Estimation and reliability of myocardial blood flow with dynamic PET



Antoine Saillant

Wolfson College

University of Oxford

A thesis submitted to the Department of Engineering Science,
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Trinity Term 2018

Acknowledgements

This thesis would not have been possible without the support and advice from many people. I would like to use these few words to express my gratitude and tell them I owe them a great debt of thanks.

First and foremost, my deepest gratitude goes to my supervisor Dr. Michael Chappell. The door to Michael's office was always open whenever I had questions, and his flexibility and patience were incredibly helpful considering the unusual set-up of my studies. I would like to thank also Dr Mark Jenkinson, Dr Jérôme Declerck, Dr Sven Zuehlsdorff and Dr Vijay Shah for providing guidance, showing me scientific excellence and steering me in the right direction whenever I needed while allowing me to have as much freedom and initiative as I wanted. I would like also to thank the experts who were involved in different parts of my work, in particular Dr Guenther Platsch, Dr Bernard Bendriem and Dr Ian Armstrong who offered great scientific and clinical enlightening on PET imaging.

This thesis is the result of a collaboration between Siemens Molecular Imaging and the University of Oxford. It would not have been possible without the support from Dr Christine Lorenz who immediately accepted to finance my doctorate when I asked. I will be forever grateful for what she has done for me and for trusting me regardless of the obvious complications of doing a PhD between Oxford and Knoxville. During my time at Siemens I have met remarkable people, but the Clinical Applications team definitely played a role in my professional development and deserves a special thanks: Chuck Hayden, Bruce Spottiswoode, Ludovic Sibille, Fei Gao and Rachid Fahmi.

I have been very lucky to find incredible friends who have encouraged me, supported me and also distracted me whenever needed: Moss Zhao, Tom Prill, Gayathri Jeyagopi, Raza Khan, Dipen Dave, Grace Gustilo and Andrew Slatton. My experience would have been for sure very different without them.

I would have not been able to do such studies without the unconditional support of my parents. *Merci* for encouraging me to the path of scientific studies. Finally, I would like to thank my fiancée Alina for supporting me during these three years and going through the difficulties of living 6,000 km apart.

Abstract

Dynamic Positron Emission Tomography (PET) enables the estimation of Myocardial Blood Flow (MBF), by modeling the temporal evolution of a radio-labeled pharmaceutical over multiple PET images acquired over time. Despite its advantages, the absolute quantification of MBF is not used routinely in the clinic due to a lack of clinical standardization. MBF estimates can be affected by patient motion, among other factors, which creates displacements between different PET frames and leads to errors in MBF calculation. Automated motion correction techniques are difficult to perform, due to the rapid change of the tracer spatial distribution throughout the PET images series, and the low resolution of PET image series. Furthermore, it is challenging to quantitatively evaluate the success of motion correction techniques, and their impact on MBF reliability is not fully understood.

In this thesis, we developed a framework for estimating MBF along with the uncertainty on the estimate. The framework relies on Bayes theory, which represents the kinetic parameters as a probability distribution, from which uncertainty measures can be extracted. If the uncertainty extracted is high the parameter studied is considered to have high variability - or low confidence - and vice versa. A Bayesian model inference technique, called Variational Bayes, was chosen due to its relatively inexpensive computations compared to other Bayesian techniques.

The reliability of parameter and uncertainty estimates were first evaluated through simulations and compared with another Bayesian model inference technique considered the "gold standard". It was found that the uncertainty metric was estimated robustly across multiple noise levels, which was necessary as motion can be seen as an additional source of noise. The uncertainty metric was then used to assess motion correction, performed manually by experienced observers. The uncertainty, calculated before and after motion correction, reduced significantly in moderate and high motion cases, demonstrating the usefulness of the method. The uncertainty was subsequently used to evaluate a motion correction algorithm, showing comparable performance to manual realignment provided by experienced users, with regards to the uncertainty metric. Finally, we investigated different situations where the use of the uncertainty metric may convey clinical information, besides assessing motion correction.

Contents

List of Figures **vii**

List of Abbreviations **ix**

1 Introduction: Precision medicine and uncertainty measurements **1**
1.1 Precision medicine and medical imaging 1
1.2 Coronary artery diseases and molecular imaging 3
1.3 The value of measuring uncertainty 4
1.4 Thesis objectives and outline 5

2 Literature review **8**
2.1 Positron Emission Tomography 8
2.1.1 Historical background of nuclear medicine 8
2.1.2 Physics principles 10
2.1.3 From raw signal to 3D images 13
2.1.4 Kinetic modeling 20
2.2 PET for coronary artery disease 23
2.2.1 Anatomy of the heart and the cardiovascular system 23
2.2.2 Cardiac studies with nuclear medicine 27
2.2.3 Cardiac studies with dynamic PET 29
2.2.4 Impact of motion in cardiac dynamic PET 35
2.3 Motion correction in dynamic PET 36
2.3.1 Motion detection 37
2.3.2 Image registration 38
2.3.3 Motion correction for PET imaging 42
2.3.4 Specificities and challenges of motion correction in PET . . . 43
2.4 Model inference and measuring uncertainty 47
2.4.1 Problem definition 48
2.4.2 Model parameter inference 50
2.4.3 Measuring parameter uncertainty 53
2.4.4 Bayesian inference 55
2.4.5 Bayes theory 55
2.5 Conclusion 60

3 Calculating kinetic parameter uncertainty with Variational Bayes 62
3.1 Introduction . . . 62
3.2 Theory . . . 63
3.2.1 Model implementation . . . 63
3.3 Methods . . . 66
3.3.1 Simulations . . . 66
3.3.2 Clinical data: cross validation between VB and MCMC . . . 71
3.4 Results . . . 73
3.4.1 Simulations . . . 73
3.4.2 Clinical data . . . 75
3.5 Discussion . . . 79
3.6 Conclusion . . . 84

4 Assessing motion correction by measuring MBF uncertainty 86
4.1 Introduction . . . 86
4.2 Methods . . . 88
4.2.1 Data . . . 88
4.2.2 PET frames realignment workflow . . . 89
4.2.3 TAC fitting and uncertainty estimation . . . 90
4.2.4 Analysis . . . 92
4.3 Results . . . 92
4.4 Discussion . . . 96
4.5 Conclusion . . . 100

5 Towards automated motion correction 101
5.1 Introduction . . . 101
5.2 Methods . . . 104
5.2.1 Motion correction framework . . . 104
5.2.2 Data . . . 108
5.2.3 Evaluation and comparison to other methods . . . 108
5.2.4 Analysis . . . 109
5.3 Results . . . 111
5.4 Discussion . . . 112
5.5 Conclusion . . . 117

6 Uncertainty and spillover effect **119**
6.1 Introduction . . . 119
6.2 Methods . . . 123
6.2.1 Experiment 1: Weighted methods to calculate flow values . . 123
6.2.2 Experiment 2: uncertainty in abnormal scans . . . 126
6.3 Results . . . 128
6.3.1 Experiment 1 . . . 128
6.3.2 Experiment 2 . . . 128
6.4 Discussion . . . 130
6.5 Conclusion . . . 134

7 Conclusions and future work **135**
7.1 Conclusions . . . 135
7.1.1 Applying a VB framework to calculate parameter uncertainty in kinetic PET . . . 136
7.1.2 Assessing PET frames realignment and motion correction . . 136
7.1.3 Correcting motion on the entire dynamic PET scan timecourse 137
7.1.4 Clinical benefit of using uncertainty . . . 138
7.2 Future work . . . 138
7.2.1 Fully automate motion correction . . . 138
7.2.2 Defining what is an "acceptable" uncertainty . . . 139
7.2.3 Assessing other factors that influence MBF estimation . . . 139
7.2.4 Extending to further applications . . . 140
7.3 Summary . . . 140

Appendices

A Clinical data: acquisition and processing **142**
A.1 Data collection . . . 142
A.2 Data acquisition and reconstruction . . . 143
A.2.1 Acquisition protocol . . . 143
A.2.2 Reconstruction protocol . . . 144
A.3 Data post-processing . . . 145
A.4 Implementation details . . . 146

References **148**

List of Figures

2.1 Example of a SPECT and PET scanner 10
2.2 PET Physics principle 11
2.3 Sinogram formation 14
2.4 Heart sinogram and reconstructed image 14
2.5 True, scatter and random coincidence events 15
2.6 Comparison of attenuation corrected and non-attenuation corrected images. 16
2.7 Schematic of attenuation correction 17
2.8 Iterative reconstruction process of a PET image 20
2.9 Example of typical Time Activity Curves 21
2.10 Example of compartmental models 21
2.11 Heart diagram 25
2.12 PET Tracer extraction fraction 28
2.13 Cardiac planes 29
2.14 From 3D segmentation to polar map 30
2.15 Tracer pattern in cardiac dynamic PET 32
2.16 One-tissue compartment model with blood volume fraction. 35
2.17 Image interpolation 39

3.1 Examples of model fitting using VB. 74
3.2 Comparison of distribution of errors for White and PET noise. . . . 74
3.3 Relative errors for estimating K_1, k_2 and S for Rest TACs. 76
3.4 Relative errors for estimating K_1, k_2 and S for Stress TACs. 77
3.5 Comparison between $f_{CI_{VB}}$ and $f_{CI_{MH}}$ 78
3.6 K1 distribution example for VB and MCMC 79
3.7 Bland Altman comparing K_1 between VB and MCMC 80
3.8 Barplot of Dice coefficient of MCMC and VB distributions 80
3.9 Influence of the spillover fraction parameter on the K_1 distribution 83

4.1 Motion alignment workflow strategy. 89
4.2 Comparison of original and motion corrected (MCAC) dataset . . . 93
4.3 Comparison of input functions and TAC before and after correction. 94

4.4 Change in uncertainty after motion correction for datasets corrected with MCAC. 94
4.5 Change in μ_{K_1} after motion correction for datasets corrected with MCAC. 95
4.6 Bland-Altmann plot of σ_{K_1} and μ_{K_1} for MC and MCAC 96
4.7 Comparison of CV_{K_1} values before and after motion correction for MCAC datasets for observer 1. 97
4.8 Comparison of uncertainties after correction of three manual realignments. 99

5.1 Temporal evolution of the tracer in the heart 102
5.2 Patient specific myocardium template 105
5.3 Frame-specific mask generation 106
5.4 Difference of Gaussians filtering 107
5.5 Comparison of uncertainty reduction between automated motion correction and manual realignment. 113
5.6 Example of intra-frame motion artifacts 115

6.1 Spillover effect 120
6.2 Relationship between uncertainty and spillover 121
6.3 Asymmetry of the myocardium 121
6.4 Effect of spillover in low perfusion area 123
6.5 Bland Altman for average methods comparison 129
6.6 Uncertainty for abnormal scans 131

A.1 Acquisition protocol 144
A.2 Myocardium segmentation 146

List of Abbreviations

1D, 2D, 3D . . One-, two- or three-Dimensional.
AC . . . . . Attenuation Correction
ASL . . . . . Arterial Spin Labeling
CAD . . . . . Coronary Artery Disease
CC . . . . . Cross-Correlation
CT . . . . . Computed Tomography
CV . . . . . Coefficient of Variation
FBP . . . . . Filtered Back Projection
FDG . . . . . FluoroDeoxyGlucose
LA . . . . . Left Atrium
LAD . . . . . Left Anterior Descending
LCX . . . . . Left Circumflex
LOR . . . . . Line Of Response
LV . . . . . Left Ventricle
OSEM . . . . . Ordered Subset Expectation Maximization
MBF . . . . . Myocardial Blood Flow
MC . . . . . Motion Correction
MCAC . . . . Motion Correction with Attenuation Correction
MCMC . . . . Monte-Carlo Markov Chain
MFR . . . . . Myocardial Flow Reserve
MH . . . . . Metropolis-Hastings
MI . . . . . Mutual Information
MLEM . . . . Maximum Likelihood Expectation Maximization
MPI . . . . . Myocardial Perfusion Imaging
MRI . . . . . Magnetic Resonance Imaging

MT . . . . . Myocardium Tissue

NLLS . . . . . Non-Linear Least Squares

PET . . . . . Positron Emission Tomography

RA . . . . . Right Atrium

RCA . . . . . Right Coronary Artery

RV . . . . . Rigt Ventricle

ROI . . . . . Region Of Interest

SNR . . . . . Signal to Noise Ratio

SPECT . . . . Single-Photon Emission Computed Tomography

SSD . . . . . Sum of Square Difference

TAC . . . . . Time-Activity Curve

VB . . . . . Variational Bayes

Publications arisen from this thesis

Journal article

- **Antoine Saillant**, Ian Armstrong, Vijay Shah, Sven Zuehlsdorff, Charles Hayden, Jerome Declerck, Kimberley Saint, Matthew Memmott, Mark Jenkinson and Michael Chappell. *"Assessing reliability of myocardial blood flow after motion correction with dynamic PET using a Bayesian framework"*, In: IEEE Trans. Med. Imaging. (In Press)

Conference proceedings

- **Antoine Saillant**, Kimberley Saint, Matthew Memmott, Ian Armstrong, Vijay Shah, Sven Zuehlsdorff, Jérôme Declerck, Mark Jenkinson, Michael Chappell. *"Estimation and reliability of myocardial blood flow after motion correction with dynamic PET using a Bayesian framework"*, In: Eur. J. Nucl. Med. Mol. Imaging 2017; 44:S119–956 - [Abstract OP-665]. Oral presentation at 2017 EANM congress.

- **Antoine Saillant**, Guenther Platsch, Vijay Shah, Sven Zuehlsdorff, Mark Jenkinson and Michael Chappell. *"Reliability of myocardial blood flow estimates in abnormal scans"*, In: J. Nucl. Med. 2018; 59:373 [Abstract 373]. Oral presentation at 2018 SNMMI congress.

- **Antoine Saillant**, Sven Zuehlsdorff, Charles Hayden, Vijay Shah and Michael Chappell. *"Reducing the Influence of Inter-user Segmentation Variability Myocardial Blood Flow Values"* JNC 2018; In Press. [ePoster 234-10]. Poster at 2018 ASNC congress.

1

Introduction: Precision medicine and uncertainty measurements

1.1 Precision medicine and medical imaging

Life expectancy of men and women has increased over the past decades [1] thanks to progresses made in medicine and science more generally. This improvement comes however with a consequent financial burden: the costs of healthcare have also increased year after year [2]. In order to maintain costs as low as reasonably possible for societies and individuals while maintaining high standard of care, a novel paradigm in healthcare has arisen: precision medicine [3, 4]. A definition of precision medicine is given by the American National Academy of Sciences: *"Precision medicine refers to the tailoring of medical treatment to the individual characteristics of each patient, encompassing the ability to classify individuals into subpopulations that differ in their susceptibility to a particular disease, in the biology and/or prognosis of those diseases that may develop, or in their response to a specific treatment"* [4]. In other words, precision medicine can be summarized by choosing the right treatment for the right disease and the right patient. If the terminology is relatively new, this notion is already commonly accepted in modern medicine: when a person needs a blood transfer, one needs to make sure that the blood coming from the donor is compatible with the receiver's.

The ambition to find the appropriate treatment relies principally on having patient data; and decision makers have started acting towards that goal. For example, in 2012 UK Prime Minister David Cameron announced the 100,000 Genomes Project to collect and read the DNA of people with cancer or rare diseases [5]. In 2015, US President Barack Obama launched the Precision Medicine Initiative, a project that aims to gather the genomes of 1 million volunteers and develop the appropriate databases to store and read this information[6].

In precision medicine, medical imaging plays a central role [7]. Whether it is through screening, early diagnosis, evaluating the response to therapy or assessing the disease recurrence, biomedical imaging provides multi-parametric information, and can measure various biomarkers. The most common types of medical imaging modalities include Computed Tomography (CT), Magnetic Resonance Imaging (MRI), ultrasound, and molecular imaging – such as Single Photon Emission Tomography (SPECT) and Positron Emission Tomography (PET). Depending on the modality and how it is used, clinicians can obtain anatomical information, for example the shape and volume of an organ, or functional information such as change in chemical composition of a tissue. PET and SPECT are two imaging techniques that require the injection of a radioactive molecule, called tracer, into the patient's body [8]. The signal emitted by the decay of the radioactive material is captured by specific cameras, and reconstructed into a 3 dimensional image, so that the tracer can be located spatially. If a dynamic scan is performed, that is if multiple 3D scans are acquired at different time intervals, the molecule can also be traced temporally. One of the advantages is indeed to be able to measure easily the evolution of the tracer over time, and to model it with a mathematical function. This process is called kinetic modelling [9]. It gives absolute quantification of a physiological phenomenon, e.g. metabolic rate constant, blood flow etc. Those biomarkers help clinicians to refine their diagnosis, and have more information than what anatomical images would provide, hence offering excellent added value in the context of precision medicine [10, 11].

1.2 Coronary artery diseases and molecular imaging

Coronary artery disease (CAD) is the reduction of blood flow to the heart muscle, named myocardium. A complete blockage of the arteries may result in cardiac tissue death (myocardial infarction – also known as heart attack). In the UK alone, coronary artery disease is the second most frequent cause of death after cancer, leading to more than 70,000 deaths in 2015 [12]. If a patient is diagnosed with coronary artery disease, multiple procedures exist for treatment, such as coronary bypass or coronary angioplasty. The question is then how to detect CADs as early as possible and with the highest accuracy.

Imaging procedures generally used to diagnose CADs are coronary angiography with a CT scan, myocardial perfusion imaging (MPI) with a SPECT scan, echocardiogram with an ultrasound probe or a contrast enhanced MR. An upcoming technique, which is increasingly accepted as clinical routine, is to measure myocardial blood flow with PET [13]. The perfusion between the blood and the myocardial tissue is captured with a dynamic scan, and the myocardial blood flow is modelled and measured from the reconstructed images. The advantage of dynamic PET over other techniques is multiple: compared to CT angiography, which is considered the gold standard for detecting CADs at the moment, a PET tracer can flow to smaller arteries, which can help for early diagnosis. Also during a CT angiography, a catheter is inserted through the blood vessel and threaded to the heart coronary artery, while a PET procedure is much less invasive, since it requires only a tracer injection[14]. The benefit of a cardiac PET over a SPECT scan is first the image resolution. The signal acquired is of a much higher signal to noise ratio (SNR), which translates into higher image quality [15]. Also it is more difficult to detect uniformly distributed blood flow reduction with a SPECT, due to the absence of absolute quantification of myocardial blood flow. It has been shown that measuring myocardial blood flow with dynamic PET helps to detect patients with balanced triple vessel disease [16, 17]. For the MRI scan, the SNR is higher than for other modalities, but it suffers from a couple of obstacles as well. Firstly,

patients with mechanical devices and coronary stents cannot be imaged with an MRI scanner, because of their metallic parts. Secondly the use of gadolinium has raised some concerns in patients with kidney failures. Finally the disadvantage of an echocardiogram over a cardiac PET scan is its poor image resolution, and the fact that is highly depends on the local user expertise [14].

1.3 The value of measuring uncertainty

As discussed in the first section, precision medicine relies on multiple sources of information, gathered from DNA tests, medical imaging or other standard medical procedures. Those parameters are then analyzed and potentially modeled with statistical tools in order to elaborate the appropriate diagnosis and treatment. Unfortunately, as for every physical measurement, these information metrics come with a certain error margin. These errors can be either systematic such as a wrong calibration of the instruments used, or random errors, as for instance an external perturbation occurring during the measurement. In either case, it is desirable to have an indicator of the reliability, or uncertainty of the metric that is being read.

There are various methods to extract uncertainty metric from a measurement[18]. A first method, which is very simple in practice, is to repeat the same experiment or measurement a number of times high enough so that statistics can be inferred from the repetition of experiments. A drawback of this method is that for the statistics to be meaningful, one needs to be sure that experiments are realized in identical conditions. This is probably an acceptable hypothesis when running simulations, but less likely to be true with clinical measurement on patients. Also for a given experiment, no inference can be made on the quality of this experiment, resulting in more or less uncertainty on the parameter inferred. However if the parameter extracted relies on a mathematical modeling of the data measured, mathematics provides us with several tools to help us measure uncertainty and quality of the measurement. A straightforward technique is to measure the distance between the modeled function and the measured data. The higher the distance is, the less likely the parameters measured are accurate. An alternative is to look at Bayes

theory, which offers a natural framework to estimate uncertainties of parameters in a model [19]. In a Bayesian framework, each parameter is represented as a probability distribution instead of a single value, from which an uncertainty metric can be drawn.

Knowing the uncertainty of model parameters can be of very high importance. Different situation such as noisy data or inappropriate model use could indeed lead to wrong estimates of the model parameters, and ultimately wrong conclusions. In the context of healthcare and precision medicine, if the information given comes along with an uncertainty metric, the physician could know how much he can trust that metric and simply ignore it, instead of reaching the wrong conclusion. Even in the best case scenario where an expert could guess, from experience, whether or not to trust a measurement, bias could be affecting his or her judgement, while an uncertainty measurement could reinforce and validate the suspicion.

1.4 Thesis objectives and outline

The estimation of myocardial blood flow (MBF) in dynamic PET can be affected by many factors, such as the choice of reconstruction method, the temporal sampling strategy, the post-processing methods and patient motion [20]. While it is accepted that motion has an impact on the measurement, it is not fully understood how the MBF measures are impacted.

The objective of this thesis is to develop a framework that can infer myocardial blood flow with an uncertainty measurement, using a Bayesian framework. The impact of motion and motion correction on the MBF uncertainties are studied, and a method for automated motion correction of dynamic cardiac scans is proposed. The work is presented in the following 5 chapters.

Chapter 2 introduces more in depth the different concepts briefly presented in this introduction. The principle of PET imaging is detailed, including physics principles, reconstruction and kinetic modelling. The chapter then examines the clinical knowledge on coronary artery disease, and how myocardial blood flow is computed from dynamic PET imaging. In this chapter we also review the literature on medical image registration and motion correction. Finally chapter 2

gives a review on uncertainty calculation with a particular focus on Bayes theory and the Variational Bayes framework, which is used in this work to measure uncertainty of myocardial blood flow.

Chapter 3 presents the implementation and validation of the Variational Bayes theory for PET kinetic modelling. The challenges arising from applying this framework to PET kinetics are discussed, and validations are made with simulations. In order to measure the quality of the uncertainty metrics computed, the framework is compared to an implementation of the Monte-Carlo Markov Chain algorithm that is regarded as "gold standard" for Bayesian inference.

Chapter 4 looks at the impact of motion on myocardial blood flow and myocardial blood flow uncertainties. Clinical datasets with different severity of motion, that have been manually corrected by three expert users, have the uncertainty of myocardial blood flow computed before and after motion correction. In particular we show that uncertainty of myocardial blood computed after motion correction significantly reduced for datasets with high motion.

Chapter 5 presents a new method to correct for motion of cardiac dynamic PET. This method uses a patient-based synthetic template to register each frame of the PET series. This chapter uses the results from the two previous chapters in order to evaluate automated motion correction algorithm of dynamic cardiac PET.

Chapter 6 explores other potential uses of uncertainty measures. We show for example that uncertainties are higher in the case of infarcted patients due to very high spillover effect, which could lead to incorrect estimation of myocardial blood flow.

Chapter 7 summarises the work presented in this manuscript and the results that were obtained. Techniques used and developed are summarised, in order to draw conclusions. Finally, this chapter provides suggestions for future work which could improve some of the methods developed in the preceding chapters.

This thesis is a contribution to the field of cardiac dynamic PET. Specifically, a Bayesian framework is used to infer the kinetic parameters uncertainty, and subsequently used to assess the quality of a motion correction of cardiac PET

dynamic images. In addition, a new motion correction framework is described to correct for motion throughout the entirety of the PET series, including early frames.

2
Literature review

This chapter presents a brief review on Positron Emission Tomography (PET), focusing on the signal formation, acquisition and reconstruction. The chapter then discusses the cardiovascular system as well as how PET imaging is used clinically for cardiac studies. It especially focuses on how the Myocardial Blood Flow (MBF) can be modeled and calculated by using a dynamic PET scan, and highlights the importance of motion correction for accurate MBF estimates. The chapter then reviews motion correction techniques and image registration theory, and surveys the motion correction techniques available in dynamic PET. The chapter concludes with the different methods to calculate uncertainties for model parameter inference. A special emphasis is given on Bayes theory, as it provides a natural mathematical framework to calculate uncertainty.

2.1 Positron Emission Tomography

2.1.1 Historical background of nuclear medicine

The development of nuclear medicine progressed with the advances in nuclear physics, chemistry and engineering throughout the 20^{th}century. It can be therefore challenging to define an exact birthdate for this field of medicine. Major steps forward were made in 1936 and 1946 [21]. In 1936, John Lawrence, an American

physicist, used for the first time an artificial radionuclide for medical application, when he used phosphorus 32 to treat a patient with leukemia [22]. The second major advance was made ten years later, when radionuclides for medical use were produced in high quantities by the Oak Ridge National Laboratory, USA [23]. The same year, a journal article reported the case of a patient with thyroid cancer metastases, whose disease progression stabilized after treatment of iodine 131 injection [24]. While it was understood that radionuclides could be injected in order to treat patients with metastases, it was only a few years later that it was possible to acquire the signal emitted by the decay of these isotopes.

In the 1950s, the first scintillation camera (also called Anger camera, the name of its inventor) was developed [21]. The system is composed of crystal material that scintillates in response to gamma radiation (photon). Photomultiplier tubes placed behind the crystal detect the scintillation and the number of events is then counted by a computer. A couple of years later those cameras were combined together, so that a two-dimensional signal could be acquired. The first scintigraphy scanner - also referred as gamma camera - was born.

In the 1980s, the development of Single Positron Emission Computed Tomography (SPECT), introduced the first three dimensional images, which was a major step forward in the development of nuclear cardiology and nuclear oncology [25]. SPECT is an imaging modality that uses a dual head gamma camera at different projection planes. Those cameras can rotate around the patient capturing the signal in 3D, as shown in Figure 2.1.

In parallel, Positron Emission Tomography (PET) saw its early development in the 1960s [26]. PET imaging relies on the emission of a pair of photons that are diametrically aligned. Unlike SPECT, where gamma cameras are in a planar position, PET detectors are positioned in circular array (gantry) (see Figure 2.1). The physics principles of PET is detailed in the next section of this chapter.

