typically 1-15).
7. The cluster grouping with the lowest BIC is taken to be the best clustered representation of the data.

Clustering can be carried out for parameter values obtained for the voxels belonging to all patients, or for the voxels belonging to patients in a specific cohort or belonging to a particular patient or even one particular scan. At different points in this work clustering has been carried out for all of these different data groups. However, when exploring the effects of Buparlisib, clustering has been carried out for all voxels together from all images collected before and after drug administration, so that the cluster-levels are the same for voxels in pre- and post-Buparlisib images, allowing the significance of changes in proportions of voxels clustered into these levels to be assessed. Likewise, when exploring the impact of depth within a tumour on fitted parameter values, clustering was carried out for voxels at all depths before analysing differences in clustering patterns according to depth.

The chi-square test of homogeneity was used to assess pre- to post-Buparlisib changes in the percentages of voxels assigned to different cluster levels for each parameter. The null hypothesis for this test is that within statistical uncertainties the same proportion of voxels lies within each cluster grouping in both pre-and post-Buparlisib images; the test therefore determines the significance of observed changes in these proportions after administration of Buparlisib. Statistical significances of $p<0.05$, 0.01, 0.001 and 0.0001 are indicated via the symbols *, **, *** and **** respectively in graphs of clustered data.

In order to assess whether on average fitted parameters increase or decrease post-Buparlisib, a weighted mean ($\bar{x}$) of the unclustered values has been calculated as:

$$\bar{x} = \frac{\sum_{i=1}^{N} w_i x_i}{\sum_{i=1}^{N} w_i} \tag{5.2}$$

where x_i is the parameter value for the i^{th} of N total values and w_i the weighting as defined in equation 5.1.

The unbiased estimate of sample variance (s^2) is:

$$s^2 = \frac{\sum_{i=1}^{N} w_i x_i{}^2 \sum_{i=1}^{N} w_i - \left(\sum_{i=1}^{N} w_i x_i\right)^2}{\left(\sum_{i=1}^{N} w_i\right)^2 - \sum_{i=1}^{N} w_i{}^2} \quad (5.3)$$

This is calculated pre- and post-Buparlisib for each parameter with the degrees of freedom (DOF) calculated as:

$$DOF = \frac{\left(\frac{s_{pre}^2}{n_{pre}} + \frac{s_{post}^2}{n_{post}}\right)^2}{\left(\frac{s_{pre}^2}{n_{pre}}\right)^2 \frac{1}{n_{pre}-1} + \left(\frac{s_{post}^2}{n_{post}}\right)^2 \frac{1}{n_{post}-1}} \quad (5.4)$$

where n_{pre} is the number of values pre-Buparlisib and n_{post} the number post-Buparlisib. The statistical significance of the difference of each weighted mean from pre- to post-Buparlisib was then assessed using a t-test for samples of unequal variance (Welch's t-test). Again, statistical significances of differences in weighted means are indicated in the tables via the symbols 'NS' (not significant), * $p<0.05$, ** $p<0.01$, *** $p<0.001$, and **** $p<0.0001$.

Using these approaches, three facets of the data are explored in this section of the work: correlations between the fitted values of different rate-constants (5.B.3); scan-to-scan, patient-to-patient, and cohort-to-cohort variations in numbers of clustering levels (5.B.4); and the impact of Buparlisib on fitted parameter values, assessed both in terms of changes in clustering patterns and in weighted means (5.B.5).

5.B.3 Results: Kinetic analysis

Figures 5.2 and 5.3 show unclustered data obtained from all analysed voxels. In Figure 5.2 TBR is plotted against $K_{1\text{-}2C}$, $k_{3\text{-}2C}$ and $k_{flux\text{-}2C}$: there is no correlation between TBR and $K_{1\text{-}2C}$ (Pearson's correlation coefficient, r=0.05), a weak correlation between TBR and $k_{3\text{-}2C}$ (r=0.53), and a strong TBR correlation with $k_{flux\text{-}2C}$ (r=0.82). In Figure 5.3 a weak negative correlation can be observed between $k_{3\text{-}2C}$ and $K_{1\text{-}2C}$ (r=-0.22)

In Figure 5.2A it can be seen that for high $K_{1\text{-}2C}$ values (greater than 0.8 ml/min/g), TBR values are quite constant at a level around unity, presumably because these voxels are not hypoxic and the rapid diffusion of FMISO from the blood into the tumour and back again effectively equilibrates tracer activity concentrations within the tumour and blood.

Figure 5.2B shows that whilst a weak positive correlation exists between $k_{3\text{-}2C}$ and TBR (higher values of TBR are associated with hypoxia), many voxels with TBRs below the 1.4 threshold also have high $k_{3\text{-}2C}$ levels. These voxels could be poorly perfused, leading to very low rates of FMISO influx which would cause TBRs to remain relatively low despite generating hypoxic conditions which would cause long-term FMISO accumulation, leading to high fitted $k_{3\text{-}2C}$ values.

Figure 5.2C shows a strong positive correlation between $k_{\text{flux-}2C}$ and TBR values determined at 4 hours post-injection. This is to be expected since $k_{\text{flux-}2C}$ is a measure of the rate of tracer uptake under equilibrium conditions and should therefore correlate strongly with TBR at late time points. A small group of voxels have high $k_{\text{flux-}2C}$ but low TBR values: these voxels turn out to have high $v_{B\text{-}2C}$ values, and therefore they would be expected to have TBR values close to 1, due to their high blood content.

Figure 5.3 shows that high $K_{1\text{-}2C}$ values are generally associated with low $k_{3\text{-}2C}$, and high $k_{3\text{-}2C}$ with low $K_{1\text{-}2C}$, as expected. In Figures 5.2 and 5.3 it is difficult to see the detail of relationships between parameter values due to the number of values plotted, and therefore in the next section clustering is used to improve clarity.

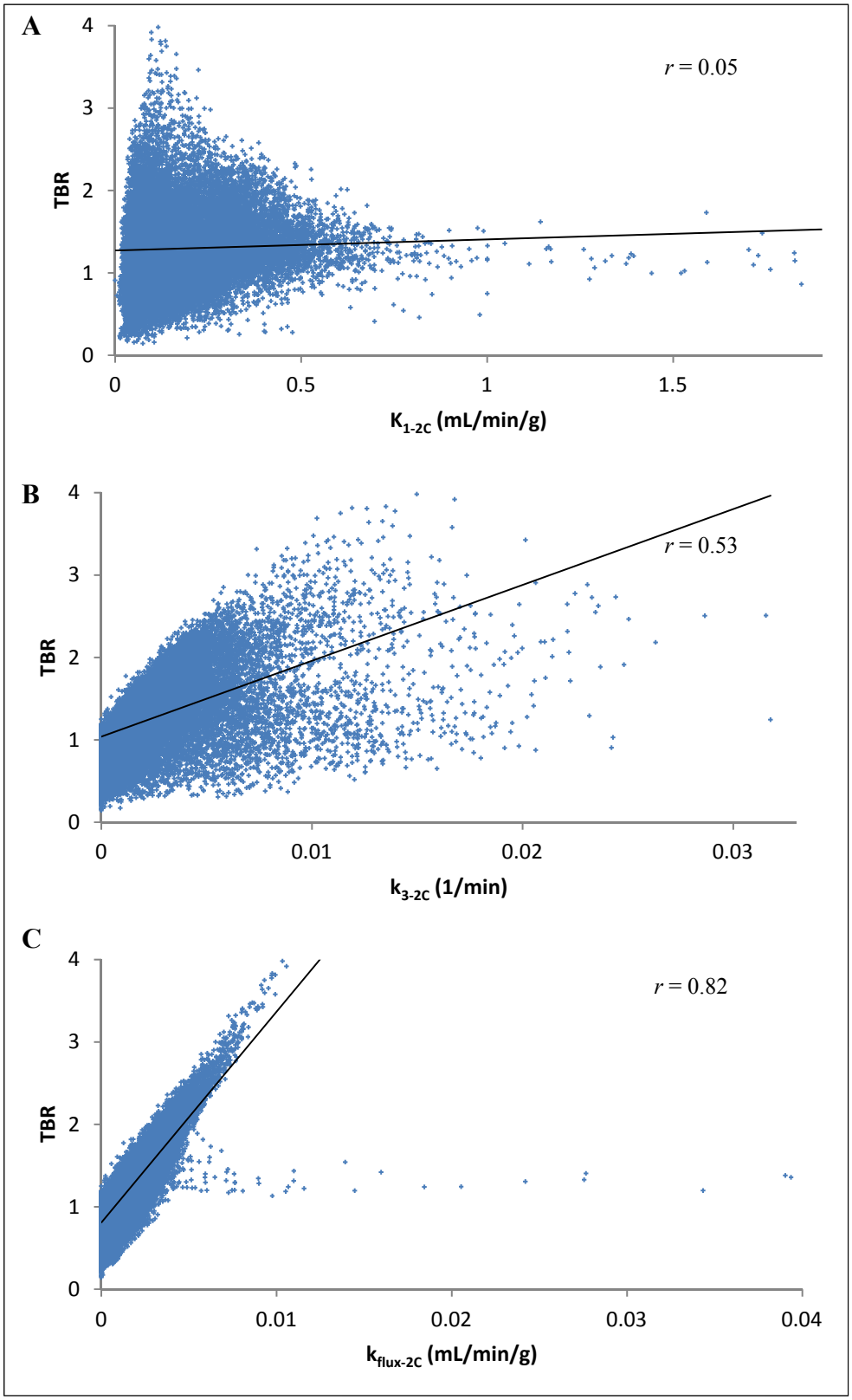


Figure 5.2 – Graph of TBR versus (A) $K_{1\text{-}2C}$, (B) $k_{3\text{-}2C}$, and (C) $k_{flux\text{-}2C}$ for all analysed voxels.

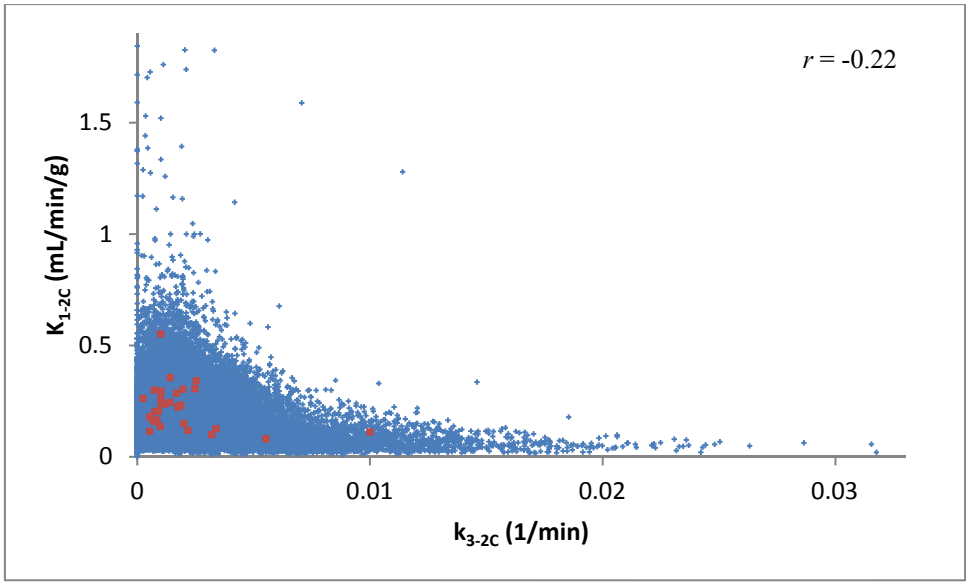


Figure 5.3 – Graph of $k_{3\text{-}2C}$ versus $K_{1\text{-}2C}$ for all analysed voxels. The red points indicate those used for the statistical simulation work (5.A.3).

The parameters derived from the kinetic analysis can be used to create parametric maps, which are spatial representations of the kinetic parameters. An example is shown in Figures 5.4A, 5.4B, and 5.6A for the parameters $v_{B\text{-}2C}$, $K_{1\text{-}2C}$, $k_{3\text{-}2C}$ respectively illustrating the same tumour slice for each.

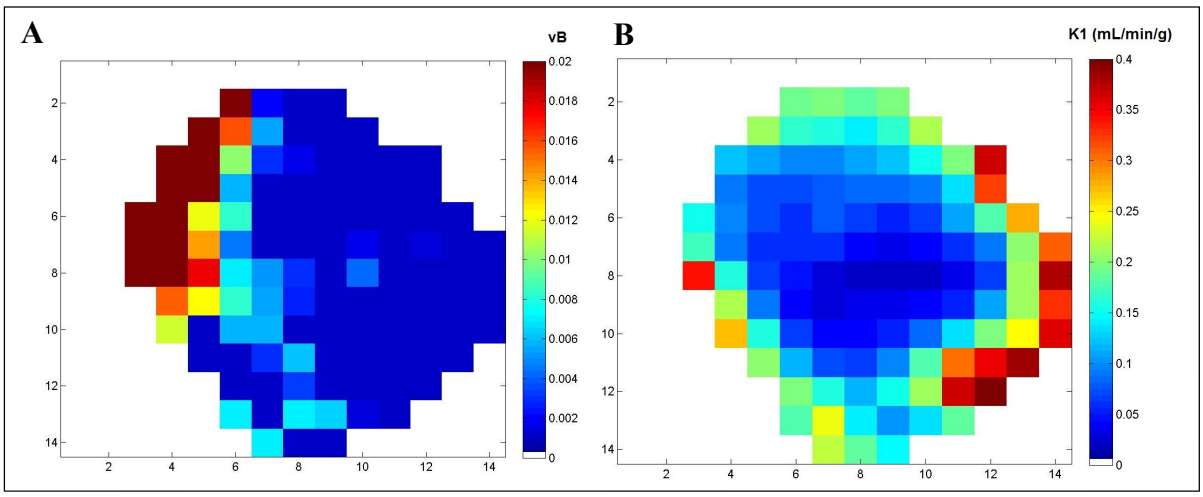


Figure 5.4 – Parametric maps obtained for Patient 6 pre-Buparlisib for (A) $v_{B\text{-}2C}$ and (B) $K_{1\text{-}2C}$.

Example TACs from this same patient are shown in Figure 5.5 for both a voxel on the edge of the tumour (Figure 5.5A and 5.5C) and in the centre of the tumour (Figure 5.5B and 5.5D).

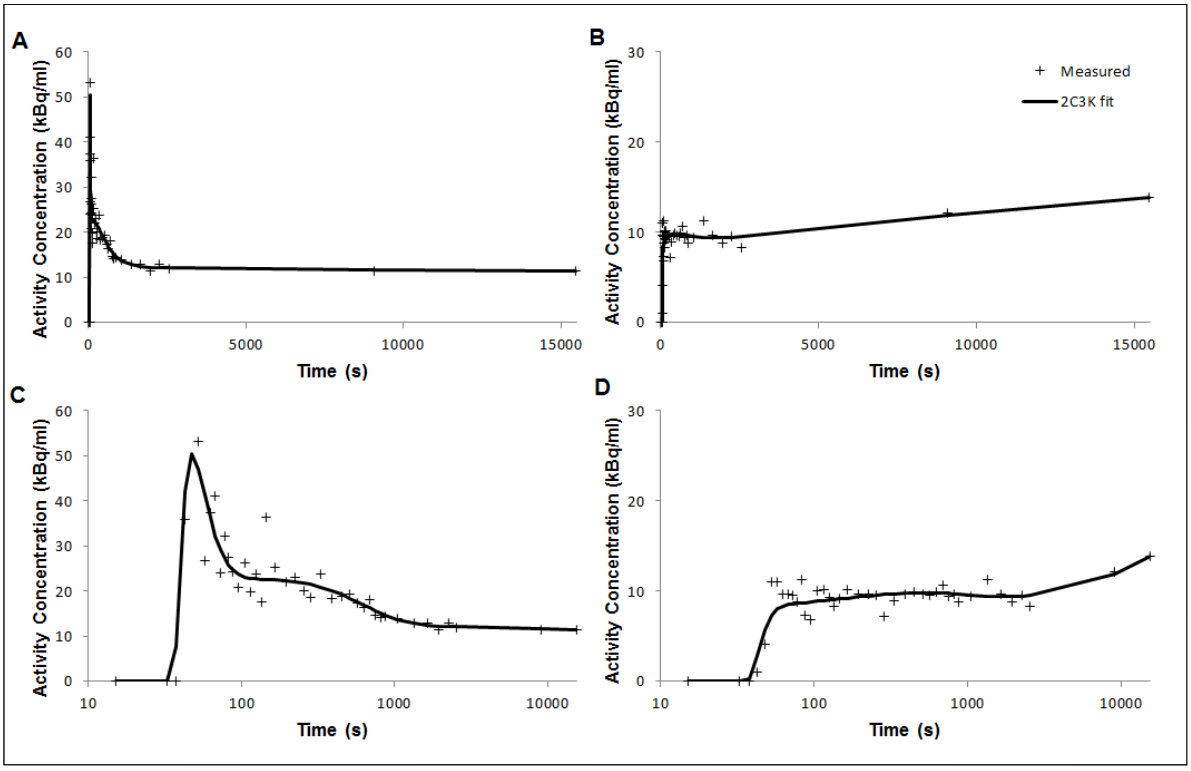


Figure 5.5 – Fits of the 2C3K model to TACs from the tumour edge (plots A and C) and tumour centre (B and D) for Patient 6 pre-Buparlisib. Time post-injection is plotted on linear (A/B) and logarithmic (C/D) scales.

5.B.4 Results: Clustering

The number of clusters into which different parameters ($v_{B\text{-}2C}$, $K_{1\text{-}2C}$, $k_{2\text{-}2C}$, $k_{3\text{-}2C}$, $k_{flux\text{-}2C}$, TBR) are grouped, when clustering is carried out separately for the voxels within each tumour on each scan, is presented in Table 5.5. Table 5.6 shows the number of clusters obtained for the same parameters when the voxels for both scans of each patient, or each cohort, or all cohorts, are included in the clustering process. The table also shows the number of voxels grouped within each clustering. Figure 5.6 shows a slice through a $k_{3\text{-}2C}$ tumour parametric map of Patient 6 obtained pre-Buparlisib, together with the clustered

version of the map, for which the clustering was performed for that scan alone. Whereas unclustered images contain a broad and continuous range of parameter data, the clustered images provide a representation in which the data is grouped into a discrete set of levels, generally over a slightly narrower range, and the number of levels and their values are chosen to best represent the raw individual voxel data, taking into consideration the individual fitted parameter values and known total uncertainties on them.

Patient /Scan	**Number of voxel TACs clustered**	**Number of cluster-levels**					
		$v_{B\text{-}2C}$	$K_{1\text{-}2C}$	$k_{2\text{-}2C}$	$k_{3\text{-}2C}$	$k_{flux\text{-}2C}$	TBR
1_{preC1}	140	2	4	5	1	1	4
1_{postC1}	135	2	5	5	1	2	3
2_{preC1}	5116	3	8	4	2	2	4
2_{postC1}	5023	4	9	6	6	8	8
3_{preC1}	1065	4	4	4	2	3	4
3_{postC1}	1102	4	5	5	4	5	6
4_{preC2}	261	4	2	2	1	1	2
4_{postC2}	268	5	4	4	1	1	3
5_{preC2}	1332	3	5	4	4	4	4
5_{postC2}	1370	3	7	5	4	6	8
6_{preC2}	1918	3	7	5	4	3	3
6_{postC2}	1803	3	9	6	4	3	4
7_{preC3}	444	4	5	4	3	4	3
7_{postC3}	344	4	5	4	3	3	2
8_{preC3}	1973	2	4	4	2	3	4
8_{postC3}	1858	2	6	4	2	3	3
9_{preC3}	259	1	2	2	1	1	2
9_{postC3}	251	2	3	2	1	1	2

Table 5.5 – Number of cluster-levels obtained when voxel-by-voxel 2C3K model fit parameter values are clustered for each scan individually.

Patient	Number of voxel TACs clustered	Number of cluster-levels					
		$v_{B\text{-}2C}$	$K_{1\text{-}2C}$	$k_{2\text{-}2C}$	$k_{3\text{-}2C}$	$k_{flux\text{-}2C}$	TBR
1_{C1}	275	2	5	6	1	1	4
2_{C1}	10139	4	9	6	3	8	6
3_{C1}	2167	4	4	5	4	5	5
4_{C2}	529	4	4	3	1	1	3
5_{C2}	2702	3	10	7	4	6	6
6_{C2}	3722	3	9	6	4	3	3
7_{C3}	788	5	6	5	3	3	3
8_{C3}	3831	2	6	5	2	3	3
9_{C3}	510	1	4	2	1	1	3
Average patient		3.1	5.2	4.2	2.6	3.0	3.8
All C1	12581	5	9	6	7	6	7
All C2	6952	7	11	8	4	5	7
All C3	5129	4	7	7	2	5	4
All C1,2,3	24662	6	14	7	4	10	7

Table 5.6 – Number of cluster-levels obtained when voxel-by-voxel 2C3K model fit parameter values are clustered for both scans of each patient, for all the scans of each cohort, and for all scans together.

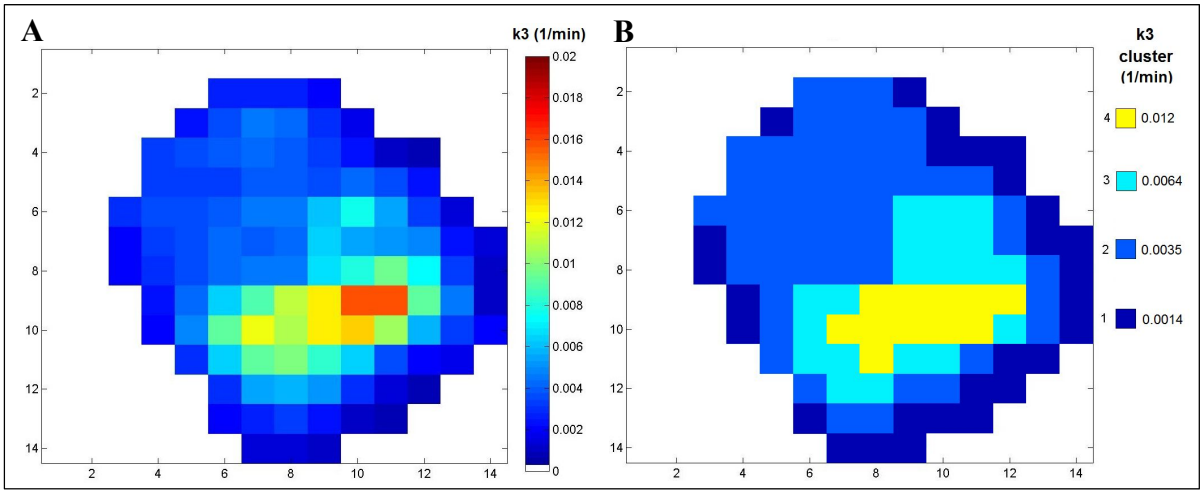


Figure 5.6 – A slice taken from the pre-Buparlisib $k_{3\text{-}2C}$ parametric map obtained for Patient 6, showing (A) unclustered, and (B) clustered $k_{3\text{-}2C}$ values. Clustering was carried out using data from both the pre- and post-Buparlisib scans of Patient 6.

Commonly FMISO images are assessed on the assumption that tissues are hypoxic within voxels having TBRs greater than a typical threshold of 1.4, as discussed in Chapter 3. Figure 5.7 is a bubble plot of the numbers of voxels assigned to $K_{1\text{-}2C}$ and $k_{3\text{-}2C}$ cluster-levels when all 24,662 voxels were clustered together. The size of each bubble represents

the number of voxels contributing to that data point, and its colour represents the proportion of those voxels having TBRs greater than 1.4 (Figure 5.7). While high percentages of voxels having TBR values >1.4 are generally associated with larger k_3 values, it can be seen that at the highest levels of $k_{3\text{-}2C}$, the fraction of voxels with TBRs greater than 1.4 is low when $K_{1\text{-}2C}$ is also low (the region within the blue box at the top left corner of the plot) – meaning that these voxels would not generally be considered hypoxic on the basis of their TBR values, despite having high $k_{3\text{-}2C}$ values. This mismatch is perhaps the result of very poor perfusion, represented by the low $K_{1\text{-}2C}$ values, fundamentally limiting the quantity of tracer available for uptake. Such mismatches have been described before (Wang *et al.*, 2010) and used as an example of the need to assess the response of each voxel via kinetic analysis as well as static uptake measures. For low $k_{3\text{-}2C}$ values almost all voxels have TBRs <1.4, as would be expected if low $k_{3\text{-}2C}$ and low TBR were both indicative of normoxia.

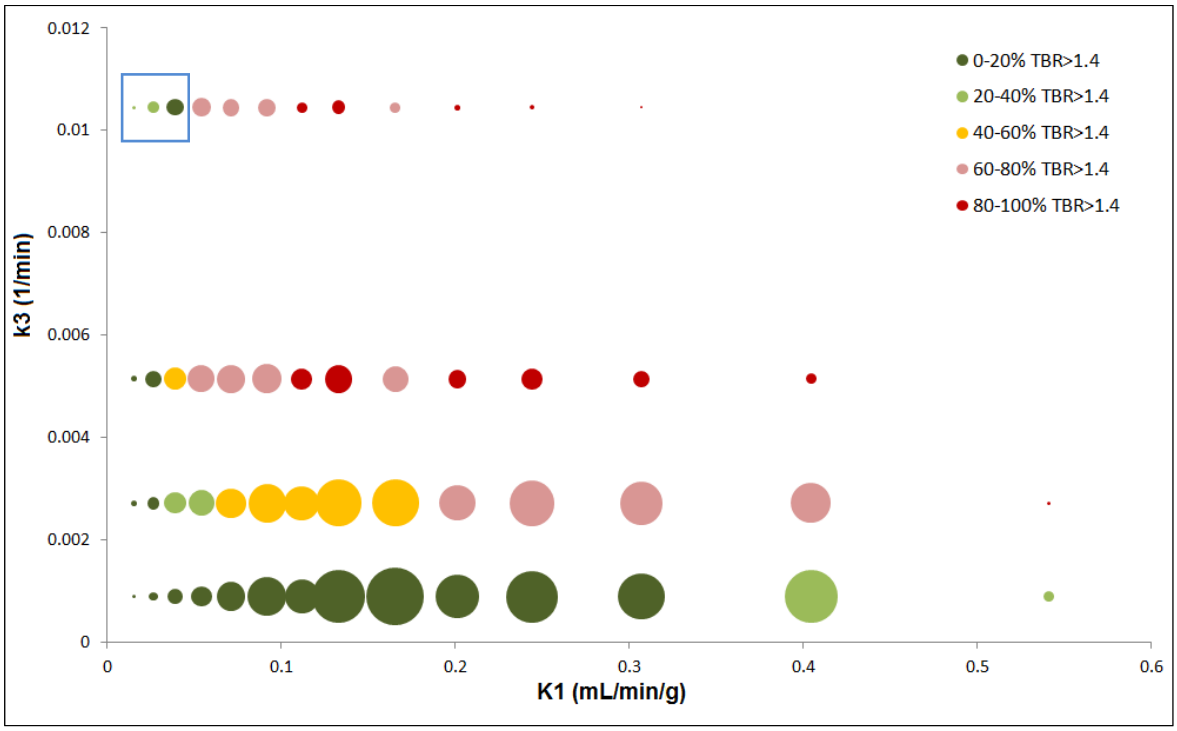


Figure 5.7 – Bubble plot showing clustered $k_{3\text{-}2C}$ versus $K_{1\text{-}2C}$; the size of the bubbles indicates the number of points in each cluster. Bubbles are coloured according to the proportion of voxels with a TBR greater than 1.4. The blue box indicates a region of the graph where $k_{3\text{-}2C}$ is high and yet there are a small number (less than 40%) of voxels with TBR>1.4.