PET and SPECT scanners acquire functional images. Functional images help the clinician understand the functioning of the body (its physiology, its metabolism) by studying the temporal and spatial distribution of the tracer interacting with the



Figure 2.1: Left: A SPECT-CT scanner. One can see the dual head gamma cameras, as well as the CT behind. The gamma cameras can be rotated and positioned with an angle to maximize the signal acquisition from a specific organ or body region.
Right: A PET-CT scanner. The PET detector are positioned in a circular ring, the patient is positioned on the bed and is moved within the detector rings for the signal acquisition. Depending on the type of the scan (organ specific or full body) the patient is placed at different bed positions.
Data courtesy: Siemens Healthineers

body. In functional images, pixels represent physiological or even cellular properties of tissues such as glucose metabolism, perfusion or binding capabilities of certain neuro receptors. Other imaging modalities such as MR or CT typically acquire anatomical images. Thus, PET and CT images can deliver complimentary clinical information on anatomical as well as molecular level. Most of PET and SPECT are commonly used as a multi-modality device with a CT scanner. The anatomical information provided by the CT image is used to estimate the attenuation that the PET or SPECT signal encounter through the patient. Attenuation correction techniques resulted in a major improvement in imaging quality [25, 26]. The first PET-CT prototype was made in 1998. Finally, about a decade ago, fully integrated PET-MR systems were made commercially available.

2.1.2 Physics principles

A brief overview of PET physics principles is given here, in depth descriptions can be found in literature [27–30].

A radioactive isotope, also called radio-isotope or radio-nucleide, is an atom whose nucleus is unstable. Such atoms decay in a process that changes their composition (number of protons and neutrons in the nucleus), so that a nucleus

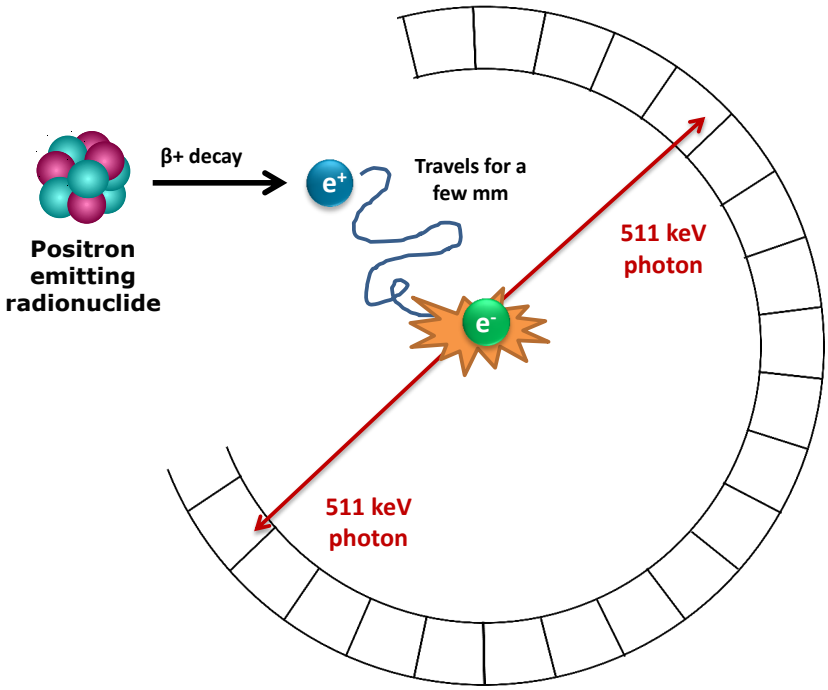


Figure 2.2: PET physics principle: Certain types of radio-isotopes, when they decay, emit a positron, which then annihilates with an electron to produce 2 photons that are 180 degrees opposed. The scanner consists of a cylinder of detectors (gantry), oriented towards the center of the circle, where the object is placed.

with a more stable state is reached. The time taken for the radioactivity of a radio-nuclide to decrease by half of its original value is called the half-life of the radio-nuclide. The radioactive activity of the radio-isotope is measured in Becquerel, which is a count rate, measured in decay per second.

Certain types of radio-nuclides are positron emitters. In such isotopes, a proton of the nucleus is transformed into a neutron and a positron. The positron is the anti-particle of the electron, i.e. they have the same mass but an opposite electrical charge. The positron is emitted with a certain kinetic energy. The positron range is defined as a finite distance the positron is required to travel within the tissue medium before undergoing annihilation. Emission ranges vary from one radio-nuclide to another. After losing its kinetic energy, the positron ultimately collides with an electron, and the two particles annihilate. In the annihilation, the mass of the electron and the positron are converted into energy and generate a pair of two 511 keV photons diametrically aligned, as shown in Figure 2.2.

PET is a molecular imaging technique, that allows clinicians to observe physiological processes in the body. Positron emitter isotopes are attached to a specific radio-pharmaceutical, also called radio-tracer or most commonly tracer. Tracers are designed to be active in a specific biological or chemical process in the body, e.g. glucose metabolism, perfusion, neuroreceptors. A PET scanner is constituted of detectors, positioned in a cylindrical array (gantry) around the patient, which detect pairs of photons, emitted by positron emitters. Consequently, PET images show the functioning of the body, through the accumulation of the tracer in tissues of interest, targeted by the radio-pharmaceutical. The two photons are measured by the PET detectors within a narrow time window (a few nanoseconds), called coincidence, and an image of the spatial distribution of the tracer concentration can be recovered. Some details about image reconstruction are given in the next section of this chapter.

The most commonly used radio-nuclides in PET are positron emitters with short half-lives (see Table 2.1) such as nitrogen 13 (^{13}N), carbon 11 (^{11}C), oxygen 15 (^{15}O), fluorine 18 (^{18}F), gallium 68 (^{68}Ga) or rubidium 82 (^{82}Rb). Those isotopes are either used directly such as ^{82}Rb (in the form of rubidium chloride solution), or assembled as radio-pharmaceutical that are normally processed by the body such as water or glucose, or incorporated to pharmaceuticals that bind to natural biological receptors such as ^{68}Ga-Prostate-Specific-Membrane-Antigen (^{68}Ga-PSMA), which is used in prostate cancer imaging. The most common radio-compound used is ^{18}F-FluoroDeoxyGlucose (^{18}F-FDG). It is a compound with a molecular structure closely similar to glucose. Another key property of PET tracers is the emission range of the radio-nuclide used. For a high emission range, the actual positron emission is located at a higher distance from the positron-electron annihilation, which increases the noise in the image and decreases its quality [31]. Emission ranges for typical radio-nuclides can be seen in Table 2.1.

Radionuclide	Half-life (in min.)	Average positron energy (in MeV)	Average emission range in water (in mm.)
^{11}C	20.4	0.39	0.85
^{13}N	10	0.50	1.15
^{15}O	2.2	0.72	1.8
^{18}F	110	0.25	0.46
^{68}Ga	68.3	0.83	2.15
^{82}Rb	1.25	1.5	4.10

Table 2.1: Properties of common positron emitters: the most important properties are the half-life and the energy of the positron when it is emitted - which correlates directly with its emission range. For a high emission range, the actual positron emission is located at a higher distance from the positron-electron annihilation.

2.1.3 From raw signal to 3D images

2.1.3.1 Data acquisition

The pair of photons emitted by the positron-electron are measured by two detectors elements of the scanner. Those elements define a line of response (LOR) for this particular annihilation event. A LOR can be characterized by the distance to the center and the azimuthal angle. The space defined by those parameters is called sinogram [32]. Thus, each annihilation event can be represented as a point in the sinogram, as shown in Figure 2.3. A sinogram of the heart is shown in Figure 2.4, together with the reconstructed image.

The measurement obtained by the pair of detectors is done with a pre-defined energy and time window, and is called a coincidence event. However all recorded events are not representing a single positron-electron annihilation. Events include true coincidence events, random events and scatter events (Figure 2.5). A true event is when the two photons are effectively detected from the annihilation occurring along the LOR. A scatter event is when two photons are detected within the same time window but one or two of the photons are scattered after the annihilation, due to their interaction with the matter. The actual event is therefore not aligned with the detected LOR. Finally a random event is when two unrelated photons, emitted from the two different annihilations are detected by the detectors within the same time window. Random and scatter events increase the background noise

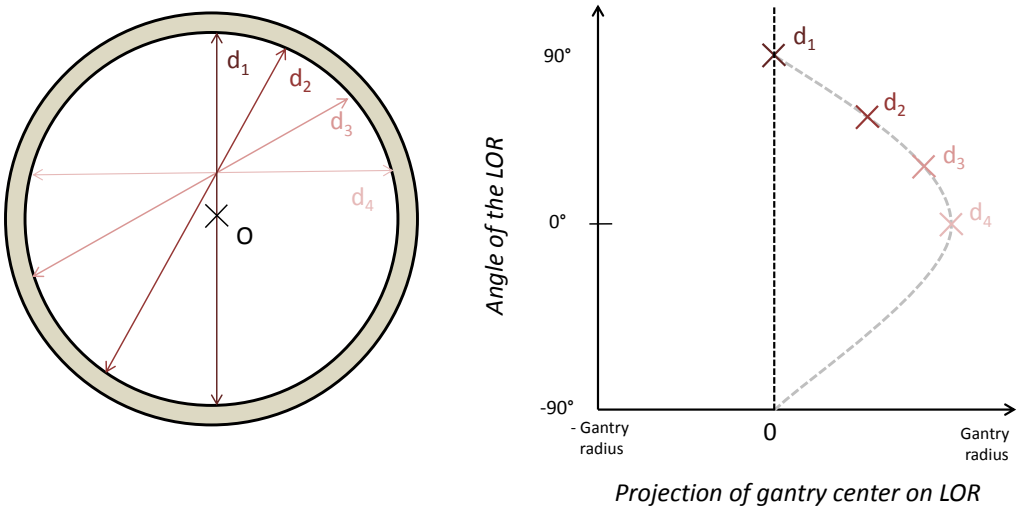


Figure 2.3: Sinogram formation. Each line of response (LOR), is characterized by its angle with the horizontal line of the gantry, and the minimum distance from the center of the gantry. The coordinates (distance,angle) of each LOR is placed into an array called the sinogram. For a high emission range, the actual positron emission is located at a higher distance from the positron-electron annihilation.

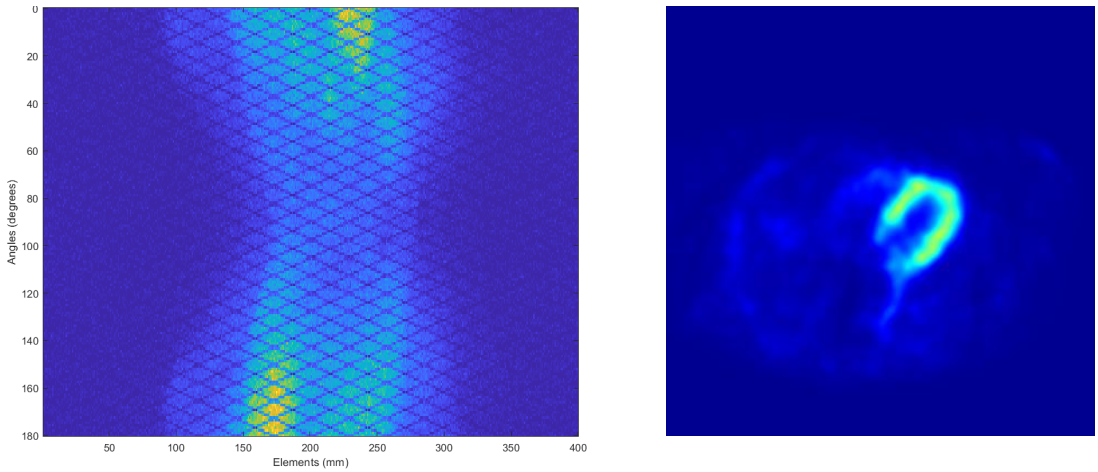


Figure 2.4: Heart sinogram (a) and reconstructed image (b). The same slice of the image is shown. The raw sinogram is displayed, as obtained from the scanner, without any corrections applied. For example, the diagonal lines are due to the gaps between the detectors, which explains the absence of recorded signal. The signal coming off the myocardium is seen in the first lines of the sinogram (0 to 10 degrees, 200 to 250mm) and the last lines (160 to 180 degrees, 150 to 200mm). The reconstructed image shows the myocardium tissue (see section 2.2.2.2.) A different windowing level is used for each image.

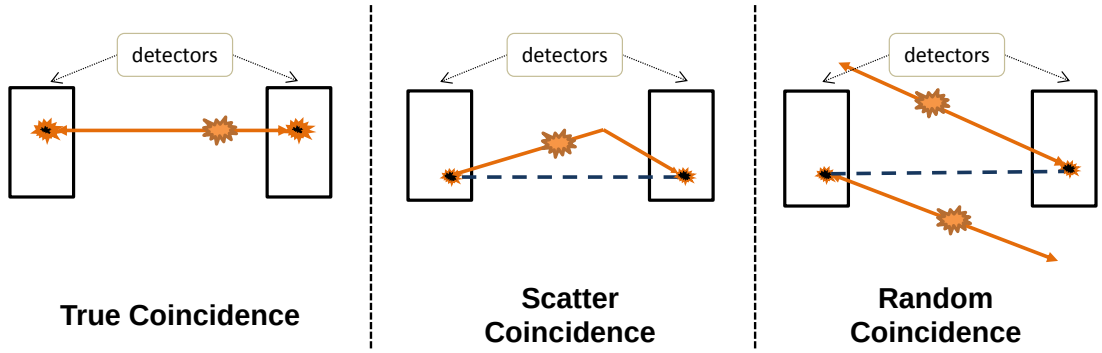


Figure 2.5: True, scatter and random coincidence events. In the case of scatter events, gamma rays interact with the matter, which means that the location of the positron-electron annihilation is not aligned with the line of response (LOR) of the detectors. It is also possible that two unrelated photons are detected by the detectors within the same time window while they are coming from two different annihilations (random events).

of the image, which results into a loss of image contrast. In order to correct for scatter events, an energy window can be applied: since scatter can result in a loss of energy, it is desirable to keep "good" events that are centered within a narrow 511 keV energy window (e.g. between 400 and 650 keV). As for random events, their proportion can be reduced by increasing the number of detectors rings (field of view) or decreasing the time coincidence window.

Finally, with recent progress in electronics, it is now possible to register the difference of the time arrival within the same coincidence window. This technique is called time of flight and allows to localize the location of the annihilation event along the LOR. Time resolution of state of the art clinical whole body systems are as short as 214ps, which corresponds to a spatial resolution of 3.2cm.

2.1.3.2 Attenuation correction

Attenuation correction is an important step in the reconstruction process, in order to obtain images of the body as accurate as possible. Photons that encounter denser material or travel further are more likely to be absorbed or scattered than the ones that pass through less dense material or travel a shorter distance. If correction for the associated attenuation is not taken into account, artifacts will occur. For example, in a non-attenuation corrected image, the activity in organs with high

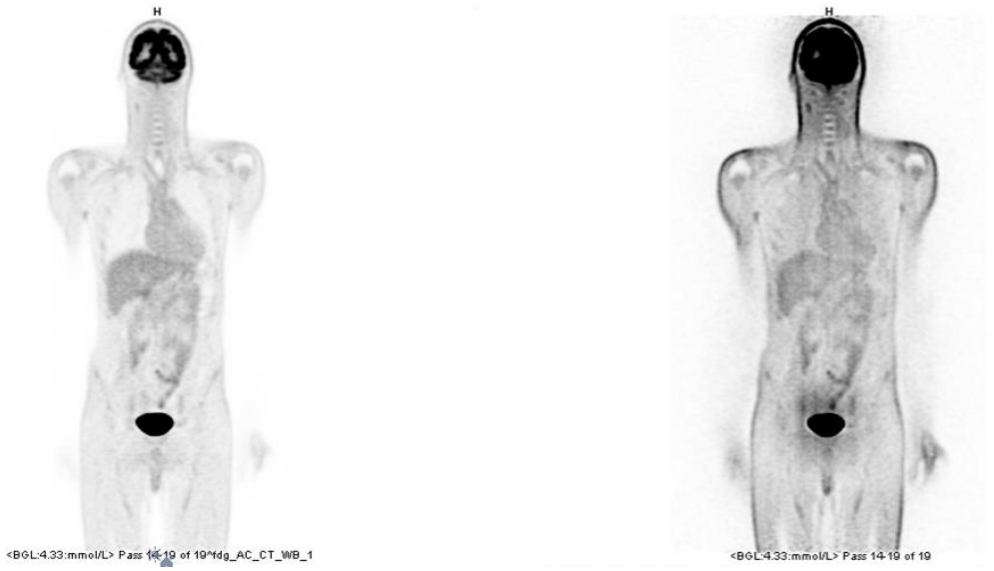


Figure 2.6: Comparison of attenuation corrected (left) and non-attenuation corrected (right) images. If the image is reconstructed without attenuation correction, some body regions with high tissue density (e.g. muscle) appear with less activity than there should be comparatively to body regions with lower tissue density (e.g. lungs). In non-attenuation corrected images, activity on the skin seems higher, as these photons travel only through air. In these images, darker regions represent higher tracer uptake, using an inverse-grey color table. Same windowing level has been used (0-5% of maximum pixel intensity).

tissue density (bone, muscle) appear reduced compared to activity in regions with lower tissue density such as the lungs since photons emitted from the lungs are less attenuated. Similarly, activity on the skin seems higher since photons travel only through air. An example of a PET scan reconstructed without and with attenuation correction can be seen in Figure 2.6.

In the simple case, where only one type of material is encountered, attenuation is corrected by calculating the probability P of a coincidence:

$$P = \exp(-\mu a)\exp(-\mu b) = \exp(-\mu(a+b)) = \exp(-\mu D) \tag{2.1}$$

Where: D is the total thickness of the body, μ is the linear attenuation coefficient in the body, and a and b are the tissue thicknesses respectively traversed by the two 511-keV photons along the LOR, as shown in Figure 2.7.

The attenuation correction factor P corrects the artifacts previously described by reducing the intensity of LOR passing through less dense tissues (small μ) or having a shorter distance to travel (small D).

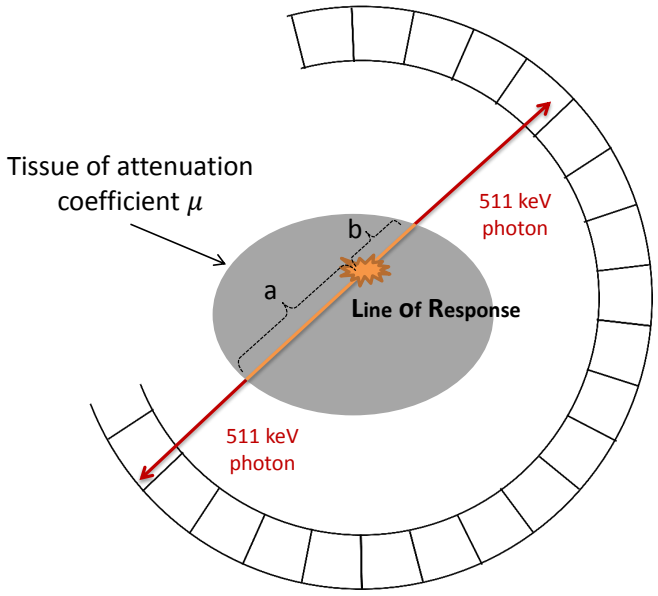


Figure 2.7: Schematic of attenuation correction: the photon's trajectory are attenuated by the material they are traversing.

In the more general case, since the human body has tissues with various densities, the correction for multiple tissue densities is as follows:

$$P = \exp(-\sum_{i}(\mu_i D_i)) \tag{2.2}$$

Where: μ_i is the attenuation coefficient in the tissue i, and D_i is the thickness of the tissue i on the line of response.

In order to correct for attenuation, a map with the spatial distribution of the coefficients μ_i for each tissue type is needed. There are various ways to estimate an attenuation map of the body, but the most commonly used in practice is to rely on information coming from an anatomical image. The most recent PET scanners are hybrid scanners (PET-CT or PET-MR) and the anatomical image of the body can be acquired, after or before the PET signal acquisition (even simultaneously in the case of PET-MR). A simple piece-wise linear function is used to map the intensity values of the anatomical scan to the attenuation map. A major assumption of estimating the attenuation map from another modality is that both images (e.g. CT and PET) are perfectly aligned. Artifacts can occur when there

is a mis-registration between the PET and CT where the attenuation correction of the signal can be mis-estimated or simply wrong.

2.1.3.3 Image reconstruction

The signal recorded into the sinogram is only the projection of the object onto the scanner detectors. To obtain a 3D image of the object scanned, a process of reconstruction is required. This problem can be formulated mathematically as an inverse problem where the objective is to obtain the definition of a three dimensional system based on its projection. A vast field of research called *"tomographic image reconstruction"* [33] tries to tackle this problem under various conditions and hypotheses. For the sake of simplicity we will discuss only two main algorithms: the filtered back-projection (FBP) and the ordered-subsets expectation maximization (OSEM).

The FBP reconstruction method relies on the mathematical theorem called projection-slice theorem. For the sake of simplicity, let us define the object to reconstruct as two dimensional object. The theorem can be generalized to higher dimensions. We represent the object by a mathematical function $f(x, y)$. The projection-slice theorem in two dimensions states that the following two operations are equal:

- Project the object $f(x, y)$ into one dimensional line, and apply a Fourier transform on that one dimensional projection.
- Apply a two dimensional Fourier transform on $f(x, y)$. Sample the line that goes through the center of the Fourier transform of the image, and parallel to the projection line.

Mathematically, let us define the operations described above. Let F_1 and F_2 be the one and two dimensional Fourier operator. P is the one dimensional projection operator and S is the slice operator of a two dimensional image, passing through its center. We have then:

$$F_1 P = S F_2 \tag{2.3}$$

Theoretically, the image $f(x, y)$ can be reconstructed by calculating the inverse Fourier transform of $F_2(f(x, y))$, defined with an infinity of projection lines. Unfortunately, there is only a finite number of projections in practice. As a consequence, a finite number of projections implies that the Fourier domain is not sampled uniformly. Backprojections from many directions results in higher line density near the origin and lower density away from the origin. Consequently low density are given too much weight compared to the high frequencies, which tend to create blurry images. The FBP algorithm, applies a filter on the one dimensional Fourier transform of the projection, in order to minimize the low frequencies and increase the high frequencies magnitude. The interested reader can refer to the projection-slice theorem and filtered back-projection algorithm [34, 35]. In the FBP algorithm, the noise of the PET signal is not modeled. Consequently, images tend to be noisy, and a smoothing filter is often applied post-reconstruction, which limits the definition of small details in the image.

It is possible to formally take into account the noise model of the PET signal, but the equations become very difficult to solve directly - even numerically. A solution is to approach the reconstructed image as an iterative process, as shown in Figure 2.8. The process is as follows: a first initial guess of the image is done (usually a blank image) and is projected to the sinogram space. That projection is then compared to the actual measured projections. The ratio between the measured projections and the estimated projections is calculated, in order to adjust the initial image estimate. This operation is repeated for a defined number of times, called iterations.

The iterative process described is usually implemented with a Maximum Likelihood Expectation-Maximization (MLEM) algorithm, but is very slow to converge. To accelerate the convergence, the OSEM algorithm groups the angular projections into different groups called subsets, and the MLEM is performed on each of the subsets. In this method, the image is updated during each sub-iteration (i.e. an iteration for a given subset) and one complete iteration will be equal to the number of subsets. Practically, this method has proved to accelerate the convergence of the algorithm.

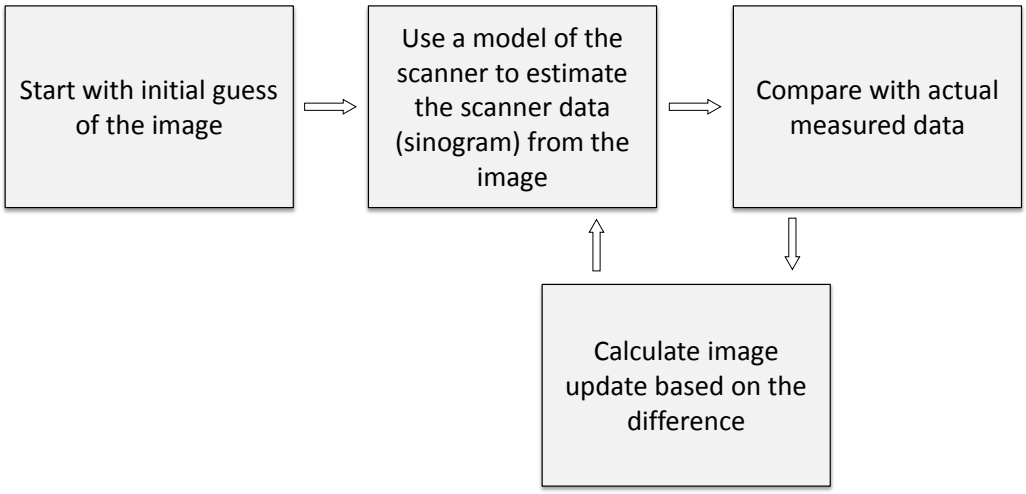


Figure 2.8: Iterative reconstruction process of a PET image. An image is initially guessed, projected and then updated from the difference with actual data.

2.1.4 Kinetic modeling

PET provides a quantitative measurement of the tracer activity (in Bq/mL) in a tissue, whether it is a Region Of Interest (ROI) or on a voxel level. Moreover if the acquisition is made at different time points over the scan, the temporal evolution of the tracer activity - called Time Activity Curve (TAC) - is measured, as shown in Figure 2.9. One parameter to calibrate when doing a dynamic PET acquisition is the duration with which each PET frame is acquired (temporal sampling). Frame duration balances the need for sufficient Signal to Noise Ratio (SNR) and required time resolution, so that variations of tracer dynamics can be detected.

The activity measured over time is directly linked with the characteristics of the tracer-tissue interactions studied (blood flow, receptor-ligands, metabolic processes) [9]. The purpose of a kinetic model is to describe in the most accurate way these characteristics. This means that a PET model tries to quantify the underlying physical, chemical and biological processes that contribute to the temporal signal measured in the tissue. The choice of the PET kinetic model depends principally on the tracer and the process we want to study. A typical class of models used in PET imaging are compartments models (Figure 2.10).

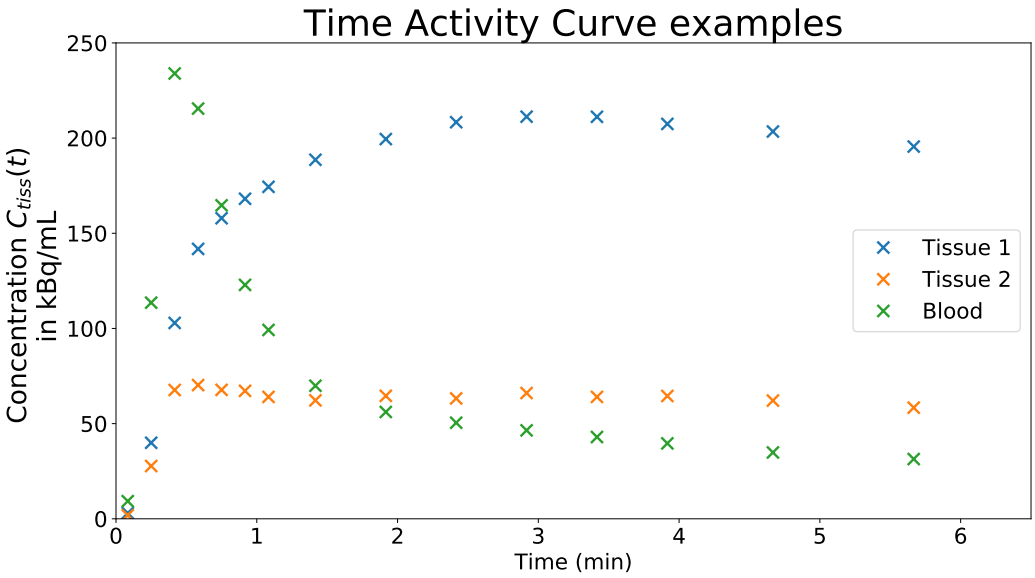


Figure 2.9: Example of typical Time Activity Curves. The concentration of the tracer in the blood and in two different tissues is plotted against the time. The aim of kinetic modeling is to find the function $f(t, K_1, ..., k_n)$ that matches as closely as possible the data points.

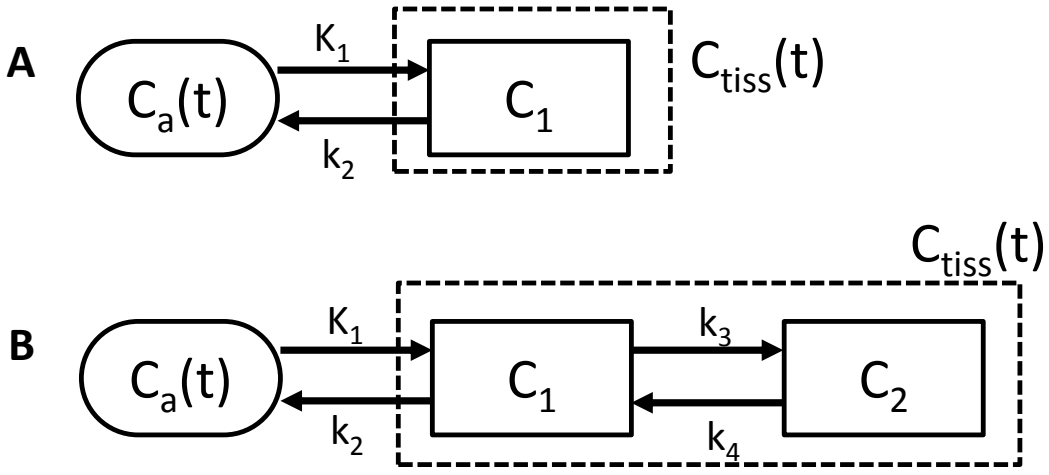


Figure 2.10: Example of compartmental models. A is a one tissue compartment model, representing a reversible tracer uptake. B is a two tissue reversible compartments model representing for example the diffusion (C_1) and the metabolism (C_2) of the tracer.

Compartmental models are the common way in PET imaging of representing the different physiological processes happening in the tissue. Each compartment models a chemical or physical state of the tracer. A model is said reversible if each compartment has bidirectional exchange with another compartment (influx and outflux). This is the case for a perfusion process, where the tracer delivered to the tissue is released back. The compartment models a irreversible process (e.g. metabolization), if there is a unidirectional exchange from one compartment to another (only one kinetic parameter). This type of model is called irreversible.

We detail here the equations for a one tissue compartment model. The principle remains the same for more complex models, i.e. with additional compartments. A one tissue compartment model is used in the case of PET perfusion studies with tracers such as ^{15}O water or ^{82}Rb, as those tracers are exchanged between the tissue and the blood, but is not processed by the tissue (metabolized). The exchange between the blood and the tissue occurs at two different transfer rates K_1 and k_2 (see Figure 2.10-A). Let $C_{tiss}(t)$ the concentration measured in the tissue and $C_a(t)$ the concentration of the tracer in the arterial blood. The following equation can be deduced:

$$\frac{dC_{tiss}}{dt} = K_1 C_a(t) - k_2 C_{tiss}(t) \tag{2.4}$$

This differential equation can be solved numerically by replacing $C_{tiss}(t)$ and $C_a(t)$ with the measured data, and calculating $\frac{dC_{tiss}}{dt}$ so that model parameters K_1 and k_2 can be inferred. The differential equation can be solved analytically [36], and $C_{tiss}(t)$ can be explicited as a function of the parameters K_1, k_2 and $C_a(t)$:

$$C_{tiss}(t) = K_1 \exp(-k_2 t) \circledast C_a(t) \tag{2.5}$$

Where: $\circledast$ is the convolution operator.

Inferring the model parameters over the analytical solution of the differential equation rather than the differential equation itself may provide a more robust and stable optimization process, especially when the number of compartments increase [37]. More details are given on model inference in the last section of this chapter.

Similarly for more complex models, multiple differential equations can be derived from the exchange between each compartment. The general solution of compartmental differential equations, has two different expressions whether the last compartment is a reversible or irreversible process. If the tracer follows a reversible model, the solution is then as follows:

$$C_{tiss}(t) = R(t) \circledast C_a(t) \qquad \text{with:} \quad R(t) = \sum_{i=1}^{n} \Phi_i \exp(-\Theta_i t) \tag{2.6}$$

Where: $\circledast$ is the convolution operator and n the number of compartments.

If the tracer follows an irreversible model, the solution is:

$$C_{tiss}(t) = R(t) \circledast C_a(t) \;\; \text{with:} \quad R(t) = \sum_{i=1}^{n-1} \Phi_i \exp(-\Theta_i t) + \Phi_n \tag{2.7}$$

The values of Φ_i and Θ_i can be then expressed as a function of the transfer parameters k_i. For example, in a one tissue compartment model, we have: $\Phi_1 = K_1$ and $\Theta_1 = k_2$ [36].

One last phenomenon to account for is the blood fraction in the measured data. Because of the spatial resolution of a PET scanner (2.5-5mm) the signal measured $C_m(t)$ is a mixture between the blood component $C_a(t)$ and the tissue component $C_{tiss}(t)$. The blood volume fraction is usually noted V_b:

$$\begin{aligned} C_m(t) &= (1 - V_b)C_{tiss}(t) + V_b C_a(t) \\ C_m(t) &= (1 - V_b)R(t) \circledast C_a(t) + V_b C_a(t) \end{aligned} \tag{2.8}$$

2.2 PET for coronary artery disease

2.2.1 Anatomy of the heart and the cardiovascular system

The heart is the organ at the center of the cardiovascular system and is composed of four chambers: the left atrium (LA), the left ventricle (LV), the right atrium (RA) and the right ventricle (RV) [38]. The heart pumps blood through two different circuits in the body. The first circuit pumps blood from the heart to the lungs

and back to heart, in order to re-oxygenate the blood. It is called the pulmonary circulation. The second circuit pumps blood from the heart to organs and muscles and back to heart, and is called the systemic circulation. De-oxygenated blood is pumped to the RA, from the inferior and superior vena cava. Blood flows through the tricuspid valve to the the right ventricle (RV). It enters then the pulmonary circuit by being pumped through the pulmonary valve, into the pulmonary artery. The pulmonary arteries irrigate the lungs, and blood is re-oxygenated by exchanging carbon dioxide with oxygen. The blood returns then back to the heart through the pulmonary veins and arrives in the LA. Blood is subsequently pumped through the mitral valve to the LV. It is finally pumped from the LV to the aorta through the aortic valve. Organs and muscle consume the oxygen supplied by the blood and produce carbon dioxide. Blood enriched with carbon dioxide returns to the right atrium of the heart. A detailed diagram of the heart is shown in Figure 2.11. The blood is pumped to the different chambers of the heart through a contraction of the heart muscle tissue, called myocardium. A contraction of the myocardium (systole) pumps the blood out of the chambers, and is followed by a relaxation (diastole), which refills the chambers with blood. There are two pairs of systoles / diastoles : the atrial systole / diastole and the ventricular systole / diastole. Each systole / diastole is synchronized, so that respective chambers (i.e. LA and RA or LV and RV) are filled and emptied with blood at the same time.

The myocardium is a muscle istelf, and therefore needs to be supplied with blood. The main artery originating from the LV (aortic root) branches out in two arteries to irrigate the heart: the left and right coronary arteries (respectively LCA and RCA). The left coronary artery bifurcates subsequently into the left anterior descending (LAD) artery and the left circumflex (LCX) artery. Each of these three different arteries artery supplies a territory of the myocardium (called coronary territory) with blood, oxygen and nutrients. The coronary arteries separate into smaller blood vessels (capillaries), which supply the myocardial tissue in oxygen. This is defined as blood perfusion.

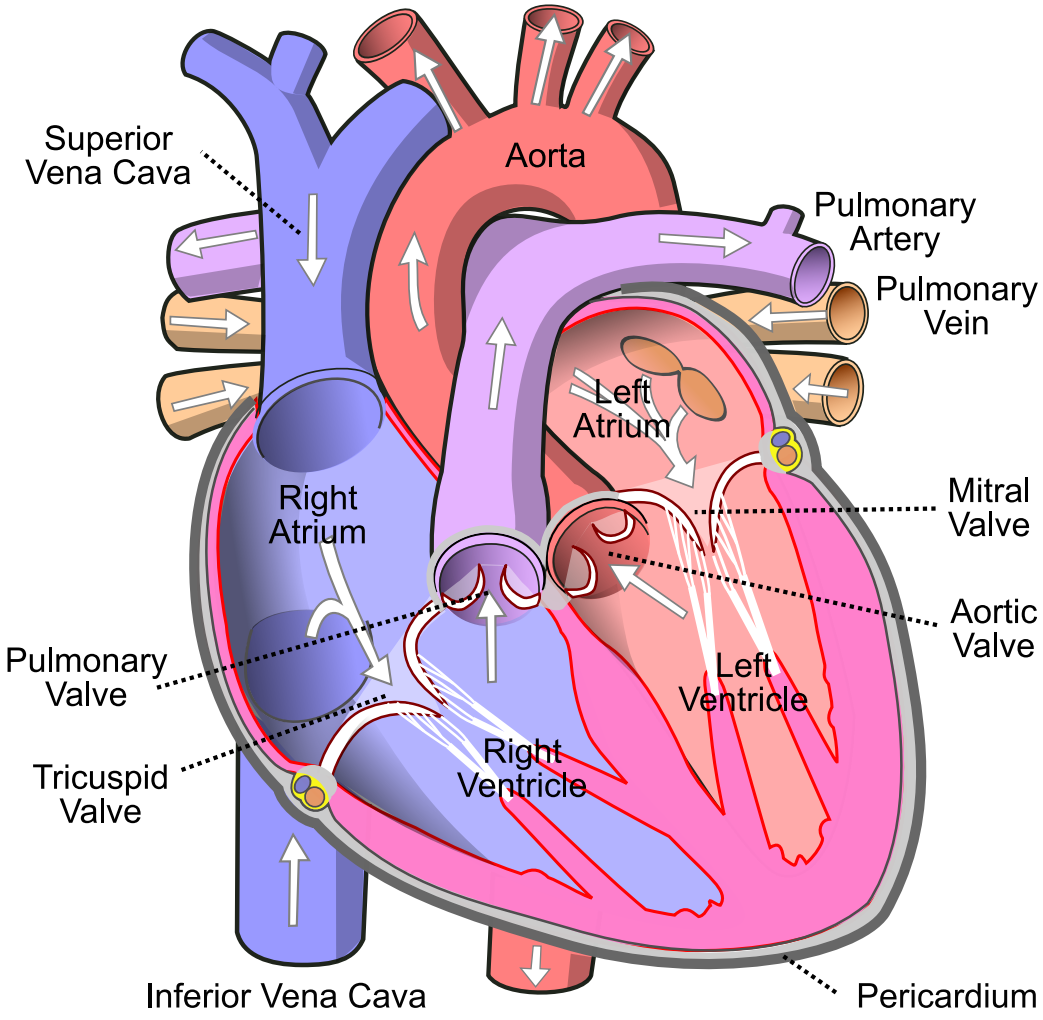


Figure 2.11: Heart diagram. The four chambers of the heart are shown: Right Atrium (RA), Right Ventricle (RV), Left Atrium (LA) and Left Ventricle (LV). The RA and RV are separated by the tricuspid valve and the LA and the LV are separated by the mitral valve. The de-oxygenated blood is pumped through the superior and inferior vena cava to the RA. It is subsequently pumped to the RV, through the tricuspid valve. From the RV, blood is pumped to the pulmonary artery through the pulmonary valve. After oxygenation in the lungs, the blood returns to the LA, through the pulmonary veins. Finally, the blood passes to the LV through the mitral valve and is pumped, through the aortic valve, to the aorta, which then separates to supply blood to different body organs and muscles. In this diagram, blue represents blood that contains mainly carbon dioxide, while red shows the blood with oxygen. Reproduced without changes from http://commons.wikimedia.org/wiki/File:Diagram_of_the_human_heart.svg under license http://creativecommons.org/licenses/by/3.0.

Coronary Artery Diseases (CAD) occur when the arteries become hardened and deposits of fatty lipids or calcium are accumulated inside the artery walls. These deposits build up to form a plaque, and narrow the diameter of the arteries. If these arteries, or the subsequent branches become partially blocked (stenosis) the blood flow is reduced, which may result in a reduction of blood perfusion. The reduction of the blood supply deprives the tissue of oxygen (ischemia), which is necessary for tissue survival and heart function. If a total blockage happens suddenly (total acute occlusion), it results in an acute myocardial infarction ("heart attack") and death if not taken care of immediately. Chronically, stenosis leads to reduced heart function, muscle damage, and eventually death of the myocardial tissue. Obesity, smoking, high blood pressure and age have been found to be major risk factors of CAD.

If the stenosis is detected early enough, there exists different procedures to improve the prognosis of the patient and restore the heart tissue. For example, if diagnosed very early, it is possible to reduce plaque build up with a healthy diet, stopping smoking and maintaining high level of physical activity. A prescription of medications can also limit or reduce the plaque formation. Finally, in case of severe CAD, clinicians may use surgery to increase the myocardial perfusion. This specialty of medicine, called interventional cardiology, may perform stenting, which consists in placing a wire mesh structure that widens and reinforces the artery. They may also perform a bypass surgery, which consists in grafting a healthy vessel from another body region, to bypass the blockage.

The role of molecular imaging is to measure the perfusion of the myocardial tissue, by assessing the distribution of the tracer in the myocardium. Cardiac PET procedures are used for [39]:

- Diagnosis of patients suspected of having CAD due to chest discomfort or other cardiac risk factors.

- Prognosis of patients with known CAD, where it is desired to know the extent of the myocardial ischemia. Cardiac PET allows clinicians to stratify patients

into subgroups with low, intermediate and high risk for cardiac events, which helps to guide medical management.

- Response to therapy, where PET provides information regarding the adequacy of revascularization.

2.2.2 Cardiac studies with nuclear medicine

2.2.2.1 Standard clinical protocol

Molecular imaging is a well established diagnostic tool to assess the cardiovascular system [20]. Whether it is a scintigraphy, a SPECT scan, a PET scan, an MR perfusion protocol or a CT perfusion scan, the clinical protocol is done usually in two acquisitions: a rest and a stress scan [20]. A rest scan consists of imaging the patient while he is at a resting state, while a stress scan is after the patient has done some physical activity. The scans are then visualized and a diagnosis is established by the cardiologist. It is necessary to perform a stress scan because the perfusion can appear normal at rest in some cases.

There are three tracers that are typically used for a MBF PET protocol: ^{82}Rb, ^{13}N ammonia and ^{15}O water [40–42]. The choice of the tracer depends on its availability, the workflow required for scanning, and physical properties such as extraction fraction. The extraction fraction is the correlation between the tracer uptake and the myocardial blood flow [43]. ^{15}O water and ^{13}N ammonia require the presence of a cyclotron on site to generate the short-lived isotopes. A cyclotron is necessary to produce most radio-isotopes, and the half lives of ^{15}O (122 sec) and ^{13}N (10 min) are too short to be produced in a different location than the clinic department performing the scan. The ^{82}Rb can be produced with a 82-Strontium (^{82}Sr) generator, as the ^{82}Rb is produced by the decay of a ^{82}Sr which has a half-life of 25 days. This property of ^{82}Rb makes this tracer more widely available. Ammonia has however a lower emission range, and a better extraction fraction, as shown in Figure 2.12. Water is mainly used in research protocols, because of its very short half-life, which makes any kind of clinical routine protocol challenging. It has however the best extraction fraction.

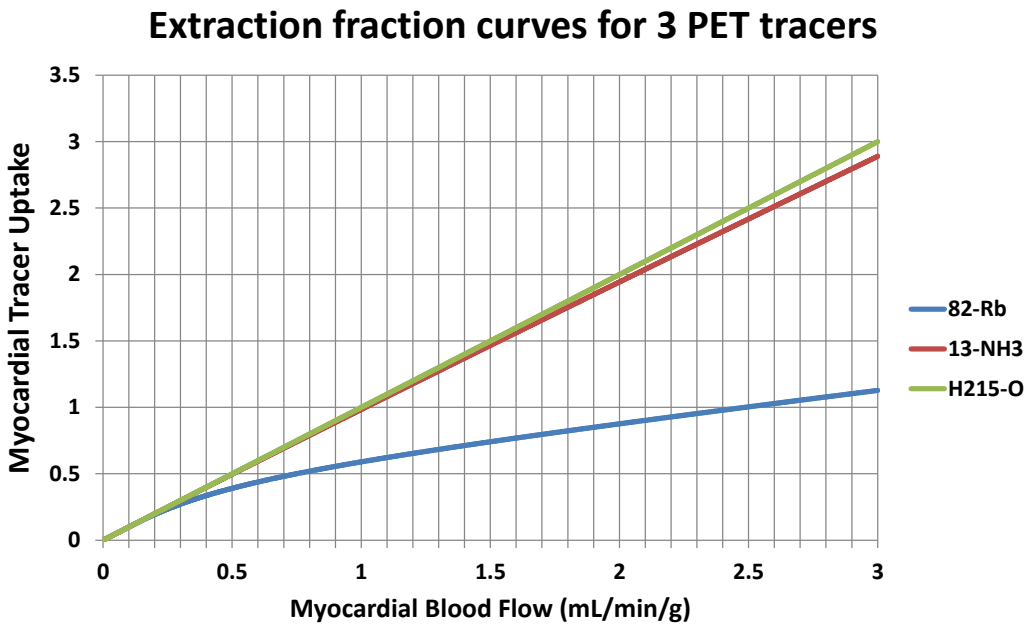


Figure 2.12: The extraction fraction of a PET tracer is the relationship between the myocardial uptake and the myocardial blood flow. It is optimal for ^{15}O water, quasi linear for ^{13}N ammonia and less linear for ^{82}Rb.

With PET protocol, since the current tracers have very short half-lives (^{82}Rb is 75 seconds and ^{13}N ammonia is 10 minutes), the stress state is usually induced by injection of a pharmacologic agent (adenosine, regadenoson...), which can be uncomfortable for the patient, and leads sometimes to substantial patient movement [44, 45].

For a PET-CT protocol, the tracer is intravenously administered when the patient is prepared on the scanner table, right before the PET acquisition. Injection is usually administered in the patient's arm vein. The CT scan can be performed before the PET scans or after. To minimize radiation exposure, usually only one CT is used for the whole procedure, hence the same attenuation map is used for both rest and stress. Each PET acquisition lasts a couple of minutes [20].

2.2.2.2 Visualisation of the heart

When a 3D scan is performed (SPECT or PET), the heart is then reoriented in three standard planes that represents the cardiac anatomy. These three planes are

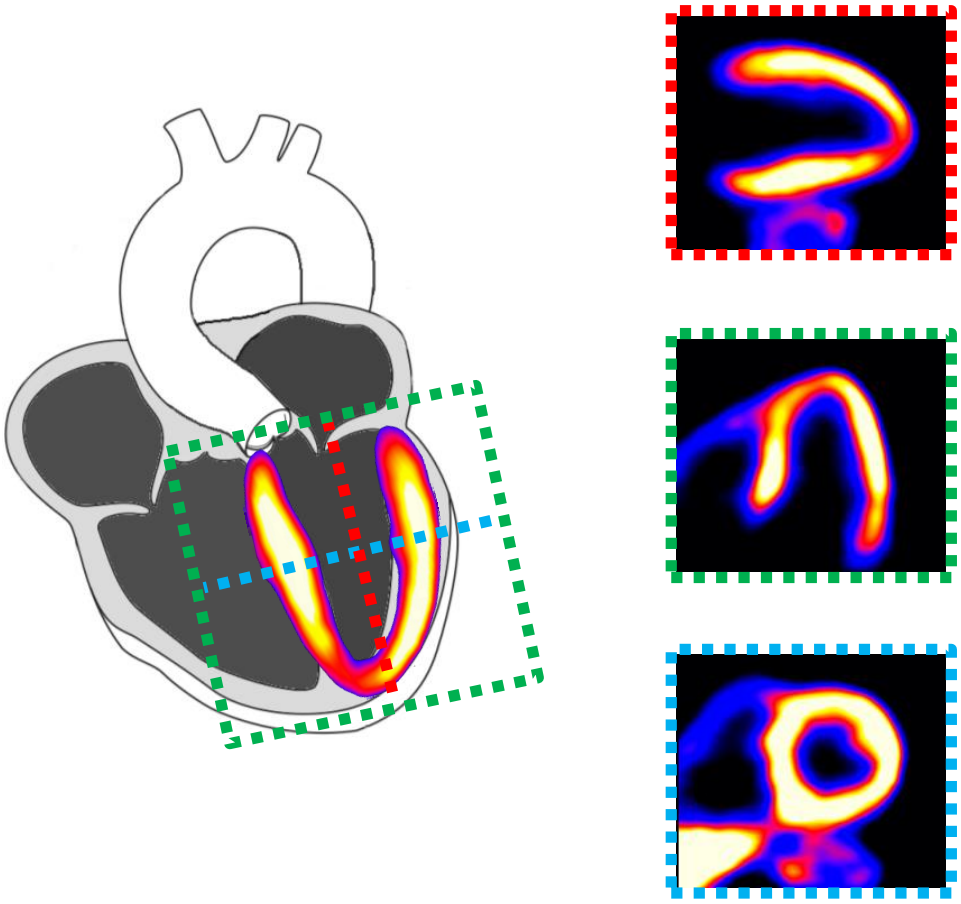


Figure 2.13: Cardiac planes: the heart is usually segmented and reoriented in the three cardiac planes. The blue line represents a plane that is a cross section of the LV. The red line represents a mid-ventricular plane perpendicular to the short axis and is called the vertical long axis. Finally the green line represents the plane orthogonal to the previous planes and is called the horizontal long axis.

respectively called the short axis, horizontal long axis and vertical long axis, as shown in Figure 2.13. In addition of the three cardiac planes, the myocardium is also segmented in 3D and converted into a 2D polar map. That polar map is used to visualize the perfusion in each coronary territory, as shown in Figure 2.14.

2.2.3 Cardiac studies with dynamic PET

Current standard protocols with scintigraphy, SPECT or PET, perform a myocardial perfusion image (MPI) acquisition. This is an established technology to help assessing coronary artery disease (CAD), and is done by acquiring and averaging the signal over a period of time. This method is semi-quantitative as the quantification of

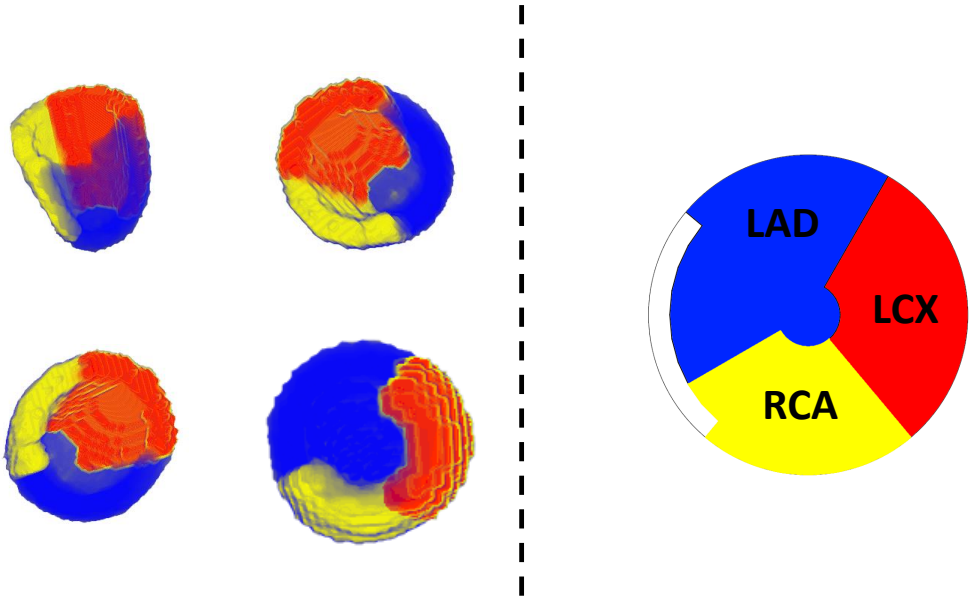


Figure 2.14: From 3D segmentation to polar map. The 3D segmentation of the myocardial tissue is transformed in a 2D polar map, so that each coronary territory can be easily visualized. Blue is the territory irrigated by the Left Anterior Descending artery (LAD). Red is the territory irrigated by the Left Circumflex artery (LCX). Yellow is the territory irrigated by the Right Coronary Artery (RCA).

the myocardial perfusion is done by normalizing intensities to the maximum intensity value in the myocardial tissue. This method can be sufficient in the case of a blood flow reduction for a single artery, where the difference of intensity uptake between normally perfused tissue and reduced flow would lead to a conclusive diagnosis. However this method could be an issue in the case of balanced blood flow reduction (triple vessel disease), where the difference in intensities between each territory is not high. PET imaging offers the possibility to have absolute quantification of the tracer activity in the tissue in Bq/mL. Kinetic modeling (see section 2.1.4) of the tracer in the myocardial tissue can be applied in cardiac dynamic PET studies, in order to compute MBF [40]. MBF is defined in unit of volume per unit of time and per unit of mass of tissue (usually mL/min/g). Additionally the Myocardial Flow

Reserve (MFR) can be calculated, by computing the stress to rest MBF ratio.