5.B.5 Results: Impact of Buparlisib on kinetics

Figure 5.8 plots the weighted means of fitted voxel parameter values in pre- and post-Buparlisib images. Table 5.7 summarises the direction of change of these weighted mean values, and indicates the statistical significance of the changes. For the parameters $K_{1\text{-}2C}$, $v_{B\text{-}2C}$, $k_{3\text{-}2C}$, and TBR, Figures 5.9 to 5.11 respectively show the fractions of tumour voxels assigned to different cluster-levels in images collected pre- and post-Buparlisib, allowing the significance of changes in cluster proportions to be assessed. The clustering was carried out for all tumour voxels in all images together, so that the cluster-level thresholds are the same for pre- and post-Buparlisib images. With the exception of $k_{3\text{-}2C}$ which clusters into only four levels, parameter cluster-levels have been partly merged for clarity, as detailed in Table 5.8, so that only four different colours are displayed in each bar chart. From Figures 5.9 to 5.12 it can be seen that there is a large amount of heterogeneity in all kinetic parameters between the three cohorts pre-Buparlisib. Due to this each cohort has been treated separately for comparisons pre- to post-Buparlisib.

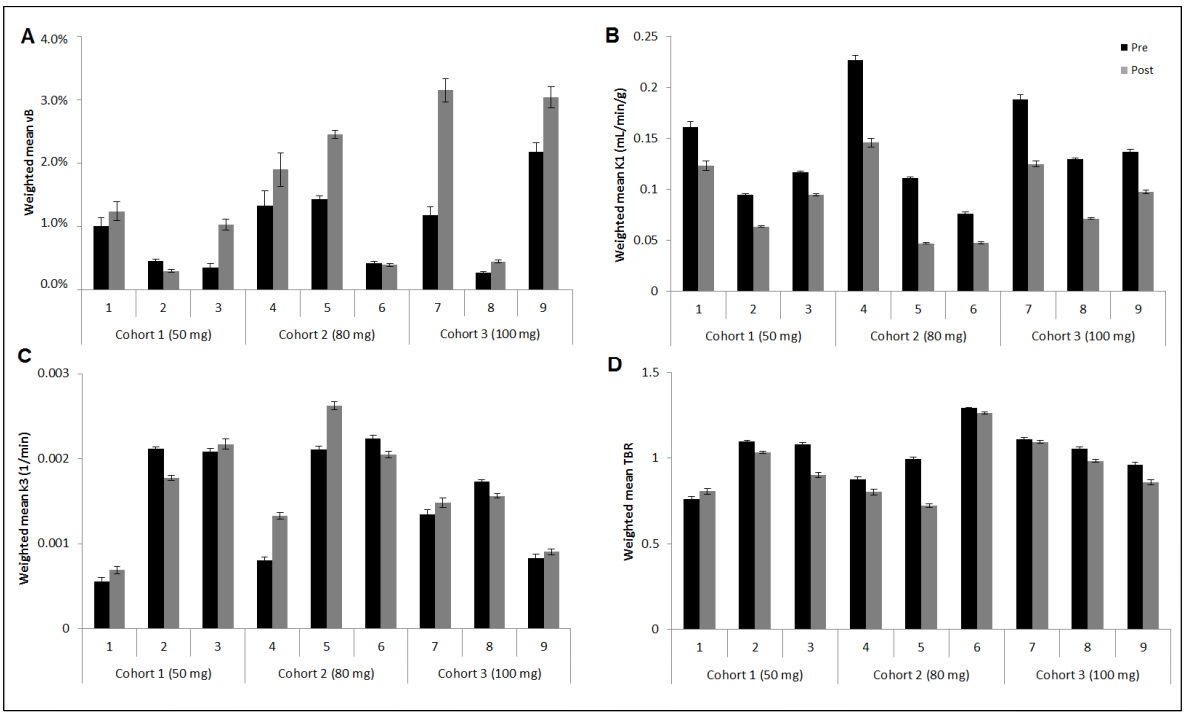


Figure 5.8 – Graphs showing weighted mean values of 2C3K kinetic parameters pre- and post-Buparlisib for (A) $v_{B\text{-}2C}$, (B) $K_{1\text{-}2C}$, (C) $k_{3\text{-}2C}$, and (D) TBR; error bars shown are the standard error on the mean.

Patients	$v_{B\text{-}2C}$	$K_{1\text{-}2C}$	$k_{3\text{-}2C}$	TBR
1_{C1}	↑NS	↓****	↑*	↑*
2_{C1}	↓****	↓****	↓****	↓****
3_{C1}	↑****	↓****	↑NS	↓****
4_{C2}	↑NS	↓****	↑****	↓***
5_{C2}	↑****	↓****	↑****	↓****
6_{C2}	↓NS	↓****	↓***	↓**
7_{C3}	↑****	↓****	↑NS	↓NS
8_{C3}	↑****	↓****	↓****	↓****
9_{C3}	↑****	↓****	↑NS	↓****
All C1	↓*	↓****	↓****	↓****
All C2	↑****	↓****	↑***	↓****
All C3	↑****	↓****	↓****	↓****
All C1,2,3	↑****	↓****	↓NS	↓****

Table 5.7 – Direction (arrows) and degree of significance (asterisks) of differences between the weighted means of fitted $v_{B\text{-}2C}$, $K_{1\text{-}2C}$, $k_{3\text{-}2C}$, and TBR parameters in images obtained before and after Buparlisib administration, listed for individual patients and for patient cohorts.

Graph Clusters	**Merged cluster levels**			
	$v_{B\text{-}2C}$	$K_{1\text{-}2C}$ (mL/min/g)	$k_{3\text{-}2C}$ (1/min)	TBR
Group 1	0.005	0.02, 0.03, 0.04, 0.05, 0.07	0.0009	0.31, 0.53, 0.76, 1.0
Group 2	0.05	0.09, 0.11, 0.13	0.003	1.4
Group 3	0.1	0.17, 0.20, 0.24	0.005	1.8
Group 4	0.18, 0.31, 0.61	0.31, 0.40, 0.54	0.01	2.4

Table 5.8 – Partial merging of the original cluster-levels for v_B, K_1, k_3, and TBR into four groupings, which are subsequently used in clustered bar charts and bullseye charts.

Figure 5.9 shows that following administration of Buparlisib the $K_{1\text{-}2C}$ cluster groups change significantly for all cohorts ($p<0.0001$). The weighted mean of $K_{1\text{-}2C}$ decreases post-Buparlisib for all cohorts and patients ($p<0.0001$) (Table 5.7). Figure 5.10 shows that post-Buparlisib, the $v_{B\text{-}2C}$ cluster groupings change significantly ($p<0.01$ for Cohort 2, and $p<0.0001$ for Cohorts 1 and 3). However, the weighted mean of $v_{B\text{-}2C}$ decreases post-Buparlisib for Cohort 1 ($p<0.05$) and increases for Cohorts 2 and 3 ($p<0.0001$) (Table 5.7). Whilst for Cohort 3 all three patients have significant increases in the weighted mean $v_{B\text{-}2C}$, for Cohort 2 there is a decrease in $v_{B\text{-}2C}$ for Patient 6 ($p>0.05$), and for Cohort 1 there is a significant decrease in $v_{B\text{-}2C}$ for Patient 2 ($p<0.0001$).

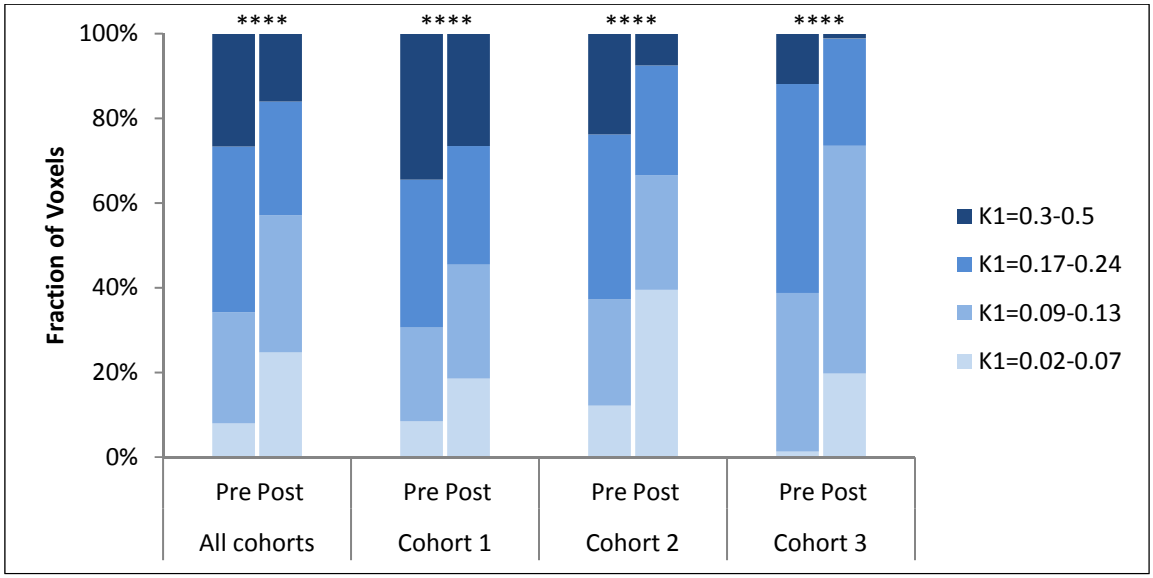


Figure 5.9 – Changes between pre- and post-Buparlisib images in percentages of fitted voxel $K_{1\text{-}2C}$ values clustered with all voxels together and by cohort. Asterisks indicate the statistical significance of changes between pre- and post-Buparlisib images.

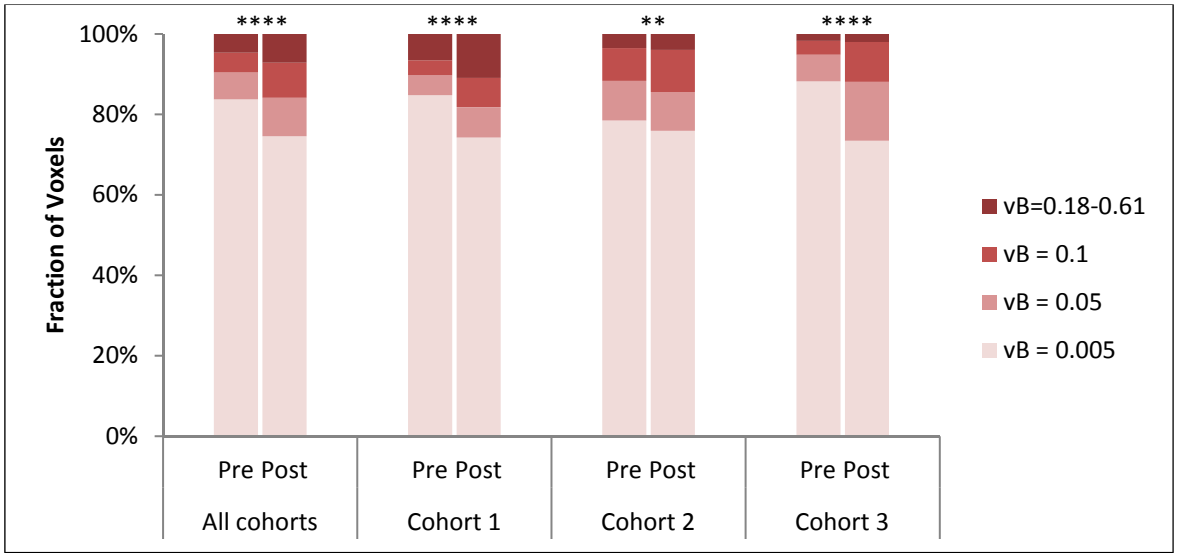


Figure 5.10 – Changes between pre- and post-Buparlisib images in percentages of fitted voxel $v_{B\text{-}2C}$ values clustered with all voxels together and by cohort. Asterisks indicate the statistical significance of changes between pre- and post-Buparlisib images.

Overall the $K_{1\text{-}2C}$ and $v_{B\text{-}2C}$ results suggests that the vasculature is becoming less permeable (leaky) in all cohorts and that the volume of blood passing through the tumour is increasing for all patients in Cohort 3 and for the majority of patients in earlier cohorts; these changes might be considered to represent a process of vascular normalisation.

Figure 5.11 shows that for each cohort there are significant changes in the patterns of $k_{3\text{-}2C}$ clustering following Buparlisib administration ($p<0.0001$). Post-Buparlisib, the number of voxels lying in the lowest $k_{3\text{-}2C}$ cluster-level ($k_{3\text{-}2C} = 0.0009$ min^{-1}) increase, as do numbers lying in the highest cluster-level ($k_{3\text{-}2C} = 0.01$ min^{-1}). These two findings run in opposite directions, which could be explained by the conjecture that the effect of Buparlisib varies with location within tumours, a hypothesis that will be explored in the next section (5.C). The weighted mean values of $k_{3\text{-}2C}$ significantly decrease for Cohorts 1 and 3 after Buparlisib ($p<0.0001$), but increase for Cohort 2 ($p<0.001$). Within each cohort the direction of these weighted mean changes varies from patient to patient: for example, in Cohorts 1 and 3 the weighted mean values of $k_{3\text{-}2C}$ actually increase for 2 out of 3 patients, despite the weighted mean for all three patients significantly decreasing. It thus appears that Buparlisib has a more complex effect on $k_{3\text{-}2C}$ than on $v_{B\text{-}2C}$ or $K_{1\text{-}2C}$, which may be due to the drug's impact on k_3 being driven by a combination of its contrasting effects on $K_{1\text{-}2C}$ and $v_{B\text{-}2C}$.

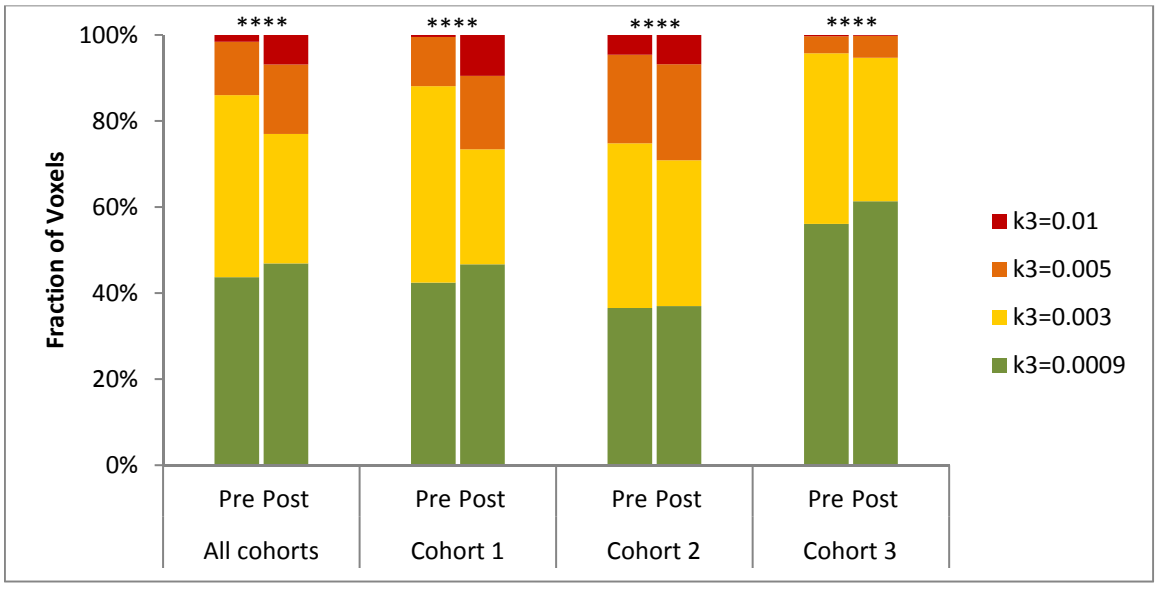


Figure 5.11 – Changes between pre- and post-Buparlisib images in percentages of fitted voxel $k_{3\text{-}2C}$ values clustered with all voxels together and by cohort. Asterisks indicate the statistical significance of changes between pre- and post-Buparlisib images.

In Figure 5.12 it can be seen that percentages of voxels assigned to different TBR cluster groups change significantly after administration of Buparlisib ($p<0.0001$). In Cohorts 1 and 2 the numbers of voxels assigned to the highest TBR cluster-level (TBR=0.24) increases post-Buparlisib, while in Cohorts 2 and 3 the number of voxels assigned to the lowest TBR cluster-level (TBR=0.3-1) also increases. Again, these findings may suggest that the effects of Buparlisib vary with location within tumours. The weighted mean TBR value across all cohorts decreases significantly post-Buparlisib ($p<0.0001$); and weighted mean TBR values also decrease for each individual patient except Patient 1. These reductions in weighted mean TBR values concord with the decreases in the proportion of tumour cells with TBR greater than 1.4 noted in Chapter 3, and might be interpreted as indicating that Buparlisib is reducing tumour hypoxia.

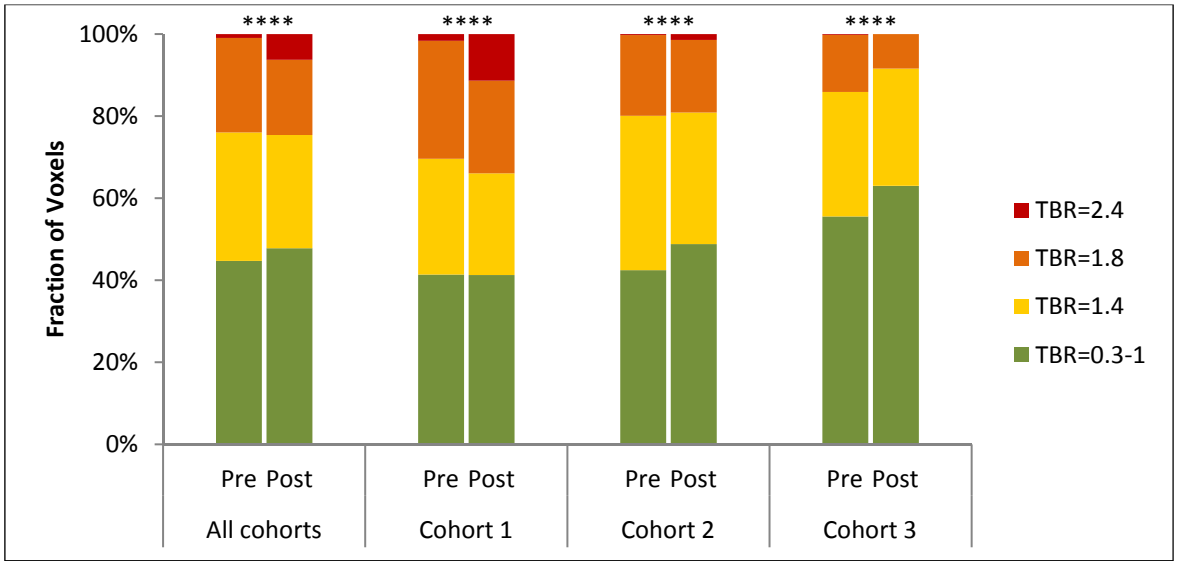


Figure 5.12 – Changes between pre- and post-Buparlisib images in percentages of fitted voxel TBR values clustered with all voxels together and by cohort. Asterisks indicate the statistical significance of changes between pre- and post-Buparlisib images.

5.B.6 Conclusions

A clustering algorithm has been used to distil the underlying information contained within parametric images from the noisy data on which they are based. Clustering all the data together yielded 6 cluster levels for $v_{B\text{-}2C}$, 14 for $K_{1\text{-}2C}$, 4 for $k_{3\text{-}2C}$, and 7 for TBR. Clustering the data for each patient separately (but including both the pre- and post-Buparlisib scan data) produced an average of 3 cluster-levels for $v_{B\text{-}2C}$, 5 for $K_{1\text{-}2C}$, 3 for $k_{3\text{-}2C}$, and 4 for TBR across the nine patients.

Following administration of Buparlisib, significant changes occurred in the clustering patterns of the $K_{1\text{-}2C}$ and $v_{B\text{-}2C}$ parameters. These changes are accompanied by a significant decrease in the overall weighted mean of $K_{1\text{-}2C}$ values, and a significant increase in the weighted means of $v_{B\text{-}2C}$ values for Cohorts 2 and 3, which is suggestive of vascular normalisation. The overall weighted mean of TBR values also falls significantly post-Buparlisib. Changes in $k_{3\text{-}2C}$ are more complex, however, with the direction of changes in weighted mean values differing both between cohorts and between individual patients within them: in Cohorts 1 and 3 the $k_{3\text{-}2C}$ weighted mean value falls significantly post-Buparlisib, whereas in Cohort 2 it significantly increases.

5.C Fitted kinetic rate-constants: variation with distance from the tumour edge

5.C.1 Introduction

It might be expected that perfusion would rise and levels of hypoxia would fall with increasing distance from the centre of tumours towards their edge. However, no work appears to have been published demonstrating this distance dependence based on PET kinetic modelling, although previous work using pCT imaging rather than PET has shown permeability and fractional blood volumes to be higher at tumour edges than at their centres (Ng *et al.*, 2007). The lack of PET studies exploring this varied effect according to distance is likely due to the fact that many of the dynamic PET scanning studies of hypoxia reported to date have been carried out for head-and-neck tumours which are relatively small in size. Several NSCLC patients from the Buparlisib trial have relatively large tumours, however, allowing the variation of the derived rate-constants based on distance from the edge of the tumour to be investigated in these patients.

5.C.2 Materials and methods

In-house MATLAB code (version R2014a, MathWorks, Natick, Massachusetts) was used to determine how far each tumour voxel (of dimension 5.5×5.5×3.3 mm^3) lies from the nearest edge of the outlined tumour volume as defined by a consultant radiologist. These distances were then separated into four categories: edge voxels on the edge of the VOI (the outermost shell of voxels), outer voxels (up to 5.5 mm from VOI edge), inner voxels (between 5.5 and 11 mm from VOI edge), and central voxels (greater than 11 mm from VOI edge). Clustering was carried out for kinetic parameter values obtained from 2C3K model fits to all 24,662 voxels, as described in 5.B.2, and then each voxel was assigned a distance category. The data was then displayed in bullseye graphs such that the central

voxels are represented by an inner ring, edge voxels by the largest diameter ring, and the other two layers represented by the intermediate rings, as illustrated in Figure 5.13.

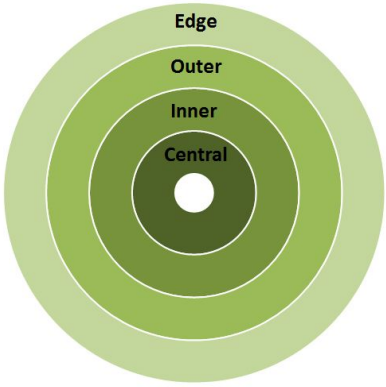


Figure 5.13 – Schematic bullseye graph showing distance categorisation.

For each distance category, the chi-square test of homogeneity was used to assess the significance of pre- to post-Buparlisib changes in the proportions of voxels assigned to each cluster level. Levels of significance $p<0.05$, 0.01, 0.001 and 0.0001 are indicated on the graphs by the symbols *, **, *** and **** respectively.

In order to assess whether the fitted parameters in each distance category increase or decrease overall post-Buparlisib, fitted weighted means of unclustered values have been calculated, as described in 5.B.2. The statistical significance of the difference in weighted mean of the parameter values in each distance category pre- to post-Buparlisib was then assessed using a t-test for samples of unequal variance (Welch's t-test).

5.C.3 Results: Variation of derived rate-constants with distance from the tumour edge

Clustered parameter values of voxels in the pre-Buparlisib image of Patient 6, categorised by distance from the tumour edge, are shown in Figure 5.14. For this patient it can be seen that the percentages of $K_{1\text{-}2C}$ and $k_{3\text{-}2C}$ values which cluster into different levels change very significantly ($p<0.0001$) between each voxel layer and the neighbouring layer lying closer to the tumour edge. Changes in the clustering patterns of $v_{B\text{-}2C}$ are similar to those seen for $K_{1\text{-}2C}$, although for $v_{B\text{-}2C}$ the changes generally have lower levels of significance. For $K_{1\text{-}2C}$, weighted mean values calculated for each distance category increase very significantly between each voxel layer and the neighbouring layer lying closer to the tumour edge ($p<0.0001$), while for $k_{3\text{-}2C}$, weighted mean values fall very significantly between the same layers (Table 5.9). For $v_{B\text{-}2C}$, weighted mean values increase insignificantly between each voxel layer and the neighbouring layer lying closer to the tumour edge ($p>0.05$). In Figure 5.15 these weighted mean values are plotted against tumour layer for $v_{B\text{-}2C}$, $K_{1\text{-}2C}$, and $k_{3\text{-}2C}$.

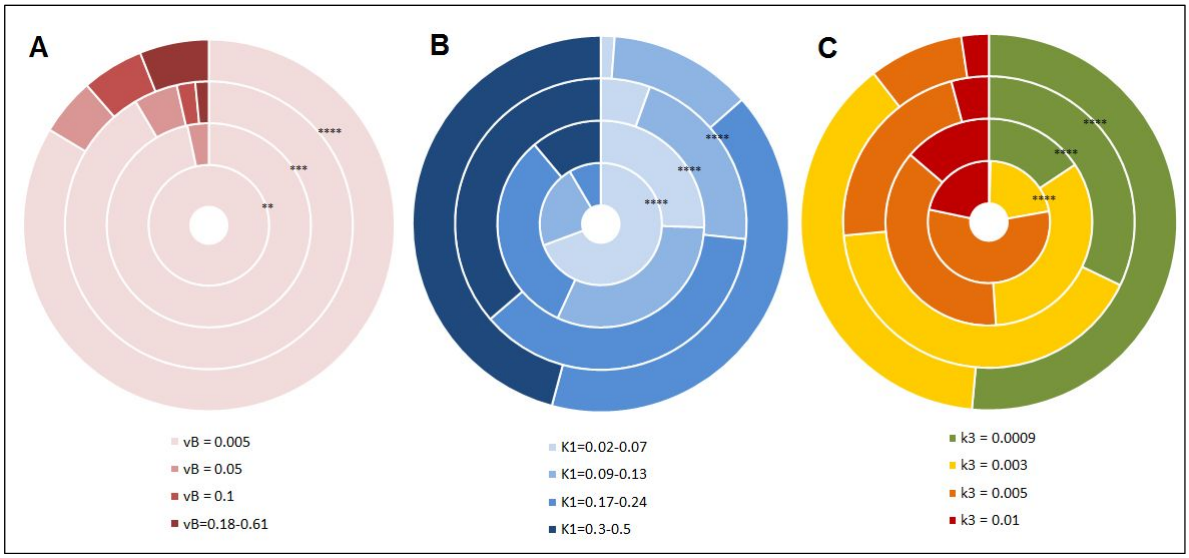


Figure 5.14 – Bullseye graphs showing distance variation of (A) $v_{B\text{-}2C}$, (B) $K_{1\text{-}2C}$, and (C) $k_{3\text{-}2C}$ for Patient 6 imaged pre-Buparlisib (tumour volume 180 mL). Asterisks indicate the statistical significance of differences in clustering patterns between adjacent distance categories.