Recent studies have shown that quantitative estimates of MBF and MFR using dynamic PET offer a superior diagnosis compared to semi-quantitative MPI [11, 13, 17, 46], especially in the case of patients with triple vessel disease. Studies have shown for example identification of triple vessel disease was significantly higher with MBF and MFR than with semi-quantitaive MPI (92% vs 46%) [47]. Another study with 120 patients [17] showed that MFR had a diagnostic sensitivity of 88% for triple vessel disease, while 60% of the patients were considered at risk from different cardiovascular tests (electrocardiogram, ejection fraction, ...), demonstrating the clinical benefit of MFR. Furthermore, quantitative measurements help tailor a treatment to each individual, by providing a better risk stratification. A study with 704 consecutive patients showed correlation between patients with abnormal MFR (<2.0) and a significantly higher cardiac event rate after one year of follow-up [48].

2.2.3.1 Tracer dynamics and processing

The tracer is intravenously administered, and arrives first in the right ventricle. It is then visible in the left ventricle. After that phase, the tracer progressively gets delivered to the myocardial tissue. A couple of minutes after the injection, the tracer uptake is only in the myocardial tissue.

In order to correctly model myocardial blood flow, an appropriate time sampling of the dynamic series framing is important [44, 49]. Because the transition between right and left ventricle happens relatively quickly (a few seconds as shown in Figure 2.15), it is necessary to sample frames with short duration. For the later part of the signal, where the tracer is perfused in the tissue, frames can have a longer duration in order to maximize signal to noise ratio.

Computation of MBF requires kinetic modeling of Time Activity Curves (TACs) of the myocardial tissue. The kinetic parameters are eventually extracted in order to compute myocardial blood flow (see next section 2.2.3.2). TACs of the tissue are sampled from the 3D segmentation of the myocardium (late phase - see Figure 2.15). As previously highlighted, kinetic modeling requires an aterial input function

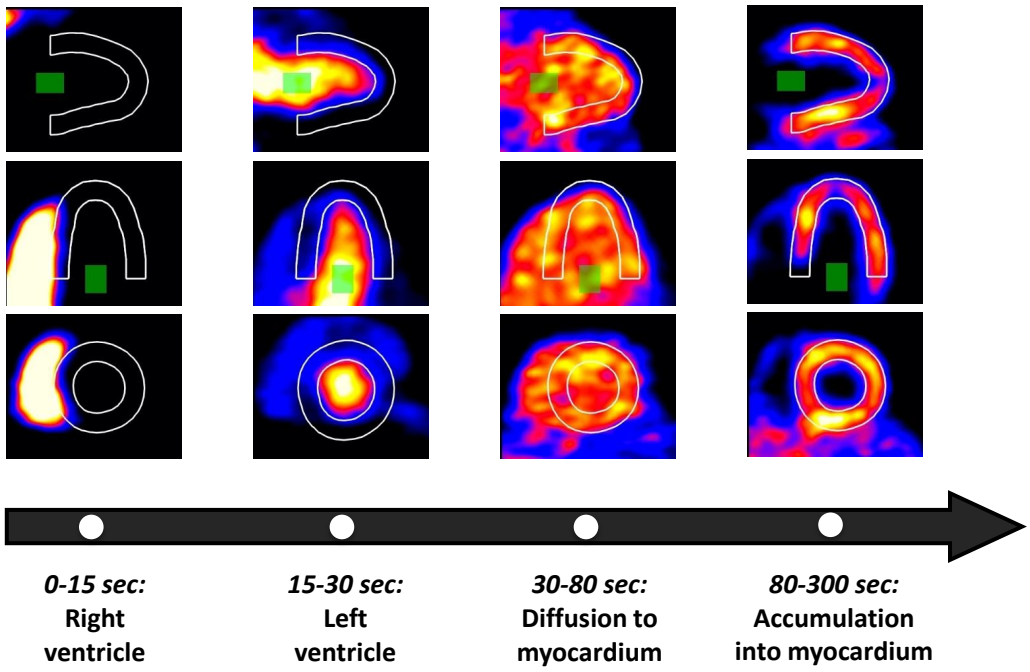


Figure 2.15: Tracer pattern in cardiac dynamic PET. The tracer is is intravenously administered in the patient, hence it is visible first in the right ventricle (RV). Typically 8 to 12 seconds later (varies on the patient), it returns from pulmonary circulation to the left ventricle. The blood, and hence the tracer, is pumped by the heart to the different organs, including the heart itself. The tracer is chosen so that it accumulates especially in the myocardial tissue. The green rectangle represents the region of interest (ROI) from which the input function $C_a(t)$ is defined. The white segmentation shows the myocardial tissue. Time Activity Curves are sampled from this segmentation, in order to calculate Myocardial Blood Flow.

(section 2.1.4), which samples the concentration of the tracer in the arterial blood. While there exists various methods to obtain the input function in general (direct arterial sampling, population based input functions) [50, 51], it is an established way to define the input function within the context of MBF processing: a region of interest is placed in the LV cavity, which can be defined with image processing methods (the LV is inside the segmentation of the myocardium, see Figure 2.15).

2.2.3.2 MBF modeling with dynamic 82-Rubidium

The choice of the tracer will influence modeling and the number of compartments used. For ^{15}O water, a one tissue compartment model is used, while for ^{13}N ammonia a two tissue irreversible model is used [9, 50].

The clinical datasets that were available for this research were acquired with a ^{82}Rb tracer, we therefore expand the details of MBF computation for this tracer. In ^{82}Rb PET protocols, the tracer is in the form of a rubidium chloride (RbCl) solution. ^{82}Rb+ has a very similar biological activity to a potassium ion (K+). Hence once in the myocardium, the rubidium ion is accumulated following the cellular Na-K exchange pathway of the myocardium cells. The more active the cells, the more tracer uptake is visible on the image.The physiological activity of the ^{82}Rb tracer follows theoretically a two tissue compartments model, as the tracer is found between the myocardial cells (extra-cellular or interstitial space) and inside the myocardial cells (intracellular space) [52, 53]. However studies have shown that due to the poor spatial and temporal resolution of ^{82}Rb dynamic PET scans, it is not possible to infer reliably the different model parameters, without strong a priori knowledge on some model parameters [53, 54]. Consequently, it was shown that a reduced one tissue compartment model, with an arterial blood fraction correction (see Figure 2.16) was suitable to model myocardial blood flow in a clinical setting [53]. While this model is a simplification of a more complex physiology, it is commonly accepted as it showed good repeatability [55] and in good concordance with other tracers, such as ^{15}O water [41]. The generic equation describing the measured concentration of the tracer $C_m(t)$ in the myocardial tissue is as follows (see section 2.1.4) :

$$C_m(t) = (1 - V_b)\Big(K_1 \exp(-k_2 t)\Big) \circledast C_a(t) + V_b C_a(t) \tag{2.9}$$

Where K_1 and k_2 are the kinetic parameters representing the exchange between the arterial blood and the tissue, and V_b is the fraction of intra-capillary blood. Values of V_b are typically found between 0.05 and 0.15 [56]. From this equation, it can be observed that the quality of the input function, will be of primary importance to estimate the kinetic parameters. One factor particularly challenging to control with ^{82}Rb is the duration of the bolus. A bolus as narrow as possible is expected, since the convolution operation acts a low pass filter. However, the age of the ^{82}Sr generator, from which ^{82}Rb is produced, will affect the duration of the bolus. As the generator ages, the quantity of ^{82}Rb to inject increases, in

order to keep the activity administered constant, ultimately decreasing the quality of the kinetic parameter estimates.

In the clinical protocol used for this project, the signal was reconstructed over multiple cardiac cycles and respiratory cycles (see Appendix A for additional details). Each frame was acquired for at least 5 seconds, and no gating procedure was used during reconstruction to correct for these effects. Consequently, the signal measured in the myocardial tissue, $C_m(t)$, is largely contaminated by the signal extracted in the blood pool, $C_a(t)$. This effect called spillover, is modelled as follows:

$$C_m(t) = (1 - S)\Big(K_1 \exp(-k_2 t)\Big) \circledast C_a(t) + SC_a(t) \tag{2.10}$$

where S is the spillover factor. Typical spillover factor found in the literature suggest values between 0.35 and 0.65 for ^{82}Rb tracer [57]. The spillover has a large impact on the quality of the MBF estimates. The relationship between the spillover and the MBF estimates is studied in chapter 6 of this thesis.

In this model, the time delay between the arterial blood and the tissue input function is set to be 0, and not inferred by a model parameter. The primary reason for that choice is due to the temporal resolution of the dynamic PET scan. With initial frames lasting only 5 seconds, while the time delay is in order of 1 to 2 cardiac cycles (i.e. 1 to 2 seconds) [58], making the time delay inference challenging [59]. The error resulting from this assumption is therefore unknown.

The extraction fraction of the ^{82}Rb tracer needs also to be accounted for, as shown in Figure 2.12. The relationship between MBF and K_1 is defined by the Renkin-Crone equation [57]:

$$K_1 = \mathrm{MBF}\Big(1 - A \exp(\frac{-B}{\mathrm{MBF}})\Big) \tag{2.11}$$

where: A defined between 0 and 1 and B positive, take different values, depending on the model chosen. A common extraction fraction correction for ^{82}Rb is the Lortie model where $A = 0.77$ and $B = 0.63$.

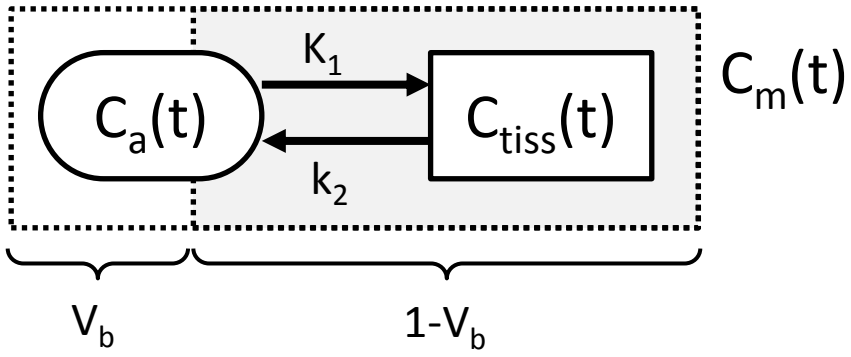


Figure 2.16: Representation of a one-tissue compartment model with blood volume fraction and reversible tracer uptake. The MBF is computed from the kinetic parameter K_1. V_b is the fraction of volume occupied by arterial blood within the tissue.

2.2.4 Impact of motion in cardiac dynamic PET

MBF and MFR computations, while providing superior diagnosis than standard MPI, are not widely adopted in the clinic, because of lack of standardization [20]. It has been reported in the literature that MBF estimates can be affected by many factors, such as the choice of reconstruction method, temporal sampling strategy, post-processing methods and patient motion [60–64].

The impact of reconstruction and acquisition parameters can be minimized with standardization and quality control [20], but patient motion cannot be avoided and is one of the most common issues with dynamic PET in clinical applications, despite careful patient preparation and coaching during the scan.

The major assumption of MBF modeling, and of kinetic modeling in general with dynamic imaging, is that the region of interest (ROI) or voxel reflects the same object or tissue over time. With motion, the region studied might not reflect the same tissue over different image frames, and consequently the final kinetic parameters are less reliable.

With respect to MBF in dynamic PET, clinical studies, with 90 [44], 225 [65] and 236 patients [66], showed that moderate or high motion can appear relatively frequently, for 30% to 40% of the scans, and are predominant in stress scans [65]. In addition to frequent occurrence, motion also has a significant impact on MBF estimates. A recent study determined, through computer simulations, that severe

patient motion (up to 2cm) could lead to 500% error in MBF measurement [67]. The repercussion of motion on clinical data has also been evaluated, demonstrating that motion can be source of error in MBF estimation [65, 66, 68]. These studies showed that motion in clinical data tend to lead to overestimated MBF values, especially in the RCA territory. This overestimation may result in false positives assessment, since higher MBF values could mislead the clinician in his diagnosis.

Additionally, motion does not only induce errors within the dynamic PET image but also between the PET series and the CT [69]. A good PET-CT alignment is important for attenuation correction purposes (see section 2.1.3.2) and can be another source of error in the MBF estimation.

2.3 Motion correction in dynamic PET

The term motion defines the physical displacement of an object. In the context of cardiac imaging, motion can describe several types of motion: the motion of the myocardium due to the heart beats (cardiac motion), the periodic chest motion due to the breathing of the patient (respiratory motion) or the patient gross motion (patient motion). Patient motion affects a dynamic PET series in two ways: it creates a misalignment of the regions of interest (ROIs) over time, which leads to inaccurate Time Activity Curves defined from these ROIs. Gross motion also creates intra-frame blurring effect: since the motion does not happen at a specific time point between two consecutive frames, the signal acquired over the duration of the frame is corrupted by the patient motion, which leads to blurred images. In this work, we specifically attempt to address the issues linked to patient motion in dynamic PET imaging, that is due to a misalignment of the PET frames.

Patient motion is a major issue for using dynamic cardiac PET in clinical settings. It happens frequently in stress scans [44] and has a significant impact on the MBF values [67]. It is therefore necessary to perform motion correction of the dynamic frames. It is possible to perform manual motion correction [65], by registering the reconstructed images but this process is time consuming and user dependent, which is why automated motion correction is preferred [65]. The process

of motion correction in dynamic medical imaging is made of two principal steps: the detection of the motion and its correction. While often treated together, it is easier to consider those steps separately, to understand the different challenges encountered when doing motion correction.

2.3.1 Motion detection

Detection of motion can be separated in two categories, based on how the motion is detected: extrinsic and intrinsic methods.

Extrinsic methods measure the motion during the scan with the use of an external device. Some methods require markers to be placed on the subject; which allows tracking the patient motion in three dimensions during an acquisition [70–72]. Other devices can be used such as a chest-belt to measure respiratory motion [73, 74]. The signal acquired from the belt is then combined with the signal from the scanner in order to take into account the motion in the reconstruction of the image. The use of external devices could be inconvenient and uncomfortable for the patients, and it may be difficult to rigidly fasten sensors, which in some situations produce inaccurate motion estimates due to slipping or sliding. Methods using stereo cameras without any markers exist [75, 76], but they require a calibration before each scan, which is unreasonable in a clinical context, where tens of patients are scanned each day. While those methods can be applied relatively easily for motion correction of brain studies [77, 78], as the movement of the brain follows the rigid movement of the skull, it would be more challenging to apply them in cardiac imaging. The movement of the chest would need to be modeled in order to translate chest movement to actual heart movement, as in [70]. Another type of extrinsic method is to rely on MRI images that have been acquired simultaneously with a hybrid PET-MR scanner [79–81]. The advantage of a hybrid PET-MR is that multiple MRI acquisitions can be made at different times so that the motion can be tracked easily. The problem is then shifted to a PET-MR registration or MR-MR registration, which tend to be easier since MRI has a greater spatial resolution, a better Signal to Noise Ratio

(SNR) and structural information to rely on. This applies however only to PET-MR hybrid systems, which are not widely available at the moment.

The major problem of all the extrinsic methods is their dependence on external devices. Intrinsic methods do not have that issue. In intrinsic methods, multiple frames are reconstructed at different time points and image registration algorithms are used to identify motion from reconstructed images [82, 83]. Image registration algorithms consist in defining images as mathematical objects, and define a function that calculate how "similar" these two objects are. The algorithm brings the moving image (source image) into alignment with the target image (see next section 2.3.2), by calculating what transformation settings (e.g. translation or rotation) maximizes the similarity function. A more formal definition of image registration algorithms is done in the next section (section 2.3.2). The drawback of these methods is their dependence on the sampling of the dynamic series. It is therefore unable to identify when the motion occurs, hence cannot correct the motion within the frame (intra-frame motion). An additional consideration must be taken when doing image based motion correction: the choice of the reference frame for the dynamic series. One can either choose to register each frame of the sequence to the previous frame [84]. Frames can also be registered to a fix image, whether it is a specific frame of the dynamic sequence [85] or a synthetic image [86].

2.3.2 Image registration

2.3.2.1 Definition

In order to define image registration formally we shall start by defining an image as a mathematical object. An image I is defined on a domain Ω, which is often a Cartesian grid in two or three dimensions. For each point x of that domain (coordinates on the grid), $I(x)$ is the intensity of the pixel (or voxel for 3D images) at that position x. Let A and B be two images defined respectively on the Cartesian grids Ω_A and Ω_B. We can define as a registration function $\mathcal{R}$, a function that maps every position from the image A to the image B. However because images are discrete objects, the source domain Ω_A needs also to be resampled to the target

domain Ω_B through an interpolation process (see Figure 2.17). The most common types of interpolation methods are [87]:

- *Nearest neighbour:* selects the value of the nearest point and does not consider the values of other neighbouring points at all.
- *Linear:* The algorithm computes a value based on linear interpolation of neighbouring point.
- *Cubic:* Instead of a line or surface, a cubic polynomial is used to interpolate between neighbouring points. This method is more accurate but more expensive computationally.
- *Spline:* A spline function is a piecewise polynomial function. Spline interpolation functions have better properties in the Fourier analysis compared to cubic interpolations, and reduce the loss of spatial resolution, but tend to be more expensive computationally.

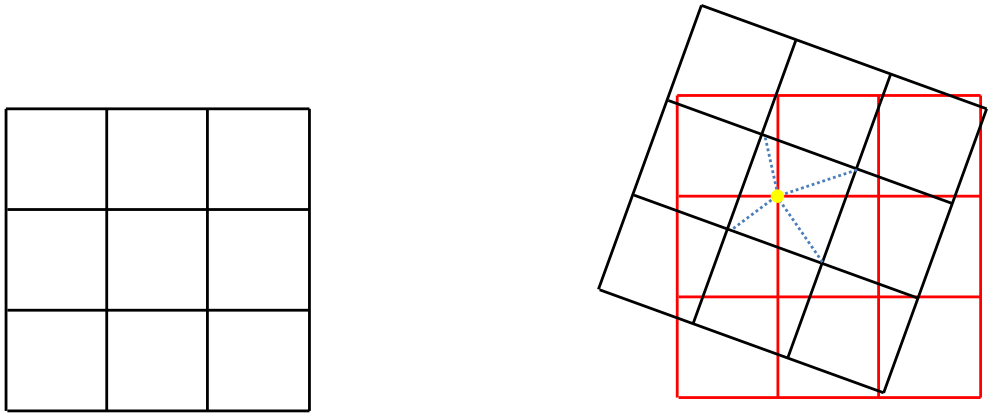


Figure 2.17: Image interpolation: because images are discrete objects, when they are transformed by a registration function, the different pixels (voxels) need to be interpolated to the target image grid.

Finally, a registration function $\mathcal{R}$ is a function that maps both position and intensity of the image A (source), to the image B (target):

$$\begin{aligned} \mathcal{R} : (\Omega_A, A) &\longrightarrow (\Omega_B, B) \\ (x_A, A(x_A)) &\longmapsto (x_B, B(x_B)) \end{aligned} \tag{2.12}$$

$\mathcal{R}$ can take different forms depending on the nature of the transformation that is desired.

2.3.2.2 Nature of registration

Registrations can be divided in two different groups: affine registrations and elastic or deformable registrations [88]. An affine registration is a registration that is composed of the following 4 components: translation, rotation, shearing and scaling. If only translation and rotation components are kept, the registration is called rigid, since it preserves the shape of the object. For 3D images, the affine registration can be represented by a 4x4 matrix :

$$M = \begin{pmatrix} m_{11} & m_{12} & m_{13} & t_1 \\ m_{21} & m_{22} & m_{23} & t_2 \\ m_{31} & m_{32} & m_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.13}$$

$\begin{pmatrix} t_1 & t_2 & t_3 \end{pmatrix}^T$ vector represents the translation part of the transformation. For a rotation, the 3x3 matrix $\left(m_{11}...m_{33}\right)$ has to be a rotation matrix. If this matrix is a 3x3 diagonal matrix it represents a scaling transformation. Skewing factors are determined by the values above and below the diagonal.

Deformable registration models are able to model local warping and displacement between the images. Those models are usually based on physical models (e.g. viscous fluid flow model or diffusion model) or interpolation theory (e.g. radial basis functions) These transformations are more powerful because they can model a large variety of deformations, but require more parameters to control. This could be challenging in a numerical optimization framework, as the robustness of these frameworks is usually better with smaller number of parameters.

2.3.2.3 Similarity Metric

For two images to be registered, the nature of the registration is not the only choice to make. It is also necessary to define a similarity metric between the two images [82, 83]. Let Sim be a similarity metric:

$$\begin{aligned} Sim : (\Omega_A, \Omega_B) &\longrightarrow \mathbb{R} \\ (A, B) &\longmapsto Sim(A, B) \end{aligned} \tag{2.14}$$

The registration process consists in finding the optimal parameters of the registration function $\mathcal{R}$, that maximizes (or minimizes depending on the definition of the similarity metric) the similarity metric:

$$\mathcal{R}_{opt} = \max_{s} \Big(Sim(\mathcal{R}(A, s), B) \Big) \tag{2.15}$$

Where: s is the parameter of the registration $\mathcal{R}$ function, e.g. the matrix components in the case of an affine transform (see equation 2.13).

Two main classes of similarity metrics can be found in the literature: landmark based and voxel intensity based metrics. In the former, key landmarks in both images are defined, usually through a segmentation algorithm, and are matched. A common algorithm landmark based match algorithm is the Iterative Closest Point algorithm (ICP) [89]. The alternative is to base the metric on matching pixel (voxel) intensities in the two images (volumes). Three main common intensity based metrics are: Sum of Squared Differences (SSD), Mutual Information (MI) and Cross-Correlation (CC). Let $\Omega = \Omega_A \cap \Omega_B$.

Sum of Squared Differences:

$$SSD(A, B) = \sum_{x \in \Omega} (A(x) - B(x))^2 \tag{2.16}$$

This metric assumes that for two images that have a similar intensity range, the smallest value of SSD will arise when images are correctly aligned. This metric is however sensitive to multiplicative factor between images.

Cross-Correlation:

$$CC(A,B) = \frac{\sum\limits_{x\in\Omega}(A(x)-\overline{A})(B(x)-\overline{B})}{\sqrt{\sum\limits_{x\in\Omega}(A(x)-\overline{A})^2}\sqrt{\sum\limits_{x\in\Omega}(B(x)-\overline{B})^2}} \tag{2.17}$$

Where: $\overline{A}$ is the mean value of the image A.

The metric is robust to multiplicative factors between the images. It is a good measure of alignment in the case of images observing similar patterns but intensity pixels are linearly correlated. For example it could be a suitable metric for a motion correction algorithm in dynamic PET imaging, since images at different time points may have similar patterns but different pixel intensities.

Mutual Information:

$$MI(A,B) = \sum_{(x,y)\in\Omega} p(A(x),B(y))\log\left(\frac{p(A(x),B(y))}{p(A(x))p(B(y))}\right) \tag{2.18}$$

Where: $p(A(x))$ (respectively $p(B(x))$) is the probability density of the voxel intensity $A(x)$ in A (respectively $B(y)$ in B). $p(A(x))$ (or $p(B(y))$) is calculated from the histogram of the image A. $p(A(x),B(y))$ is the joint probability density of the couple $A(x),B(y)$ in the joint-histogram of A and B.

Mutual information is an information theory measure of the statistical dependence between two random variables. It can be qualitatively considered as a measure of how well one image explains the other. MI is robust to multiplicative factors between images as well as non-linear dependencies. It is therefore suitable for registration of the same subject acquired with different modalities, where dependencies between tissues might be more complex than a linear correlation, like in PET. [90].

2.3.3 Motion correction for PET imaging

Once motion has been detected, the choice can be made to correct for it (the other option could also be to simply ignore it if it is considered not significant). This can be performed either before or after the reconstruction [71].

In the ideal case, motion correction needs to be applied before or during the reconstruction process, so that no artifacts is due to the reconstruction algorithm. The raw PET signal, that is the signal recorded in sinogram space, needs to be corrected so that the artifacts that have possibly occurred during the reconstruction are removed. This is especially true for periodic motion such as respiratory or cardiac motion. After modifying the sinograms, the attenuation correction step needs also to be modified, since there is a potential mismatch between the PET image and the attenuation map [69]. This can be especially true for patient motion, which can be of a high magnitude. For extrinsic methods, it is very common to apply motion correction before or during the reconstruction, since the motion tracking signal is usually coupled with the reconstruction machine [77, 91].

However for the intrinsic methods, raw data is often not available. It is therefore important to know if the mismatch due to the attenuation correction is significant or not. In that case the images are resampled, through an interpolation function, with the registration function parameters that have been found optimal.

2.3.4 Specificities and challenges of motion correction in PET

2.3.4.1 General case

Motion correction for PET imaging is a very active and vast area of research in biomedical engineering. Most of the research is based on extrinsic methods, whether they rely on a camera in the case of head motion [77, 91], a chest-belt for respiratory gated motion [74, 92] or MR sequences when acquired on a hybrid PET-MR scanner [79–81]. It is desirable to develop methods that can be seamlessly integrated into a clinical service. This typically mandates to reduce the assumptions about hardware, workflow and software. Additionally for our research project, only raw data (sinograms) were available. Consequently we review principally the main challenges specific to intrinsic methods.

It is very challenging to perform an automated motion correction on the full dynamic cardiac PET series because of the tracer dynamics, at least with intrinsic

methods. This is because the spatial distribution of the tracer changes rapidly over the time course and the tracer uptake seen is very different from one frame to another during the early frames (see section 2.15), which makes the task of matching images features or intensities difficult, as two consecutive frames may have a very different tracer uptake.

Most of the instrinsic motion correction techniques focus on images that are acquired tens of minutes (depending on the tracer and the application) after the injection, where the tracer has a more stable uptake. The objective of motion correction in that case is not to calculate more reliable and accurate kinetic parameters, but to remove the intra-frame motion or blur [93]. Images are first reconstructed as a dynamic sequence and then, once corrected, are summed together to provide a better 3D static image [84]. Several methods have been proposed to correct for brain motion correction, using different similarity measures and reference frames [94, 95]. Those methods define a frame-to-frame registration framework but do not take advantage of the tracer kinetic information. More recently a novel method suggested to incorporate temporal information of the PET tracer to the similarity measure function, joining kinetic modeling and frame-to-frame registration [96]. This method showed smaller registration errors and improved kinetic parameter estimates based on simulations results, when compared with image registration only results, motivating the integration of spatio-temporal information in motion correction framework for PET data.

When using intensity based similarity metrics for motion correction framework in dynamic medical images, a study has highlighted some artefacts that occur [84]. Due to the discrete nature of the images, local optima can appear during the calculation of the similarity metric, especially when using nearest neighbour or linear interpolations. Therefore finding the global maximum tends to be very difficult and may result in wrong or sub-optimal registration. To circumvent this problem, it is recommended to apply pre-processing on the images, such as resampling one of the two images, or removing the image background [84, 97].

2.3.4.2 Dynamic cardiac PET

Few studies are available on automated motion correction of dynamic cardiac PET. Performing automated motion correction of full dynamic cardiac PET is challenging as the tracer uptake visible in the different frames evolves rapidly from one frame to another, with for example high left ventricle tracer concentration in the early frames, to predominant myocardial tissue uptake in the late frames, as shown in Figure 2.15. Most studies that look at the impact of motion on Myocardial Blood Flow in PET used a manual correction for the early frames [65], or did not correct the early frames [44]. Some authors report automated motion correction of the late frames [44, 67, 98], using an intensity based registration, on the frames where the myocardial tissue is principally visible. Because the image structure is similar from one frame to another, image registration techniques can help correct for motion automatically.

One study offers a solution to address automated motion correction for the early frames [86]. The method suggests to define two different templates based on the segmentation of a long static image (3 to 5 minutes), taken towards the end of the dynamic PET series (5 minutes after beginning of acquisition), were the tracer is in an equilibrium or wash-out phase, in the myocardial tissue. The different templates include a blood pool template (i.e. Left Ventricle), and a myocardial tissue template. Those two templates are used as a reference to register the dynamic series against, by matching each frame of the PET series to a corresponding template. For frames that do not have any corresponding template (e.g. right ventricle), no registration is applied. While this study looked at the impact of motion correction on the TACs qualitatively, the motion correction framework did not incorporate any temporal information to guide the registration process. Moreover this studies does not explain how to match each frame with its corresponding template.

2.3.4.3 Motion correction in cardiac Magnetic Resonance Imaging

While the physics behind MRI and PET are different, image based motion correction of the heart faces similar challenges with the change of signal between the different heart chambers and myocardial tissue. Consequently ideas used in MR techniques

could be applied or translated to dynamic PET imaging to overcome the rapid change of the spatial distribution of the tracer.

A study addressed the registration of cardiac MR T1 mapping, by combining model fitting of the magnetization curves with deformable registrations [99]. Instead of considering the problem of frames registration, followed by fitting, the authors developed a method that treats motion correction as a joint problem of image registration and temporal model fitting. The method is an iterative algorithm, where synthetic images based on the results obtained from fitting magnetization are generated, and then registered to the MR frames with a deformable registration algorithm. This technique helps to guide the registration for frames with a limited magnetization signal and also a limited contrast. However, in cardiac T1 mapping, the temporal function that is used to fit the MR frames is tissue agnostic. It is the same function (mono-exponential decay) is fitted for every voxel of the heart. In the case of cardiac dynamic PET, the different tissues follow different models, and the TACs of the LV cavity are used to model the perfusion of the tracer in the myocardium.

In addition to motion correction in T1 mapping, some studies have addressed the motion correction of MR perfusion [100, 101], also called dynamic contrast MR. Cardiac MR perfusion is similar to MBF PET, in the sense that it relies on the injection of a contrast agent (Gadolinium) to measure MBF. The underlying idea of these two methods was to segment the heart in different subregions (LV, RV, myocardium), from a specific frame. A kinetic model is then applied to the different regions to generate synthetic images. These images are then used subsequently as references to register the different MR frames. Both algorithms showed promising qualitative results, where the TACs extracted from the myocardium before and after the registrations were shown to be smoother. While the idea is certainly valid for MR, changes would be needed to apply it to PET, as structural images are visible in some MR frames, even in the absence of contrast agent, unlike PET which is purely functional and no features of the heart can be seen.

2.3.4.4 Assessing the success of motion correction

It remains challenging to evaluate quantitatively motion correction accuracy on clinical data, because the ground truth is unavailable. Simulations - with well defined ground truth - can be performed in order to establish the performance of motion correction [67], but cannot capture systematically what is happening in actual clinical data. Several groups have evaluated the impact of motion correction on clinical images with generic image quality metrics, such as Signal to Noise Ratio (SNR) [102, 103]. Using this approach, an increase of the SNR calculated on the summed image of the dynamic series could demonstrate better image quality, this being associated with success in motion correction [104]. Since this metric is derived from a static image, it cannot characterize the effect of motion correction on dynamic data specifically, which is important for MBF quantification because it relies on features within the dynamic series. Hence a common assessment of the quality of motion correction remains a visual evaluation by a clinical expert [44, 67, 98], which is not objective, being subject to experience and interpretation of the observer. Moreover, these types of metrics do not relate directly with an objective assessment on uncertainty of kinetic parameters and MBF directly. It is recommended that motion correction should be performed, but currently there is no quantitative way to determine whether motion correction provides more reliable MBF estimates in an individual patient.

2.4 Model inference and measuring uncertainty

A model is a mathematical function that is matched as closely as possible to measured data, through a process called inference. Models help us to understand physical, chemical or biological behaviors [105]. In the context of medicine and biology they assist in refining and reporting findings, by providing a quantitative measurement to a visual and qualitative observation. However when using models, one needs to be aware of its limitations: a model simplifies the reality by representing a system (in our case the myocardial blood flow) in a mathematical way. For a given

model, its parameters need to be adjusted, through an optimization process, with the values that fit measured data the best. However, even if we assume that the optimization process converged to the mathematical optimal solution, one needs to be careful when interpreting the results. For example an excessively complex model, with too many parameters, might not be robust and would not be generalizable to many similar systems and data (overfitting), while an oversimplified simple model might introduce bias in the way we interpret the data (underfitting). Finally the model is a simplification of the actual system, and if perturbations modify the system behavior so that the model does not reflect the reality, interpretation might be incorrect. Our work will focus mainly on the latter aspect: we assume that the model used (a one tissue compartment) is suitable for describing the myocardial blood flow with a ^{82}Rb tracer, but we cannot consider that images systematically adhere to the model assumptions (e.g. presence of motion).

2.4.1 Problem definition

2.4.1.1 Notations

Let $\mathcal{M}_{PET}$ be a PET kinetic model parametrized with a set of M parameters $p = \{p_1, ..., p_M\}$. The parameters p are the kinetic parameters S, $K_1, ..., k_M$. $\mathcal{M}_{PET}$ is called a forward model because it starts with the cause (kinetic parameters) to model the results (measured data). Furthermore $\mathcal{M}_{PET}$ is nonlinear, because results do not depend linearly on, at least some of, the parameters. In our case of MBF modeling with dynamic ^{82}Rb PET, the PET kinetic model is represented by the following equation, as detailed in section 2.1.4.

$$C_m(t) = (1 - S)\Big(K_1 \exp(-k_2 t)\Big) \circledast C_a(t) + SC_a(t) \tag{2.19}$$

The measured signal over the N data points is denoted $y = \{y_1, ..., y_N\}$. In our case, y is a time activity curve extracted from the dynamic PET data. The objective of model inference is to find the parameters p, so that the distance between $C_m(t, p)$ and y is minimal.

2.4.1.2 Noise representation

A model is only a mathematical representation of a complex system and the measured data differs from the model function through noise measurements. This difference between $C_m(t)$ and y is due to the statistical noise coming from the random nature of decay events, also amplified with scanner's detection and acquisition system as well as the reconstruction process. Therefore in the mathematical modeling of the system, the noise of the system needs to be accounted for.

The statistical modeling of the noise in reconstructed PET dynamic images is a challenging task because the noise process in reconstructed images is not well defined. The positron emission itself is well modeled by a Poisson distribution [106]. However the scanner's detection system and other electronic devices add their own contribution to the initial Poisson distribution. Moreover the reconstruction process distorts the noise distribution in the reconstructed images, which makes it challenging to derive an analytical formula. Some studies of kinetic modeling with dynamic PET modeled the statistical noise in the reconstructed image by modified white noise [107–109]. It was modeled by a Normal distribution, whose variance σ^2 depends on the frame duration, and is proportional to the activity in the image:

$$\sigma^2 = \alpha^2 \frac{y(t_i)}{\Delta t_i} \tag{2.20}$$

Where $y(t_i)$ is the averaged activity in Bq/mL for frame i of the dynamic sequence, with mid-time t_i in seconds, Δt_i is the i-th frame length and α is a constant determining the noise level, that depends on the scanner and the dose injected.

While this type of noise representation is commonly used in simulations, it can be challenging to incorporate this formulation in the optimization process, e.g, by weighting each frame by their duration. This has been found to make the parameter inference process unstable. Furthermore, other studies advocated for the use of a gamma distribution or a normal distribution to represent the noise process in PET reconstructed images [110], while another study reported that the

difference between different weighting approaches are not statistically significant for a one compartment model [107].

In the case of white noise distribution, we denote the standard deviation of the Gaussian distribution ϕ^{-1} (ϕ is called precision). We also denote $\Theta = \{p, \phi\}$.

2.4.2 Model parameter inference

2.4.2.1 Non-linear least squares fitting

The objective of model parameters inference is to find the parameters p_{opt} of the model that match the measured data the most closely. A classic method for inferring non-linear forward model parameters is to use a Non-Linear Least Squares (NLLS) optimization. The aim is to define a cost-function or a distance between the model function (in our case $C_m(t)$) and the measured data (in our case y). A classic choice is to use the chi-square function χ^2 [111], and is defined as follows:

$$\chi^2(\Theta) = \sum_{i=1}^{N} \Big(\phi_i(y_i - C_m(t_i, p))\Big)^2 \tag{2.21}$$

Where: N is the number of time points and $C_m(t_i, p)$ is the value of the compartment model function at time t_i and for the parameters value p, ϕ_i is the estimated noise measurement between the model and the measurement at t_i, and y_i is the time point of the TAC at time t_i. ϕ_i is often unknown a priori and is not needed for numerical optimization of the function. It is hence set to 1 during the minimization process and estimated a posteriori if needed (see section 2.4.3.2).

The χ^2 function is minimized through an optimization algorithm, such as a gradient-descent. To find a local minimum using gradient descent, one takes an initial guess (starting point) at the parameters of the function to be minimized, and calculates the negative of the gradient of the function at that point. Then a new point is defined by taking a step proportional to the negative gradient. After a few iterations, a convergence point is reached to a minimum of the function.

2.4.2.2 Spectral analysis

The NLLS method is a generic and well-established model inference scheme. It suffers however of a lack of robustness for certain non-linear models. While the χ^2 function has a global minimum, it could also have, depending on the model, several local minima. This is problematic with a gradient descent scheme where the algorithm might not be able to overcome those local minima.

The spectral analysis method [112–114] tries to overcome those problems with reformulating the problem in a linear way, by pre-calculating certain functions. As detailed in section 2.1.4, the general equation that represents a compartment model with n compartments and spillover correction S is as follows:

$$C_m(t,p) = (1-S)R(t) \circledast C_a(t) + SC_a(t) \qquad \text{with:} \quad R(t) = \sum_{i=1}^{n} \Phi_i \exp(-\Theta_i t) \tag{2.22}$$

The idea behind spectral analysis is to define a set of basis functions that are linear in the parameter estimates, in order to calculate a linear least squares solution. Let $M \in \mathbb{N}$ the number of functions that are constructed. Usually, M is chosen relatively high ($M = 100$ or $M = 1,000$). Multiple functions that follow the same expression are constructed:

$$C_m(t,p) = (1-S)\sum_{j=0}^{M} \Phi_j B_j(t) + SC_a(t) \tag{2.23}$$

Where $B_j(t) = \exp(-\Theta_j t) \circledast C_a(t)$. Θ_j is sampled in a range of physiologically plausible values: $\Theta_{min} < \Theta_j < \Theta_{max}$, to cover an appropriate spectral range from the slowest possible event of the tracer in the tissue up to a value appropriate to transient phenomena (e.g., the passage of activity through the tissue vasculature) [114]. The basis functions $B_j(t)$ are pre-calculated for different values of Θ_j, so that the term $\exp(-\Theta_j t) \circledast C_a(t)$ becomes a fixed value. Φ_j are defined as positive parameters. In practice, only a few components with $\Phi_j > 0$ are found after the numerical minimization process, defining what is called the kinetic spectrum of the tracer in the tissue. A common practice to apply to the spectrum defined

below, is to use an information criterion (typically Akaike or Bayesian Information Criterion), to minimize the number of components that contribute to the signal, and reduce the sensitivity of the method to noise.

$$\text{spectrum} = \underset{\Phi_j, j\in\{1,M\}}{\operatorname{argmin}} \|y - C_m(t,p)\|^2 = \|y - (1-S)\sum_{j=0}^{M} \Phi_j B_j(t) + SC_a(t)\|^2 \quad (2.24)$$

The advantage of using spectral analysis is that the convergence is guaranteed to a global minimum because the model is linearized. Spectral analysis is also advantageous in a voxel-wise analysis. In a classic compartment model approach, the number of compartments is fixed for every voxel, and may not be appropriate for certain region of interest where too many compartments are not required. In spectral analysis, the best mathematical fit is provided, and the interpretation of the number of components is done a posteriori by studying the kinetic spectrum of the tracer. As explained in [114]: the terms for Θ_j very large become proportional to $C_a(t)$ and can be seen as "high-frequency" components. Terms with $\Theta_j = 0$ or close to 0, become proportional to $\int C_a(t)$ and can be viewed as the "low-frequency" component, that is, accounting for slower kinetics and limiting to its irreversible trapping in the tissue. Components with intermediate values of Θ_j represent tissue compartments that exchange material directly or indirectly with the plasma. The number of intermediate values of Θ_j corresponds to the number of identifiable tissue compartments within the volume of interest. From the kinetic spectrum, some physiological meaning on the tracer can be explicited, e.g. the transport of tracer from plasma to tissue, which is the kinetic parameter K_1, is given by the sum of spectrum components amplitudes:

$$K_1 = \sum_{j=0}^{M} \Phi_j \quad (2.25)$$

However this method may not be as accurate as a full non-linear least squares, since the basis functions are fixed a priori. Furthermore, in spectral analysis only macro-parameters can be retrieved, that is the relationship between Φ_i and Θ_i to the individual kinetic parameters k_i are lost. Finally, from the indistinguishability

theorem in [113] *"any two plasma input models, either reversible or irreversible, with a total of N tissue compartments are indistinguishable"* meaning that different compartmental arrangements may have the same kinetic spectrum.

In the context of myocardial blood flow modeling, the fitted TACs are extracted from a segmentation of the myocardium tissue. For this ROI, the type of model is known (considering the information about the tracer injected is also given). Thus in myocardial blood flow modeling, a spectral analysis is not necessary and compartmental model approach is usually preferred.

2.4.3 Measuring parameter uncertainty

Different methods exist to calculate uncertainty measurements on the parameters derived from a model. We review here two methods that are an alternative to Bayesian inference, which is further detailed in the next section.

2.4.3.1 Frequentist method

In the frequentist strategy, an experiment is repeated multiples times in the same conditions to estimate variability or uncertainty in the estimated parameters. By compiling the results of these experiments, statistics can be calculated, from which confidence intervals can be drawn. This technique is typically used by doing simulations that are designed to be as close as possible to the actual system. This approach has been used for example to measure the impact of the Renkin-Crone parameters on the MBF variability [60]. With this approach, the conditions of the experiment need to be reproduced exactly, in order to derive meaningful uncertainty measurements. In the context of studying the impact of motion on kinetic parameters, in patient studies, it is not possible to know from a single scan what influence motion correction has on the kinetic parameters, unless a scan is repeated multiple times, which is unrealistic.

2.4.3.2 Standard error calculation

When doing a non-linear least squares (NLLS) model fitting, it is possible to calculate the variance-covariance matrix of the model parameters [115]. By considering that the square of the residuals follows a gaussian distribution, the variance-covariance matrix Σ can be calculated with:

$$\Sigma^{-1} = (J^T J)^{-1} \frac{\sum_{i=1}^{N} \left(y_i - C_m(t_i, p_{opt})\right)^2}{M - N} \tag{2.26}$$

Where: N is the number of time points, M is the number of the model parameters, J is the Jacobian of the function C_m taken at the parameter p_{opt} and p_{opt} is the parameter value that provides the best fit.

This approach is commonly used to calculate the standard error of a parameter from the difference between the model fit and the data. The standard error of a given parameter is simply defined as the square root of the variance parameter, extracted from the variance-covariance matrix. This method has been used in the context of PET kinetic modeling to calculate the covariance of the parameters in a generic PET model [116].

A drawback of this method is that the error is calculated only once the algorithm has converged, but it can be challenging in practice to know when convergence has been reached [115].

As an alternative to the NLLS method, it is possible to use the Median Absolute Deviation (MAD) to fit the model $C_m(t, p)$ to the observed data points y [117, 118]. The MAD calculates the median of the absolute deviations between the model and the fit:

$$MAD = ||y - C_m(t, p)| - \text{median}(|y - C_m(t, p)|)| \tag{2.27}$$

The advantage of the MAD resides in its robustness to outlier data points.

2.4.4 Bayesian inference

2.4.5 Bayes theory

A natural and elegant way to calculate uncertainties for model parameters is to appeal to probability theory and Bayes theorem. In a Bayesian inference framework, each parameter is represented by a probability distribution, instead of a single value, from which an uncertainty metric can be derived [119].

The objective is to infer the posterior distribution $P(\Theta \mid y, \mathcal{M}_{PET})$, which is the conditional probability of having the parameters Θ from the observed y. Bayes theory updates the posterior probability according to:

$$P(\Theta \mid y, \mathcal{M}_{PET}) = \frac{P(y \mid \Theta, \mathcal{M}_{PET})P(\Theta \mid \mathcal{M}_{PET})}{P(y \mid \mathcal{M}_{PET})} \tag{2.28}$$

Where:

- The prior $P(\Theta \mid \mathcal{M}_{PET})$ is a probability distribution on the parameters capturing prior knowledge of their value before any new data has been considered, and is often constructed from literature or other data sources.
- The likelihood $P(y \mid \Theta, \mathcal{M}_{PET})$ is the probability of observing y given a set of kinetic parameters p and a noise measurement ϕ, and is computed directly from the model and the observation of the data, the TAC y in the case of PET kinetics.
- The evidence $P(y \mid \mathcal{M}_{PET})$ is the distribution of the observed data, marginalized over the parameters p, $P(y \mid \mathcal{M}_{PET}) = \int P(y \mid \Theta, \mathcal{M}_{PET})P(\Theta \mid \mathcal{M}_{PET})d\Theta$. The calculation of the evidence $P(y \mid \mathcal{M}_{PET})$ requires the computation of an integral which is often intractable [120].

In our case, a single model $\mathcal{M}_{PET}$ is used to fit the TAC y, that is a one compartment model. Therefore, the notation about the model dependency can be omitted, as it is implicitly assumed.

$$P(\Theta \mid y) = \frac{P(y \mid \Theta)P(\Theta)}{P(y)} \tag{2.29}$$

As mentioned earlier, the calculation of the evidence requires the computation of an integral which is often intractable. Multiple frameworks exist in the literature to estimate the posterior distribution without computing the evidence, and two main approaches are reviewed here. The first class of methods is Markov Chain Monte-Carlo (MCMC) [119, 121] algorithms and the second is Variational Bayes [120].

2.4.5.1 Markov Chain Monte-Carlo

Markov Chain Monte-Carlo is a class of algorithms that sample a probability distribution by constructing a Markov Chain. After a certain number of iterations, the Markov Chain converges to the posterior distribution. A common MCMC algorithm used to sample the posterior distribution of the Bayes equation is Metropolis-Hastings (MH).

To detail the MH algorithm, let us define θ a Markov Chain. $\theta^{(t)}$ is the value of the chain at the step t. The objective is to sample $P(\Theta \mid y)$ where the only inputs are $P(\Theta)$ (the priors), y (the measured data), and G the number of iterations, and n_0 the burn-in rate. The algorithm is detailed in alg. 1.

Algorithm 1 Metropolis-Hastings algorithm

Initialize the Markov-Chain $\theta^{(0)}$
for $t : 1..G$ **do**
  Sample a new parameter value θ_{prop} (usually from a normal distribution, centered on $\theta^{(t)}$)
  Calculate the posterior ratio $r = \dfrac{P(\theta_{prop} \mid y)}{P(\theta^{(t)} \mid y)} = \dfrac{P(y \mid \theta_{prop})P(\theta_{prop})}{P(y \mid \theta^{(t)})P(\theta^{(t)})}$
  Draw a random number u between 0 and 1.
  if $u \in [0, r]$ **then**
    $\theta^{(t+1)} \leftarrow \theta_{prop}$
  else
    $\theta^{(t+1)} \leftarrow \theta^{(t)}$
  end if
end for
Discard the first n_0 elements of θ (burn-in).

If G is sufficiently large, then θ samples the distribution $P(\Theta \mid y)$. Another key parameter is the width of the distribution from which θ_{prop} is sampled: if it is set too large, then the proposal will be too far from an acceptable value and θ_{prop} will

always be rejected. On the contrary if it is too small, then the proposal will always be accepted and the number of iterations to explore the whole domain will not be enough. A good distribution has a balanced ratio of acceptance/rejection. A common improvement to the MH consists in monitoring the acceptance/rejection ratio and update the width of that distribution accordingly.

The advantage of the MH algorithm is that it does not require any assumption on the model, the priors or the parameters. For that reason, MH is considered the "gold-standard", since the Markov Chain will sample the true posterior distribution very closely (providing there is enough iterations).

To derive an uncertainty measure from the posterior distribution defined with the Markov chain, one can calculate a confidence interval. Since there is no analytic formulation of the distribution, percentiles of the sampled distribution can be calculated to define the lower and upper bound of a confidence interval. For example, by calculating the 2.5% and 97.5% percentile, one can define the 95% confidence interval. The wider the interval, the less confidence there is in the parameter. The best estimate for the parameter can be chosen as the mean or the median of the distribution.

The major issue with the MCMC family algorithm is their convergence speed. While the Markov Chain will ultimately converge to the posterior distribution, the required number of iterations to converge in practice is very high, leading to slow computations. While it could be acceptable for certain applications, it can be easily impracticable for kinetic modeling in PET, especially for voxel-wise analysis, since the fitting algorithm would need to be run potentially millions of times [122].

2.4.5.2 Variational Bayes

Variational Bayes uses a different approach to estimate the posterior distribution. The idea behind variational inference methods is to define a distribution $Q(\Theta)$ to approximate the true posterior $P(\Theta|y)$. $Q(\Theta)$ is optimized using theory of calculus of variations and optimization [123, 124]. The objective is to search

over the distribution space $Q(\Theta)$ that minimizes the Kullback-Leibler divergence between $Q(\Theta)$ and $p(\theta|y)$:

$$KL\Big(Q(\Theta) \mid\mid P(\Theta|y)\Big) = \int Q(\Theta) \log \big(\frac{p(\Theta \mid y)}{Q(\Theta)}\big) d\Theta \tag{2.30}$$

One can see that when $P(\Theta|y) = Q(\Theta)$ the distance is minimal. Because the calculation of $P(\Theta|y)$ is intractable, the problem can be reformulated by maximizing the free energy:

$$F = \int Q(\Theta) \log \big(\frac{P(y \mid \Theta)P(\Theta)}{Q(\Theta)}\big) d\Theta \tag{2.31}$$

In order to maximize the free energy, different assumptions need to be made on the distribution $Q(\Theta)$. The goal of those assumptions is to make the inference tractable while providing a close approximation of the true posterior. The first hypothesis is to apply the mean field approximation, which assumes that the parameters can be divided into blocks of parameters that are independent. For the case of PET kinetics, we assume that the noise parameter ϕ is independent from the kinetic parameters p [125]. It is a reasonable assumption since the noise magnitude should not depend on the kinetic parameters but on external factors (acquisition device, reconstruction protocol, etc...)

$$Q(\Theta) = Q_p(p)Q_\phi(\phi) \tag{2.32}$$

A second assumption is made in order to make the distribution tractable: the prior distributions are chosen to be conjugate to the likelihood [124]. The prior is conjugate to the likelihood if and only if the resulting posterior has the same parametric distribution as the prior. The use of conjugate priors allows us for an explicit formulation of the distribution $Q(\Theta)$. Practical implementations of VB usually take priors that belong to the conjugate-exponential family.

A recent study applied the VB framework to a non-linear forward model [120]. In this work, the difference between the model and the data is modeled by a white noise distribution. Consequently, since the likelihood on the parameters is Gaussian, the priors chosen for the model parameters p are a MultiVariate Normal

(MVN) distribution, as the conjugate prior of a Gaussian likelihood is Gaussian. Similarly the likelihood on the noise parameter is Gaussian with known mean (a white noise is a 0 mean Gaussian distribution), thus the prior for the noise precision ϕ is a Gamma distribution:

$$\begin{aligned} P(p) &= MVN(p, \mu_0, \Sigma_0^{-1}) \\ P(\phi) &= \Gamma(\phi, s, c) \end{aligned} \tag{2.33}$$

where: μ_0 is the prior mean vector of the parameters p, Σ_0^{-1} is the prior variance-covariance matrix of the parameters p, s is the prior shape on the gamma distribution and c the prior scale on the gamma distribution. Each kinetic parameter is consequently represented by a marginal distribution on $p_k, k \in \{1, ..., M\}$ which follows a normal distribution $\mathcal{N}(\mu_k; \sigma_k^2)$ of respective mean and standard deviation $(\mu_k; \sigma_k)$. The mean μ_k may be taken as the best estimate for the parameter p_k, while the standard deviation σ_k is associated with the uncertainty in the measurement. Equally, one can draw credible intervals for the parameters; for example, the 95% confidence interval CI_{95} of a parameter p_k is defined as:

$$CI_{95}(p_k) = [\mu_k - 1.96\sigma_k; \mu_k + 1.96\sigma_k] \tag{2.34}$$

Another way to look at the uncertainty is to compute the coefficient of variation $CV(p_k)$:

$$CV(p_k) = \frac{\sigma_k}{\mu_k} \tag{2.35}$$

The advantage of the coefficient of variation is that it represents a unitless metric that allows comparison across different datasets. Similarly a higher CV implies a higher uncertainty on the parameter p_k.

In general applying a VB inference scheme to a non-linear forward model – which is the case for a one tissue compartment model – will produce intractable optimization. By doing a first order Taylor expansion around the mode of the posterior distribution, the posterior distribution can be optimized iteratively. The interested reader can find details of the implementation of the VB algorithm with non-linear forward model here [120].

The advantage of Variational Bayes lies in its relatively inexpensive computations and fast convergence. This method has already been successfully applied to kinetic modeling in PET [122], Dynmic Contrast Enhanced MRI (DCE-MRI) [126] and Arterial Spin Labeling (ASL) [127], to compute parametric images. A disadvantage of VB is the necessity of using conjugate priors. The fact that the priors used a normal distribution, imply that that parameters can take any values in $\mathbb{R}$. This can be problematic in the context of kinetic modeling in PET where some parameters are strictly positive (K_1, k_2) or constrained to be between 0 and 1 (V_b). Those limitations are discussed in the next chapter where we review the implementation of a PET kinetic model inference using VB.