Distance categories	$v_{B\text{-}2C}$	$K_{1\text{-}2C}$	$k_{3\text{-}2C}$
Centre to inner	↑NS	↑****	↓****
Inner to outer	↑NS	↑****	↓****
Outer to edge	↑NS	↑****	↓****

Table 5.9 – Direction of differences in weighted means between sequential distance categories pre-Buparlisib for Patient 6 imaged pre-Buparlisib; asterisks indicate the statistical significance of differences in weighted mean values between adjacent distance categories.

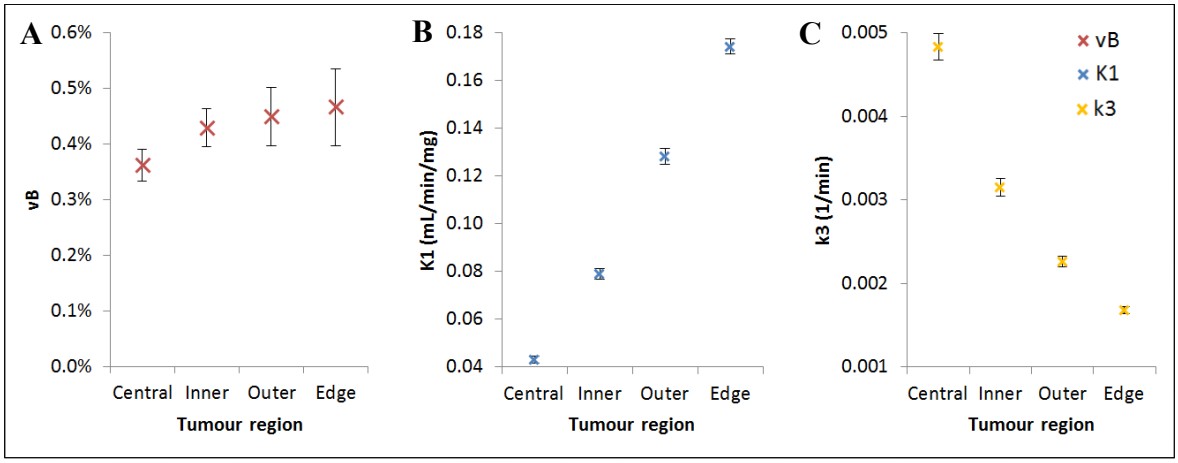


Figure 5.15 – Variation with distance from the tumour edge of weighted mean values of (A) $v_{B\text{-}2C}$, (B) $K_{1\text{-}2C}$, and (C) $k_{3\text{-}2C}$ parameters for Patient 6 imaged pre-Buparlisib. Error bars shown are standard error on the mean.

Clustered parameter values of all voxels in the pre-Buparlisib images, categorised by distance from the tumour edge, are shown in Figure 5.16. Significant differences ($p<0.001$) are seen between each sequential distance category in the proportions of voxels assigned to each cluster grouping.

Moving outwards from the tumour centre, weighted mean $k_{3\text{-}2C}$ values decrease very significantly between each voxel layer and the neighbouring layer lying closer to the tumour edge ($p<0.0001$), reflecting the changes seen for Patient 6. For $K_{1\text{-}2C}$, weighted mean values increase very significantly ($p<0.0001$) between each layer and the neighbouring layer lying further from the tumour centre until the outer layer is reached, when the weighted mean $K_{1\text{-}2C}$ then falls very significantly between this layer and the edge. For $v_{B\text{-}2C}$, weighted mean values rise significantly between each voxel layer and the

neighbouring layer lying closer to the tumour edge ($p<0.001$). In Figure 5.17 these weighted mean values are plotted against tumour layer for $v_{B\text{-}2C}$, $K_{1\text{-}2C}$, and $k_{3\text{-}2C}$.

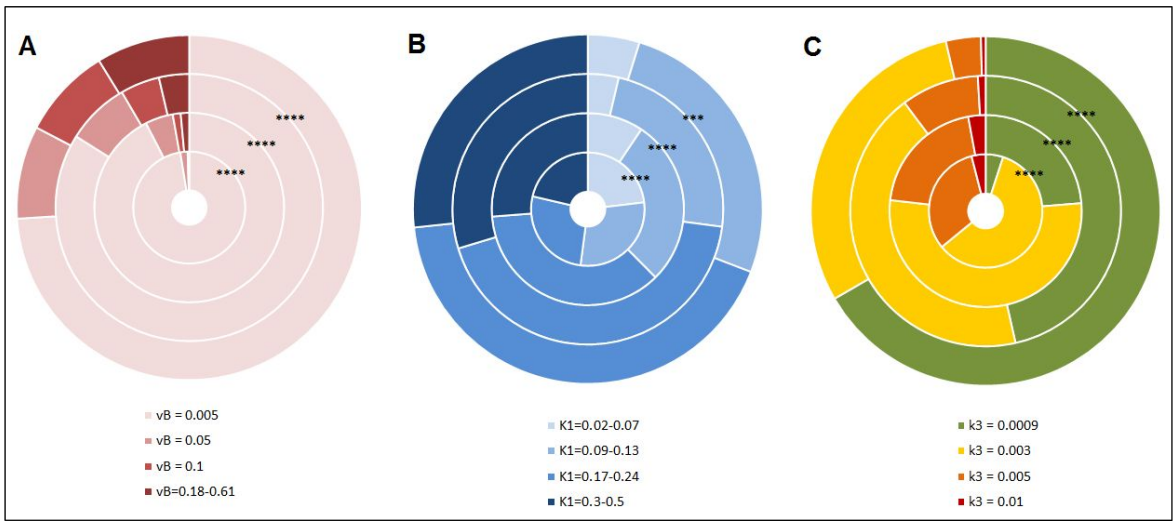


Figure 5.16 – Distance variation of the clustering of (A) $v_{B\text{-}2C}$, (B) $K_{1\text{-}2C}$, and (C) $k_{3\text{-}2C}$ values obtained from all pre-Buparlisib images. Asterisks indicate statistical significances of differences between adjacent distance categories.

Distance categories	$v_{B\text{-}2C}$	$K_{1\text{-}2C}$	$k_{3\text{-}2C}$
Centre to inner	↑***	↑****	↓****
Inner to outer	↑****	↑****	↓****
Outer to edge	↑***	↓***	↓****

Table 5.10 – Direction of differences in weighted means between sequential distance categories pre-Buparlisib for all patients; asterisks indicate statistical significances of differences between adjacent distance categories.

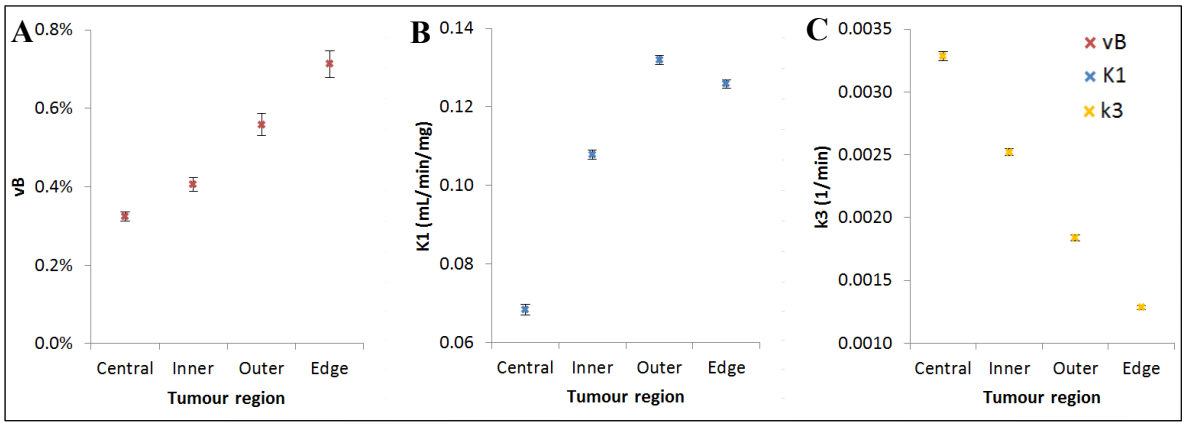


Figure 5.17 – Variation with distance from the tumour edge of weighted mean values of (A) $v_{B\text{-}2C}$, (B) $K_{1\text{-}2C}$, and (C) $k_{3\text{-}2C}$ parameters for all patients pre-Buparlisib. Error bars shown are standard error on the mean.

5.C.4 Results: Influence of distance from the tumour edge on the impact of Buparlisib on FMISO uptake kinetics

In Figures 5.18-20 bullseye graphs representing percentages of voxels clustered into different data groups are shown for the parameters $v_{B\text{-}2C}$, $K_{1\text{-}2C}$, and $k_{3\text{-}2C}$ respectively. The data included in the graphs represents all the tumour voxels imaged, but is split between voxels in pre- and post-Buparlisib images. The data is presented for all patients together, and split by cohort. From Figures 5.18 to 5.20 it can be seen that there is a large amount of heterogeneity in all kinetic parameters between the three cohorts pre-Buparlisib. Due to this each cohort has been treated separately for comparisons pre- to post-Buparlisib.

The clustering of parameter values was carried out before the splitting of the data, thereby ensuring consistent cluster-levels across the different graphs. Figure 5.21 presents summary bullseye graphs showing changes in weighted mean values for $v_{B\text{-}2C}$, $K_{1\text{-}2C}$, and $k_{3\text{-}2C}$ for each distance category, calculated for individual cohorts and for all patients together.

Changes in clustering patterns are seen after administration of Buparlisib, which for $v_{B\text{-}2C}$ and $K_{1\text{-}2C}$ are very significant ($p<0.0001$) for the majority of distance categories and cohorts (Figures 5.18 and 5.19). Weighted mean values of $K_{1\text{-}2C}$ (Figure 5.21B) decrease significantly after Buparlisib administration for all cohorts and distance categories ($p<0.0001$) except for Cohort 1 central voxels, for which $K_{1\text{-}2C}$ decreases insignificantly ($p>0.05$). Weighted mean values of $v_{B\text{-}2C}$ (Figure 5.21A) increase overall post-Buparlisib, and for Cohorts 2 and 3 these increases are significant for all but one distance category ($p<0.05$). For Cohort 1 $v_{B\text{-}2C}$ falls insignificantly post-Buparlisib ($p>0.05$) in all but the central voxels, where $v_{B\text{-}2C}$ decreases significantly post-Buparlisib ($p<0.01$). These findings match the conclusions of the overall analysis (not split by distance category) (5.B.5), but here it can be seen that vascular normalisation (lower $K_{1\text{-}2C}$, higher $v_{B\text{-}2C}$)

occurs throughout tumour volumes in Cohorts 2 and 3, while in Cohort 1 there is less evidence of vascular normalisation.

For $k_{3\text{-}2C}$ the response to Buparlisib is less coherent, as can be seen in Figure 5.20. The percentages of voxels assigned to different $k_{3\text{-}2C}$ cluster-levels change significantly after administration of Buparlisib in all distance categories when all patients are considered together. However, changes for some distance categories are insignificant in Cohorts 2 and 3 when considered separately. Generally the patterns of change seen are complex, with a tendency for the percentages of voxels lying within both the lowest cluster-level and the highest two cluster-levels to increase post-Buparlisib.

Overall, the weighted mean value of $k_{3\text{-}2C}$ (Figure 5.21C) decreases very significantly for edge category voxels post-Buparlisib and decreases insignificantly in the outer layer, but rises significantly in the inner and central layers. For individual cohorts the picture is more mixed. For Cohort 1 the weighted mean $k_{3\text{-}2C}$ value falls significantly in the edge and outer layers post-Buparlisib, but increases significantly in the central layer. For Cohort 2 the weighted mean $k_{3\text{-}2C}$ rises significantly in all layers after administration of Buparlisib, while for Cohort 3 it rises significantly in the inner and central layers but falls significantly in the outer and edge layers. Generally, then, weighted mean $k_{3\text{-}2C}$ values fall in regions lying close to the tumour edge following administration of Buparlisib, but rise in more central regions.

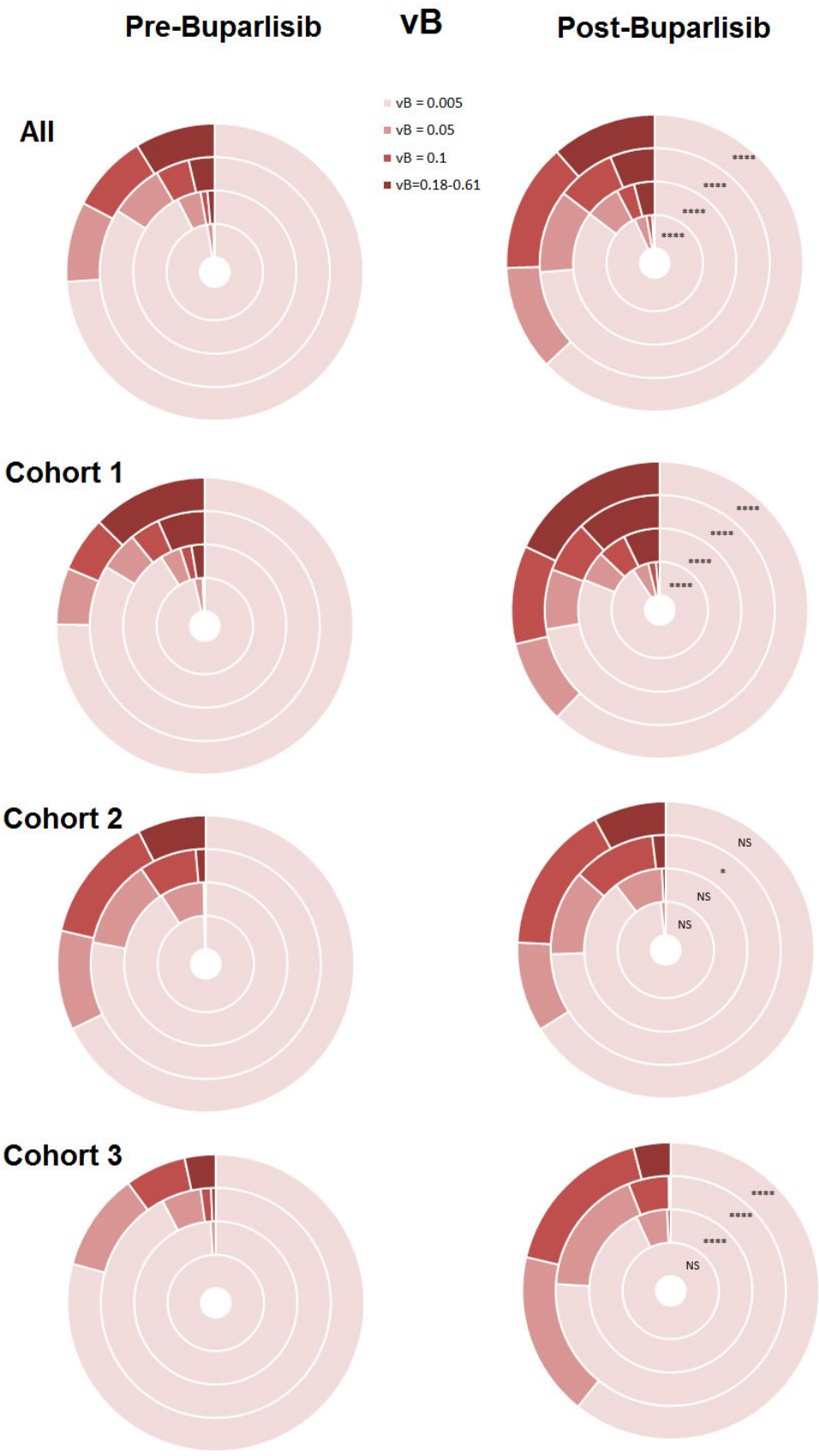


Figure 5.18 – Bullseye graphs showing pre- versus post-Buparlisib $v_{B\text{-}2C}$ changes in fractions of voxels clustered into different cluster-levels, both for all patients together (top) and by cohort. Asterisks indicate statistical significances of pre- versus post-Buparlisib differences.

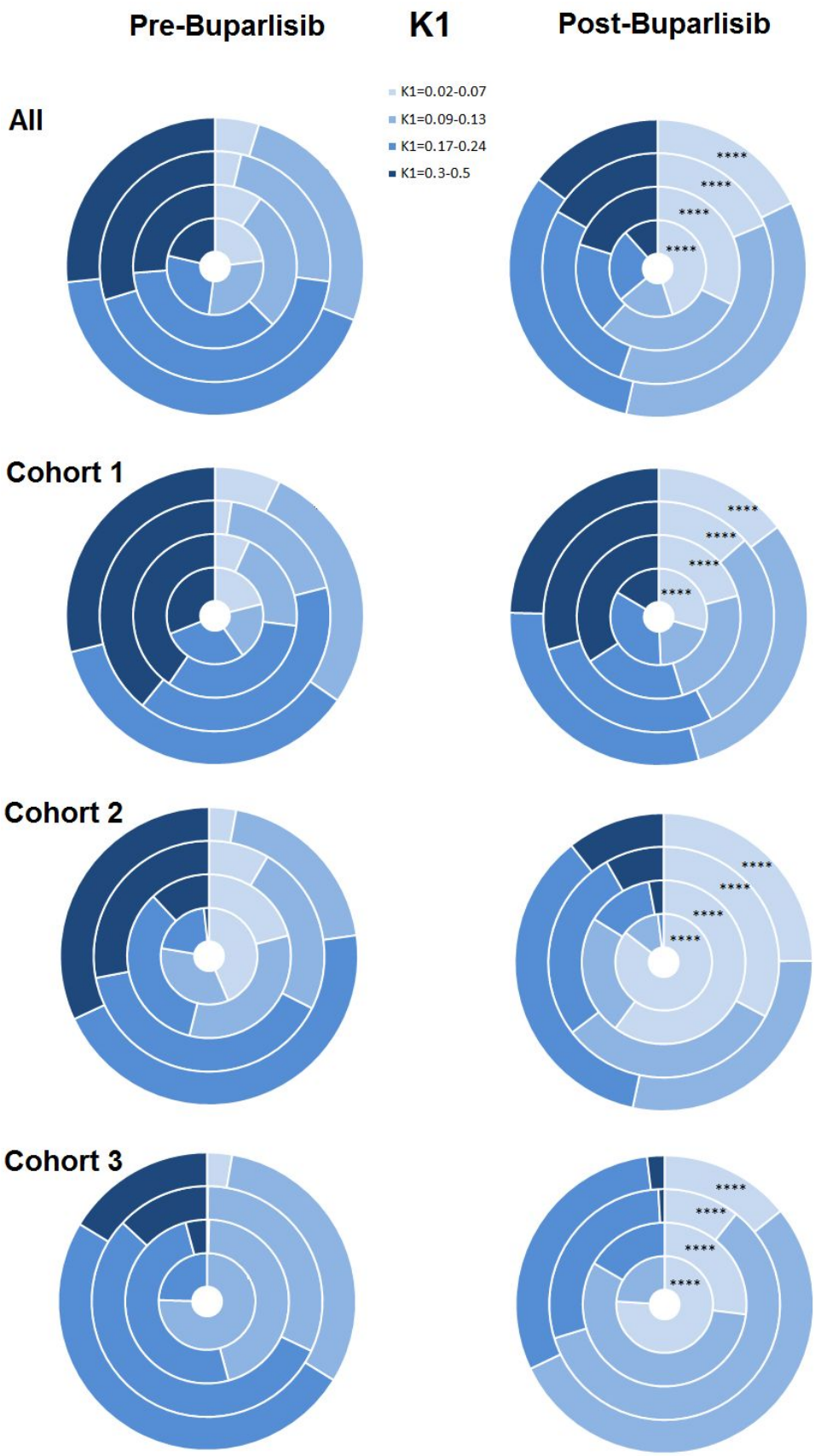


Figure 5.19 – Bullseye graphs showing pre- versus post-Buparlisib $K_{1\text{-}2C}$ changes in fractions of voxels clustered into different cluster-levels, both for all patients together (top) and by cohort. Asterisks indicate statistical significances of pre- versus post-Buparlisib differences.

Figure 5.20 – Bullseye graphs showing pre- versus post-Buparlisib $k_{3\text{-}2C}$ changes in fractions of voxels clustered into different cluster-levels, both for all patients together (top) and by cohort. Asterisks indicate statistical significances of pre- versus post-Buparlisib differences.

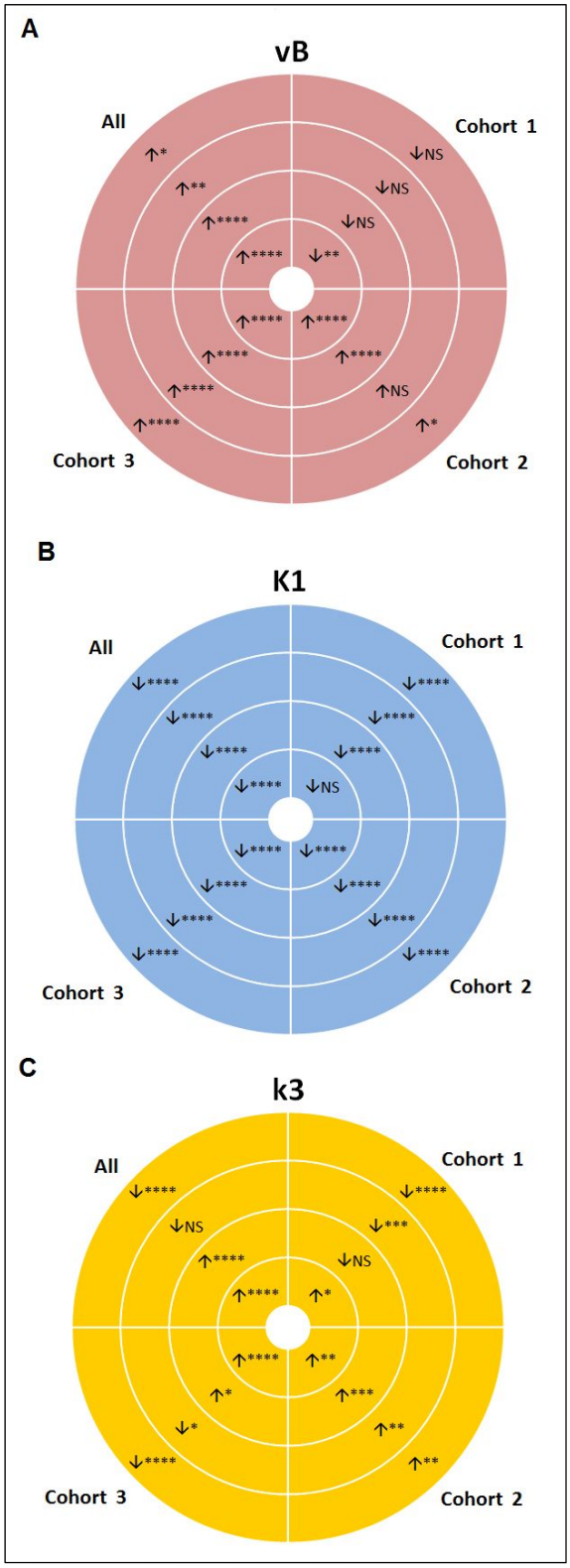


Figure 5.21 – Bullseye graphs showing the direction and significance of changes in weighted mean values for (A) $v_{B\text{-}2C}$, (B) $K_{1\text{-}2C}$, and (C) $k_{3\text{-}2C}$ following administration of Buparlisib. Asterisks indicate the degree of statistical significance of the pre- versus post-Buparlisib changes for each distance category.

5.C.5 Conclusions

Significant variation has been found between the distribution of parameter values in voxels lying at different distances from the tumour edge, with weighted mean values of $v_{B\text{-}2C}$ and $K_{1\text{-}2C}$ generally increasing towards the tumour edge and $k_{3\text{-}2C}$ decreasing towards the tumour edge.

After administration of Buparlisib, significant changes have been observed in the weighted means of $v_{B\text{-}2C}$ (increasing for Cohorts 2 and 3) and of $K_{1\text{-}2C}$ (decreasing for all cohorts) throughout the tumour volume. The response of $k_{3\text{-}2C}$ to Buparlisib is distance dependent: weighted mean values decrease in edge and outer distance categories and increase in inner and central categories when all voxels are considered together. This overall pattern is also seen for Cohorts 1 and 3, but for Cohort 2 the weighted mean values of $k_{3\text{-}2C}$ increase significantly post-Buparlisib for voxels in all distance categories.

5.D Overall conclusions of the voxel-by-voxel kinetic analysis

In this Chapter it has been shown that fits of the irreversible 2C3K two-compartment model to TACs obtained for individual tumour voxels provide estimates of rate-constants which have lower total uncertainties than the fitted parameter values obtained for any of the other models studied. A comprehensive voxel-by-voxel analysis of the uptake kinetics of FMISO has been carried out using the 2C3K model, studying changes in FMISO kinetics following administration of Buparlisib, and variation of the kinetics with depth within tumours.

Variations in the weighted mean values of $v_{B\text{-}2C}$ and $K_{1\text{-}2C}$ have been shown, with both generally increasing significantly from the tumour centre towards the tumour edge and the weighted mean values of $k_{3\text{-}2C}$ decreasing significantly towards the tumour edge.

Following administration of Buparlisib, a significant decrease in the weighted mean of $K_{1\text{-}2C}$ was observed for all cohorts, and a significant increase in the weighted mean of $v_{B\text{-}2C}$ for Cohorts 2 and 3, which is suggestive of vascular normalisation. These changes in $K_{1\text{-}2C}$ and $v_{B\text{-}2C}$ post-Buparlisib have been shown to occur throughout the tumour volume.

Variation within the weighted mean values of $k_{3\text{-}2C}$ is more complex, with the direction of change differing both between cohorts and between individual patients within them: in Cohorts 1 and 3 the $k_{3\text{-}2C}$ weighted mean falls significantly post-Buparlisib, whereas in Cohort 2 it increases significantly. The response of $k_{3\text{-}2C}$ to Buparlisib appears to be distance dependent: weighted mean values decrease in edge and outer distance categories and increase in inner and central categories, when all voxels are considered together.

6 Perfusion CT

6.1 Abstract

A study is presented of the analysis of pCT data from the Buparlisib trial. A first pass pCT technique was performed on all the patients pre- and post-Buparlisib. The pCT data was analysed using a commercial software model (AATH) as well as an alternative model, the Extended Toft's model (ETM), which models the blood volume in a similar way to PET kinetic modelling.

There is a decrease in blood flow (BF) post-Buparlisib for Cohorts 2 and 3, significant for Cohort 2. There is a significant decrease in $v_{B\text{-}pCT\text{-}AATH}$ for all patients post-Buparlisib.

A good correlation (r=0.75) has been shown between K_1 as measured from PET kinetic modelling and BF×E (the equivalent parameter) as measured by pCT analysed by the commercial software model. Using the alternative pCT model a good correlation (r=0.69) is seen between $v_{B\text{-}pCT\text{-}ETM}$ and $v_{B\text{-}PET}$.

6.2 Introduction

6.2.1 Perfusion CT analysis

There are a variety of methods to analyse pCT data, based on the relationship between the concentration of iodine within vessels and tissue being linearly proportional to changes in measured CT attenuation. Iodine is injected as a bolus through a venous cannula and its passage through the patient recorded in order to measure tissue perfusion, a technique first proposed in 1980 (Axel, 1980).

Several commercial manufacturers have implemented different kinetic models to analyse this data; GE systems use the Adiabatic Approximation to the Tissue Homogeneity (AATH) model (St Lawrence and Lee, 1998). This is essentially a deconvolution technique which calculates the impulse residue function (IRF) from the arterial and tissue TACs (Petralia *et al.*, 2010). The IRF represents the response of each tissue voxel that would have resulted from an ideal (instantaneous) injection of the contrast agent: an example is shown in Figure 6.1 with its associated equation 6.1. The algorithm then calculates various perfusion parameters from this: blood flow (BF), blood volume ($v_{\text{B-pCT-AATH}}$), mean transit time (MTT), and permeability surface area product (PS). BF is the flow rate through vasculature in tissue, $v_{\text{B-pCT-AATH}}$ the volume of blood within vasculature, MTT the time from artery to vein, and PS the rate-constant for flux between the plasma and interstitial space. $v_{\text{B-pCT-AATH}}$ is calculated by multiplying BF by MTT, using the central volume principal (Meier and Zierler, 1954).

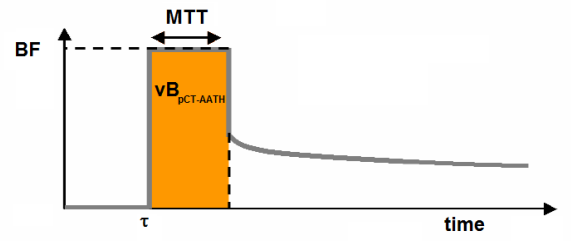


Figure 6.1 – Example AATH IRF.

$$IRF(t) = \begin{cases} 0 & t < \tau \\ BF & \tau < t < MTT + \tau \\ BF \times E(1 - e^{-(t-MTT-\tau)PS/BF}) & MTT + \tau < t \end{cases} \quad (6.1)$$

where E is the extraction fraction (defined in equation 6.4), and τ the contrast arrival delay.

An alternative perfusion model is the Extended Toft's model (ETM) (Tofts, 1997; Sourbron and Buckley, 2013). This is more constrained than the AATH model, giving three perfusion-related parameters rather than four: the volume transfer constant k_{trans}, vascular volume $v_{B\text{-}pCT\text{-}ETM}$, and interstitial volume v_e. It models the vasculature in a similar way to PET kinetic modelling with no mean transit time, meaning that the input function appears as a sharper peak unblurred by the transit time. The IRF equation for ETM is:

$$IRF(t) = \begin{cases} 0 & t < \tau \\ \delta(t - t_0)v_{B-pCT-ETM} + k_{trans}e^{-\frac{(t-t_0)k_{trans}}{v_e}} & \tau < t \end{cases} \quad (6.2)$$

There has been some work on the reproducibility of perfusion parameters using various models. Work by Ng on lung tumours showed large variability, with changes of 33-148% required for the different parameters to be significantly greater than the test-retest variability (Ng *et al.*, 2011). However, this was performed on a 16-slice CT, whereas a 12 patient study on a 64-slice CT had smaller errors with within-subject coefficient of variations of 9.3%, 16.4%, 11.2% and 14.9% for $v_{B\text{-}pCT\text{-}AATH}$, BF, MTT and PS respectively (Larici *et al.*, 2014). Using the published Bland-Altman plots, changes of 26%, 43%, 27%, and 35% are required for a change to be significantly greater than the test-retest variability (Larici *et al.*, 2014).

6.2.2 Perfusion in PET

The rate-constant derived from PET kinetic modelling, K_1, is associated with blood flow (BF) by

$$K_1 = E \times BF \tag{6.3}$$

where E is the extraction fraction defined by considering the capillary as a cylinder as (Renkin, 1959; Crone, 1963):

$$E = 1 - \exp(-PS/BF) \tag{6.4}$$

where P is the permeability and S the surface area (and PS the permeability surface area product). For highly permeable tracers the product PS is much greater than BF giving an extraction fraction close to 1: these tracers are useful for measuring blood flow and are flow-limited. Alternatively, where BF is greater than PS (with K_1 therefore tending to PS), tracers are used to measure permeability and are permeability-limited.

It is thus possible to associate the measurement of flow calculated from a pCT scan (BF) with K_1 derived from PET kinetic modelling if the PS for the tracer, in this case FMISO, is known.

The permeability can be estimated using (Laidler, 1978; Cussler, 1997):

$$P = DK/\Delta x \tag{6.5}$$

where D is the diffusivity of misonidazole shown to be 5.5×10^{-11} m^2 s^{-1} for a multi cellular membrane (Cowan *et al.*, 1996), K the partition fraction of 0.40 for FMISO (Grunbaum *et al.*, 1987), and Δx the width of capillary membrane, 0.2-1 μm (Marieb, 2005; Wieser *et al.*, 2005). This gives a permeability in the range of $2.2–11\times10^{-5}$ m s^{-1}. The centre of this range is 6.6×10^{-5} m s^{-1} = 4.0×10^{-3} m min^{-1}.

The capillary surface area to volume ratio is given by:

$$S/V = 2\pi rL/\pi r^2 L = 2/r \tag{6.6}$$

where L is the capillary length and r the capillary radius of approximately 10 μm (Marieb, 2005) giving $S/V = 2\times10^5$ m^{-1}.

In pCT analysis, BF is measured in mL/100g/min, and so with units of mL/g/min for K_1:

$$K_1 = E \times \frac{BF}{100} = \left(1 - \exp\left(-\frac{100PS}{BF}\right)\right)\frac{BF}{100} = \left(1 - \exp\left(-\frac{80}{BF}\right)\right)\frac{BF}{100} \tag{6.7}$$

is the expected relationship between K_1 derived from PET kinetic modelling and BF from pCT. A single estimate is used for FMISO PS here, because the PS from the pCT analysis, calculated for each individual patient, relates to the iodine contrast agent and generally requires a delayed imaging phase to achieve an accurate measurement (Miles, 2003; Miles *et al.*, 2012).

This chapter will investigate the changes in derived perfusion parameters post-Buparlisib and the correlation of these parameters with the corresponding ones from PET kinetic modelling.

6.3 Materials and methods

6.3.1 Perfusion CT imaging protocol

Following the FMISO PET/CT scan at 4 hours post-injection, the pCT imaging protocol takes place as detailed in Figure 6.2. Patients are imaged in a supine position with their arms raised (if possible), using a GE Discovery 690 PET/CT scanner (GE Healthcare, Milwaukee, USA). A pre-contrast CT is performed for localisation of the region that the pCT scan will be performed on. During the pCT scan the patient is instructed to hold their breath for as long as possible (at inspiration) and to breathe out very slowly if they need to.

The pCT scan takes an image each second for 45 s with the venous CT taking place immediately afterwards, over the same volume as the pre-contrast CT.

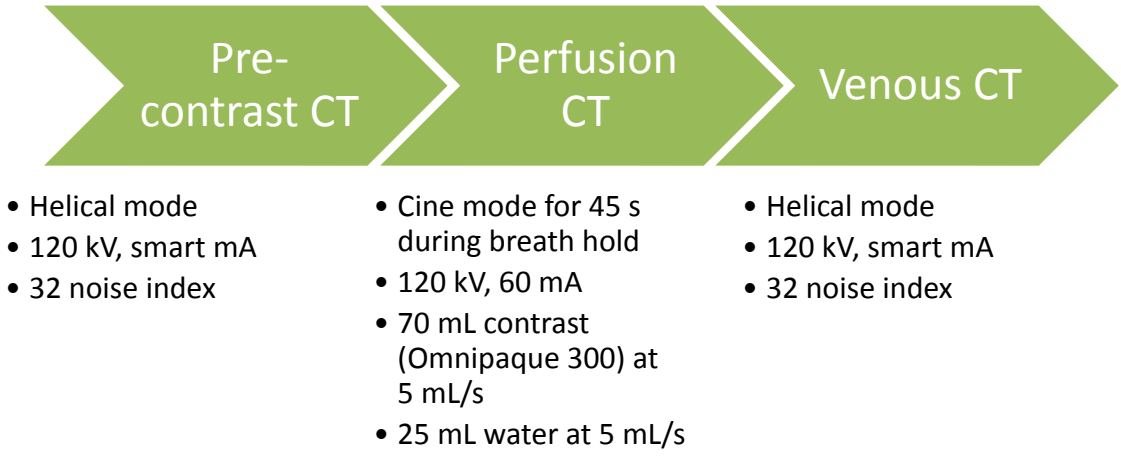


Figure 6.2 – Details of pCT imaging protocol.

Whilst the pre-contrast CT can have a large axial coverage, the pCT is limited to 4 cm axially due to high temporal resolution (one image per second) and the CT detector width. For larger tumours this means that the entire tumour is not fully covered by the pCT scan, as shown in Figure 6.3. Prior to patient scanning a consultant radiologist defines a reproducible CT landmark that can be used to maximise the amount of viable tumour included in the pCT coverage.

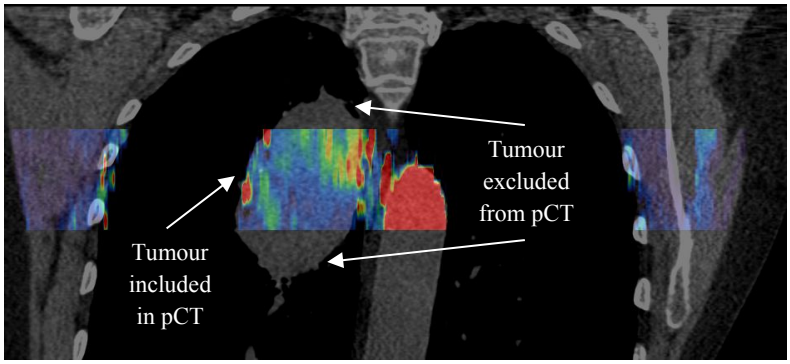


Figure 6.3 – Coronal slice of Patient 5; the section in colour indicates the volume over which the pCT was performed and the underlying greyscale CT is from the pre-contrast CT.

6.3.2 Perfusion CT image analysis

Once the pCT images have been acquired they are processed as shown in Figure 6.4.

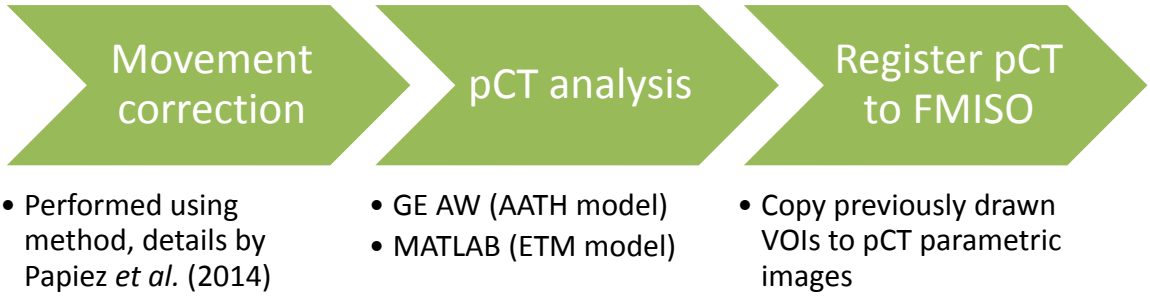


Figure 6.4 – Detail of pCT analysis protocol.

A motion correction step is used to pre-process the pCT data as it is very difficult for the patient to maintain a breath hold for 45 s. Movement during the scan is an issue for pCT images as the patient is in an extreme position (full inspiration) and the voxel size is very small in the x- and y-directions (0.7 mm), which is why Oxford's engineering department has developed software that corrects for this movement (Papież *et al.*, 2014). The motion-corrected pCT data is then processed voxel-by-voxel using the commercial software GE Perfusion 4D (GE Healthcare, Milwaukee, USA) which generates pCT parametric images for BF, $v_{\text{B-pCT-AATH}}$, MTT, and PS. Additionally the ETM model was fitted to the data by collaborators (Oxford Institute of Biomedical Engineering) using MATLAB code (version R2014a, MathWorks, Natick, Massachusetts) which generated pCT result images for k_{trans}, $v_{\text{B-pCT-ETM}}$, and v_{e}. The input function for both models was taken from the mean of a ROI drawn in the centre of the descending aorta.

The pre-contrast CT (hard registered to the pCT result images) was rigidly registered (also allowing for manual adjustment if required) to the CT of the dynamic FMISO PET/CT scan using the Hermes Hybrid Viewer software (Hermes Medical AB, Sweden). For tumours where the 4 cm axial coverage did not include the entire tumour the PET kinetic

analysis was repeated over the volume that was within the pCT, as per the method detailed in 4.3.2, so that the PET kinetic modelling and pCT results could be compared. The previously defined tumour VOIs were copied from the dynamic PET study to the pCT result images, voxel values from the pCT parametric images were exported, and then the mean values from the imaged tumour for each of the pCT parameters were recorded. A two-tailed paired t-test was used on these mean values to assess the change due to Buparlisib in each cohort.

6.4 Results

6.4.1 Perfusion CT data

The tumour pCT result values for each patient pre- and post-Buparlisib are shown in Table 6.1, analysed using the AATH model.

Cohort	**Patient**	**AATH Model**				**Volume of tumour imaged using pCT** (mL)	**Fraction of the PET-imaged tumour volume** (%)
		BF (mL/ 100g/ min)	$v_{B\text{-}pCT\text{-}AATH}$ (mL/ 100g)	MTT (s)	PS (mL/ 100g /min)		
Cohort 1 (50 mg Buparlisib)	1_{preC1}	42.0	3.6	8.6	14.3	13	100
	1_{postC1}	42.6	2.6	6.2	14.8	13	100
	2_{preC1}	88.4	9.0	11.2	8.7	140	27
	2_{postC1}	118.5	8.0	8.3	7.4	140	29
	3_{preC1}	64.2	4.0	6.3	9.2	74	70
	3_{postC1}	60.2	3.8	5.8	7.6	74	75
Cohort 2 (80 mg Buparlisib)	4_{preC2}	102.2	7.2	6.4	7.8	29	100
	4_{postC2}	84.3	4.7	6.2	6.0	27	100
	5_{preC2}	55.8	4.0	8.1	5.7	93	68
	5_{postC2}	36.2	3.0	10.1	6.0	100	71
	6_{preC2}	106.2	6.5	7.8	6.0	87	49
	6_{postC2}	94.5	6.2	6.1	6.8	84	48
Cohort 3 (100 mg Buparlisib)	7_{preC3}	79.5	7.7	7.8	7.1	29	65
	7_{postC3}	69.2	5.9	8.1	9.6	22	58
	8_{preC3}	69.1	7.6	10.1	12.7	115	61
	8_{postC3}	35.7	3.1	9.3	11.1	111	61
	9_{preC3}	58.6	4.8	8.2	8.7	40	100
	9_{postC3}	47.4	3.2	6.9	9.0	37	100

Table 6.1 – Mean pCT parameters for the AATH and ETM model for patients' tumours pre- and post-Buparlisib.

6.4.2 Response to Buparlisib

Figure 6.5 shows BF results for an example axial slice pre- to post-Buparlisib of an AATH model pCT parametric image, showing a decrease in BF post-Buparlisib. The AATH model pCT parameters pre- and post-Buparlisib are shown in Figures 6.6 and 6.7.

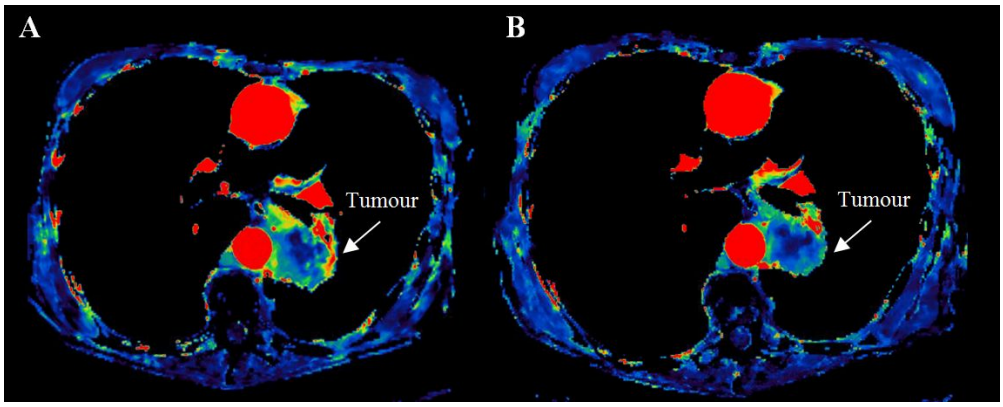


Figure 6.5 – pCT BF results for Patient 7 (A) pre-, and (B) post-Buparlisib, showing a decrease in BF. Colour scale with values 0-200 ml/100g/min with red indicating values above 200.

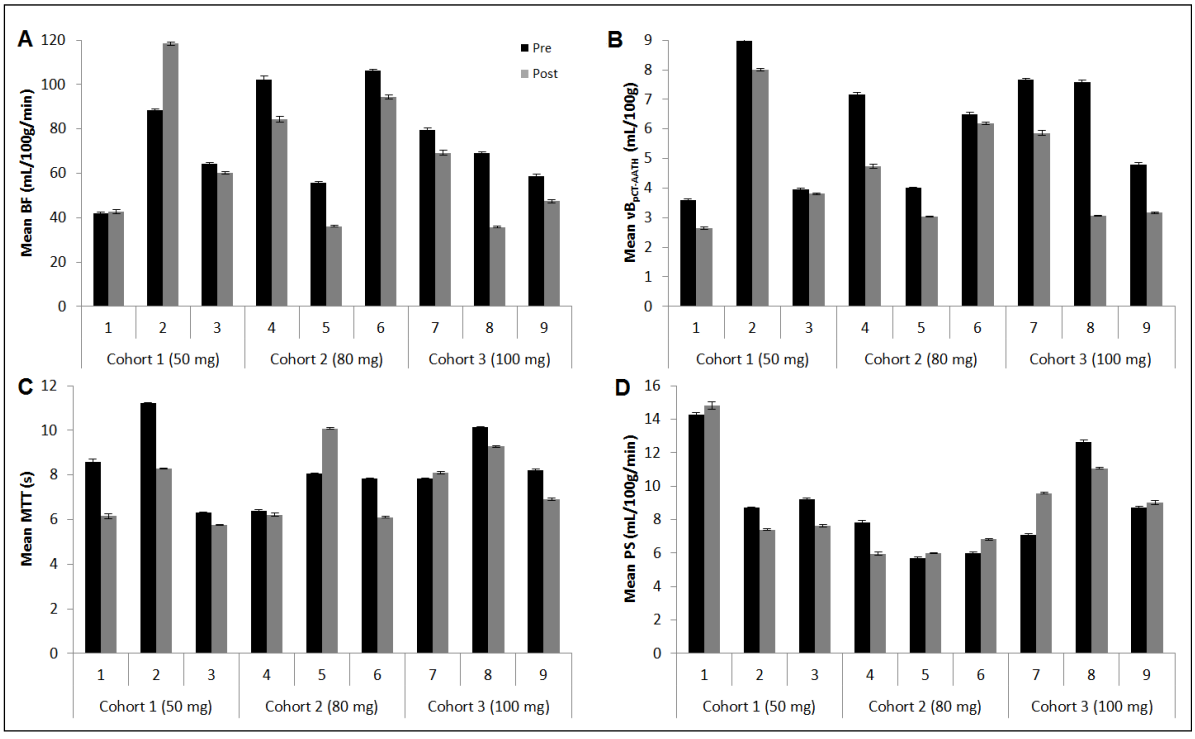


Figure 6.6 – Mean AATH pCT parameter values in (A) BF, (B) $v_{B\text{-}pCT\text{-}AATH}$, (C) MTT, and (D) PS pre- and post-Buparlisib. Error bars are the standard error on the mean.

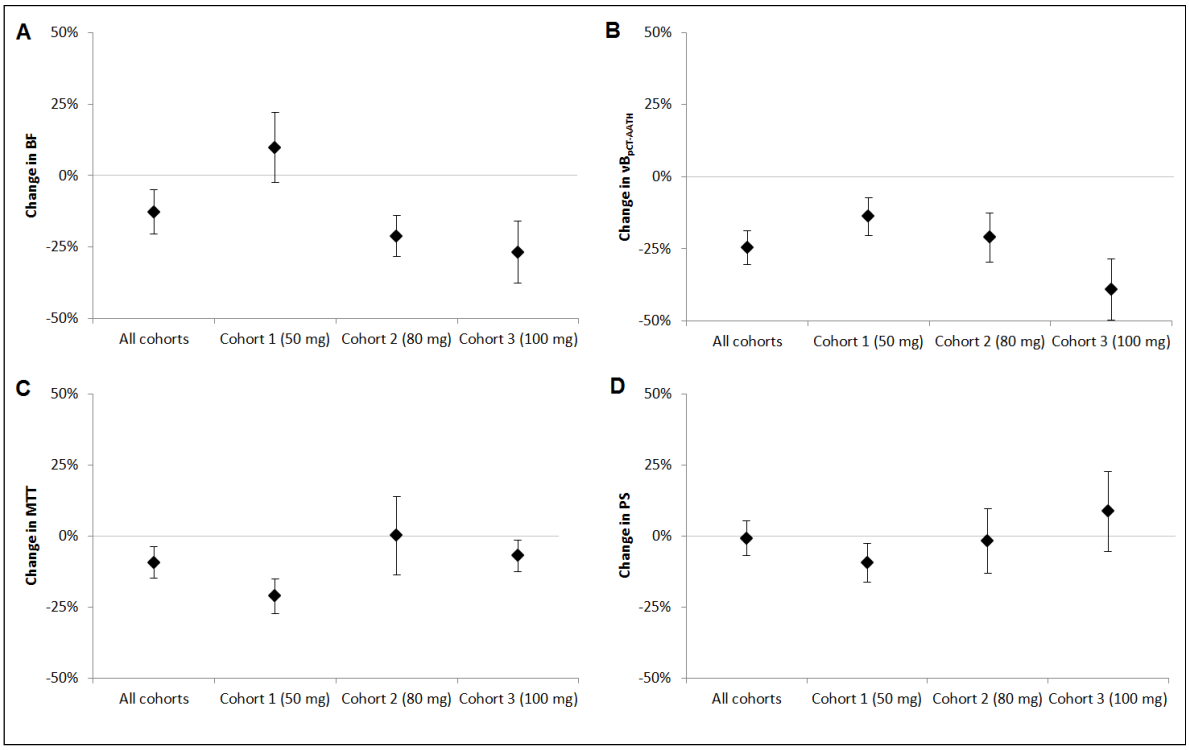


Figure 6.7 – Mean percentage change post-Buparlisib for all cohorts and each cohort in (A) BF, (B) $v_{B\text{-}pCT\text{-}AATH}$, (C) MTT, and (D) PS. Error bars are the standard error on the mean.

The statistical significance of differences from pre- to post-Buparlisib for each perfusion parameter are summarised in Table 6.2. There are no significant differences between cohorts of the mean cohort change in each perfusion parameter ($p>0.05$). A paired t-test was chosen to assess Buparlisib response as the standard error on the perfusion parameters is considerably less than the patient-to-patient parameter variability.

Patients	BF	$v_{B\text{-}pCT\text{-}AATH}$	MTT	PS
All	-	0.009	-	-
C1	-	-	-	-
C2	0.02	-	-	-
C3	-	-	-	-

Table 6.2 – Statistical differences from pre- to post-Buparlisib considering all patients together and each cohort individually. Dashes indicate $p>0.05$.

For both Cohorts 2 and 3 the mean BF for each tumour is reduced post-Buparlisib (significantly for Cohort 2), but for Cohort 1 there is a mixed response (Figure 6.6A); whereas the K_1 from whole tumour PET kinetic modelling decreased for all cohorts. The mean cohort BF change is higher for Cohort 1 and lower for Cohorts 2 and 3 (Figure 6.7A).

For all patients $v_{B\text{-}pCT\text{-}AATH}$ decreases post-Buparlisib (Figure 6.6B); this is the reverse of the results from the PET kinetic analysis, which found $v_{B\text{-}PET}$ increases in all patients, and may be due to the differences between the way AATH (pCT) and 3C5K (PET) model the blood (explored in section 6.4.3). The $v_{B\text{-}pCT\text{-}AATH}$ decrease is higher for Cohort 3 than Cohorts 1 or 2 (Figure 6.7B).

For MTT and PS there is a mixed change from the patients post-Buparlisib (Figure 6.6C and D); there are no directly corresponding PET parameters for these. The mean cohort change graphs show that post-Buparlisib there is minimal (and not significant) change in MTT and PS (Figure 6.7C and D). Generally accurate PS measurement requires a longer perfusion acquisition (at least 120 s with sparser temporal sampling after 45 s) rather than just the first pass protocol used here (Miles, 2003; Miles *et al.*, 2012).

6.4.3 Correlation with FMISO PET kinetics

Figure 6.8 shows the relationship between K_1 as measured from PET kinetic modelling and the equivalent combination of parameters (BF×E, as detailed in 6.2.2, with the PS for FMISO estimated as 80 mL/100g/min) as measured using pCT and analysed by the AATH model. This shows a good correlation between the two independent measurements of K_1 (Pearson's r coefficient, r=0.75). Measured PS values from fits to the pCT data (Table 6.1) are less than the 80 mL/100g/min PS value estimated for FMISO (section 6.2.2), potentially due to differences in permeability for FMISO and the iodine CT contrast agent, in addition to errors in the pCT PS estimation due to the first pass technique used (Miles, 2003; Miles *et al.*, 2012).