2.5 Conclusion

In this chapter we reviewed the basic principles of PET imaging from signal formation to image reconstruction. We detailed how dynamic PET can help to compute Myocardial Blood Flow (MBF), through kinetic modeling. The MBF computation is however challenging and frequently subject to patient motion artifacts, which is the reason it is desirable to use automated motion correction techniques. We surveyed the basic principles of motion correction in PET, with an emphasis on image registration, and detailed the different methods available for automated motion correction in PET and other modalities. In the recent literature, there was as yet has been no studies that estimated the impact of motion correction on the kinetic parameters or to assess the quality of motion correction on dynamic features. This motivates the need to provide uncertainty measurements alongside with the MBF metric. For that purpose we examined the theory of uncertainty measurement for model parameter inference. Bayesian inference provides a natural framework for that purpose. After detailing two main theories for Bayesian inference, MCMC and Variational Bayes (VB), the VB method was chosen for calculating MBF and its uncertainty. This choice was motivated by the relatively inexpensive computations of VB compared to MCMC, and by the previous successful use of VB in medical imaging applications (ASL, DCE-MRI and PET). The implementation specificities

of a one tissue compartment model in a VB framework inference model is discussed in the next chapter, as well as the quality of the uncertainty metrics.

3

Calculating kinetic parameter uncertainty with Variational Bayes

3.1 Introduction

The quantification of myocardial blood flow (MBF) in Positron Emission Tomography (PET) is subject to various sources of parameters, such as the reconstruction options, the temporal sampling of the PET series, and artifacts, such as patient motion, also called gross motion. For that reason, it is desirable to provide MBF parameters accompanied by an uncertainty metric, in order to know what is the variability of the MBF parameter for the given data and the model used.

As presented in the previous chapter, Variational Bayes (VB) methods can provide an estimate of the uncertainty for each parameter of the model and have been successfully applied in a number of Magnetic Resonance Imaging (MRI) applications as well as dynamic brain PET kinetic modeling. In this chapter we apply the VB framework to a one tissue compartment PET model, used in the context of 82-Rb dynamic cardiac PET.

This chapter assesses the feasibility of applying VB methods for PET kinetic modeling, by performing simulation studies and applying VB methods on clinical datasets. In particular, the purpose of this work is to evaluate the robustness of

both (1) the kinetic parameters estimated by VB methods and (2) the uncertainty of the flow parameter, as the flow is the parameter of main clinical interest.

As outlined in the previous chapter, specific challenges arise when applying VB methods to PET MBF and some specific questions with regards to model implementation have to be answered. Those challenges are summarized as follows:

1. Noise model: The representation of the noise in PET is different than in dynamic MR since the nature of the acquisition is different leading to a different noise process. This noise model is not implemented in the existing VB methods.

2. Parameter constraints: The transfer rate parameters in PET kinetic models are physiologically defined to be positive and the spillover fraction is constrained to be between 0 and 1. Existing VB methods do not apply upper or lower bounds on parameters other than by Gaussian prior settings.

These two aspects are first evaluated through computer simulations of realistic Time-Activity Curves (TACs), since ground truth is known and relative errors to the ground truth can be calculated. The quantitative analysis of the uncertainty estimate is challenging. While a qualitative interpretation is straightforward - a higher uncertainty translates into higher variability of the studied parameter - quantifying in absolute terms the reliability of the uncertainty measure requires more elaborate methodologies. Therefore the quality of the uncertainty metrics is estimated through computer simulations and clinical datasets by comparing the results obtained with another Bayesian inference scheme, the Metropolis-Hastings algorithm from the Markov-Chain Monte-Carlo (MCMC) family, considered as the "gold standard" in Bayesian methods.

3.2 Theory

3.2.1 Model implementation

Designing and implementing an algorithm to perform model parameter estimation requires to have a stable process and obtain reliable results. One challenging task,

especially in non-linear model fitting is to set boundaries for the parameters, while interfering as little as possible with the convergence process. As a reminder, the function $C_m(t)$ to be fitted to the observed Time Activity Curve (TAC) is as follows:

$$C_m(t) = (1 - S)C_{tiss}(t) + SC_a(t) \quad \text{with:} \quad C_{tiss}(t) = K_1 \exp(-k_2 t) \circledast C_a(t) \quad (3.1)$$

where $C_a(t)$ is the concentration of the tracer in the arterial blood; K_1 and k_2 are kinetic parameters, S is the spillover factor, and $\circledast$ is the convolution operator.

In the case of a one-tissue compartment model, the two transfer parameters K_1 and k_2 are strictly positive and the spillover fraction parameter S (proportion of arterial blood in the measured data) is between 0 and 1.

As detailed previously, our implementation of VB approximates the posterior distribution of the kinetic parameters p_i for $i \in \{1, N\}$, where N is the number of parameters ($N = 3$ for a one-tissue compartmental model), by a normal distribution with respective μ_i and standard deviation σ_i:

$$p_i \sim \mathcal{N}(\mu_i, \sigma_i) \quad (3.2)$$

μ_i is taken as the best estimated parameter value of p_i and σ_i represents the measure of uncertainty in that estimate.

With these assumptions, each parameter can take values on $\mathbb{R}$, which could violate the conditions defined previously, since the parameters have upper and/or lower boundaries.

Two different models are studied in that chapter, derived from the equation 3.1. The first model, called $\mathcal{M}_1$ was a straightforward implementation of the equation 3.1. With the model $\mathcal{M}_1$, each parameter follows a normal distribution, which is easier to interpret to a certain extent. However it suffers from its inability to restrict a parameter to a domain definition, with our implementation of VB where priors are chosen to follow a Gaussian distribution, and hence do not have hard boundaries.

A second approach is to use estimate the parameter p_i with a reparametrization function. For that method, a strictly increasing function h is chosen such as:

$$h : \mathbb{R} \longrightarrow \mathcal{I} \tag{3.3}$$

$$x_i \longrightarrow h(x_i) = p_i \tag{3.4}$$

where $\mathcal{I}$ is the definition interval of the parameter p_i; x_i is a hyperparameter that is estimated by the VB method; and h is a strictly increasing function.

With such reparametrization, the parameter p_i is naturally restricted to its domain definition. The disadvantage of this approach is that the normal distribution on p_i is lost, since the VB methods estimate the probabilities on the hyperparameter x_i instead of p_i. However, while the distribution densities are distorted with such reparametrization, the maximum (mode) of the density functions are preserved as the function used is strictly increasing over $\mathbb{R}$. Consequently the best estimated value of p_i can be obtained by applying the reparametrization function h on the mean of the hyper-parameter x_i. Similarly statistical indicators (median, quartile, decile) and confidence intervals can be estimated by calculating those indicators on the hyperparameter and subsequently applying the reparametrization function h on the hyperparameter.

A second model $\mathcal{M}_2$ was implemented, where the parameters k_2 and S are fitted through a reparametrization, as follows:

$$S = \frac{1}{\pi}\arctan(u) + \frac{1}{2} \tag{3.5}$$

$$k_2 = \exp(v) \tag{3.6}$$

where v and u are hyper-parameters that can take values on $\mathbb{R}$. These functions were chosen as they satisfy the boundary conditions on S (restricted between 0 and 1) and k_2 (strictly positive).

The parameter K_1 is not reparametrized, since it is the parameter of principal clinical interest, in direct relation with the MBF parameter, and keeping the normal distribution representation facilitates the uncertainty interpretation calculated by the VB methods. However, if the estimates of K_1 were not robust it would be possible to apply a reparametrization function, similarly than for k_2.

3.3 Methods

3.3.1 Simulations

The purpose of simulations was to obtain TACs that are similar to what can be found in clinical data, while defining the ground truth parameters against which the VB methods can be evaluated objectively.

3.3.1.1 Data generation

We used a range of TACs representing typical Stress and Rest curves. Parameters were randomly sampled from an interval of likely values of Stress and Rest kinetic parameters. The intervals were defined as follows, based on literature [57] and observations of our clinical data:
Intervals for kinetic parameters representing Stress data:

$$K_1 \in [0.8, 1.7] \tag{3.7}$$

$$k_2 \in [0.15, 0.3] \tag{3.8}$$

Intervals for kinetic parameters representing Rest data:

$$K_1 \in [0.35, 0.8] \tag{3.9}$$

$$k_2 \in [0.05, 0.15] \tag{3.10}$$

The spillover fraction S was sampled from the following interval:

$$S \in [0.35, 0.65] \tag{3.11}$$

Five values were sampled for the pair (K_1, k_2) for each Stress and Rest cases. Five values of S were sampled, and used for both Stress and Rest case. This choice was made in order to study the relationship between S and the kinetic parameters K_1 and k_2. As such, for a given value of S two different pairs of (K_1, k_2) were compared. Furthermore, the spillover S parameter does not depend on the Stress or Rest protocol, but on the definition of the regions of interests (ROIs) from which the TACs are sampled, as well as the resolution of the scanner, which justifies the

choice of the same parameter range for S in Stress and Rest. S is indeed primarily due to blood contamination of the tissue signal because of cardiac and respiratory motion, and its magnitude is similar in Stress and Rest protocols.

In total, 10 different TACs were generated. The ground truth values are summarized in the table 3.1.

Table 3.1: Values of kinetic parameters used for simulations.

	K_1 (mL/min/g)	k_2 (1/min)	S (unitless)
TAC1_Rest	0.58	0.05	0.47
TAC2_Rest	0.69	0.12	0.45
TAC3_Rest	0.46	0.11	0.44
TAC4_Rest	0.56	0.08	0.57
TAC5_Rest	0.36	0.1	0.39
TAC1_Stress	1.13	0.24	0.47
TAC2_Stress	0.92	0.22	0.45
TAC3_Stress	1.33	0.19	0.44
TAC4_Stress	0.95	0.17	0.57
TAC5_Stress	1.58	0.26	0.39

The TACs generated were a sequence of time points, associated with a given activity, representing the mid-frame time of the frame. The activity at a given point was taken as the mean activity integrated over the time duration of the frame. The frames had the following duration: 1x10 seconds, 8x5 seconds, 3x10 seconds, 2x20 seconds and 4x60 seconds, which is the framing strategy used in a clinical workflow for PET cardiac studies and for the clinical data used in this thesis. The blood input function (BIF), corresponding to $C_a(t)$ in equation 3.1, was extracted from clinical data, using a commercial PET cardiology software (syngo.PET Myocardial Blood Flow - Siemens Medical Solutions USA, Inc., Malvern, PA, USA)[128].

Noise was added to the TACs in order to simulate realistic data. In order to validate the influence of the noise modeling on the errors in the parameters estimates, two different types of noise were used.

First, Gaussian noise was assumed, where the variance of the Gaussian noise was modeled as follows:

$$\sigma^2 = \alpha^2 \frac{y(t_i)}{\Delta t_i} \tag{3.12}$$

where $y(t_i)$ is the simulated noise free TAC at frame i in Bq/mL, of mid-time t_i in seconds, Δt_i is the i-th frame length and α is a constant determining the noise level. This noise model is referred as "PET noise" in this chapter.

In addition, a white noise of fixed variance was taken:

$$\sigma^2 = \tau\alpha^2 \tag{3.13}$$

where τ is a constant that represents the average activity of all frames over time and is set to 1,000 Bq/mL/s, this value was chosen to be of similar magnitude to the ratio $\frac{y(t_i)}{\Delta t_i}$ on average.

For each noise level 1000 realisations were generated. α was varied to have 15 values: $\{10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150\}$ and each noise level is denoted $NL = \{1, ..., 15\}$. NL values between 10 and 15 correspond to typical values for polar map TACs in Rb-82 dynamic PET. Since the clinical datasets were reconstructed iteratively with an OSEM method [129], a non-negativity constraint was added to the simulated TACs. Potential noise realizations that generated negative activities were discarded and regenerated until a positive value was obtained. Since the clinical data were decay corrected, the decay of the tracer was not included in our noise modeling.

Both kinetic model $\mathcal{M}_1$ and $\mathcal{M}_2$ were fitted to the simulated TACs of each noise realisation (white noise and PET noise, which is the adaptive white noise in equation 3.12) with the implementation of VB that is detailed in [120].

3.3.1.2 Analysis

From the VB algorithm, two sets of parameters need to be evaluated, the mean μ_i of the kinetic parameters p_i and the standard deviation σ_i.

In order to evaluate the robustness of the two model implementations $\mathcal{M}_1$ and $\mathcal{M}_2$, we calculated the number of times at least one parameter was outside of its domain definition, or if it was above 10 times the ground truth value, which would indicate that the algorithm diverged.

To evaluate the reliability of the mean estimate $\hat{\mu}_i$ of the parameter p_i (i.e. K_1, k_2 or S), the relative error ΔE for each realization of each noise level was calculated as follows:

$$\Delta E = \frac{\hat{\mu}_i - p_{i_O}}{p_{i_O}} \tag{3.14}$$

where p_{i_O} is the ground truth value of the parameter p_i. For example, K_{1_O} is the ground truth value of the parameter K_1. If a reparametrization function was used, the relative error is calculated as follows:

$$\Delta E = \frac{h(\hat{x}_i) - p_{i_O}}{p_{i_O}} \tag{3.15}$$

where $\hat{x}_i$ is the mean estimated value of the hyperparameter used; and h is the reparametrization function.

In order to evaluate the difference of noise models, a two-sample Kolmogorov-Smirnov test was run on the distribution of relative errors calculated for each noise model.

To evaluate the reliability of the uncertainty estimate $\widehat{\sigma}_{p_i}$ from the VB algorithm, 95% confidence intervals $\widehat{CI}_{VB}(p_i)$ on the p_i distribution were used. Those intervals were calculated for each TAC at each noise level. In order to estimate the reliability of those confidence intervals, a frequentist approach was used: we counted, for each noise realization of a given noise level, the number of times p_{i_O}, the ground truth value of p_i, fell into the estimated confidence intervals $\widehat{CI}_{VB}(p_i)$. This frequency is denoted $f_{\widehat{CI}_{VB}}$:

$$f_{\widehat{CI}_{VB}}(p_i) = \frac{1}{n}\sum_{j=1}^{n} \delta^j \tag{3.16}$$

Where

$$\delta^j = \begin{cases} 1 & if\ p_{i_O} \in \widehat{CI}^j_{VB}(p_i) \\ 0 & if\ p_{i_O} \notin \widehat{CI}^j_{VB}(p_i) \end{cases} \quad ; \tag{3.17}$$

j is the j$^{\text{th}}$random noise realization of a given noise level;

$$\widehat{CI}^j_{VB}(p_i) = [\widehat{\mu}^j_i - 1.96\widehat{\sigma}^j_i; \widehat{\mu}^j_i + 1.96\widehat{\sigma}^j_i]; \tag{3.18}$$

n is the number of noise realizations and $(\widehat{\mu}_i^j, \widehat{\sigma}_i^j)$ are the respective estimated mean and standard deviation values of the p_i density distribution for the j$^{\text{th}}$TAC of a given noise level. If the uncertainty estimate from the VB method is acceptable then the value $f_{\widehat{CI}_{VB}}$ should be constant across the different noise levels. As the noise increases, $\widehat{\sigma}_i$ will increase as well, but the proportion of true values within the confidence interval should remain the same. If the frequency $f_{\widehat{CI}_{VB}}$ is 100% for every noise realization the uncertainty given by VB is overestimated, i.e. $\widehat{\sigma}_i$ is too high. Similarly if $f_{\widehat{CI}_{VB}}$ is substantially low, then the uncertainty provided by VB is underestimated, i.e. $\widehat{\sigma}_i$ is too low. If a reparametrization function is used, the lower and upper bounds of the 95% confidence interval are extracted from the hyperparameter, and then the reparametrization function calculates the lower and upper bounds of the confidence interval on the parameter p_i.

The same method was then applied to confidence intervals drawn from an implementation of the Metropolis-Hastings algorithm (CI) $\widehat{CI}_{MH}(p_i)$. The Metropolis-Hastings chain was run with 15,000 iterations with the first 5,000 discarded as burn-in. The same prior distributions were used as for VB. The frequency that the ground truth fell into the estimated confidence intervals $\widehat{CI}_{MH}(p_i)$ was calculated:

$$f_{\widehat{CI}_{MH}}(p_i) = \frac{1}{n}\sum_{j=1}^{n}\delta^j \tag{3.19}$$

Where

$$\delta^j = \begin{cases} 1 & if\ p_{i_O} \in \widehat{CI}^j_{MH}(p_i) \\ 0 & if\ p_{i_O} \notin \widehat{CI}^j_{MH}(p_i) \end{cases} \quad ; \tag{3.20}$$

j is the j-th random noise realization of a given noise level; and n is the number of noise realizations.

The 95% confidence interval $\widehat{CI}^j_{MH}$ was directly defined from the posterior distribution sampled from the Metropolis Hastings algorithm.

Finally, the difference ϵ between the frequency $f_{\widehat{CI}_{VB}}$ and $f_{\widehat{CI}_{MH}}$ metric was used to compare the uncertainties estimated by VB and MH methods:

$$\epsilon = f_{\widehat{CI}_{VB}} - f_{\widehat{CI}_{MH}}$$

3.3.2 Clinical data: cross validation between VB and MCMC

To further validate the mean and standard deviation estimates of the parameter K_1, the estimates between VB and MCMC have been compared on clinical data. We detail the methods for the analysis of the K_1 parameter, because K_1 is the parameter in direct relation with the Myocardial Blood Flow (MBF), as detailed in the previous chapter (see section 2.2.3.2). A similar approach could be used for studying other parameters of interest.

3.3.2.1 Data

In order to compare the results of VB and MCMC, datasets with as little artifacts as possible were considered from the pool of data presented in Appendix A. This means that for this experiment datasets were (1) motion-free and (2) with normal perfusion patterns, in order to maximize the signal to noise ratio. The influence of motion on K_1 reliability is detailed in the next chapter, and the case of low perfusion regions is discussed in the chapter 6 of this manuscript. In total, 12 pairs Stress/Rest datasets have been selected from the pool of datasets available.

The list of datasets used is as follows: 14_S0_R0; 17_S0_R0; 21_S0_R0; 22_S0_R0; 26_S0_R0; 34_S0_R0; 38_S0_R0; 39_S0_R0; 51_S0_R0; 58_S0_R0; 64_S0_R0; 66_S0_R0;

The reconstruction parameters as well as pre-processing of these datasets are detailed in the Appendix A. For a given dataset, 540 TACs were extracted in 4 different groups: the whole myocardium (Tot) and the 3 coronary territories (LAD, LCX, RCA).

3.3.2.2 Analysis

Comparing probability distributions is challenging [130], especially since MCMC does not provide an analytical formulation of the probability density. One solution is to compare specific statistical indicators of the two distributions. We compared the mean and median of the MCMC sampled distribution with the mean of the distribution estimated from VB, which is also the median for the Gaussian distribution used.

We compared the MCMC and VB methods by first computing each indicator (e.g. mean or median) of the parameter distribution for a given method (e.g. MCMC). The results within each TAC group (Tot, LAD, LCX, RCA) were then averaged over the group studied. The outcome was a single average value of that statistical indicator over the set of TACs in that group. The same process is repeated for the second method (e.g. VB) and the results are compared, across the whole set of data, by computing the individual difference between the average TAC group values obtained with the two methods.

To compare the uncertainty parameter extracted from VB with the distribution sampled from MCMC, we calculated the dice coefficient of the 95% confidence intervals derived respectively from the VB method and the MCMC method. This indicator was chosen since the uncertainty extracted from the MCMC method is not expicit, unlike for the VB method where the standard deviation σ_{K_1} gives a direct measure of the uncertainty on K_1. The uncertainty metrics are therefore compared with an indirect measurement, and the dice coefficient over the 95% confidence intervals was chosen. Let g_{VB} be the probability density of the parameter K_1 estimated for a given TAC, and g_{MH} the probability density sampled from the Metropolis-Hastings algorithm. Let CI_{VB} the 95% confidence interval derived from VB, and CI_{MH} from Metropolis-Hastings. We define the interval $\mathcal{D}$ as the intersection of the two intervals CI_{MH} and CI_{VB}:

$$\mathcal{D} = CI_{VB} \cap CI_{MH} \tag{3.21}$$

The dice coefficient DSC of the 95% confidence intervals between VB and MH is then defined as follows:

$$DSC = \frac{2\int_{\mathcal{D}} \min(g_{VB}(x), g_{MH}(x))dx}{\int_{CI_{VB}} g_{VB}(x)dx + \int_{CI_{MH}} g_{MH}(x)dx} \tag{3.22}$$

By definition $\int_{CI_{VB}} g_{VB}(x)dx = \int_{CI_{MH}} g_{MH}(x)dx = 0.95$, hence the dice coefficient can be simplified to:

$$DSC = \frac{\int_{\mathcal{D}} \min(g_{VB}(x), g_{MH}(x))dx}{0.95} \tag{3.23}$$

The dice coefficient is calculated for each TAC in each dataset, and then averaged by group of TACs (Tot, LAD, LCX, RCA).

3.4 Results

3.4.1 Simulations

The model $\mathcal{M}_1$ was found to diverge for high noise realization of some Stress cases (TAC4_Stress and TAC5_Stress), up to 10% of TACs for noise level $NL = 15$ could not be estimated reliably. The model $\mathcal{M}_2$ was found to be robust for all noise realizations of every TAC. For that reason, only results from model $\mathcal{M}_2$ are reported.

Figure 3.1 shows representative simulated noisy samples of TAC3_Rest at two different noise levels. As expected, the noisier curve leads to a wider $\widehat{K_1}$ distribution.

Figure 3.2 shows the distribution of relative errors ΔE for the K_1 parameter for a noise level $NL = 15$ for both noise models (PET and white noise). For each TAC, the distribution of relative errors are similar, with similar medians and quartiles. The two-sample Kolmogorov-Smirnov test showed statistical differences (p-value < 0.001) only for TAC5_Stress and TAC4_Rest.

Figures 3.3 and 3.4 show the relative errors ΔE for the three parameters estimates K_1, k_2 and S. The relative error was robust to different noise levels for the parameters K_1 and S and to a lesser extent for k_2. The average error was at most -26.6% (k_2 for TAC4_Stress at noise level $NL = 15$), determined by

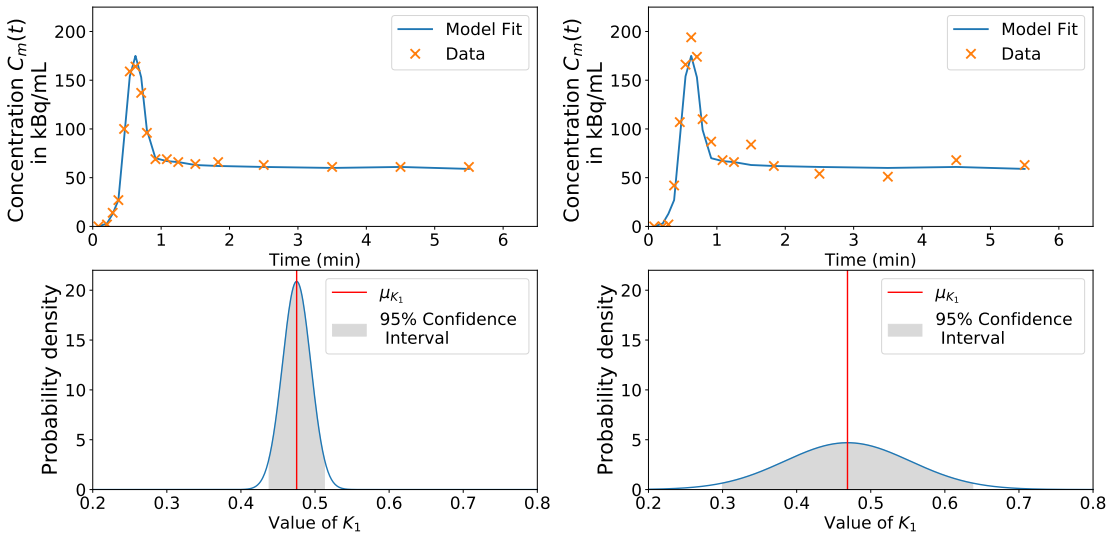


Figure 3.1: Examples of model fitting using VB. The top row shows two graphs representing simulated TAC3_Rest with different levels of noise (respectively 5 and 15). Noise level of 15 is what is typically seen in clinical data. The bottom row shows the impact on the estimated distribution of the parameter K_1. The graph on the bottom left, has a narrower distribution, with $\widehat{\sigma}_{K_1}$=0.019 mL/min/g for noise level $NL = 5$ and $\widehat{\sigma}_{K_1}$=0.084 mL/min/g for noise level $NL = 15$. Hence the noisier curve leads to higher uncertainty and less reliability.

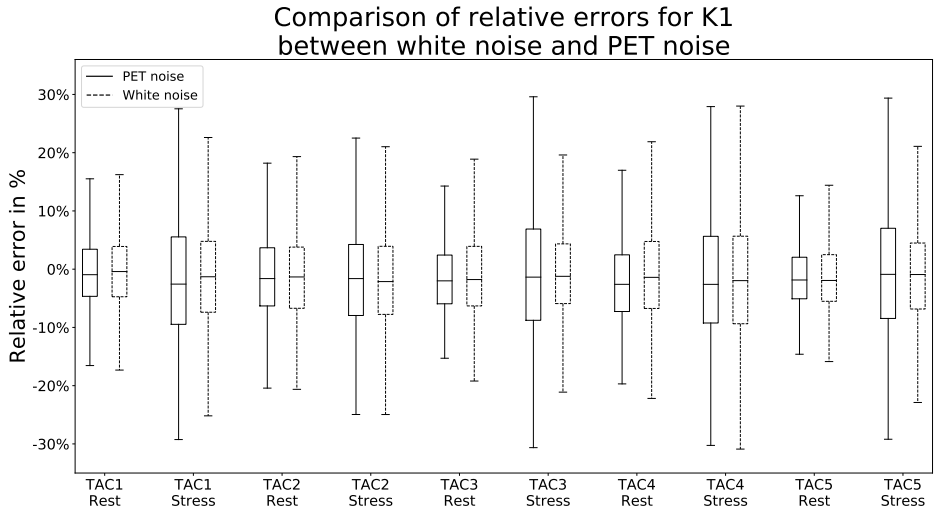


Figure 3.2: Comparison of distribution of relative errors ΔE for White and PET noise, for the maximum noise level $NL = 15$. For each TAC, the distribution of errors of the PET noise is similar to White noise. The median is represented by the middle of the box, the first and third quartiles by the edges, and first and last decile by the whiskers

averaging, separately for each parameter and for each TAC in case of noise level $NL = 15$. The parameter with the highest relative error was k_2, especially in the Rest cases (see Figure 3.3 a,b,c,d,e). The parameters K_1 and S had smaller relative errors than k_2. As the noise increases, a bias was observed for the parameter k_2. The distribution of relative errors for this parameter shows some underestimation, e.g. the mean relative error in TAC4_Stress was -26.6% for k_2 at noise level $NL = 15$. The S and k_2 parameters were well estimated for the different TACs at every noise level, with an average error at most of 1.0% for S (TAC4_Stress for $NL = 15$) and of -3.5% for K_1 (TAC5_Rest).