In comparison, there is a poor correlation between $v_{B\text{-}PET}$ (measured from PET kinetic modelling) and $v_{B\text{-}pCT\text{-}AATH}$ (r=0.32), shown in Figure 6.9.

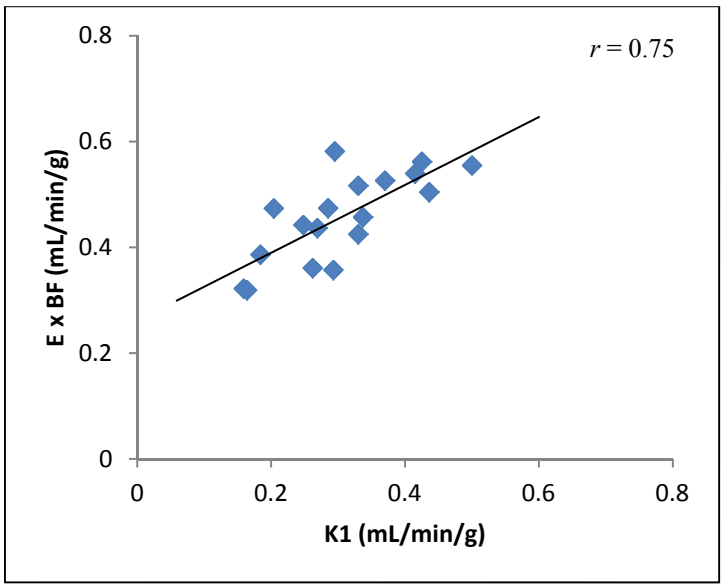


Figure 6.8 – Graph of K_1 (measured from PET) vs BF×E (measured from AATH pCT).

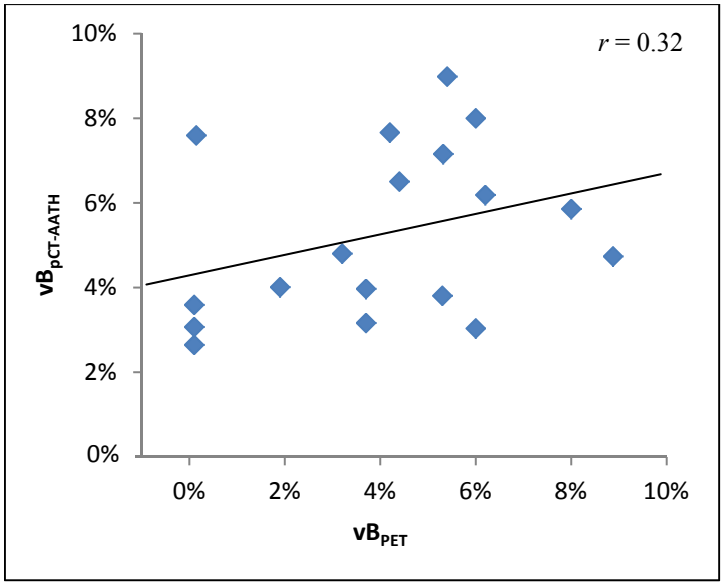


Figure 6.9 – Graph of $v_{\text{B-PET}}$ versus $v_{\text{B-pCT-AATH}}$.

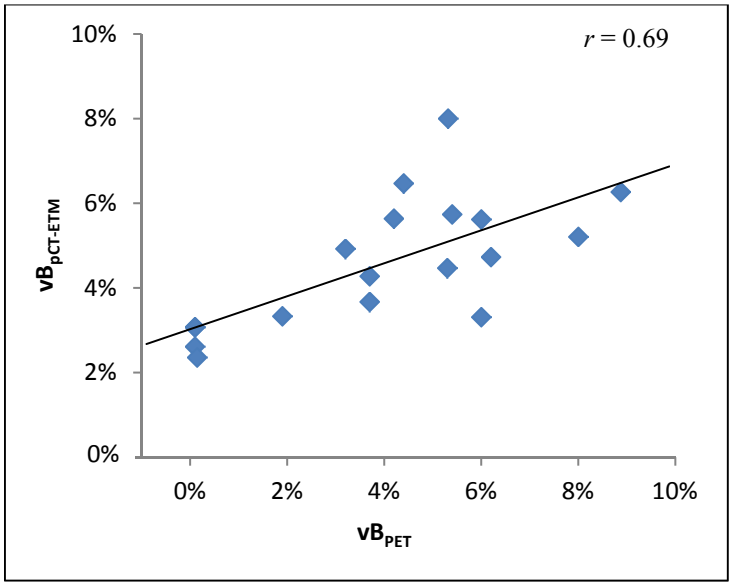


Figure 6.10 – Graph of $v_{\text{B-PET}}$ versus $v_{\text{B-pCT-ETM}}$.

The poor correlation between $v_{B\text{-}PET}$ and $v_{B\text{-}pCT\text{-}AATH}$ (r=0.32) might be expected as the pCT AATH model considers the blood within each voxel in a different way to the PET kinetic model; whereas the AATH model IRF has an extended top-hat element, in the PET kinetic model and in ETM this is replaced by a delta function (6.2.1). A comparison of $v_{B\text{-}pCT\text{-}ETM}$ and $v_{B\text{-}PET}$ yields an improved correlation (r=0.69, Figure 6.10). However, there are still non-significant decreases in $v_{B\text{-}pCT\text{-}ETM}$ post-Buparlisib for five patients (Figure 6.11), whereas no patients had decreased $v_{B\text{-}PET}$ post-Buparlisib. The remaining differences to PET kinetic results is possibly due to inherent errors (noise, partial volume effect) in the pCT or PET data. Future planned work will investigate perfusion-derived parameters from dPET and pCT correlated with histology to assess the optimum model for each modality.

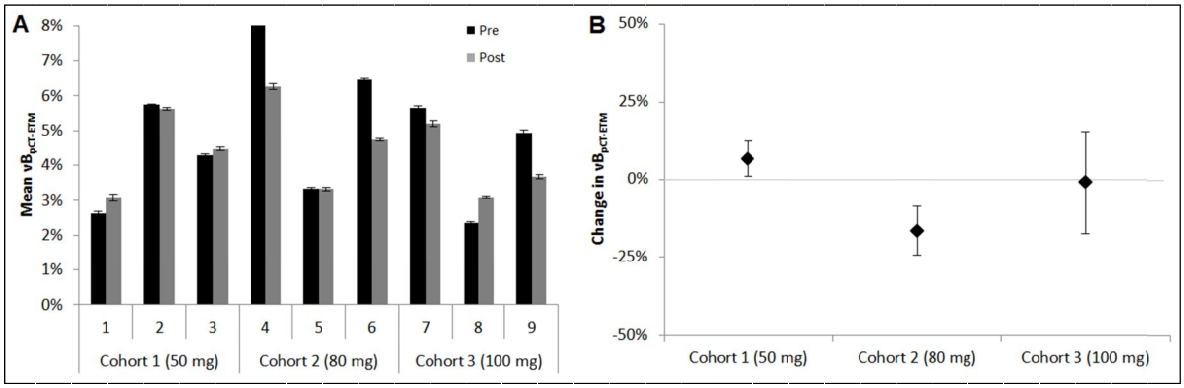


Figure 6.11 – (A) Mean ETM pCT parameter values and (B) mean percentage change in each cohort in vascular volume. Error bars are the standard error on the mean.

6.5 Conclusion

There is a decrease in BF post-Buparlisib for Cohorts 2 and 3, which is significant for Cohort 2. There is a significant decrease in $v_{B\text{-}pCT\text{-}AATH}$ for all patients post-Buparlisib. There are minimal and insignificant changes in MTT and PS post-Buparlisib.

A strong correlation (r=0.75) has been shown between K_1 as measured from PET kinetic modelling and BF×E (the equivalent parameter) as measured by pCT. This could mean that PET kinetic analysis could be used instead of pCT to measure the flux of tracer from the blood (K_1); this would provide a greater axial coverage (pCT is limited to 4 cm whilst PET is 15 cm) and would reduce the radiation dose to the patient (by not performing the relatively high dose pCT). A poor correlation (r=0.32) was shown between $v_{B\text{-}pCT\text{-}AATH}$ and $v_{B\text{-}PET}$, likely due to the different way the blood is modelled in the AATH and PET kinetic models. Using a pCT model (ETM) that models the blood in the same way as for PET kinetic modelling gives a higher correlation (r=0.69) between $v_{B\text{-}pCT\text{-}ETM}$ and $v_{B\text{-}PET}$.

Overall the pCT analysis shows BF decreasing in concordance with the PET kinetic analysis (Chapter 4). However, both $v_{B\text{-}pCT\text{-}AATH}$ and $v_{B\text{-}pCT\text{-}ETM}$ also decrease for the majority of patients whereas $v_{B\text{-}PET}$ increases. Future work will investigate the optimum pCT model by testing the information criteria for fits of multiple models and performing statistical simulations. In addition, perfusion derived parameters (from pCT and dPET) will be correlated with histopathology data from which a direct measurement of blood volume will be possible.

7 Conclusions

This thesis has examined FMISO and pCT imaging data from the first nine patients scanned as part of the Buparlisib trial in order to determine the optimum methods to analyse the data and to assess the impact of Buparlisib on these NSCLC patients. A summary of results from this work are detailed in section 7.1.

Many trials combining drug and RT fail in the clinical setting despite yielding promising pre-clinical data (Bradley et al., 2015; Chinot et al., 2014; Crosby et al., 2013). Although the causes for this are multi-factorial, most studies have failed to stratify patients likely to benefit from treatment and have not undertaken initial proof-of-concept studies to ascertain whether the chosen drug induces the same effects within patients as were observed pre-clinically. This thesis and associated imaging analysis have acted as a proof-of-concept for Buparlisib. The imaging analysis methods developed here will also be applied in future trials (such as ATOM, described in 7.2). In future, imaging-based treatment decisions could stratify patients to ensure they received the optimum treatment regime, thereby improving outcomes from trials combining drug and RT.

Regarding hypoxia post-Buparlisib this thesis has shown a significant decrease in $V_{TBR1.4}$ (hypothesised hypoxic volume) for Cohorts 2 and 3 compared to Cohort 1. Using the kinetic model found to be best for whole tumour FMISO analysis (3C5K), there was no significant change in k_5 (estimated FMISO intracellular binding rate-constant associated with cellular oxygen concentration) but its value did increase for Patients 1-5 (in Cohorts 1 and 2) and decrease for Patients 6-8 (in Cohorts 2 and 3). Applying the kinetic model found to be best for voxel-by-voxel FMISO kinetic analysis (2C3K), it was seen that the change in $k_{3\text{-}2C}$ (analogous to k_5 in the 3C5K model) appears to be distance dependent: values decrease in edge ($p<0.0001$) and outer ($p>0.05$) distance categories and increase in inner ($p<0.0001$) and central ($p<0.0001$) categories.

When the relevant parameters of the 3C5K PET kinetic models were utilised to investigate perfusion post-Buparlisib, the whole tumour analysis showed a significant decrease in K_1 (rate-constant of FMISO from blood to interstitium) and increase in $v_{B\text{-}PET}$ (tumour fractional blood volume), suggestive of vascular normalisation. The voxel-by-voxel analysis found that the same changes had occurred throughout the different tumour volume distance categories. However, the pCT results, using the standard AATH model, show that both BF (associated with K_1) and $v_{B\text{-}pCT\text{-}AATH}$ decrease. This difference between changes in v_B values obtained using the two imaging techniques (which appear to be driven by the model rather than the modality) will be investigated in planned future work.

7.1 Chapter summaries

In Chapter 2 a study was performed to assess methods of performing respiratory motion correction on a PET/CT system, using a moving phantom. It was seen that the impact of regular respiratory motion on PET images can be substantially reduced using gating or the QFreeze system. QFreeze provided small but significant advantages over gating in our experiments at clinically typical noise levels ($p<0.005$).

Chapter 3 examined the imaging protocols used to assess changes in hypoxia and perfusion after a week of Buparlisib through FMISO PET and pCT imaging. A study was performed to assess changes in the hypothesised hypoxic volume ($V_{TBR1.4}$) by analysing FMISO images at four hours post-injection. The $V_{TBR1.4}$ is the volume of voxels within the tumour which have a TBR greater than 1.4. The $V_{TBR1.4}$ volumes were seen to decrease after a week of daily administration of Buparlisib for doses exceeding 50 mg, with significant differences between the $V_{TBR1.4}$ in Cohort 1 and that from either Cohort 2 or 3 (both $p=0.02$).

Chapter 4 described a study to find the optimum model to describe FMISO time-courses in advanced stage NSCLC using information criteria and statistical simulations, which would allow the impact of Buparlisib on derived parameters to be assessed. It was shown that an irreversible three compartment (3C5K) model is optimal with a precision of 10% for fitted values of v_B and k_5. In this model v_B is the tumour fractional blood volume**,** K_1 the rate-constant of FMISO from blood to interstitium**,** and k_5 the intracellular FMISO binding rate-constant**.** The cellular oxygen concentration is more directly associated with the estimation of the FMISO intracellular binding rate-constant than FMISO uptake at a late time point. Using the 3C5K model to assess whole tumour changes it was observed that, independent of cohort, K_1 significantly decreased ($p=0.001$) and v_B significantly increased ($p=0.02$) post-Buparlisib, a finding which is consistent with vascular normalisation seen in pre-clinical studies. The change in whole tumour k_5 values was not significant across all nine patients with varying changes post-Buparlisib ($p>0.05$).

Chapter 5 sought to find the optimum model to describe FMISO time-courses voxel-by-voxel in advanced stage NSCLC using information criteria and statistical simulations. This allowed the change in derived rate-constants as distance increased from the tumour edge to be investigated as well as the impact of Buparlisib to be assessed at the voxel level. It was shown an irreversible two compartment (2C3K) model is the most accurate and precise at describing the voxel-by-voxel tumour FMISO time-courses by advanced stage NSCLC with a precision of 39% for fitted values of $v_{B\text{-}2C}$, 21% for $K_{1\text{-}2C}$, and 37% for $k_{3\text{-}2C}$. In this model $v_{B\text{-}2C}$ is the tumour fractional blood volume**,** $K_{1\text{-}2C}$ the rate-constant of FMISO from blood to cell**,** and $k_{3\text{-}2C}$ the intracellular FMISO binding rate-constant. It was shown that $v_{B\text{-}2C}$ and $K_{1\text{-}2C}$ weighted mean values generally increase significantly from the tumour centre towards the tumour edge (both $p<0.001$) with $k_{3\text{-}2C}$ weighted mean values decreasing significantly towards the tumour edge ($p<0.0001$).

Using the 2C3K model, post-Buparlisib a significant decrease in the weighted mean of $K_{1\text{-}2C}$ ($p<0.0001$) was observed for all cohorts, and a significant increase in the weighted mean of $v_{B\text{-}2C}$ for Cohorts 2 and 3 ($p<0.05$), which is suggestive of vascular normalisation. These changes in $K_{1\text{-}2C}$ and $v_{B\text{-}2C}$ post-Buparlisib were shown to occur throughout the tumour volume. However, there is a more complex change in the weighted mean values of $k_{3\text{-}2C}$: in Cohorts 1 and 3 the $k_{3\text{-}2C}$ weighted mean value falls significantly (both $p<0.0001$) post-Buparlisib, whereas in Cohort 2 it significantly increases ($p<0.001$). The response of $k_{3\text{-}2C}$ to Buparlisib appears to be distance dependent: the weighted mean values decrease in edge ($p<0.0001$) and outer ($p>0.05$) distance categories and increase in inner ($p<0.0001$) and central ($p<0.0001$) categories.

Chapter 6 analysed the pCT data to assess changes in derived parameters post-Buparlisib and to assess similarities between pCT parameters and the perfusion parameters derived from PET kinetic analysis. An algorithm implemented in commercial software was used (AATH) to calculate BF, $v_{B\text{-}pCT\text{-}AATH}$, MTT, and PS. BF is the flow rate through vasculature in tissue, $v_{B\text{-}pCT\text{-}AATH}$ the volume of blood within vasculature, MTT the time from artery to vein, and PS the rate-constant for flux between the plasma and interstitial space. A decrease in BF post-Buparlisib was observed for Cohorts 2 ($p=0.02$) and 3 ($p>0.05$) and increase for Cohort 1 ($p>0.05$), whereas a decrease was seen in all cohorts for K_1 (the corresponding PET kinetic parameter). A significant decrease for all patients was seen in $v_{B\text{-}pCT\text{-}AATH}$ ($p<0.01$), whereas an increase was seen in $v_{B\text{-}PET}$.

A strong correlation ($r=0.75$) was shown between K_1 and E×BF. A similar correlation ($r=0.69$) was shown between $v_{B\text{-}PET}$ and $v_{B\text{-}pCT\text{-}ETM}$ (ETM models the blood volume in the same way as for PET kinetic modelling), with a weaker correlation ($r=0.32$) between $v_{B\text{-}PET}$ and $v_{B\text{-}pCT\text{-}AATH}$.

7.2 Future work

Additional patients (a minimum of six) are to be recruited for the Buparlisib trial. This will allow more data to be considered when assessing the impact of Buparlisib both on a whole tumour and voxel-by-voxel level and provide validation of the results presented here.

With the exception of the good correlation between K_1 as measured from PET and BF as measured from pCT there are no independent (non-imaging) measures of perfusion or hypoxia in the current trial. A new trial 'Pre-operative window of opportunity study of the effects of atovaquone on hypoxia in non-small cell lung carcinoma' (ATOM) is about to open investigating the impact of atovaquone on NSCLC patients using FMISO, pCT, and MRI imaging. ATOM will recruit surgical NSCLC patients and so allow a comparison to be made between the clinical imaging, using the analysis methods developed in this thesis, and the histological data.

The pCT data presented in this thesis has been analysed using a commercial system and one other pCT model. Future work will investigate the benefits of combining kinetic analysis and movement correction (in this work these were performed sequentially) as well as exploring other pCT models. Similarly another ongoing DPhil project is researching performing PET kinetic modelling and movement correction concurrently.

ATOM will have fifteen patients with two FMISO and pCT imaging procedures 1-2 weeks apart with no intervention, and therefore in addition to providing histological data the new trial will also allow a reproducibility assessment to be made on the derived parameters. This will assist with analysing the Buparlisib imaging trial data as well as guiding future trials.

References

Adams, M.C., T.G. Turkington, J.M. Wilson, and T.Z. Wong. 2010. A systematic review of the factors affecting accuracy of SUV measurements. *AJR. Am. J. Roentgenol.* 195:310–20. doi:10.2214/AJR.10.4923.

Akaike, H. 1974. A new look at the statistical model identification. *Autom. Control. IEEE Trans.* 19:716–723.

Akin, O., S.B. Brennan, D.D. Dershaw, M.S. Ginsberg, M.J. Gollub, H. Schöder, D.M. Panicek, and H. Hricak. 2012. Advances in oncologic imaging: update on 5 common cancers. *CA. Cancer J. Clin.* 62:364–93. doi:10.3322/caac.21156.

Akinleye, A., P. Avvaru, and M. Furqan. 2013. Phosphatidylinositol 3-kinase (PI3K) inhibitors as cancer therapeutics. *J Hematol Oncol*. 6:1–17.

Axel, L. 1980. Cerebral blood flow determination by rapid-sequence computed tomography: theoretical analysis. *Radiology*. 137:679–86. doi:10.1148/radiology.137.3.7003648.

Bendell, J.C., J. Rodon, H. a. Burris, M. De Jonge, J. Verweij, D. Birle, D. Demanse, S.S. De Buck, Q.C. Ru, M. Peters, M. Goldbrunner, and J. Baselga. 2012. Phase I, dose-escalation study of BKM120, an oral pan-class I PI3K inhibitor, in patients with advanced solid tumors. *J. Clin. Oncol.* 30:282–290. doi:10.1200/JCO.2011.36.1360.

Bentzen, S.M. 2005. Theragnostic imaging for radiation oncology: dose-painting by numbers. *Lancet Oncol.* 6:112–7. doi:10.1016/S1470-2045(05)01737-7.

Berg, M., and K. Soreide. 2012. EGFR and Downstream Genetic Alterations in KRAS/BRAF and PI3K/AKT Pathways in Colorectal Cancer — Implications for Targeted Therapy. *Discov. Med.* 14:207–214.

Bertoldo, A., P. Peltoniemi, V. Oikonen, J. Knuuti, P. Nuutila, and C. Cobelli. 2001. Kinetic modeling of [(18)F]FDG in skeletal muscle by PET: a four-compartment five-rate-constant model. *Am. J. Physiol. Endocrinol. Metab.* 281:E524–36.

Bettinardi, V., L. Presotto, E. Rapisarda, M. Picchio, L. Gianolli, and M.C. Gilardi. 2011. Physical performance of the new hybrid PET/CT Discovery-690. *Med. Phys.* 38:5394–411. doi:10.1118/1.3635220.

Bittner, M.-I., N. Wiedenmann, S. Bucher, M. Hentschel, M. Mix, W. a. Weber, and A.-L. Grosu. 2013. Exploratory geographical analysis of hypoxic subvolumes using 18F-MISO-PET imaging in patients with head and neck cancer in the course of primary chemoradiotherapy. *Radiother. Oncol.* 108:511–516. doi:10.1016/j.radonc.2013.06.012.

Boellaard, R., A. Van Lingen, and A.A. Lammertsma. 2001. Experimental and Clinical Evaluation of Iterative Reconstruction (OSEM) in Dynamic PET : Quantitative Characteristics and Effects on Kinetic Modeling. 42:808–817.

Bollineni, V.R., G.S.M. a. Kerner, J. Pruim, R.J.H.M. Steenbakkers, E.M. Wiegman, M.J.B. Koole, E.H. de Groot, a. T.M. Willemsen, G. Luurtsema, J. Widder, H.J.M. Groen, and J. a. Langendijk. 2013. PET Imaging of Tumor Hypoxia Using 18F-Fluoroazomycin Arabinoside in Stage III-IV Non-Small Cell Lung Cancer Patients. *J. Nucl. Med.* 54:1175–1180. doi:10.2967/jnumed.112.115014.

Bradley, J., W.L. Thorstad, S. Mutic, T.R. Miller, F. Dehdashti, B. a Siegel, W. Bosch, and R.J. Bertrand. 2004. Impact of FDG-PET on radiation therapy volume delineation in non-small-cell lung cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 59:78–86. doi:10.1016/j.ijrobp.2003.10.044.

Bradley, J.D., R. Paulus, R. Komaki, G. Masters, G. Blumenschein, S. Schild, J. Bogart, C. Hu, K. Forster, A. Magliocco, V. Kavadi, Y.I. Garces, S. Narayan, P. Iyengar, C. Robinson, R.B. Wynn, C. Koprowski, J. Meng, J. Beitler, R. Gaur, W. Curran, and H. Choy. 2015. Standard-dose versus high-dose conformal radiotherapy with concurrent and consolidation carboplatin plus paclitaxel with or without cetuximab for patients with stage IIIA or IIIB non-small-cell lung cancer (RTOG 0617): a randomised, two-by-two factorial p. *Lancet. Oncol.* 16:187–99. doi:10.1016/S1470-2045(14)71207-0.

Brown, J.M. 1979. Evidence for acutely hypoxic cells in mouse tumours, and a possible mechanism of reoxygenation. *Br. J. Radiol.* 52:650–6. doi:10.1259/0007-1285-52-620-650.

Bruehlmeier, M., U. Roelcke, P.A. Schubiger, and S.M. Ametamey. 2004. Assessment of Hypoxia and Perfusion in. *J. Nucl. Med.* 45:1851–1859.

Burnham, K.P. 2004. Multimodel inference: understanding AIC and BIC in model selection. *Sociol. Methods Res.* 33:261–304. doi:10.1177/0049124104268644.

Cancer Research UK. 2016. No Title.

Carlin, S., and J.L. Humm. 2012. PET of hypoxia: current and future perspectives. *J. Nucl. Med.* 53:1171–4. doi:10.2967/jnumed.111.099770.

Casciari, J., M. Graham, and J. Rasey. 1995. A modeling approach for quantifying tumor hypoxia with [F-18]fluoromisonidazole PET time-activity data. *Med. Phys.* 22:1127–1139.

Chapman, J.D. 1979. Hypoxic sensitizers--implications for radiation therapy. *N. Engl. J. Med.* 301:1429–32. doi:10.1056/NEJM197912273012606.

Chen, L., Z. Zhang, H.C. Kolb, J.C. Walsh, J. Zhang, and Y. Guan. 2012. ^{18}F-HX4 hypoxia imaging with PET/CT in head and neck cancer: a comparison with ^{18}F-FMISO. *Nucl. Med. Commun.* 33:1096–102. doi:10.1097/MNM.0b013e3283571016.

Cherry, S.R., J.A. Sorenson, and M.E. Phelps. 2012. Physics in Nuclear Medicine. Elsevier Health Sciences. 544 pp.

Chinot, O.L., W. Wick, W. Mason, R. Henriksson, F. Saran, R. Nishikawa, A.F. Carpentier, K. Hoang-Xuan, P. Kavan, D. Cernea, A.A. Brandes, M. Hilton, L. Abrey, and T. Cloughesy. 2014. Bevacizumab plus radiotherapy-temozolomide for newly diagnosed glioblastoma. *N. Engl. J. Med.* 370:709–22. doi:10.1056/NEJMoa1308345.

Cobelli, C., and J.J. DiStefano. 1980. Parameter and structural identifiability concepts and ambiguities: a critical review and analysis. *Am. J. Physiol.* 239:R7–24.

Cobelli, C., D. Foster, and G. Toffolo. 2001. Tracer Kinetics in Biomedical Research: From Data to Model. Kluwer Academic, New York.

Cowan, D.S., K.O. Hicks, and W.R. Wilson. 1996. Multicellular membranes as an in vitro model for extravascular diffusion in tumours. *Br. J. Cancer. Suppl.* 27:S28–31.

Crone, C. 1963. The Permeability of Capillaries in Various Organs as Determined by Use of the “Indicator Diffusion” Method. *Acta Physiol. Scand.* 58:292–305. doi:10.1111/j.1748-1716.1963.tb02652.x.

Crosby, T., C.N. Hurt, S. Falk, S. Gollins, S. Mukherjee, J. Staffurth, R. Ray, N. Bashir, J.A. Bridgewater, J.I. Geh, D. Cunningham, J. Blazeby, R. Roy, T. Maughan, and G. Griffiths. 2013. Chemoradiotherapy with or without cetuximab in patients with oesophageal cancer (SCOPE1): a multicentre, phase 2/3 randomised trial. *Lancet. Oncol.* 14:627–37. doi:10.1016/S1470-2045(13)70136-0.

Cussler, E.L. 1997. Diffusion. Cambridge University Press.

Deniaud-Alexandre, E., E. Touboul, D. Lerouge, D. Grahek, J.-N. Foulquier, Y. Petegnief, B. Grès, H. El Balaa, K. Keraudy, K. Kerrou, F. Montravers, B. Milleron, B. Lebeau, and J.-N. Talbot. 2005. Impact of computed tomography and 18F-deoxyglucose coincidence detection emission tomography image fusion for optimization of conformal radiotherapy in non-small-cell lung cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 63:1432–41. doi:10.1016/j.ijrobp.2005.05.016.

Devesa, S.S., F. Bray, A.P. Vizcaino, and D.M. Parkin. 2005. International lung cancer trends by histologic type: male:female differences diminishing and adenocarcinoma rates rising. *Int. J. Cancer*. 117:294–9. doi:10.1002/ijc.21183.

van Elmpt, W., D. De Ruysscher, A. van der Salm, A. Lakeman, J. van der Stoep, D. Emans, E. Damen, M. Öllers, J.-J. Sonke, and J. Belderbos. 2012. The PET-boost randomised phase II dose-escalation trial in non-small cell lung cancer. *Radiother. Oncol.* 104:67–71. doi:10.1016/j.radonc.2012.03.005.

Erdi, Y.E., K. Rosenzweig, A.K. Erdi, H. a Macapinlac, Y.C. Hu, L.E. Braban, J.L. Humm, O.D. Squire, C.S. Chui, S.M. Larson, and E.D. Yorke. 2002. Radiotherapy treatment planning for patients with non-small cell lung cancer using positron emission tomography (PET). *Radiother. Oncol.* 62:51–60.

Eschmann, S., and F. Paulsen. 2005. Prognostic impact of hypoxia imaging with 18F-misonidazole PET in non-small cell lung cancer and head and neck cancer before radiotherapy. *J. Nucl. Med.* 46:253–260.

Feng, D., S.C. Huang, and X. Wang. 1993. Models for computer simulation studies of input functions for tracer kinetic modeling with positron emission tomography. *Int. J. Biomed. Comput.* 32:95–110.

Fleming, I.N., R. Manavaki, P.J. Blower, C. West, K.J. Williams, a L. Harris, J. Domarkas, S. Lord, C. Baldry, and F.J. Gilbert. 2014. Imaging tumour hypoxia with positron emission tomography. *Br. J. Cancer*. 112:238–250. doi:10.1038/bjc.2014.610.

Fokas, E., J.H. Im, S. Hill, S. Yameen, M. Stratford, J. Beech, W. Hackl, S.-M. Maira, E.J. Bernhard, W.G. McKenna, and R.J. Muschel. 2012. Dual inhibition of the PI3K/mTOR pathway increases tumor radiosensitivity by normalizing tumor vasculature. *Cancer Res.* 72:239–48. doi:10.1158/0008-5472.CAN-11-2263.

Gagel, B., P. Reinartz, C. Demirel, H.J. Kaiser, M. Zimny, M. Piroth, M. Pinkawa, S. Stanzel, B. Asadpour, K. Hamacher, H.H. Coenen, U. Buell, and M.J. Eble. 2006. [18F] fluoromisonidazole and [18F] fluorodeoxyglucose positron emission tomography in response evaluation after chemo-/radiotherapy of non-small-cell lung cancer: a feasibility study. *BMC Cancer*. 6:51. doi:10.1186/1471-2407-6-51.

Gagel, B., P. Reinartz, E. Dimartino, M. Zimny, M. Pinkawa, P. Maneschi, S. Stanzel, K. Hamacher, H.H. Coenen, M. Westhofen, U. Büll, and M.J. Eble. 2004. pO(2) Polarography versus positron emission tomography ([(18)F] fluoromisonidazole, [(18)F]-2-fluoro-2'-deoxyglucose). An appraisal of radiotherapeutically relevant hypoxia. *Strahlentherapie und Onkol. Organ der Dtsch. Röntgengesellschaft ... [et al]*. 180:616–22. doi:10.1007/s00066-004-1229-y.

Gambhir, S.S., M. Schwaiger, S.C. Huang, J. Krivokapich, H.R. Schelbert, C.A. Nienaber, and M.E. Phelps. 1989. Simple noninvasive quantification method for measuring myocardial glucose utilization in humans employing positron emission tomography and fluorine-18 deoxyglucose. *J. Nucl. Med.* 30:359–66.

Ganin, A., and S. Wollenweber. 2012. Q.Freeze White Paper.

Goldstraw, P., J. Crowley, K. Chansky, D.J. Giroux, P.A. Groome, R. Rami-Porta, P.E. Postmus, V. Rusch, and L. Sobin. 2007. The IASLC Lung Cancer Staging Project: proposals for the revision of the TNM stage groupings in the forthcoming (seventh) edition of the TNM Classification of malignant tumours. *J. Thorac. Oncol.* 2:706–14. doi:10.1097/JTO.0b013e31812f3c1a.

Graves, E.E., A. Maity, and Q.-T. Le. 2010. The tumor microenvironment in non-small-cell lung cancer. *Semin. Radiat. Oncol.* 20:156–63. doi:10.1016/j.semradonc.2010.01.003.

Gross, M.W., U. Karbach, K. Groebe, a J. Franko, and W. Mueller-Klieser. 1995. Calibration of misonidazole labeling by simultaneous measurement of oxygen tension and labeling density in multicellular spheroids. *Int. J. Cancer*. 61:567–73.

Grunbaum, Z., S.J. Freauff, K.A. Krohn, D.S. Wilbur, S. Magee, and J.S. Rasey. 1987. Synthesis and characterization of congeners of misonidazole for imaging hypoxia. *J. Nucl. Med.* 28:68–75.

Gunn, R.N., A.A. Lammertsma, S.P. Hume, and V.J. Cunningham. 1997. Parametric imaging of ligand-receptor binding in PET using a simplified reference region model. *Neuroimage*. 6:279–87. doi:10.1006/nimg.1997.0303.

Guo, H., R. Renaut, K. Chen, and E. Reiman. 2009. FDG-PET parametric imaging by total variation minimization. *Comput. Med. Imaging Graph.* 33:295–303. doi:10.1016/j.compmedimag.2009.01.005.

Gupta, A., D. Soto, M. Feldman, J. Goldsmith, R. Mick, S. Hahn, M. Machtay, R. Muschel, and W.G. McKenna. 2004. Signaling Pathways in NSCLC as a Predictor of Outcome and Response to Therapy. *Lung*. 182. doi:10.1007/s00408-004-0310-8.

Hall, R. 1971. Vascular injuries resulting from arterial puncture of catheterization. *Br. J. Surg.* 58:513–6.

Hanley, A., J. Mcneil, and D. Ph. 1982. The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve. *Radiology*. 143:29–36.

Higgins, G.S., D.R. McGowan, L. Collins, N. McGregor, and S. Love. 2014. BKM120 Trial Protocol.

Höckel, M., C. Knoop, K. Schlenger, B. Vorndran, E. Baußnann, M. Mitze, P.G. Knapstein, and P. Vaupel. 1993. Intratumoral pO2 predicts survival in advanced cancer of the uterine cervix. *Radiother. Oncol.* 26:45–50. doi:10.1016/0167-8140(93)90025-4.

Höckel, M., and P. Vaupel. 2001. Tumor hypoxia: definitions and current clinical, biologic, and molecular aspects. *J. Natl. Cancer* …. 93.

Horsman, M.R., L.S. Mortensen, J.B. Petersen, M. Busk, and J. Overgaard. 2012. Imaging hypoxia to improve radiotherapy outcome. *Nat. Rev. Clin. Oncol.* 9:674–87. doi:10.1038/nrclinonc.2012.171.

Hu, M., L. Xing, D. Mu, W. Yang, G. Yang, L. Kong, and J. Yu. 2013. Hypoxia imaging with 18F-fluoroerythronitroimidazole integrated PET/CT and immunohistochemical studies in non-small cell lung cancer. *Clin. Nucl. Med.* 38:591–6. doi:10.1097/RLU.0b013e318279fd3d.

Hudson, H.M., and R.S. Larkin. 1994. Ordered Subsets of Projection Data. 13:601–609.

Hueting, R., V. Kersemans, B. Cornelissen, M. Tredwell, K. Hussien, M. Christlieb, A.D. Gee, J. Passchier, S.C. Smart, J.R. Dilworth, V. Gouverneur, and R.J. Muschel. 2014. A comparison

of the behavior of (64)Cu-acetate and (64)Cu-ATSM in vitro and in vivo. *J. Nucl. Med.* 55:128–34. doi:10.2967/jnumed.113.119917.

IND. 2013. An Investigational Positron Emission Tomography (PET) radiopharmaceutical for injection and intended for use as an in vivo diagnostic for imaging hypoxia in tumors.

Jemal, A., F. Bray, M.M. Center, J. Ferlay, E. Ward, and D. Forman. 2011. Global cancer statistics. *CA. Cancer J. Clin.* 61:69–90. doi:10.3322/caac.20107.

Kallinowski, F., R. Zander, M. Hoeckel, and P. Vaupel. 1990. Tumor tissue oxygenation as evaluated by computerized-po2-histography. *Int. J. Radiat. Oncol.* 19:953–961. doi:10.1016/0360-3016(90)90018-F.

Kelly, C.J., and M. Brady. 2006. A model to simulate tumour oxygenation and dynamic [18F]-Fmiso PET data. *Phys. Med. Biol.* 51:5859–73. doi:10.1088/0031-9155/51/22/009.

Kim, I.-A., S.-S. Bae, A. Fernandes, J. Wu, R.J. Muschel, W.G. McKenna, M.J. Birnbaum, and E.J. Bernhard. 2005. Selective Inhibition of Ras, Phosphoinositide 3 Kinase, and Akt Isoforms Increases the Radiosensitivity of Human Carcinoma Cell Lines. *Cancer Res.* 65:7902–7910. doi:10.1158/0008-5472.CAN-05-0513.

Kinahan, P.E., and J.W. Fletcher. 2010. Positron emission tomography-computed tomography standardized uptake values in clinical practice and assessing response to therapy. *Semin. Ultrasound. CT. MR.* 31:496–505. doi:10.1053/j.sult.2010.10.001.

Koh, W.J., K.S. Bergman, J.S. Rasey, L.M. Peterson, M.L. Evans, M.M. Graham, J.R. Grierson, K.L. Lindsley, T.K. Lewellen, and K.A. Krohn. 1995. Evaluation of oxygenation status during fractionated radiotherapy in human nonsmall cell lung cancers using [F-18]fluoromisonidazole positron emission tomography. *Int. J. Radiat. Oncol. Biol. Phys.* 33:391–8. doi:10.1016/0360-3016(95)00170-4.

Koh, W.J., J.S. Rasey, M.L. Evans, J.R. Grierson, T.K. Lewellen, M.M. Graham, K.A. Krohn, and T.W. Griffin. 1992. Imaging of hypoxia in human tumors with [F-18]fluoromisonidazole. *Int. J. Radiat. Oncol. Biol. Phys.* 22:199–212. doi:10.1016/0360-3016(92)91001-4.

Krohn, K. a, J.M. Link, and R.P. Mason. 2008. Molecular imaging of hypoxia. *J. Nucl. Med.* 49 Suppl 2:129S–48S. doi:10.2967/jnumed.107.045914.

Laidler, K. 1978. Physical chemistry with biological applications. The Benjamin/Cummings Publishing Co., Menlo Park CA [etc.].

Larici, A.R., L. Calandriello, M. Amato, R. Silvestri, A. Del Ciello, F. Molinari, C. De Waure, M.L. Vita, G. Carnassale, and L. Bonomo. 2014. First-pass perfusion of non-small-cell lung cancer (NSCLC) with 64-detector-row CT: A study of technique repeatability and intra- and interobserver variability. *Radiol. Medica.* 119:4–12. doi:10.1007/s11547-013-0300-0.

Le, Q.-T., E. Chen, A. Salim, H. Cao, C.S. Kong, R. Whyte, J. Donington, W. Cannon, H. Wakelee, R. Tibshirani, J.D. Mitchell, D. Richardson, K.J. O'Byrne, A.C. Koong, and A.J. Giaccia. 2006. An evaluation of tumor oxygenation and gene expression in patients with early stage non-small cell lung cancers. *Clin. Cancer Res.* 12:1507–14. doi:10.1158/1078-0432.CCR-05-2049.

Lee, S.H., H.S. Kim, W.S. Park, S.Y. Kim, K.Y. Lee, S.H. Kim, J.Y. Lee, and N.J. Yoo. 2002. Non-small cell lung cancers frequently express phosphorylated Akt; an immunohistochemical study. *APMIS.* 110:587–592. doi:10.1034/j.1600-0463.2002.11007811.x.

Lee, S.T., and A.M. Scott. 2007. Hypoxia positron emission tomography imaging with 18f-fluoromisonidazole. *Semin. Nucl. Med.* 37:451–61. doi:10.1053/j.semnuclmed.2007.07.001.

Li, F., J. Joergensen, A. Hansen, and A. Kjaer. 2014. Kinetic modeling in PET imaging of hypoxia. *Am J Nucl Med Mol Imaging*. 4:490–506.

Lin, Z., J. Mechalakos, S. Nehmeh, H. Schoder, N. Lee, J. Humm, and C.C. Ling. 2008. The influence of changes in tumor hypoxia on dose-painting treatment plans based on 18F-FMISO positron emission tomography. *Int. J. Radiat. Oncol. Biol. Phys.* 70:1219–28. doi:10.1016/j.ijrobp.2007.09.050.

Ling, C.C., J. Humm, S. Larson, H. Amols, Z. Fuks, S. Leibel, and J. a Koutcher. 2000. Towards multidimensional radiotherapy (MD-CRT): biological imaging and biological conformality. *Int. J. Radiat. Oncol. Biol. Phys.* 47:551–60.

Liu, D., A. Chalkidou, D.B. Landau, P.K. Marsden, and J.D. Fenwick. 2014. 18F-FLT uptake kinetics in head and neck squamous cell carcinoma: a PET imaging study. *Med. Phys.* 41:041911. doi:10.1118/1.4868462.

Liu, P., H. Cheng, T.M. Roberts, and J.J. Zhao. 2009. Targeting the phosphoinositide 3-kinase pathway in cancer. *Nat. Rev. Drug Discov.* 8:627–44. doi:10.1038/nrd2926.

Logan, J., J.S. Fowler, N.D. Volkow, A.P. Wolf, S.L. Dewey, D.J. Schlyer, R.R. MacGregor, R. Hitzemann, B. Bendriem, and S.J. Gatley. 1990. Graphical analysis of reversible radioligand binding from time-activity measurements applied to [N-11C-methyl]-(-)-cocaine PET studies in human subjects. *J. Cereb. Blood Flow Metab.* 10:740–7. doi:10.1038/jcbfm.1990.127.

Lohith, T.G., T. Kudo, Y. Demura, Y. Umeda, Y. Kiyono, Y. Fujibayashi, and H. Okazawa. 2009. Pathophysiologic correlation between 62Cu-ATSM and 18F-FDG in lung cancer. *J. Nucl. Med.* 50:1948–53. doi:10.2967/jnumed.109.069021.

Lubberink, M., R. Boellaard, A.P. van der Weerdt, F.C. Visser, and A.A. Lammertsma. 2004. Quantitative Comparison of Analytic and Iterative Reconstruction Methods in 2- and 3-Dimensional Dynamic Cardiac 18F-FDG PET. *J. Nucl. Med.* 45:2008–2015.

Lusted, L. 1971. Decision-making studies in patient management. *N. Engl. J. Med.* 284:416.

Machleder, H.I., J.P. Sweeney, and W.F. Barker. 1972. Pulseless arm after brachial-artery catheterisation. *Lancet (London, England)*. 1:407–9.

Mah, K., C.B. Caldwell, Y.C. Ung, C.E. Danjoux, J.M. Balogh, S.N. Ganguli, L.E. Ehrlich, and R. Tirona. 2002. The impact of (18)FDG-PET on target and critical organs in CT-based treatment planning of patients with poorly defined non-small-cell lung carcinoma: a prospective study. *Int. J. Radiat. Oncol. Biol. Phys.* 52:339–50.

Maity, A., and E.J. Bernhard. 2010a. Modulating tumor vasculature through signaling inhibition to improve cytotoxic therapy. *Cancer Res.* 70:2141–5. doi:10.1158/0008-5472.CAN-09-3615.

Maity, A., and E.J. Bernhard. 2010b. Modulating tumor vasculature through signaling inhibition to improve cytotoxic therapy. *Cancer Res.* 70:2141–5. doi:10.1158/0008-5472.CAN-09-3615.

Marieb, E.N. 2005. Human Anatomy and Physiology. Pearson Education, Limited.

McGowan, D.R., R.E. Macpherson, K.M. Bradley, J.D. Fenwick, F. V Gleeson, and G.S. Higgins. 2016. 18 F-Misonidazole PET-CT scan detection of occult bone metastasis. *Thorax*. 71:97. doi:10.1136/thoraxjnl-2015-207400.

McKenna, W.G., R.J. Muschel, A.K. Gupta, S.M. Hahn, and E.J. Bernhard. 2003. The RAS signal transduction pathway and its role in radiation sensitivity. *Oncogene*. 22:5866–75. doi:10.1038/sj.onc.1206699.

Meier, P., and K.L. Zierler. 1954. On the Theory of the Indicator-Dilution Method for

Measurement of Blood Flow and Volume. *J Appl Physiol.* 6:731–744.

Mesina, C.T., R. Boellaard, G. Jongbloed, A.W. Van Der Vaart, and A.A. Lammertsma. 2003. Experimental evaluation of iterative reconstruction versus filtered back projection for 3D [15 O] water PET activation studies using statistical parametric mapping analysis. 19:1170–1179. doi:10.1016/S1053-8119(03)00075-2.

Miles, K. a. 2003. Perfusion CT for the assessment of tumour vascularity: which protocol? *Br. J. Radiol.* 76:S36–S42. doi:10.1259/bjr/18486642.

Miles, K. a, T.-Y. Lee, V. Goh, E. Klotz, C. Cuenod, S. Bisdas, a M. Groves, M.P. Hayball, R. Alonzi, and T. Brunner. 2012. Current status and guidelines for the assessment of tumour vascular support with dynamic contrast-enhanced computed tomography. *Eur. Radiol.* 22:1430–41. doi:10.1007/s00330-012-2379-4.

Morimoto, T., H. Ito, A. Takano, Y. Ikoma, C. Seki, T. Okauchi, K. Tanimoto, A. Ando, T. Shiraishi, T. Yamaya, and T. Suhara. 2006. Effects of image reconstruction algorithm on neurotransmission PET studies in humans: Comparison between filtered backprojection and ordered subsets expectation maximization. *Ann. Nucl. Med.* 20:237–243. doi:10.1007/BF03027437.

Muzi, M., D.A. Mankoff, J.R. Grierson, J.M. Wells, H. Vesselle, and K.A. Krohn. 2005. Kinetic modeling of 3’-deoxy-3'-fluorothymidine in somatic tumors: mathematical studies. *J. Nucl. Med.* 46:371–80.

Muzi, M., F. O’Sullivan, D.A. Mankoff, R.K. Doot, L.A. Pierce, B.F. Kurland, H.M. Linden, and P.E. Kinahan. 2012. Quantitative assessment of dynamic PET imaging data in cancer imaging. *Magn. Reson. Imaging*. 30:1203–15. doi:10.1016/j.mri.2012.05.008.

National Cancer Intelligence Network. 2015. Survival by stage. 3rd ed.

Nehmeh, S. a, Y.E. Erdi, K.E. Rosenzweig, H. Schoder, S.M. Larson, O.D. Squire, and J.L. Humm. 2003. Reduction of respiratory motion artifacts in PET imaging of lung cancer by respiratory correlated dynamic PET: methodology and comparison with respiratory gated PET. *J. Nucl. Med.* 44:1644–8.

Nehmeh, S. a, N.Y. Lee, H. Schröder, O. Squire, P.B. Zanzonico, Y.E. Erdi, C. Greco, G. Mageras, H.S. Pham, S.M. Larson, C.C. Ling, and J.L. Humm. 2008. Reproducibility of intratumor distribution of (18)F-fluoromisonidazole in head and neck cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 70:235–42. doi:10.1016/j.ijrobp.2007.08.036.

Ng, C.S., A.G. Chandler, W. Wei, E.F. Anderson, D.H. Herron, C. Charnsangavej, and R. Kurzrock. 2011. Reproducibility of perfusion parameters obtained from perfusion CT in lung tumors. *AJR. Am. J. Roentgenol.* 197:113–21. doi:10.2214/AJR.10.5404.

Ng, Q.-S., V. Goh, J. Milner, A.R. Padhani, M.I. Saunders, and P.J. Hoskin. 2007. Acute tumor vascular effects following fractionated radiotherapy in human lung cancer: In vivo whole tumor assessment using volumetric perfusion computed tomography. *Int. J. Radiat. Oncol. Biol. Phys.* 67:417–24. doi:10.1016/j.ijrobp.2006.10.005.

Nordsmark, M., M. Overgaard, and J. Overgaard. 1996. Pretreatment oxygenation predicts radiation response in advanced squamous cell carcinoma of the head and neck. *Radiother. Oncol.* 41:31–9.

Okamoto, S., T. Shiga, K. Yasuda, Y.M. Ito, K. Magota, K. Kasai, Y. Kuge, H. Shirato, and N. Tamaki. 2013. High reproducibility of tumor hypoxia evaluated by 18F-fluoromisonidazole PET for head and neck cancer. *J. Nucl. Med.* 54:201–7. doi:10.2967/jnumed.112.109330.

Ozgül, M.A., G. Kirkil, E.C. Seyhan, E. Cetinkaya, G. Ozgül, and M. Yüksel. 2013. The maximum

standardized FDG uptake on PET-CT in patients with non-small cell lung cancer. *Multidiscip. Respir. Med.* 8:69. doi:10.1186/2049-6958-8-69.

Papież, B.W., M.P. Heinrich, J. Fehrenbach, L. Risser, and J. a Schnabel. 2014. An implicit sliding-motion preserving regularisation via bilateral filtering for deformable image registration. *Med. Image Anal.* 18:1299–311. doi:10.1016/j.media.2014.05.005.

Parvizi, N., J.M. Franklin, D.R. McGowan, E.J. Teoh, K.M. Bradley, and F. V. Gleeson. 2015. Does a novel penalized likelihood reconstruction of 18F-FDG PET-CT improve signal-to-background in colorectal liver metastases? *Eur. J. Radiol.* 84:1–6. doi:10.1016/j.ejrad.2015.06.025.

Patlak, C.S., and R.G. Blasberg. 1985. Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. Generalizations. *J. Cereb. Blood Flow Metab.* 5:584–90. doi:10.1038/jcbfm.1985.87.

Petralia, G., L. Bonello, S. Viotti, L. Preda, G. d'Andrea, and M. Bellomi. 2010. CT perfusion in oncology: how to do it. *Cancer Imaging*. 10:8–19. doi:10.1102/1470-7330.2010.0001.

Picard, R.R., and R.D. Cook. 1984. Cross-validation of regression models. *J. Am. Stat. Assoc.* 79:575–583.

Piert, M., H.-J. Machulla, M. Picchio, G. Reischl, S. Ziegler, P. Kumar, H.-J. Wester, R. Beck, A.J.B. McEwan, L.I. Wiebe, and M. Schwaiger. 2005. Hypoxia-specific tumor imaging with 18F-fluoroazomycin arabinoside. *J. Nucl. Med.* 46:106–13.

PMOD. 2012. PMOD Kinetic Modeling (PKIN) User's Guide.

Postema, E.J., A.J.B. McEwan, T.A. Riauka, P. Kumar, D.A. Richmond, D.N. Abrams, and L.I. Wiebe. 2009. Initial results of hypoxia imaging using 1-alpha-D: -(5-deoxy-5-[18F]-fluoroarabinofuranosyl)-2-nitroimidazole (18F-FAZA). *Eur. J. Nucl. Med. Mol. Imaging*. 36:1565–73. doi:10.1007/s00259-009-1154-5.

Powis, G., and L. Kirkpatrick. 2004. Hypoxia inducible factor-1alpha as a cancer drug target. *Mol. Cancer Ther.* 3:647–54.

Prekeges, J.L., J.S. Rasey, Z. Grunbaum, and K.H. Krohn. 1991. Reduction of fluoromisonidazole, a new imaging agent for hypoxia. *Biochem. Pharmacol.* 42:2387–95.

Prezzi, D., A. Khan, and V. Goh. 2015. Perfusion CT imaging of treatment response in oncology. *Eur. J. Radiol.* 1–6. doi:10.1016/j.ejrad.2015.03.022.

Qayum, N., R.J. Muschel, J.H. Im, L. Balathasan, C.J. Koch, S. Patel, W.G. McKenna, and E.J. Bernhard. 2009. Tumor vascular changes mediated by inhibition of oncogenic signaling. *Cancer Res.* 69:6347–54. doi:10.1158/0008-5472.CAN-09-0657.

Raftery, A. 1995. Bayesian model selection in social research. *Sociol. Methodol.* 25:111–163.

Rami-Porta, R., D. Ball, J. Crowley, D.J. Giroux, J. Jett, W.D. Travis, M. Tsuboi, E. Vallières, and P. Goldstraw. 2007. The IASLC Lung Cancer Staging Project: proposals for the revision of the T descriptors in the forthcoming (seventh) edition of the TNM classification for lung cancer. *J. Thorac. Oncol.* 2:593–602. doi:10.1097/JTO.0b013e31807a2f81.

Rasey, J.S., W.-J. Koh, J.R. Grierson, Z. Grunbaum, and K.A. Krohn. 1989. Radiolabeled fluoromisonidazole as an imaging agent for tumor hypoxia. *Int. J. Radiat. Oncol. Biol. Phys.* 17:985–991. doi:10.1016/0360-3016(89)90146-6.

Reinders, A.A.T.S., A.M.J. Paans, B.M. de Jong, J.A. den Boer, and A.T.M. Willemsen. 2002. Iterative versus filtered backprojection reconstruction for statistical parametric mapping of

PET activation measurements: a comparative case study. *Neuroimage*. 15:175–81. doi:10.1006/nimg.2001.0963.

Renkin, E.M. 1959. Transport of potassium-42 from blood to tissue in isolated mammalian skeletal muscles. *Am. J. Physiol. -- Leg. Content*. 197:1205 –1210.

Samuels, Y., Z. Wang, A. Bardelli, N. Silliman, J. Ptak, S. Szabo, H. Yan, A. Gazdar, S.M. Powell, G.J. Riggins, J.K. V Willson, S. Markowitz, K.W. Kinzler, B. Vogelstein, and V.E. Velculescu. 2004. High frequency of mutations of the PIK3CA gene in human cancers. *Science*. 304:554. doi:10.1126/science.1096502.

Seppenwoolde, Y., H. Shirato, K. Kitamura, S. Shimizu, M. van Herk, J. V Lebesque, and K. Miyasaka. 2002. Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 53:822–34.

Shaw, R.J., and L.C. Cantley. 2006. Ras, PI(3)K and mTOR signalling controls tumour cell growth. *Nature*. 441:424–30. doi:10.1038/nature04869.