Figure 3.5 shows the quantities $f_{CI_{VB}}$ and $f_{CI_{MH}}$, representing the frequency with which the ground truth value of K_1 fell into the 95% confidence interval estimated by the different inference approaches. For the Rest TACs (TAC1 to TAC5), both methods gave reliable estimates of the 95% confidence intervals (respectively CI_{MH} for MCMC and CI_{VB} for VB); the values were between 94% and 96% for both VB and MCMC algorithm. For the Stress TACs, both methods tended to underestimate the size of the confidence intervals, especially for TAC1_Stress, TAC3_Stress and TAC5_Stress the average values being respectivel 91.1%, 88.7% and 86.2% for VB and 90.1%, 87.6% and 84.7% for MCMC. The difference between $f_{CI_{VB}}$ and $f_{CI_{MH}}$, ϵ, was small for the Rest TACs with an average difference of -0.07%, 0.8%, -0.1%, -0.14% and -0.29% for TAC1_Rest to TAC5_Rest. Similar findings were observed for the Stress TACs with an average difference of 0.96%, 0.98%, 1.01%, 0.67% and 1.45%, for TAC1_Stress to TAC5_Stress. For all TACs the results did not vary substantially across the different noise levels, as shown in Figure 3.5.

3.4.2 Clinical data

Figure 3.6 shows an example of two distributions of K_1, estimated by the VB method and the MCMC algorithm. Both distributions have similar probability densities, with less than 5% difference in the means, and a high overlap of the 95% confidence interval.

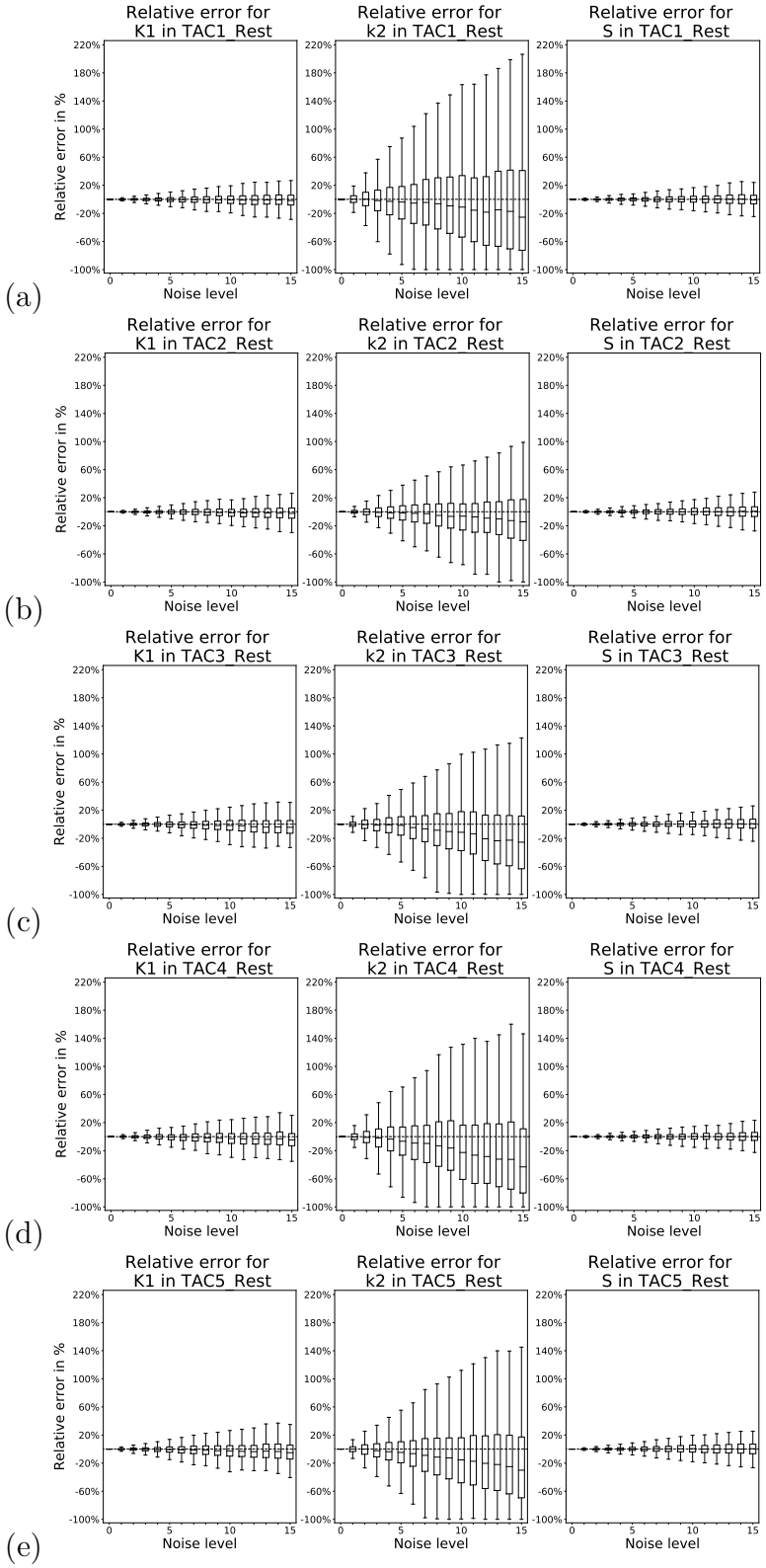


Figure 3.3: Relative errors for estimating K_1, k_2 and S for Rest TACs. The best estimates of parameters $\widehat{K_1}$ and $\widehat{S}$ show robustness to different noise levels. The best estimates of parameter $\widehat{k_2}$ show robustness to different noise levels, but is more sensitive to noise than the two other parameters. The box plot represents the distribution of the realizations for a given noise level. The median is represented by the middle of the box, the first and third quartiles by the edges, and first and last decile by the whiskers.

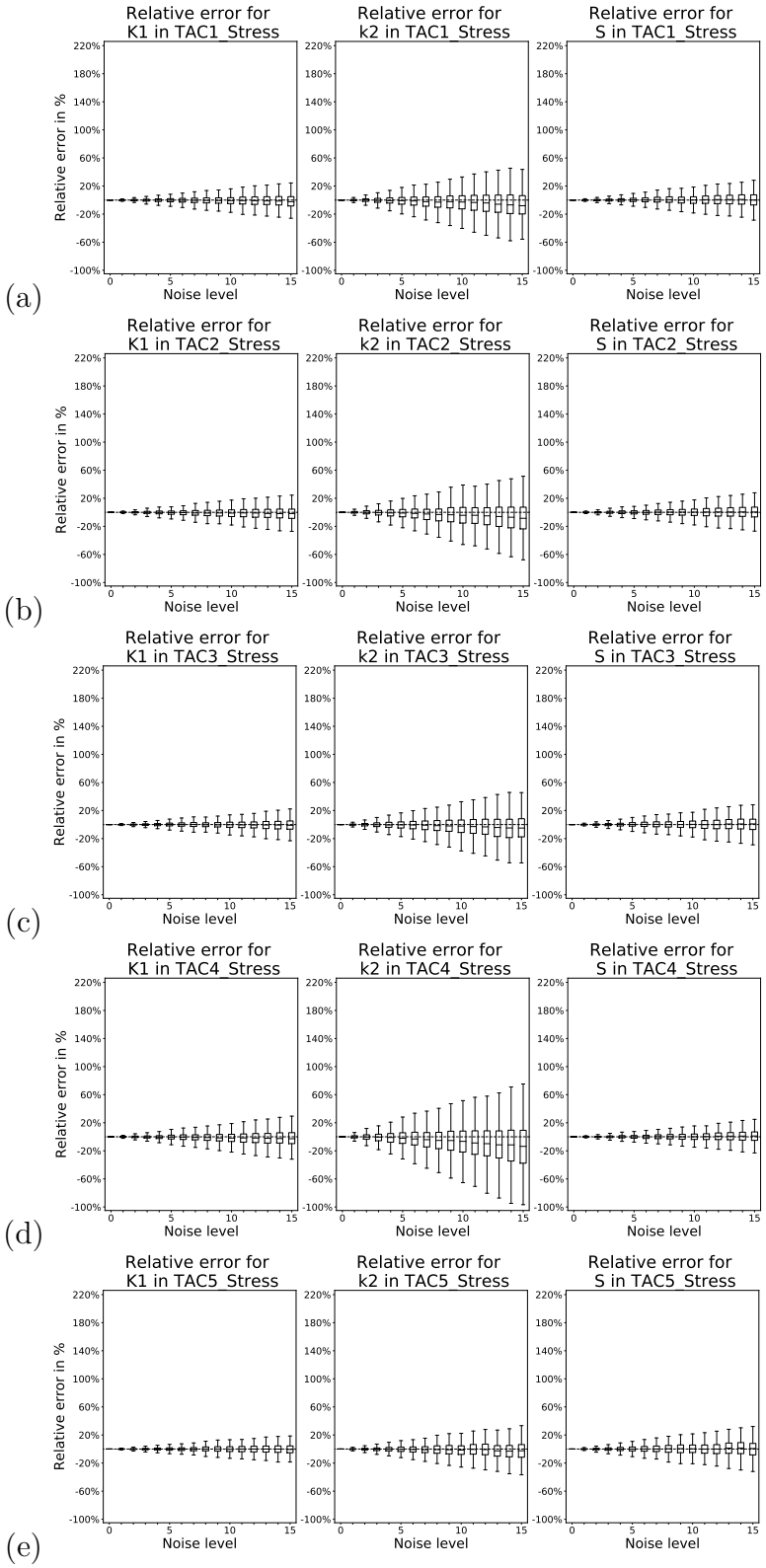


Figure 3.4: Relative errors ΔE for estimating K_1, k_2 and S for Stress TACs. Similarly than for Rest cases, the best estimates of parameters $\widehat{K_1}$ and $\widehat{S}$ show robustness to different noise levels. The best estimates of parameter $\widehat{k_2}$ show robustness to different noise levels, but are more sensitive to noise than the two other parameters.

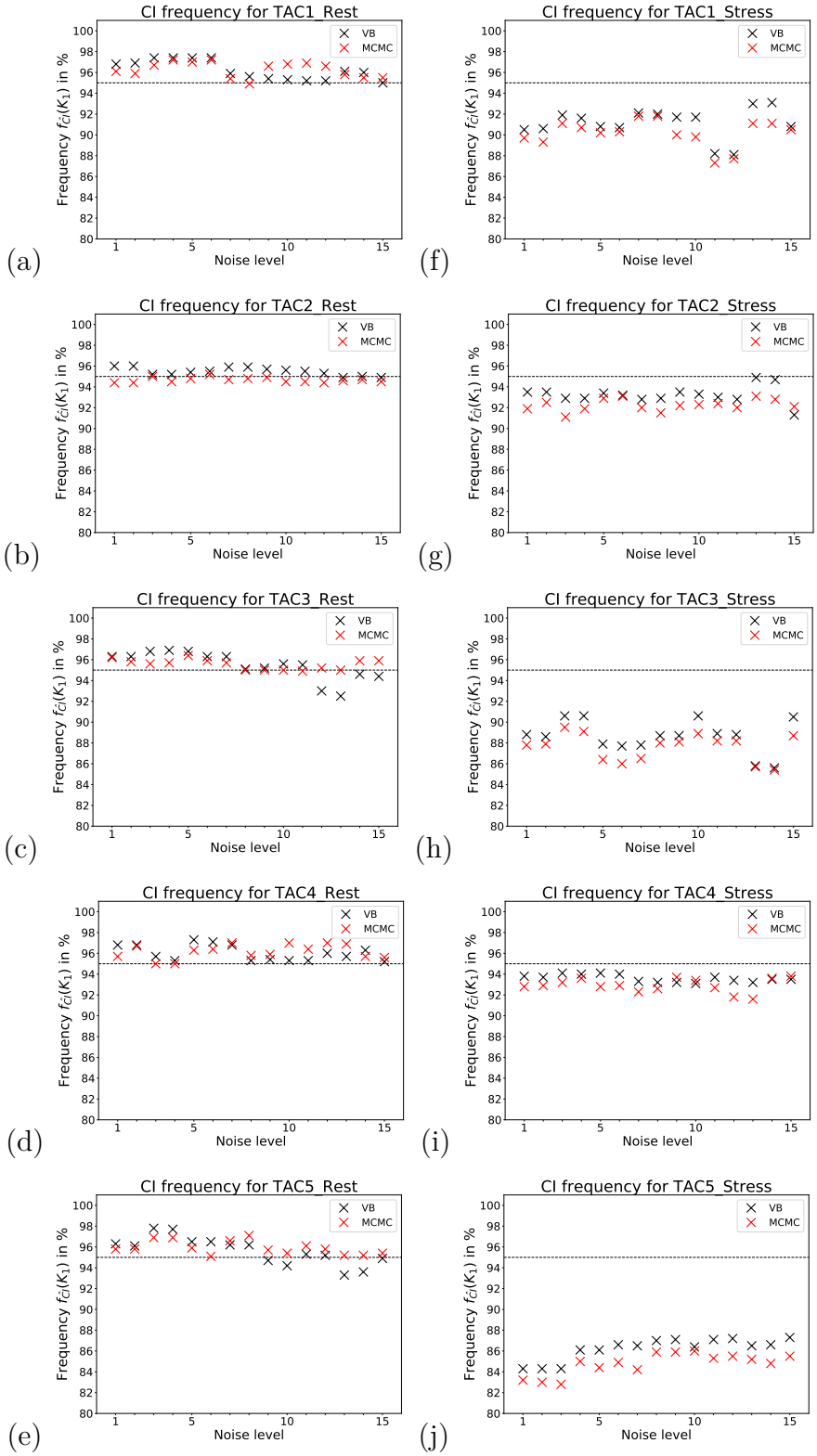


Figure 3.5: Comparison between $f_{CI_{VB}}$ and $f_{CI_{MH}}$ for the different noise level and each simulated TAC. The dashed line in the graphs shows the 95% value. The difference was small for all TACs, showing that the 95% confidence interval was estimated with similar performance for the VB algorithm and for MCMC.

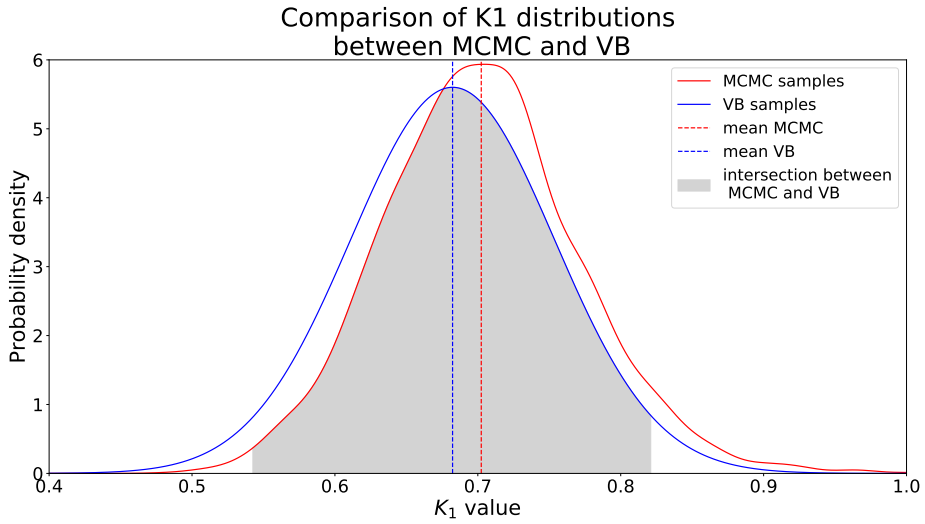


Figure 3.6: K1 distribution example for VB and MCMC. Both distributions are estimated for the same TAC extracted from a clinical dataset. Both distributions have similar shapes, showing a good agreement between each method. The area in grey represents the intersection of the 95% confidence intervals of the two distributions. The dice coefficient for this example is 0.78.

Figure 3.7 compares the mean of the K_1 distributions estimated by MCMC and VB, through a Bland-Altman plot. For the four TAC groups, MCMC tended to provide higher estimates on average than for VB. That difference was limited, with about 5% difference on average, and 95% of the datasets are within 10% of difference. No statistical differences were found between the median and the mean of the MCMC distributions, with less than 1% difference on average.

Figure 3.8 compiles the results of the dice coefficients for each group of TACs and datasets. Dice coefficients were between 0.75 and 0.8 on average, with consistent values between groups of TACs. The standard deviation was small, below 0.05, indicating a consistent metric for every TAC.

3.5 Discussion

In this chapter, methods to evaluate the reliability of PET kinetic parameters alongside uncertainty measurements were developed and presented. To estimate kinetic parameters of dynamic PET along with an uncertainty measurement, the

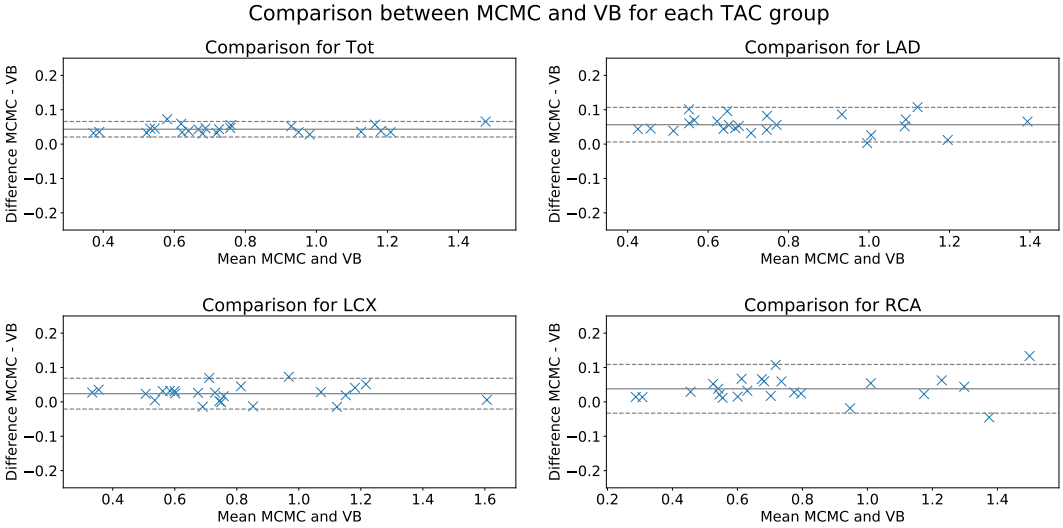


Figure 3.7: Bland-Altman comparing the mean of K_1 distribution estimated with the VB methods and MCMC, for the four groups of TACs. The solid line shows the average difference across datasets, and the dashed lines represent the 95% confidence interval. A good agreement can be observed between the two methods, with less than 10% difference. MCMC tended to have higher mean estimates than VB.

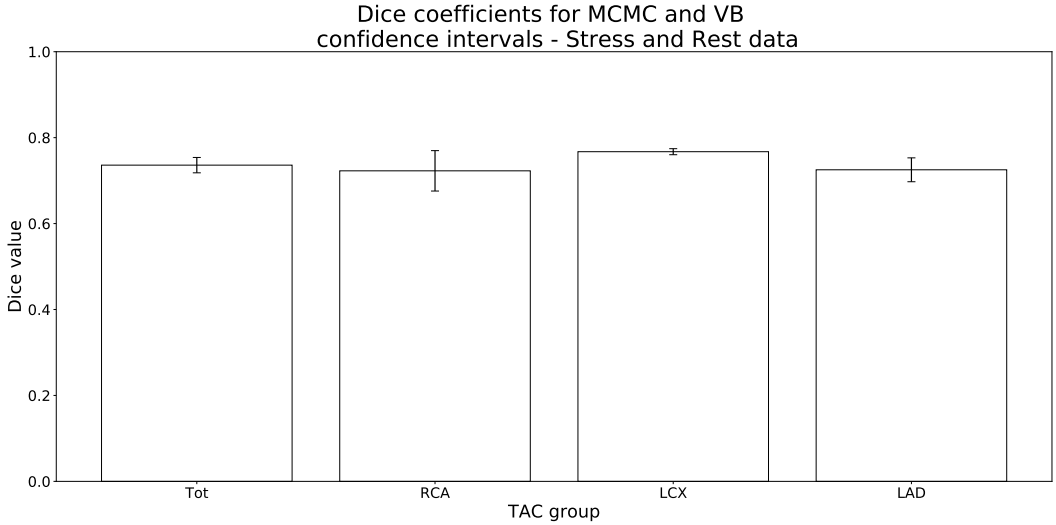


Figure 3.8: Barplot of dice coefficient of MCMC and VB distributions. The bar shows the average dice coefficient for the different datasets, while the error bars show the standard deviation. For each group of TACs the dice coefficient was between 0.70 and 0.80 on average, with standard deviations below 0.05.

Variational Bayes framework has been chosen due to its fast convergence and relatively inexpensive computations.

Two different implementations of the one tissue compartment model was implemented. The first implementation $\mathcal{M}_1$ was a straightforward implementation of the model $C_m(t)$ (see equation 3.1), while the implementation $\mathcal{M}_2$ embedded a reparametrization of the parameters k_2 and S which ensured that k_2 and S would be restricted to their definition domain. It was found that $\mathcal{M}_1$ had unreliable parameter estimates for up to 10% for high level of noise realizations, while $\mathcal{M}_2$ was found to be systematically reliable. This aspect is of primary importance in the context of this thesis, since the objective of using the VB method is to compare the uncertainty of kinetic parameters before and after motion correction, and TACs extracted from high motion datasets could be considered with higher noise than motion-free datasets. Furthermore, the fact that K_1 was reliably estimated showed that it is not necessary to apply a reparametrization function for K_1.

The VB method was initially evaluated using simulations, on a range of TACs with known ground truth parameters. The three different kinetic parameters K_1, k_2 and S were found to be well estimated at all noise levels. Errors on each kinetic parameter were found to be slightly different. The average error on S was found to be minimal with less than 1.0% on average across different noise levels and TACs. Errors on K_1 were a little higher with errors around 3% for high noise levels. Finally the errors on k_2 showed more substantial underestimation of the parameter in some cases of high noise levels. This is most likely explained by the non-linearity of the model: the additive noise is generated on the ground truth TAC, which does not necessarily translate into linear noise in the kinetic parameters and can introduce some bias in parameter estimations.

In simulations, uncertainty estimates of the kinetic parameters were found to be well estimated at all noise levels: that is, while the uncertainty increased progressively with an increase in noise, as shown in Figure 3.5, the proportion of simulations where the true value fell within the confidence interval were similar from one noise level to another. The estimates from VB were very close to the confidence

intervals derived from an MCMC implementation, which is normally regarded as the "gold standard" for Bayesian inference, albeit usually computationally impractical for routine application to parametric estimation in medical images (see appendix A.4). To evaluate the validity of the uncertainty measures arrived at by Bayesian inference we compared the predictions from VB and MCMC to the ensemble confidence interval from Monte Carlo simulation, i.e. a frequentist measure. For the simulated Rest cases and some Stress cases (TAC2_Stress and TAC4_Stress) there was a close agreement, with values of $f_{CI_{VB}}$ and $f_{CI_{MH}}$ close to 95%. However, for some of the Stress cases both Bayesian methods similarly underestimated the confidence interval, with values of $f_{CI_{VB}}$ and $f_{CI_{MH}}$ around 85% to 90%. Given the agreement between VB and MCMC this is unlikely to be as a result of the approximations made in the VB algorithm. However, it may well be reflective of the nature of the non-linear kinetic model employed resulting in a deviation between the confidence interval as might be defined using a Bayesian or frequentists approach. The three Stress TACs where the confidence interval seem to be underestimated are the cases where the K_1 ground truth values were the highest, consistent with the hypotheses that this phenomenon is explained by the non-linear model or the difference of definition between Bayesian and frequentist approach.

The VB methods were compared to MCMC on 12 pairs of Stress/Rest datasets. A good agreement was found between the average of each K_1 mean values for VB and MCMC, with around 5% of difference. However this difference was found to be systematically positive for every cases, indicating that MCMC tended to have higher estimates of K_1 mean parameter than VB. This difference is potentially explained by the difference of approach of VB methods compared to MCMC. First, in our implementation of the VB methods, normal distributions are assumed for the kinetic parameters, while MCMC does not impose any form for the posterior distributions. A second difference comes from the positive bias, of our MCMC implementation: by setting up boundary conditions (i.e. non-negativity) on the kinetic parameters, the sampled distribution would have only non-negative samples. Thus in the sampling step of a given size, a new sampled value has a higher probability of

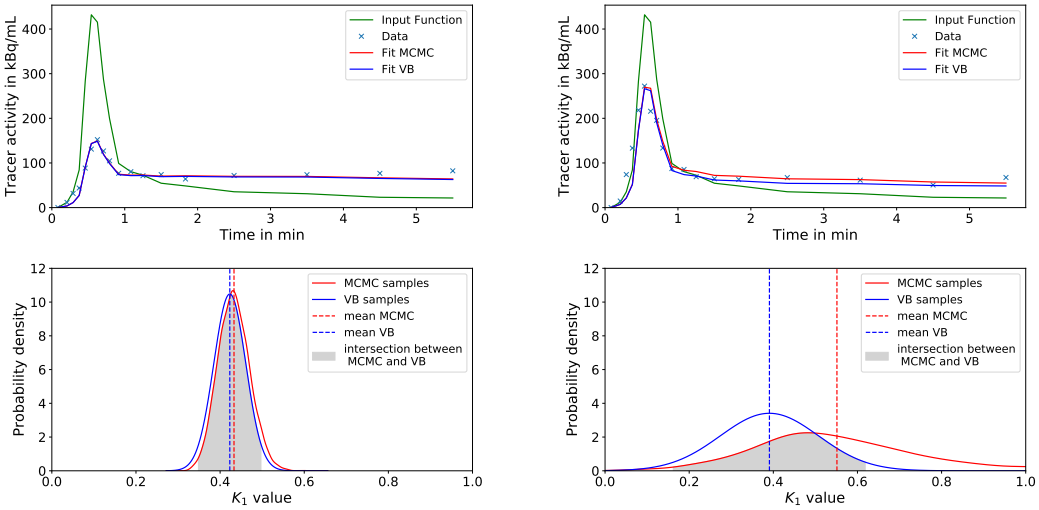


Figure 3.9: Influence of the spillover fraction parameter on the K_1 distribution: The top row shows two TACs extracted from a single dataset (14_S0_R0_Rest). The left graph has an estimated $S = 0.31$ while the right graph has an estimated $S = 0.62$. The resulting probability distributions of K_1 estimated by MCMC and VB are shown on the bottom row. For a low S the distributions are very similar: symmetrical and centered on the mean. For a higher S the uncertainty is higher (distributions are wider), and the MCMC distribution is asymmetrical: the maximum of the distribution does not coincide with the mean.

being accepted if it is higher than the current value, whereas if the new sampled value is negative it will be systematically rejected, by design of the algorithm. The non-negativity constraint could create non-symmetrical distributions for TACs in which the uncertainty in K_1 is high, which is typically the case for TACs with higher S. This case is shown in Figure 3.9. The reason why higher S values tend to lead to higher uncertainties, comes from the model definition itself (see equation 3.1). If S is high, the signal coming from the tissue $C_{tiss}(t)$ is lower, which leads to higher uncertainties on the parameter studied. The consequences of higher K_1 uncertainties when S is high is studied and discussed more extensively in chapter 6.