Sokoloff, L., M. Reivich, C. Kennedy, M.H. Des Rosiers, C.S. Patlak, K.D. Pettigrew, O. Sakurada, and M. Shinohara. 1977. The [14C]deoxyglucose method for the measurement of local cerebral glucose utilization: theory, procedure, and normal values in the conscious and anesthetized albino rat. *J. Neurochem.* 28:897–916.

Søndergaard, H.M., M.M. Madsen, K. Boisen, M. Bøttcher, O. Schmitz, T.T. Nielsen, H.E. Bøtker, and S.B. Hansen. 2007. Evaluation of iterative reconstruction (OSEM) versus filtered back-projection for the assessment of myocardial glucose uptake and myocardial perfusion using dynamic PET. *Eur. J. Nucl. Med. Mol. Imaging*. 34:320–9. doi:10.1007/s00259-006-0198-z.

Soret, M., S.L. Bacharach, and I. Buvat. 2007. Partial-volume effect in PET tumor imaging. *J. Nucl. Med.* 48:932–45. doi:10.2967/jnumed.106.035774.

Sourbron, S.P., and D.L. Buckley. 2013. Classic models for dynamic contrast-enhanced MRI. *Class. Model. Dce-Mri*. 26:1004–1027. doi:10.1002/nbm.2940.

Søvik, A., E. Malinen, and D.R. Olsen. 2009. Strategies for biologic image-guided dose escalation: a review. *Int. J. Radiat. Oncol. Biol. Phys.* 73:650–8. doi:10.1016/j.ijrobp.2008.11.001.

Span, P.N., and J. Bussink. 2015. Biology of Hypoxia. *Semin. Nucl. Med.* 45:101–109. doi:10.1053/j.semnuclmed.2014.10.002.

St Lawrence, K.S., and T.Y. Lee. 1998. An adiabatic approximation to the tissue homogeneity model for water exchange in the brain: II. Experimental validation. *J. Cereb. Blood Flow Metab.* 18:1378–1385. doi:10.1097/00004647-199812000-00012.

Swanson, K.R., G. Chakraborty, C.H. Wang, R. Rockne, H.L.P. Harpold, M. Muzi, T.C.H. Adamsen, K. a Krohn, and A.M. Spence. 2009. Complementary but distinct roles for MRI and 18F-fluoromisonidazole PET in the assessment of human glioblastomas. *J. Nucl. Med.* 50:36–44. doi:10.2967/jnumed.108.055467.

*Teoh, E.J., D.R. *McGowan, R.E. Macpherson, K.M. Bradley, F. V. Gleeson, E.J. Teoh, D.R. McGowan, R.E. Macpherson, K.M. Bradley, F. V. Gleeson, E.J. *Teoh, D.R. *McGowan, R.E. Macpherson, K.M. Bradley, and F. V. Gleeson. 2015. Phantom and clinical evaluation of the Bayesian penalized likelihood reconstruction algorithm Q.Clear on an LYSO PET/CT system. *J. Nucl. Med.* 56:1447–1453. doi:10.2967/jnumed.115.159301.

Teoh, E.J., D.R. McGowan, K.M. Bradley, E. Belcher, E. Black, and F. V. Gleeson. 2015. Novel penalised likelihood reconstruction of PET in the assessment of histologically verified small pulmonary nodules. *Eur. Radiol.* doi:10.1007/s00330-015-3832-y.

Thomlinson, R.H., and L.H. Gray. 1955. The histological structure of some human lung cancers and the possible implications for radiotherapy. *Br. J. Cancer*. 9:539–549. doi:10.1038/bjc.1955.55.

Thorndyke, B., a. Koong, and L. Xing. 2008. Reducing respiratory motion artifacts in radionuclide imaging through retrospective stacking: A simulation study. *Linear Algebra Appl.* 428:1325–1344. doi:10.1016/j.laa.2007.06.019.

Thorwarth, D., S.M. Eschmann, F. Paulsen, and M. Alber. 2005. A kinetic model for dynamic [18F]-Fmiso PET data to analyse tumour hypoxia. *Phys. Med. Biol.* 50:2209–24. doi:10.1088/0031-9155/50/10/002.

Thorwarth, D., S.-M. Eschmann, J. Scheiderbauer, F. Paulsen, and M. Alber. 2005b. Kinetic analysis of dynamic 18F-fluoromisonidazole PET correlates with radiation treatment outcome in head-and-neck cancer. *BMC Cancer*. 5:152. doi:10.1186/1471-2407-5-152.

Thureau, S., P. Chaumet-Riffaud, R. Modzelewski, P. Fernandez, L. Tessonnier, L. Vervueren, F. Cachin, a. Berriolo-Riedinger, P. Olivier, H. Kolesnikov-Gauthier, O. Blagosklonov, B. Bridji, a. Devillers, L. Collombier, F. Courbon, E. Gremillet, C. Houzard, J.M. Caignon, J. Roux, N. Aide, I. Brenot-Rossi, K. Doyeux, B. Dubray, and P. Vera. 2013. Interobserver Agreement of Qualitative Analysis and Tumor Delineation of 18F-Fluoromisonidazole and 3'-Deoxy-3'-18F-Fluorothymidine PET Images in Lung Cancer. *J. Nucl. Med.* 54:1543–1550. doi:10.2967/jnumed.112.118083.

Tofts, P.S. 1997. Modeling tracer kinetics in dynamic Gd-DTPA MR imaging. *J. Magn. Reson. Imaging*. 7:91–101. doi:10.1002/jmri.1880070113.

Toulany, M., and H.P. Rodemann. 2013. Potential of Akt mediated DNA repair in radioresistance of solid tumors overexpressing erbB-PI3K-Akt pathway. *Transl. Cancer Res.* 2:190–202. doi:10.3978/j.issn.2218-676X.2013.04.09.

Trinkaus, M.E., R. Blum, D. Rischin, J. Callahan, M. Bressel, T. Segard, P. Roselt, P. Eu, D. Binns, M.P. MacManus, D. Ball, and R.J. Hicks. 2013. Imaging of hypoxia with 18F-FAZA PET in patients with locally advanced non-small cell lung cancer treated with definitive chemoradiotherapy. *J. Med. Imaging Radiat. Oncol.* 57:475–81. doi:10.1111/1754-9485.12086.

Turkheimer, F., R.M. Moresco, G. Lucignani, L. Sokoloff, F. Fazio, and K. Schmidt. 1994. The use of spectral analysis to determine regional cerebral glucose utilization with positron emission tomography and [18F]fluorodeoxyglucose: theory, implementation, and optimization procedures. *J. Cereb. Blood Flow Metab.* 14:406–22. doi:10.1038/jcbfm.1994.52.

Vaupel, P., K. Schlenger, C. Knoop, and M. Höckel. 1991. Oxygenation of human tumors: evaluation of tissue oxygen distribution in breast cancers by computerized O2 tension measurements. *Cancer Res.* 51:3316–22.

Vera, P., P. Bohn, A. Edet-Sanson, A. Salles, S. Hapdey, I. Gardin, J.-F. Ménard, R. Modzelewski, L. Thiberville, and B. Dubray. 2011. Simultaneous positron emission tomography (PET) assessment of metabolism with ^{18}F-fluoro-2-deoxy-d-glucose (FDG), proliferation with ^{18}F-fluoro-thymidine (FLT), and hypoxia with ^{18}fluoro-misonidazole (F-miso) before and during radiotherapy in patients wit. *Radiother. Oncol.* 98:109–16. doi:10.1016/j.radonc.2010.10.011.

Wack, L.J., D. Mönnich, W. van Elmpt, C.M.L. Zegers, E.G.C. Troost, D. Zips, and D. Thorwarth. 2015. Comparison of [18F]-FMISO, [18F]-FAZA and [18F]-HX4 for PET imaging of hypoxia - a simulation study. *Acta Oncol.* 54:1370–7. doi:10.3109/0284186X.2015.1067721.

Wald, A., and J. Wolfowitz. 1940. On a test whether two samples are from the same population.

Ann. Math. Stat. 11:147–162. doi:10.1214/aoms/1177731909.

Walsh, J.C., A. Lebedev, E. Aten, K. Madsen, L. Marciano, and H.C. Kolb. 2014. The Clinical Importance of Assessing Tumor Hypoxia: Relationship of Tumor Hypoxia to Prognosis and Therapeutic Opportunities. *Antioxid. Redox Signal.* 21:1516–1554. doi:10.1089/ars.2013.5378.

Wang, W., J.-C. Georgi, S. Nehmeh, M. Narayanan, T. Paulus, M. Bal, J. O'Donoghue, P.B. Zanzonico, C.R. Schmidtlein, N.Y. Lee, and J.L. Humm. 2009. Evaluation of a compartmental model for estimating tumor hypoxia via FMISO dynamic PET imaging. *Phys. Med. Biol.* 54:3083–99. doi:10.1088/0031-9155/54/10/008.

Wang, W., N.Y. Lee, J.-C. Georgi, M. Narayanan, J. Guillem, H. Schöder, and J.L. Humm. 2010. Pharmacokinetic analysis of hypoxia (18)F-fluoromisonidazole dynamic PET in head and neck cancer. *J. Nucl. Med.* 51:37–45. doi:10.2967/jnumed.109.067009.

van der Weerdt, A.P., L.J. Klein, R. Boellaard, C.A. Visser, F.C. Visser, and A.A. Lammertsma. 2001. Image-derived input functions for determination of MRGlu in cardiac (18)F-FDG PET scans. *J. Nucl. Med.* 42:1622–1629.

Wieser, E., D. Strohmeyer, H. Rogatsch, W. Horninger, G. Bartsch, and P. Debbage. 2005. Access of tumor-derived macromolecules and cells to the blood: an electron microscopical study of structural barriers in microvessel clusters in highly malignant primary prostate carcinomas. *Prostate*. 62:123–32. doi:10.1002/pros.20129.

Wollenweber, S.D., S. Member, G. Gopalakrishnan, K. Thielemans, and R.M. Manjeshwar. 2012. Evaluation of the Accuracy and Robustness of a Motion Correction Algorithm for PET Using a Novel Phantom Approach. 59:123–130.

Yang, D., S. Wallace, A. Cherif, C. Li, M. Gretzer, E. Kim, and D. Podoloff. 1995. Development of F-18 labelled Fluroerythronitroimidazole as a PET Agent for Imaging Tumour Hypoxia. *Radiology*. 795–800.

Yuan, T.L., and L.C. Cantley. 2008. PI3K pathway alterations in cancer: variations on a theme. *Oncogene*. 27:5497–510. doi:10.1038/onc.2008.245.

Zasadny, K.R., and R.L. Wahl. 1993. Standardized uptake values of normal tissues at PET with 2-[fluorine-18]-fluoro-2-deoxy-D-glucose: variations with body weight and a method for correction. *Radiology*. 189:847–850.

Zegers, C.M.L., W. Van Elmpt, R. Wierts, B. Reymen, H. Sharifi, M.C. Öllers, F. Hoebers, E.G.C. Troost, R. Wanders, A. Van Baardwijk, B. Brans, J. Eriksson, B. Windhorst, F.M. Mottaghy, D. De Ruysscher, and P. Lambin. 2013. Hypoxia imaging with [18F]HX4 PET in NSCLC patients: Defining optimal imaging parameters. *Radiother. Oncol.* 109:58–64. doi:10.1016/j.radonc.2013.08.031.

Zips, D., K. Zöphel, N. Abolmaali, R. Perrin, A. Abramyuk, R. Haase, S. Appold, J. Steinbach, J. Kotzerke, and M. Baumann. 2012. Exploratory prospective trial of hypoxia-specific PET imaging during radiochemotherapy in patients with locally advanced head-and-neck cancer. *Radiother. Oncol.* 105:21–8. doi:10.1016/j.radonc.2012.08.019.

Appendix

Phantom and Clinical Evaluation of the Bayesian Penalized Likelihood Reconstruction Algorithm Q.Clear on an LYSO PET/CT System

Eugene J. Teoh*[1,2], Daniel R. McGowan*[2,3], Ruth E. Macpherson[1], Kevin M. Bradley[1], and Fergus V. Gleeson[1,2]

[1]Department of Radiology, Churchill Hospital, Oxford University Hospitals NHS Trust, Oxford, United Kingdom; [2]Department of Oncology, University of Oxford, Old Road Campus Research Building, Oxford, United Kingdom; and [3]Radiation Physics and Protection, Churchill Hospital, Oxford University Hospitals NHS Trust, Oxford, United Kingdom

Q.Clear, a Bayesian penalized-likelihood reconstruction algorithm for PET, was recently introduced by GE Healthcare on their PET scanners to improve clinical image quality and quantification. In this work, we determined the optimum penalization factor (beta) for clinical use of Q.Clear and compared Q.Clear with standard PET reconstructions. **Methods:** A National Electrical Manufacturers Association image-quality phantom was scanned on a time-of-flight PET/CT scanner and reconstructed using ordered-subset expectation maximization (OSEM), OSEM with point-spread function (PSF) modeling, and the Q.Clear algorithm (which also includes PSF modeling). Q.Clear was investigated for β (B) values of 100–1,000. Contrast recovery (CR) and background variability (BV) were measured from 3 repeated scans, reconstructed with the different algorithms. Fifteen oncology body ^{18}F-FDG PET/CT scans were reconstructed using OSEM, OSEM PSF, and Q.Clear using B values of 200, 300, 400, and 500. These were visually analyzed by 2 scorers and scored by rank against a panel of parameters (overall image quality; background liver, mediastinum, and marrow image quality; noise level; and lesion detectability). **Results:** As β is increased, the CR and BV decreases; Q.Clear generally gives a higher CR and lower BV than OSEM. For the smallest sphere reconstructed with Q.Clear B400, CR is 28.4% and BV 4.2%, with corresponding values for OSEM of 24.7% and 5.0%. For the largest hot sphere, Q.Clear B400 yields a CR of 75.2% and a BV of 3.8%, with corresponding values for OSEM of 64.4% and 4.0%. Scorer 1 and 2 ranked B400 as the preferred reconstruction in 13 of 15 (87%) and 10 of 15 (73%) cases. The least preferred reconstruction was OSEM PSF in all cases. In most cases, lesion detectability was highest ranked for B200, in 9 of 15 (67%) and 10 of 15 (73%), with OSEM PSF ranked lowest. Poor lesion detectability on OSEM PSF was seen in cases of mildly ^{18}F-FDG–avid mediastinal nodes in lung cancer and small liver metastases due to background noise. Conversely, OSEM PSF was ranked second highest for lesion detectability in most pulmonary nodule evaluation cases. The combined scores confirmed B400 to be the preferred reconstruction. **Conclusion:** Our phantom measurement results demonstrate improved CR and reduced BV when using Q.Clear instead of OSEM. A β value of 400 is recommended for oncology body PET/CT using Q.Clear.

Received Apr. 16, 2015; revision accepted Jul. 7, 2015.
For correspondence or reprints contact: Eugene J. Teoh, Department of Radiology, Churchill Hospital, Old Rd., Headington, Oxford OX3 7LE, United Kingdom.
E-mail: eugene.teoh@oncology.ox.ac.uk
*Contributed equally to this work.
Published online Jul. 9, 2015.

Key Words: positron emission tomography; image reconstruction; Bayesian penalized likelihood; NEMA; image quality; optimization

J Nucl Med 2015; 56:1447–1452
DOI: 10.2967/jnumed.115.159301

PET/CT is commonly performed in the staging of malignancy and in determining disease response, most commonly with ^{18}F-FDG, although one of the limitations of PET is its relatively low, 5-mm, spatial resolution (*1*), which leads to failure to detect small lesions due to an underestimation of tracer uptake. Additionally, the use of semiquantitative analyses for differentiating benign from malignant disease with ^{18}F-FDG, for instance, in pulmonary nodules and mediastinal lymph nodes in patients with lung cancer, is partially limited in its acceptance because there is no agreed standardized uptake value (SUV) for their differentiation (*2*). This is in part because technical factors such as the image reconstruction methodologies used affect the accuracy and reproducibility of SUV measurements (*3,4*).

Recently, GE Healthcare introduced a penalized-likelihood iterative PET reconstruction in their commercial software, termed Q.Clear, which is available on GE Healthcare PET/CT scanners. It includes point-spread function (PSF) modeling (*5*) and controls the noise through the use of a penalty term. Although penalized-likelihood algorithms have been in existence since 1987 (*6*), their clinical use has so far been limited. This Q.Clear algorithm includes a relative difference penalty (*7*), which is a function of the difference between neighboring voxels and a function of their sum (*8*). This penalty function acts as a noise suppression term and is controlled by a penalization factor (termed β), which is the only user-input variable to the algorithm with the γ factor in the penalty function set to 2. Modified block sequential regularized expectation maximization is used as an optimizer for this Q.Clear algorithm, which, because of the penalty function, allows an effective convergence to be achieved in images, potentially providing a more accurate SUV (*8,9*). This is in contrast to ordered-subset expectation maximization (OSEM) reconstructions, which have to be stopped before contrast convergence to prevent image noise increasing excessively (*5,10*). Accordingly, Q.Clear has been shown to significantly improve signal-to-noise in clinical scans, compared with OSEM (*11–14*), particularly in small faintly avid abnormalities (*15,16*).

The aim of this study was to determine the optimum penalization factor (β) for clinical use; this was achieved by masked clinical scoring of a subset of β values, which were determined from phantom studies.

MATERIALS AND METHODS

Phantom

A National Electrical Manufacturers Association (NEMA) image-quality phantom (*17*) was scanned on a Discovery 690 scanner (GE Healthcare), with the 4 smallest spheres filled with an activity concentration in a 4-to-1 ratio to the background activity concentration. The phantom was scanned following the NEMA procedure (*17*), repeating the scan 3 times. Images were reconstructed using our center's standard time-of-flight (ToF) OSEM protocol (2 iterations, 24 subsets, 6.4-mm gaussian filter), a ToF OSEM PSF protocol (3 iterations, 24 subsets, 2-mm gaussian filter) (*18*), and Q.Clear (number of iterations is variable and dependent on the sinogram size) over a range of β values (100, 200, up to 1,000).

The data were analyzed using the NEMA analysis tool (GE Healthcare) to determine contrast recovery (CR) and background variability (BV) for each sphere (j) and reconstruction in addition to the residual lung error (LE) using methods defined by the NEMA standard (*17*), shown in Equations 1–3.

$$CR_j = \frac{\frac{C_{H,j}}{C_{B,j}} - 1}{\frac{a_H}{a_B} - 1} \times 100\% \quad \text{Eq. 1}$$

$$BV_j = \frac{SD_j}{C_{B,j}} \times 100\% \quad \text{Eq. 2}$$

$$LE_i = \frac{C_{lung,i}}{C_{B,i}} \times 100\% \quad \text{Eq. 3}$$

Where $C_{H,j}$ is the average counts within a region of interest (ROI) drawn on each sphere j on the central PET slice, $C_{B,j}$ the average background counts for ROIs of the same size, a_H the activity concentration in the hot spheres, a_B the activity concentration in the background, SD_j the SD of the background ROI counts, $C_{lung,i}$ the average counts for an ROI drawn in the lung insert on slice i, and $C_{B,i}$ the average background counts for slice i (*17*). LE is then given by the average of LE_i for all PET slices. Contrast-to-noise ratio was defined as CR/BV. Paired *t* tests were used to test the significance of the differences between the reconstructions.

Clinical Evaluation

Case Selection. Informed consent is not necessary for retrospective reviews of this nature in our institution. Fifteen ^{18}F-FDG PET/CT scans were retrospectively selected. The scans were selected to represent a range of pathology, all with prior histologic confirmation: subcentimeter pulmonary nodules, mildly ^{18}F-FDG–avid mediastinal lymph nodes in non–small cell lung cancer, and small liver metastases. These were obtained between March 2011 and November 2013 for pulmonary nodule evaluation (5 scans), staging of non–small cell lung cancer (5 scans), and staging of colorectal cancer with known liver metastases (5 scans). There were 9 men and 6 women, with a weight range of 37–106 kg and a median weight of 72 kg.

^{18}F-FDG PET/CT scans were acquired on a 3-dimensional mode ToF Discovery 690 PET/CT system with lutetium yttrium orthosilicate (LYSO) crystals (GE Healthcare). The patients were required to fast for at least 6 h before their scan. Their blood glucose was measured before intravenous injection of ^{18}F-FDG (4 MBq/kg). Imaging commenced 90 min after injection and covered the skull base to upper thighs. The PET images were acquired under normal tidal respiration for 4 min per bed position. The CT was obtained using a pitch of 0.984, 120 kV, and automA with a noise index of 25.

Reconstruction and Analyses. PET images were reconstructed using 3 different algorithms, each of which used the CT scan for attenuation correction and the same normalization correction factors with scatter and randoms corrected as has been previously described (*19–21*). The standard PET reconstruction algorithm used at our center is ToF OSEM (VPFX; GE Healthcare), used with 2 iterations, 24 subsets, and 6.4-mm filter. The sinograms generated at the time of scanning were retrospectively processed using a ToF OSEM PSF protocol (3 iterations,

TABLE 1
Mean CR and BV for Hot and Cold Spheres

Sphere diameter (mm)	CR (%)						BV (%)						Residual LE (%)
	10	13	17	22	28	37	10	13	17	22	28	37	
OSEM*	24.7 (2.9)	41.5 (3.4)	50.1 (2.7)	64.4 (1.3)	67.2 (0.8)	78.5 (0.7)	5.0 (0.3)	4.6 (0.2)	4.2 (0.2)	4.0 (0.1)	3.7 (0.1)	3.5 (0.1)	9.1 (0.3)
OSEM PSF†	42.7 (3.9)	63.3 (5.3)	66.4 (3.2)	76.9 (1.9)	75.8 (0.8)	85.2 (0.6)	6.1 (0.4)	5.3 (0.3)	4.7 (0.3)	4.2 (0.2)	3.9 (0.1)	3.5 (0.1)	6.2 (0.4)
B100‡	49.5 (3.9)	68.2 (5.5)	70.2 (3.4)	79.9 (2.7)	82.4 (0.9)	90.6 (0.5)	6.1 (0.4)	5.3 (0.3)	4.7 (0.3)	4.3 (0.2)	3.9 (0.2)	3.6 (0.1)	2.7 (0.2)
B200	40.0 (3.7)	62.4 (5.2)	67.6 (3.4)	78.3 (2.4)	79.7 (0.6)	88.9 (0.4)	5.1 (0.5)	4.6 (0.3)	4.2 (0.1)	4.0 (0.1)	3.7 (0.1)	3.5 (0.2)	2.9 (0.2)
B300	33.3 (3.4)	57.3 (4.8)	64.9 (3.4)	76.7 (2.2)	77.3 (0.5)	87.4 (0.3)	4.7 (0.4)	4.4 (0.2)	4.1 (0.1)	3.8 (0.1)	3.6 (0.1)	3.5 (0.2)	3.0 (0.2)
B400	28.4 (3.2)	52.7 (4.6)	62.4 (3.3)	75.2 (2.2)	75.2 (0.4)	86.1 (0.2)	4.5 (0.3)	4.2 (0.3)	4.0 (0.1)	3.8 (0.2)	3.6 (0.2)	3.4 (0.2)	3.1 (0.3)
B500	24.7 (2.9)	48.8 (4.3)	59.9 (3.2)	73.7 (2.0)	73.3 (0.4)	84.8 (0.2)	4.3 (0.3)	4.1 (0.3)	3.9 (0.2)	3.7 (0.2)	3.5 (0.2)	3.4 (0.2)	3.2 (0.2)
B600	21.9 (2.7)	45.4 (4.0)	57.6 (3.0)	72.3 (2.0)	71.6 (0.4)	83.5 (0.2)	4.2 (0.3)	4.0 (0.3)	3.8 (0.2)	3.7 (0.2)	3.5 (0.2)	3.4 (0.2)	3.5 (0.3)
B700	19.6 (2.5)	42.3 (3.9)	55.4 (3.0)	71.0 (1.9)	70.0 (0.4)	82.4 (0.3)	4.1 (0.4)	3.9 (0.3)	3.8 (0.3)	3.7 (0.2)	3.5 (0.2)	3.4 (0.2)	3.6 (0.3)
B800	17.8 (2.3)	39.6 (3.6)	53.4 (2.9)	69.7 (1.9)	68.5 (0.4)	81.3 (0.2)	4.0 (0.4)	3.9 (0.3)	3.8 (0.3)	3.6 (0.3)	3.5 (0.3)	3.5 (0.2)	4.0 (0.2)
B900	16.2 (2.2)	37.2 (3.5)	51.6 (2.8)	68.5 (1.8)	67.1 (0.4)	80.3 (0.2)	4.0 (0.3)	3.9 (0.3)	3.8 (0.3)	3.6 (0.2)	3.6 (0.2)	3.5 (0.2)	4.4 (0.3)
B1,000	15.0 (2.0)	35.1 (3.3)	49.8 (2.7)	67.3 (1.8)	65.9 (0.4)	79.3 (0.2)	3.9 (0.4)	3.8 (0.3)	3.7 (0.3)	3.6 (0.3)	3.5 (0.2)	3.5 (0.2)	4.8 (0.3)

*OSEM (2 iterations, 24 subsets, 6.4-mm filter).
†OSEM PSF (3 iterations, 24 subsets, 2-mm filter).
‡B*n* indicates a Q.Clear reconstruction with a β value of *n*.
Hot spheres had diameters of 10, 13, 17, and 22 mm and cold spheres 28 and 37 mm, and mean residual LE as defined by NEMA (*17*). Values in parentheses are SDs.

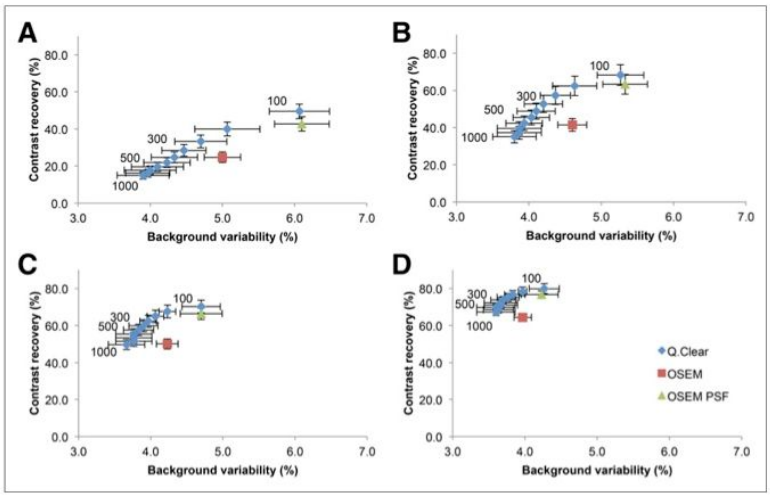


FIGURE 1. Graphs show mean CR and BV for hot spheres with diameters of 10 (A), 13 (B), 17 (C), and 22 mm (D) (cold spheres in Supplemental Fig. 1). These are shown for OSEM (2 iterations, 24 subsets, 6.4-mm filter), OSEM PSF (3 iterations, 24 subsets, 2-mm filter), and Q.Clear (β = 100–1,000, as labeled on the points). Error bars shown are 1 SD.

24 subsets, 2-mm filter) and the new Q.Clear reconstruction algorithm for penalization factors (β): 200, 300, 400, and 500.

Visual analyses of the OSEM, OSEM PSF, and Q.Clear PET images, 6 reconstructions per case, were performed by 2 consultants (designated scorer 1 and 2) with double accreditation in clinical radiology and nuclear medicine, and 11 and 3 y consultant experience, respectively. Images were viewed on an Advantage Workstation (AW4.6; GE Healthcare). The reconstructions were labeled A–F in a randomized order, with the CT component available for image fusion. Cases were reviewed sequentially, and the reconstructions were ranked (from 1 to 6) according to 6 image quality (IQ) parameters: overall IQ (1, excellent; 5, worst), background liver IQ (1, excellent; 5, worst), background mediastinum IQ (1, excellent; 5, worst), background marrow IQ (1, excellent; 5, worst), noise level (1, minimal; 5, unacceptable), and lesion detectability (1, excellent; 5, poor).

Scorers also indicated their most and least preferred reconstruction for each case. Interrater agreement on ranking within each of the 6 IQ parameters was assessed using Cohen's κ statistic. The proportions of the highest- and lowest-ranked reconstructions were calculated for each parameter. Alongside the highest frequencies of the most and least preferred reconstruction indicated by the scorers, scores by both scorers for all parameters across the cases were summated for each reconstruction to confirm this quantitatively.

Statistical analyses were performed using SPSS Statistics 22.0 (IBM Corp.). κ values were interpreted using the guidelines laid out by Landis and Koch (*22*).

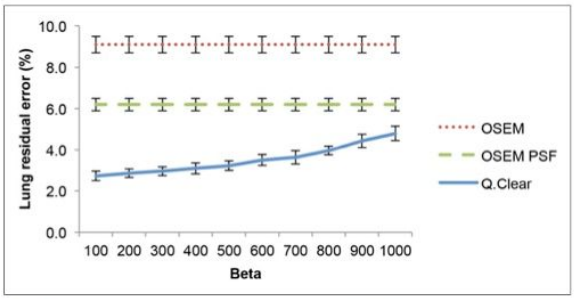


FIGURE 2. Graph of mean residual LE for OSEM (2 iterations, 24 subsets, 6.4-mm filter), OSEM PSF (3 iterations, 24 subsets, 2-mm filter), and Q.Clear (β = 100–1,000). Error bars shown are 1 SD.

RESULTS

The results for the phantom study are summarized in Table 1, Figures 1–3, and Supplemental Figure 1 (supplemental materials are available at http://jnm.snmjournals.org). As the noise penalization factor (β) was increased, the CR and BV decreased for all spheres sizes whereas the LE increased (Table 1). Q.Clear reconstructions generally gave higher CR and lower BV than OSEM (Fig. 1; Supplemental Fig. 1). Q.Clear (independent of β) had a significantly ($P < 0.0005$) reduced LE, compared with both OSEM and OSEM PSF reconstruction (Fig. 2).

In the clinical evaluation, there was moderate to substantial agreement between the 2 scorers on ranking the IQ parameters (Table 2). In most cases (87% and 73% by scorers 1 and 2, respectively), both scorers chose B400 (B*n* indicates a Q.Clear reconstruction with a β value of *n*) as their most preferred reconstruction. The least preferred reconstruction was OSEM PSF in all cases. This was confirmed by the combined scores for each reconstruction: 570, 792, 494, 258, 262, and 862 (OSEM, B200, B300, B400, B500, and OSEM PSF, respectively).

For all of the individual IQ parameters, with the exception of lesion detectability, B400 and B500 were the most consistently highest-ranked and OSEM PSF the lowest-ranked reconstruction (Table 2).

For lesion detectability, both scorers ranked B200 the highest in all lung nodule cases and in most of the mediastinal node cases (scorer 1, 5/5 cases; scorer 2, 4/5 cases). There was more variation in ranking the liver metastases cases, in which B400 and B200 were ranked highest in 3 of 5 cases by scorers 1 and 2, respectively (Fig. 4). OSEM PSF was the lowest-ranked construction for all the liver metastases cases (Fig. 4) and most of the mediastinal node cases (scorer 1, 3/5 cases; scorer 2, 5/5 cases) because of the degree of inherent noise in these respective organs. In contrast, OSEM PSF ranked second highest for lesion detectability in most of the lung nodule cases (scorer 1, 5/5 cases; scorer 2, 4/5 cases) (Fig. 5), which is explained by low noise levels observed in the lungs.

DISCUSSION

The aim of this study was to determine the optimum penalization factor (β) for the clinical use of Q.Clear. The phantom work was used to narrow the choice of β values used in the clinical part of the investigation and to investigate the properties of the new Q.Clear reconstruction.

Most Q.Clear reconstructions have a higher CR and lower BV than OSEM, suggesting that Q.Clear is an improved reconstruction (Table 1; Fig. 1). Although the increased CR with Q.Clear can be partly accounted for by the effective convergence achieved in

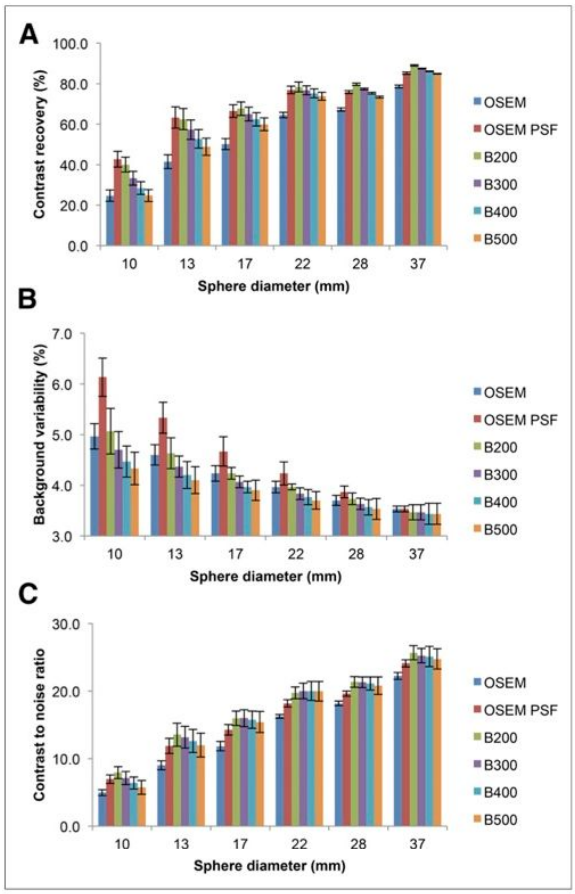


FIGURE 3. Graphs showing mean CR (A), BV (B), and contrast-to-noise ratio (C) for all spheres in NEMA IQ phantom (10, 13, 17, and 22 mm were hot and 28 and 37 mm cold). These are shown for OSEM (2 iterations, 24 subsets, 6.4-mm filter), OSEM PSF (3 iterations, 24 subsets, 2-mm filter), and Q.Clear (β = 200–500). Error bars shown are 1 SD.

Q.Clear (*8,9*), another reason to consider, especially for the small spheres, is the inclusion of PSF modeling into the Q.Clear algorithm. To consider this, OSEM PSF has also been included in the results. When Q.Clear is compared with OSEM PSF, some Q.Clear reconstructions have a higher CR and some lower, depending on the β value. With the exception of B100, Q.Clear reconstructions have lower BV than OSEM PSF. There is no statistically significant difference ($P > 0.05$) between BV from OSEM PSF or B100. There is therefore a choice to be made between increased CR (from OSEM PSF) and decreased BV (from Q.Clear). Ideally these points on the graph would lie in the top left of each figure (Fig. 1).

In our center, OSEM PSF is not used because of the noise seen in the clinical images (*23*), and so the B100 reconstruction can be discounted for the clinical part of the investigation because of its similarity to OSEM PSF. For high β values, it is possible to get a lower CR than OSEM, which is not ideal, despite the lower BV. For the smallest sphere (10 mm), there is a significant increase in CR for B100–400 ($P < 0.003$), no significant difference for B500 ($P = 0.3$), and a significant decrease in CR for B600–1,000 ($P < 0.001$) (Fig. 1A), confirming that it is appropriate to examine Q. Clear in the B200–500 range as that is where the main improvements can be seen, compared with standard OSEM reconstructions.

Considering the reconstructions assessed as part of the clinical investigation (OSEM, OSEM PSF, B200–B500) for all spheres, OSEM PSF and Q.Clear had higher CR than OSEM (Fig. 3A), which was significant ($P < 0.003$) for all but B500 for the smallest sphere. For the 2 smallest sphere sizes, OSEM PSF had the highest CR whereas for the remaining 4 sphere sizes B200 had the highest (Fig. 3A). The significances of these differences (between CR OSEM PSF and B200) for the 6 spheres from smallest to largest was a P value of 0.02, 0.0007, 0.006, 0.03, 0.001, and 0.0007, respectively. The absolute mean BV was less for Q.Clear than OSEM; however, this was not significant ($P > 0.05$) (Fig. 3B). OSEM PSF had higher BV than OSEM, and this was significant ($P < 0.03$) for all but the largest sphere ($P = 0.3$) (Fig. 3B). The contrast-to-noise ratio for all spheres was higher for Q.Clear than OSEM, which was significant ($P < 0.05$) for all but B500 for the largest and smallest sphere (both $P = 0.08$) (Fig. 3C). There was no statistical difference ($P > 0.05$) between the CNR from Q.Clear and OSEM PSF except for B200 for the 2 smallest spheres ($P < 0.02$).

The visual analysis of OSEM, OSEM PSF, and Q.Clear reconstructions of clinical scans mirrors the results of the phantom study. Overall, the IQ of reference organs and noise level was judged to be the best in Q.Clear reconstructions, specifically between B400 and B500, which produced a relatively smooth and

TABLE 2
Clinical Evaluation of Interrater Agreement and Individual IQ Parameter Rankings

Parameter	Highest-ranked reconstruction (% of cases)		Lowest-ranked reconstruction (% of cases)		Agreement	κ (95% confidence interval)	*P*
	Scorer 1	Scorer 2	Scorer 1	Scorer 2			
Overall IQ	B400 (87%)	B400 (73%)	OSEM PSF (100%)	OSEM PSF (100%)	Moderate	0.43 (0.30–0.55)	<0.001
Background liver	B500 (93%)	B400 (73%)	OSEM PSF (100%)	OSEM PSF (100%)	Moderate	0.60 (0.47–0.72)	<0.001
Background mediastinum	B500 (93%)	B400 (87%)	OSEM PSF (100%)	OSEM PSF (100%)	Moderate	0.55 (0.42–0.66)	<0.001
Background marrow	B500 (93%)	B500 (53%)	OSEM PSF (100%)	OSEM PSF (100%)	Substantial	0.72 (0.61–0.82)	<0.001
Noise level	B500 (93%)	B500 (47%)	OSEM PSF (100%)	OSEM PSF (100%)	Moderate	0.60 (0.47–0.71)	<0.001
Lesion detectability	B200 (67%)	B200 (73%)	OSEM PSF (53%)	OSEM PSF (67%)	Moderate	0.59 (0.45–0.70)	<0.001

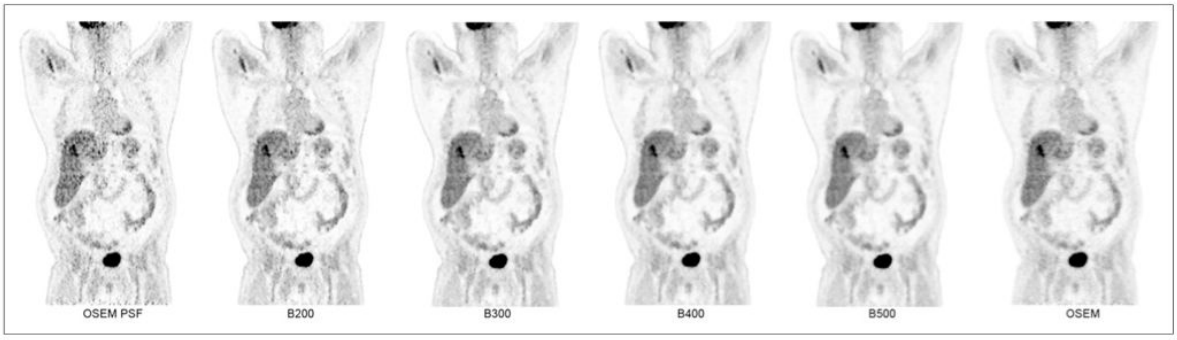


FIGURE 4. Coronal PET images demonstrating ^{18}F-FDG–avid liver metastasis across OSEM and Q.Clear reconstructions. Smooth homogeneous appearance of background liver on Q.Clear reconstruction renders metastasis more conspicuous than OSEM PSF. Conversely, there is also risk of false-positive findings for small foci on OSEM PSF reconstruction because of high level of background noise. All images are displayed on SUV scale 0–6.

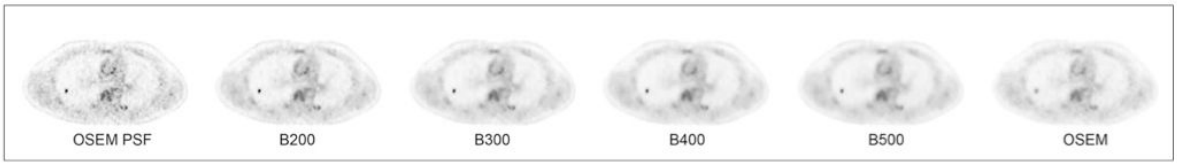


FIGURE 5. Axial PET images demonstrating ^{18}F-FDG–avid 7-mm right upper lobe lung nodule. This was mildly ^{18}F-FDG–avid on OSEM, but degree of uptake was shown to be higher using Q.Clear reconstruction. Despite similar degree of uptake on OSEM PSF, there was still low level of background noise from within lungs, which was reduced even on B200. Improvement in background noise within mediastinum, compared with OSEM PSF, is also illustrated. All images are displayed on SUV scale 0–6.

homogeneous appearance of background structures (Figs. 4 and 5). OSEM PSF had the lowest ranking in these IQ parameters, where the degree of noise in background structures had been previously deemed unsuitable for use in our center. These results collectively reflect the aforementioned differences in BV between the various reconstructions.

CR was highest in most spheres using B200, followed by OSEM PSF, which was reflected in lesion detectability being ranked highest in the lung nodule cases on B200 and visually appearing similar to OSEM PSF (Fig. 5; Table 2). However, OSEM PSF and B200 were the lower ranked reconstructions in the cases of mildly ^{18}F-FDG–avid mediastinal lymph nodes in non–small cell lung cancer and small liver metastases. This can be explained by the degree of inherent noise in the mediastinum and liver (Fig. 4), in comparison to the lung.

In most cases, B400 was selected as the most preferred reconstruction by both scorers. Considering the observed fall in CR beyond B500, this to some extent again corroborates the findings of the phantom study, suggesting that B400 would be the most appropriate choice that strikes a balance between optimizing CR and image noise levels.

The use of Q.Clear may have implications for the assessment and quantification of treatment response in cancer, in particular the improved detection of small foci, as demonstrated by the high CR and low BV in phantom work for B400. This may be particularly useful where abnormalities are surrounded by a degree of background tracer uptake (e.g., liver and mediastinum) and may increase the detection and diagnostic confidence of abnormalities in these areas.

There are several limitations to consider in this report. The number of cases and scorers used is small, and it is possible that increasing either or both of these variables may alter the preferred β value. It was also not possible to mask the scorers to the 2 different reconstruction algorithms, because Q.Clear has a visual appearance different from OSEM PSF at all β values, and this may have introduced bias. Additionally, this analysis is restricted to oncologic ^{18}F-FDG PET/CT; neurologic PET/CT would require different parameters as is the case for standard OSEM. Caution should also be exercised where SUV thresholds for the diagnosis of likely malignant involvement are used in clinical practice as these may have to be adjusted if Q.Clear is adopted.

CONCLUSION

Q.Clear subjectively improves image quality and increases CR and decreases BV in phantom studies, compared with standard OSEM reconstructions. A β value of 400 appears to be the optimal value in ^{18}F-FDG oncology PET/CT using the new Q.Clear reconstruction algorithm.

DISCLOSURE

The costs of publication of this article were defrayed in part by the payment of page charges. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734. A portion of this work was funded by a National Institute for Health Research (NIHR) Healthcare Scientist Chief Scientific Officer Doctoral Award. The research was supported by the NIHR Oxford Biomedical Research Centre Programme. This paper presents independent research funded by the NIHR. The views expressed are those of the authors and not necessarily those of the NHS, NIHR, or U.K. Department of Health. No other potential conflict of interest relevant to this article was reported.

ACKNOWLEDGMENTS

We thank Alice Asseraf and Hugo Arques at GE Healthcare for their assistance in the processing of sinograms using Q.Clear.

REFERENCES

1. Maffione AM, Grassetto G, Rampin L, et al. Molecular imaging of pulmonary nodules. *AJR.* 2014;202:W217–W223.
2. Kwee TC, Cheng G, Lam MG, Basu S, Alavi A. SUVmax of 2.5 should not be embraced as a magic threshold for separating benign from malignant lesions. *Eur J Nucl Med Mol Imaging.* 2013;40:1475–1477.
3. Adams MC, Turkington TG, Wilson JM, Wong TZ. A systematic review of the factors affecting accuracy of SUV measurements. *AJR.* 2010;195:310–320.
4. Kinahan PE, Fletcher JW. Positron emission tomography-computed tomography standardized uptake values in clinical practice and assessing response to therapy. *Semin Ultrasound CT MR.* 2010;31:496–505.
5. GE Healthcare White Paper: Q.Clear. GE Healthcare website. http://www3.gehealthcare.com/~/media/documents/us-global/products/pet-ct/whitepaper/q%20clear/ge-healthcare-white-paper_qclear.pdf. 2014. Accessed August 12, 2015.
6. Geman S, McClure DE. Statistical methods for tomographic image reconstruction. *Bull Int Stat Inst.* 1987;52:5–21.
7. Nuyts J, Beque D, Dupont P, Mortelmans L. A concave prior penalizing relative differences for maximum-a-posteriori reconstruction in emission tomography. *IEEE Trans Nucl Sci.* 2002;49:56–60.
8. Asma E, Ahn S, Ross SG, Chen A, Manjeshwar RM. Accurate and consistent lesion quantitation with clinically acceptable penalized likelihood images [abstract]. *2012 IEEE Nucl Sci Symp Med Imaging Conf Rec (NSS/MIC).* 2012:4062–4066.
9. Ahn S, Fessler JA. Globally convergent image reconstruction for emission tomography using relaxed ordered subsets algorithms. *IEEE Trans Med Imaging.* 2003;22:613–626.
10. Jaskowiak CJ, Bianco JA, Perlman SB, Fine JP. Influence of reconstruction iterations on ^{18}F-FDG PET/CT standardized uptake values. *J Nucl Med.* 2005;46:424–428.
11. Parvizi N, Franklin JM, McGowan DR, Teoh EJ, Bradley KM, Gleeson FV. Does a novel penalized likelihood reconstruction of ^{18}F-FDG PET-CT improve signal-to-background in colorectal liver metastases? *Eur J Radiol.* June 29, 2015 [Epub ahead of print].
12. Sah B, Veit-Haibach P, Delso G, et al. Clinical evaluation of a new block sequential regularized expectation maximization (BSREM) reconstruction algorithm in PET/CT studies [abstract]. *J Nucl Med.* 2014;55(suppl 1):2097.
13. Passalaqua S, Kappadath S, Branch D, et al. Qualitative and quantitative evaluation of regularized PET image reconstruction [abstract]. *J Nucl Med.* 2014;55(suppl 1):579.
14. Ma H, Asma E, Ahn S, et al. Clinical evaluation of penalized likelihood reconstruction in whole-body PET studies. *Eur J Nucl Med Mol Imaging.* 2013;40:S109.
15. Teoh EJ, McGowan DR, Bradley KM, Belcher E, Black E, Gleeson FV. Novel penalised likelihood reconstruction of PET in the assessment of histologically verified small pulmonary nodules. *Eur Radiol.* May 20, 2015 [Epub ahead of print].
16. Teoh EJ, McGowan DR, Belcher E, et al. SUVmax assessment of mediastinal nodes in non-small cell lung cancer using a novel penalised likelihood PET reconstruction. *Eur J Nucl Med Mol Imaging.* 2014;41:S397.
17. National Electrical Manufacturers Association. *NEMA NU-2-2012 Performance Measurement of Positron Emission Tomography.* Rosslyn, VA: NEMA; 2013.
18. Kenny KE, McGowan DR. Effect of the GE SharpIR PET reconstruction on contrast convergence and measured SUV values in the NEMA image quality phantom and patients. *Eur J Nucl Med Mol Imaging.* 2013;40:S209–S210.
19. Stearns CW, McDaniel DL, Kohlmyer SG, Arul PR, Geiser BP, Shanmugam V. Random coincidence estimation from single event rates on the discovery ST PET/CT scanner [abstract]. *2003 IEEE Nucl Sci Symp Conf Rec.* 2003;5:3067–3069.
20. Wollenweber SD. Parameterization of a model-based 3-D PET scatter correction. *IEEE Trans Nucl Sci.* 2002;49:722–727.
21. Iatrou M, Manjeshwar R, Ross SG, Thielemans K, Stearns CW. 3D implementation of scatter estimation in 3D PET [abstract]. *2006 IEEE Nucl Sci Symp Conf Rec.* 2006;4:2142–2145.
22. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–174.
23. Rahmim A, Qi J, Sossi V. Resolution modeling in PET imaging: Theory, practice, benefits, and pitfalls. *Med Phys.* 2013;40:064301.

IMAGES IN THORAX

^{18}F-Misonidazole PET-CT scan detection of occult bone metastasis

Daniel R McGowan,[1,2] Ruth E Macpherson,[3] Kevin M Bradley,[3] John D Fenwick,[1] Fergus V Gleeson,[1,3] Geoff S Higgins[1]

[1]Department of Oncology, University of Oxford, Oxford, UK
[2]Radiation Physics and Protection, Oxford University Hospitals NHS Trust, Oxford, UK
[3]Department of Radiology, Oxford University Hospitals NHS Trust, Oxford, UK

Correspondence to
Daniel R McGowan, Department of Oncology, University of Oxford, Oxford OX3 7DQ, UK; daniel.mcgowan@oncology.ox.ac.uk

Received 8 June 2015
Revised 10 August 2015
Accepted 14 August 2015
Published Online First 8 September 2015

A 67-year-old man with locally advanced non-small cell lung cancer entered a phase I trial combining a novel phosphoinositide 3-kinase inhibitor (BKM120) with palliative radiotherapy.[1] The trial uses ^{18}F-misonidazole (FMISO) positron emission tomography (PET)-CT imaging to assess any changes in tumour hypoxia. FMISO is a tracer currently only used in the research setting.

A pretreatment FMISO PET-CT scan identified an occult, asymptomatic scapular metastasis (figure 1) undetectable on routine CT imaging (figure 2A). This was subsequently confirmed on MRI imaging. The patient was referred for palliative chemotherapy upon completing the trial. To our knowledge, this is the first published report of FMISO PET-CT detecting an occult metastasis. The FMISO image (figure 1) also shows hypoxia within mediastinal nodes. The enlarged nodes were seen on the routine CT imaging (figure 2B) unlike the occult scapular metastasis.

FMISO is reduced and retained in hypoxic tissue and so allows the identification of hypoxic volumes. Tumour hypoxia is associated with marked resistance to radiotherapy and poor clinical outcomes.[2] There is significant interest in using FMISO PET-CT to monitor changes to tumour hypoxia arising from drug or radiotherapy treatment. FMISO images also identify hypoxic volumes of the tumour that may benefit from boosting the radiation dose. This image illustrates that small tumour volumes may harbour clinically relevant regions of hypoxia.

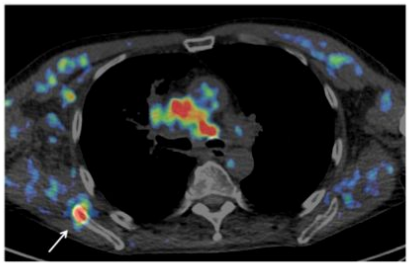


Figure 1 ^{18}F-Misonidazole positron emission tomography (PET) image fused with the corresponding CT displayed on a tumour-to-blood ratio (TBR) colour scale. Red regions depict a TBR >1.4, indicating hypoxia, and no visible PET depicts a TBR <1, indicating normoxia.[3] This axial image shows the unexpected hypoxic bone metastasis indicated by the white arrow in addition to hypoxic nodal disease in the mediastinum.

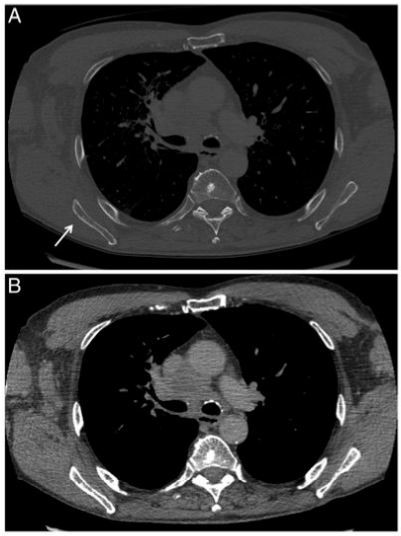


Figure 2 CT contrast-enhanced image reformatted using 1.25 mm axial thickness for (A) a bone reconstruction showing that the bone metastasis visible on ^{18}F-misonidazole (FMISO) positron emission tomography (PET) is not visible on CT and (B) standard reconstruction showing the enlarged mediastinal nodal disease shown to be hypoxic on the FMISO PET.

Contributors All authors wrote, revised and approved the final manuscript.

Funding The trial was supported by Cancer Research UK (CRUKD/12/016), the National Institute of Health Research (NIHR), Health Education England (HEE), the Oxford ECMC, the CRUK EPSRC Oxford Cancer Imaging Centre, the CRUK Oxford Centre and Novartis.

Competing interests None declared.

Patient consent Obtained.

Ethics approval Oxford REC.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 University of Oxford. A CR-UK Phase I Study of BKM120 in Patients With Non-small Cell Lung Cancer (NSCLC) Receiving Thoracic Radiotherapy. http://www.clinicaltrials.gov/ct2/show/NCT02128724 NLM Identifier: NCT02128724
2 Moeller BJ, Richardson RA, Dewhirst MW. Hypoxia and radiotherapy: opportunities for improved outcomes in cancer treatment. *Cancer Metastasis Rev* 2007;26:241–8.
3 Koh WJ, Rasey JS, Evans ML, *et al*. Imaging of hypoxia in human tumors with [F-18]fluoromisonidazole. *Int J Radiat Oncol Biol Phys* 1992;22:199–212.

To cite: McGowan DR, Macpherson RE, Bradley KM, *et al*. *Thorax* 2016;**71**:97.