The dice coefficient computed for each TAC of each clinical dataset showed a good agreement between the two methods, with values between 0.70 and 0.80. This was comparable to dice coefficients computed for simulated TACs of noise levels $NL \in \{10, 15\}$ (data not shown). While it is difficult to evaluate whether 0.75 of dice coefficient is an indicator of a "small" distance between two distributions, the

dice coefficient was similar across all group of TACs, with little variability, showing a consistent agreement between MCMC and VB.

A potential limitation of this study is the choice of a white noise model to fit typical PET Time Activity Curves. It was found in the literature that while it is conceptually more accurate to model noise in PET TACs with an adaptive Gaussian noise (so-called PET noise in this chapter), the differences between white noise and PET noise were not statistically significant [107]. This finding was highlighted in Figure 3.2 where the distribution of relative errors for both noise models were showed to be similar. Furthermore, implementing an optimization framework with a weighted time points approach as suggested in [122], has shown to be less stable and robust [111]. For those reasons, we believe that the choice of a white noise model for the fitting is acceptable.

Overall the VB method was found to estimate accurately and reliably the kinetic parameters together with their uncertainty metric. Some differences were observed between VB and MCMC on clinical datasets, however the simulations results showed good accuracy of the VB methods in the kinetic parameter estimates.

3.6 Conclusion

In this chapter we evaluated the robustness of the VB algorithm on a PET kinetic model used in dynamic PET for MBF estimation. We investigated the specificity of the VB algorithm for PET kinetic modeling, specifically how to constrain the kinetic parameters to their domain definition with a reparametrization of some parameters and validated the use of a white noise estimation for PET kinetics. We also evaluated the robustness of the uncertainty metric estimated by VB, through simulations and by comparing to a "gold standard" Bayesian method.

It was found that the VB algorithm is able to estimate robustly and consistently the estimates $(\widehat{\mu}_{K_1}; \widehat{\sigma}_{K_1})$. The robustness of the estimates $(\widehat{\mu}_{K_1}; \widehat{\sigma}_{K_1})$ (Figure **??** and Figure 3.5) is of particular importance for the application to motion correction, because motion can be seen as another source of noise, albeit potentially more

complex in practice than the white noise assumption made in the simulations. The relationship between uncertainty and motion is studied in the next chapter.

4

Assessing motion correction by measuring MBF uncertainty

4.1 Introduction

We have seen in the literature review that motion correction of dynamic PET is necessary to perform accurate MBF estimation. With motion, the region of interest (ROI) studied might not reflect the same tissue over different PET frames, and consequently the final estimate of MBF is less reliable. Moreover, patient motion, or gross motion, is non-periodic, unlike respiratory or cardiac motion. Thus it can be difficult to predict its impact on MBF estimates. Clinical studies, with 90 [44], 225 [65] and 236 patients [66], showed that moderate or high motion can appear relatively frequently, for 30% to 40% of the scans, and is more common in stress scans than in rest scans [65]. In addition to frequent occurrence, motion also has a significant impact on MBF estimates. A recent study [67] determined, through computer simulations, that severe patient motion (up to 2cm) could lead to 500% error in MBF measurement. The impact of motion on clinical data has also been evaluated, demonstrating that motion can be a source of error in MBF estimation [65, 66]. Additionally, motion does not only induce errors within the dynamic PET image but also between the PET series and the CT. A good PET/CT

alignment is important for attenuation correction purposes and can be another source of error in the MBF estimation [69].

Motion correction is necessary, but it remains challenging to quantitatively evaluate its accuracy on clinical data, because the ground truth is unavailable. Simulations - with defined ground truth - can be performed in order to establish the performance of motion correction algorithms [67], but cannot capture systematically everything that is happening in actual clinical data. Some studies have evaluated the impact of motion correction on clinical images with generic image quality metrics, such as Signal to Noise Ratio (SNR) [102, 103]. Using this approach, an increase of the SNR calculated on the summed image of the dynamic series could demonstrate better image quality, this being associated with success in motion correction [104]. Since this metric is derived from a static image, it cannot characterize the effect of motion correction on dynamic data specifically, which is important for MBF quantification because it relies on features within the dynamic series. Moreover, these types of metrics do not relate to an objective assessment on uncertainty of kinetic parameters and MBF directly. It is recommended that motion correction should be performed before or during reconstruction of the dynamic PET data [61, 67], but currently there is no quantitative way to determine whether motion correction provides more reliable MBF estimates in an individual patient.

To estimate uncertainties of parameters in a model, we can appeal to Bayes theory, which offers a natural framework to estimate uncertainties of parameters in a model [19]. In a Bayesian framework, each parameter is represented as a probability distribution instead of a single value, from which an uncertainty metric can be drawn. In the previous chapter, we have seen that the Variational Bayes method, which approximates the posterior distribution defined by the Bayes equation using variational theory, provide reliable kinetic parameter estimates for MBF modeling. Particularly, we have shown that the uncertainty metric provided by VB was robustly estimated at different noise levels, which is important in the context of motion correction, since motion can be seen as another source of noise.

This chapter assesses the motion correction of dynamic cardiac PET data, by evaluating the uncertainty of the MBF estimate, derived from the VB framework. Specifically, the uncertainty of the MBF estimate is calculated, from clinical data with multiple degrees of motion, before and after manual realignment performed by experienced observers. This chapter also investigates the difference in performing motion correction before or after reconstruction of the PET data.

4.2 Methods

4.2.1 Data

Only stress images were considered for this study, since motion is more frequent and pronounced in stress than rest. The patient data were categorized into 4 different groups based on the amount of motion, according to the method in [44]. The motion categorization was done as follows: group 0 was characterized by no motion, group 1 contained light motion (less than half the width of the left ventricular (LV) myocardial wall), group 2 contained moderate motion, i.e. greater than half the width of the LV myocardial wall but could be corrected partially by automated motion correction algorithm of syngo.PET MBF [128], and group 3 contained severe motion cases, same as 2 but could not be corrected by automated motion correction algorithm of syngo.PET MBF (visually assessed by review of the dynamic PET series post-correction). Ten datasets of each group were selected. Additional information on data acquisition can be found in appendix A. The list of datasets used was :

- Group 3 - high motion: 05_S3_R2; 11_S3_R2; 13_S3_R1; 76_S3_R1; 79_S3_R0; 86_S3_R0; 88_S3_R2; 97_S3_R0; 100_S3_R0 and 101_S3_R3.
- Group 2 - moderate motion: 08_S2_R0; 18_S2_R0; 20_S2_R0; 29_S2_R0; 37_S2_R0; 57_S2_R0; 62_S2_R0; 72_S2_R0; 78_S2_R0 and 81_S2_R0.
- Group 1 - light motion: 24_S1_R0; 27_S1_R0; 32_S1_R0; 35_S1_R0; 42_S1_R0; 48_S1_R0; 50_S1_R0; 56_S1_R0; 63_S1_R0 and 68_S1_R0.

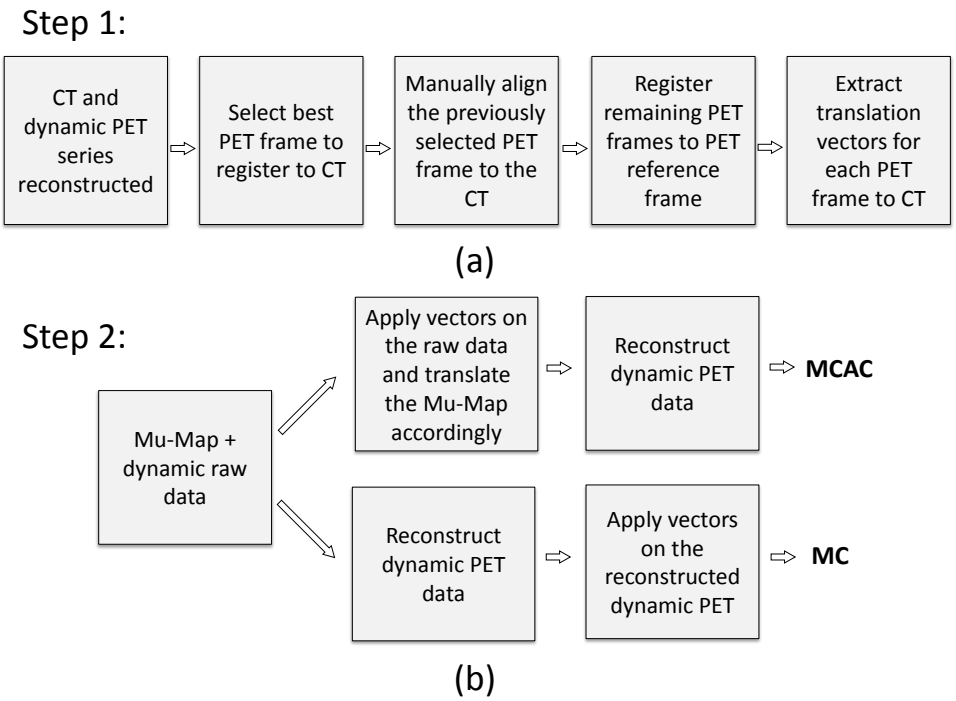


Figure 4.1: Motion alignment workflow strategy.
a) Step 1: Workflow used for finding motion vectors of each PET frame. b) Step 2: Difference of motion alignment strategy: either before or after the reconstruction. Motion Corrected with Attenuation Correction images (MCAC) were corrected before the reconstruction, while Motion Corrected (MC) were corrected after.

- Group 0 - no motion: 09_S0_R3; 14_S0_R0; 17_S0_R0; 21_S0_R0; 22_S0_R0; 28_S0_R0; 30_S0_R0; 34_S0_R0; 38_S0_R0 and 39_S0_R0.

Additional information about image acquisition protocol and reconstruction parameters can be found in appendix A.

4.2.2 PET frames realignment workflow

Frames were manually realigned by three experienced observers in nuclear cardiology. The alignment process consisted of two steps, Fig. 4.1. This process is referred as motion correction in this study.

The first step consisted of manually estimating the motion, see Fig. 4.1a: For each dataset, one of the four last frames of the PET series was manually aligned to the CT image and acted as a reference. For a given subject, every observer used the same PET frame to align with the CT image. Once this reference PET frame was

aligned with the CT, all other PET frames were manually aligned to the reference PET frame. Three dimensional translations alone were used to correct for motion. The ranges of motion vectors applied are detailed in Table 4.1.

The second step consisted of applying the motion realignment vectors from step 1, see Fig. 4.1b. For the purposes of this study this was done in two different ways:

1. A dynamic series was created by applying the motion vectors to the raw data. The attenuation correction map, or Mu-Map, which is a tissue density map, generated from the estimated density of the tissue from the CT image, was shifted for each frame, in order to keep the CT and the PET frame aligned. By shifting the Mu-Map, the attenuation correction was specific for each frame. This series is referred as Motion Corrected with Attenuation Correction (MCAC).

2. An additional dynamic series was created from the original series, by reconstructing the dynamic series and then applying the motion vectors defined by users. This series is referred as Motion Corrected (MC).

MC and MCAC images had the same motion realignment vectors; the only difference was the stage in the reconstruction in which the motion realignment was applied. Consequently, the two images differed in intensity, due to the different reconstruction methods.

4.2.3 TAC fitting and uncertainty estimation

Time-activity curves were obtained with the syngo.PET MBF software application [128]. The default automatic motion correction was disabled for the processing of each of the three series Original, MC and MCAC. For each dataset 540 TACs were extracted, and regrouped into 4 categories: the 3 coronary territories (LAD, LCX and RCA) and the whole myocardium (Tot). Each TAC was fitted individually with a one-tissue compartment model, where the two kinetic parameters K_1 and

	Light motion (group 1)	Moderate motion (group 2)	Severe motion (group 3)
Range for average translation (mm) for user 1	[0.8;1.7;2.1]	[1.4;3.2;5.7]	[3.3;4.8;8.2]
Range for average translation (mm) for user 2	[1.7;2.4;4.6]	[1.7;4.0;5.5]	[3.6;5.2;10.2]
Range for average translation (mm) for user 3	[1.4;2.2;3.4]	[1.7;3.7;7.6]	[3.1;5.0;9.0]

Table 4.1: Range of motion vectors for the three users for every dataset of each motion group. The ranges given are for relative frame realignment (i.e frame-to-frame). Average relative alignments across all PET frames are computed for each dataset in each motion category. The intervals indicate the minimum, median and maximum values in a motion category. Ranges correlate with the qualitative motion categorization of the datasets.

k_2 were extracted as well as the spillover factor S. As a reminder, K_1 is expressed as an increasing function of MBF:

$$K_1 = \mathrm{MBF}\Big(1 - A\exp(\frac{-B}{\mathrm{MBF}})\Big) \tag{4.1}$$

where A and B are two positive numbers. Therefore, in order to assess the reliability of the realignment applied to the PET frames, the uncertainty on the K_1 parameter was calculated, with the VB methods detailed in the previous chapter. K_1 was used here as a surrogate for MBF.

For each TAC, the VB method estimates the mean μ_{K_1} and standard deviation σ_{K_1} of the parameter K_1 approximate posterior distribution. These values were averaged, per patient, for the group of TACs considered, that is either the whole myocardium (Tot) or one of the three coronary territory (LAD, LCX and RCA). The number of TACs considered for each group was 487, 239, 132 and 116 for Tot, LAD, LCX and RCA respectively. While each territory was independent, the result for the whole myocardium will be influenced by the results of each coronary territory.

4.2.4 Analysis

The impact of motion correction on the kinetic parameter K_1 was evaluated by calculating the normalized difference of K_1 estimates as follows:

$$\Delta\mu_{K_1} = \frac{\mu_{K_{1_C}} - \mu_{K_{1_O}}}{\mu_{K_{1_O}}} \tag{4.2}$$

where: $\mu_{K_{1_C}}$ is the average of μ_{K_1} for the group of TACs studied (Tot, LAD, LCX or RCA), after correction and $\mu_{K_{1_O}}$ is the parameter average before correction.

The effectiveness and reliability of a motion correction strategy was evaluated by computing the normalized difference of the estimated uncertainty of K_1:

$$\Delta\sigma_{K_1} = \frac{\sigma_{K_{1_C}} - \sigma_{K_{1_O}}}{\sigma_{K_{1_O}}} \tag{4.3}$$

where: $\sigma_{K_{1_C}}$ is the averaged uncertainty on K_1, for the group of TACs studied, after correction and $\sigma_{K_{1_O}}$ is the uncertainty on K_1 before correction. If $\Delta\sigma_{K_1} < 0$ the correction has reduced the uncertainty of K_1, i.e. motion correction resulted in a higher confidence in K_1. A negative change means that the motion corrected image provides a more reliable estimate of K_1 than the original series. A paired t-test was performed in order to evaluate the statistical difference between $\sigma_{K_{1_C}}$ and $\sigma_{K_{1_O}}$.

To compare the effectiveness of the MC method with the MCAC, the standard deviation σ_{K_1} and the mean value μ_{K_1} of each dataset was computed for each observer. Similarly a paired t-test was computed between $\sigma_{K_{1_{MC}}}$ and $\sigma_{K_{1_{MCAC}}}$ as well as between $\mu_{K_{1_{MC}}}$ and $\mu_{K_{1_{MCAC}}}$, where the subscript indicates the respective MC and MCAC corrections described in Figure 4.1. Finally, to evaluate the effectiveness of motion correction in comparison to motion free datasets, the coefficient of variation CV_{K_1} was computed for each dataset after correction and compared to group 0 datasets.

4.3 Results

Figure 4.2 shows an example of frames realignment for a severe motion case (group 3). Most of the motion was found to be in the vertical long axis direction), affecting

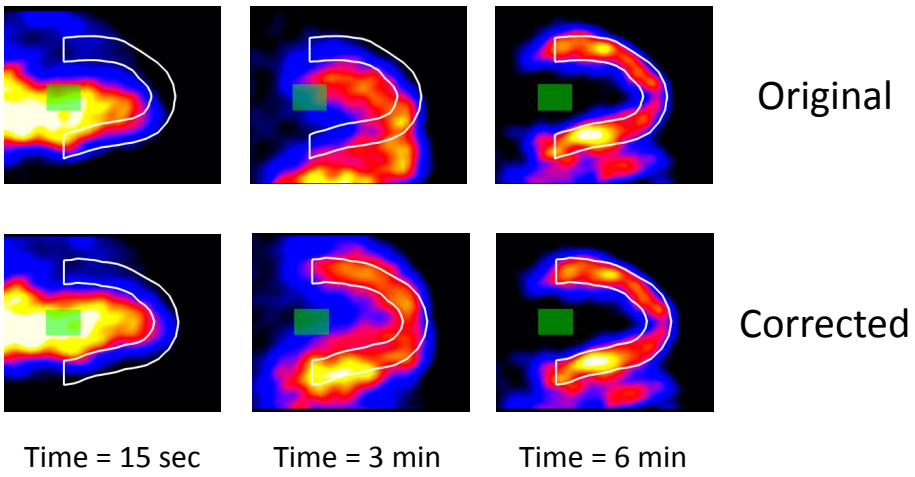


Figure 4.2: Comparison of original and motion corrected dataset, for a severe motion case (group 3). Each row shows three frames at representative time points of the dynamic PET series. The top row shows the original dataset and the bottom row shows the dataset as corrected by one observer. For the original dataset, the segmentation of the myocardial tissue (white contour), defined from the last frame, is misaligned with earlier frames. The ROI sampling the arterial blood activity (green rectangle) in the left ventricle does not capture the intensity peak in the case of the original dataset. A better alignment of the various image structures can be observed in the corrected dataset. The windowing level is adapted for each frame of the PET series but is kept the same between the original and corrected dataset of a given time point.

principally the LAD and RCA coronary territories. In the case of high motion, the segmentation of the myocardial tissue is not only misaligned with the myocardium in earlier frames but also with the left ventricle activity in the very early frames (from 0 to 30 sec). An example of the effect of the motion correction on the input function and the TACs can be seen in Figure 4.3.

Overall a reduction in $\Delta\sigma_{K_1}$ after MCAC correction for the data in the whole myocardium was observed in the three motion groups, Figure 4.4. For motion group 3, the mean reduction for the whole myocardium was 45.3% on average for the three observers. The biggest change was seen for the RCA with a reduction in uncertainty of 63.6%. Similar results were observed for motion group 2 - a reduction of 40.5% was observed on average for the whole myocardium for the three observers. The RCA territory showed the largest decrease at 56.0% on average. As expected, because there was less motion than in groups 2 and 3, changes were smaller for group 1, but the territory most affected remained the RCA with a mean

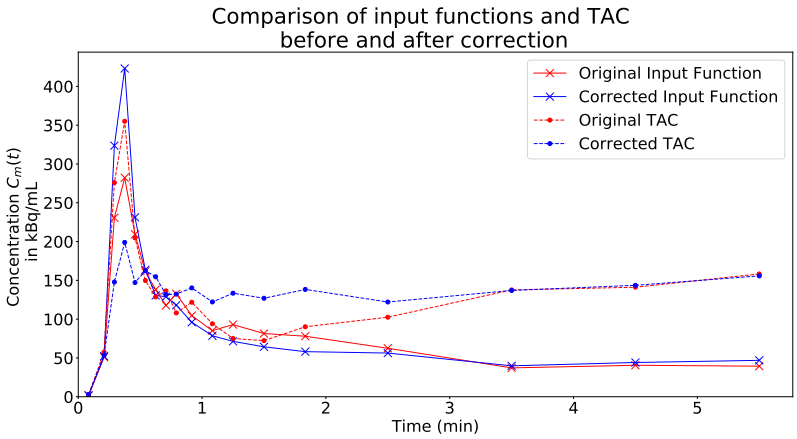


Figure 4.3: Comparison of input functions and TAC before and after correction, for a severe motion case (group 3). The TAC represented is extracted from the RCA territory. The input function of the original dataset has a lower peak, indicating the ROI sampling the arterial input function does not capture the full activity of the injected tracer. Artifacts on the later time points of the original TAC can be observed, as the activity between 1 and 3 min is lower than for time points between 3 and 6 min, whereas the TAC extracted from the corrected dataset is approximately constant.

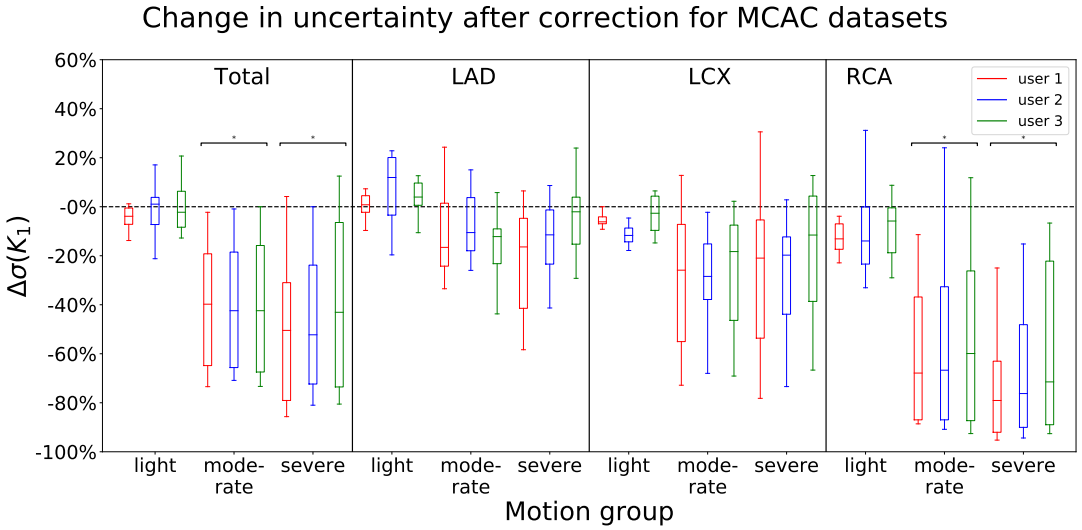


Figure 4.4: Change in uncertainty after motion correction for datasets corrected with MCAC. The value $\Delta\sigma_{K_1}$ is plotted for the whole myocardium (Total) as well as for each territory. A global decrease can be seen for each motion group. The difference was statistically significant for Total and RCA territory in moderate and severe motion groups (p-value < 0.05). The reduction in uncertainty was comparable between users (median and quartiles are similar).

reduction of 10.1%. For every observer, statistical differences between $\sigma_{K_{1_{MCAC}}}$ and $\sigma_{K_{1_O}}$ were found (p-value < 0.05) for both moderate and severe motion in the whole myocardium and the RCA.

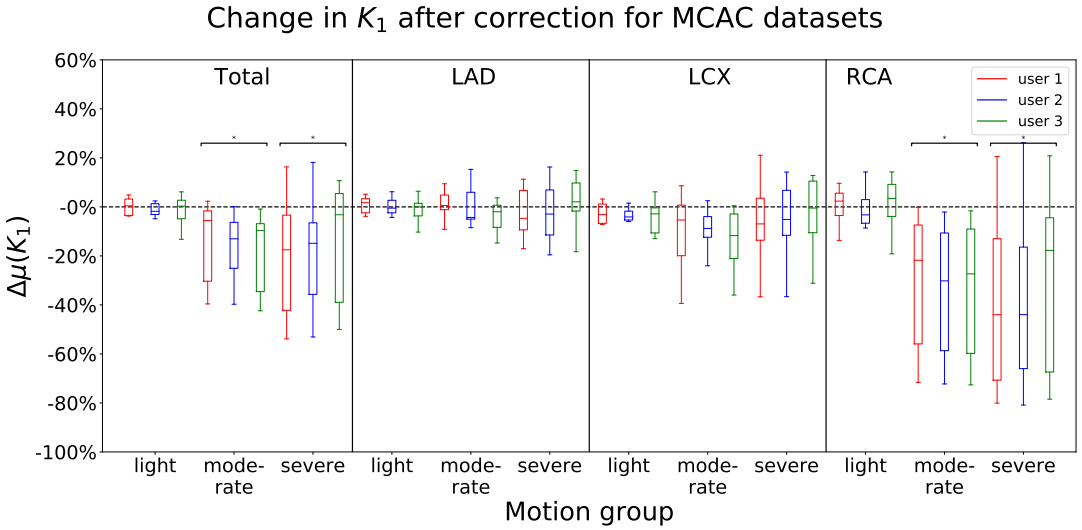


Figure 4.5: Change in μ_{K_1} after motion correction for datasets corrected with MCAC. The value μ_{K_1} is plotted for the whole myocardium (Total) as well as for each territory. The difference was statistically significant for Total and RCA territory in moderate and severe motion groups (p-value < 0.05).

Figure 4.5 shows the change in K_1 values after motion correction. The highest changes were observed for severe and moderate motion. The territory most affected by motion was the RCA, with up to 90% reduction in K_1. Statistical differences (p-value < 0.05) were found for severe and moderate motion datasets in the RCA territory as well as the whole myocardium.

The distribution of $\sigma_{K_1 MC}$ and $\sigma_{K_1 MCAC}$ were found to be very similar, as shown in Figure 4.6. For the 3 motion groups, the average values were indistinguishable for the whole myocardium (p-value>0.05). Similar conclusions were observed when comparing μ_{K_1} after MC and MCAC, (p-value>0.05).

Figure 4.7 details the values of CV_{K_1} for observer 1 before and after MCAC. As expected, before correction CV_{K_1} was largest in groups with greatest motion. The distribution of CV_{K_1} values after MCAC was very similar across all motion groups with, for example, mean values for the total territory of 10.3% in group 0 and 8.0% for group 1, 11.5% for group 2 and 10.4% for group 3. Similar results were found for individual territories. While there were some individual differences for a given dataset between the three observers (data not shown), no major difference were found in CV_{K_1} between the three observers. Similar observations made about

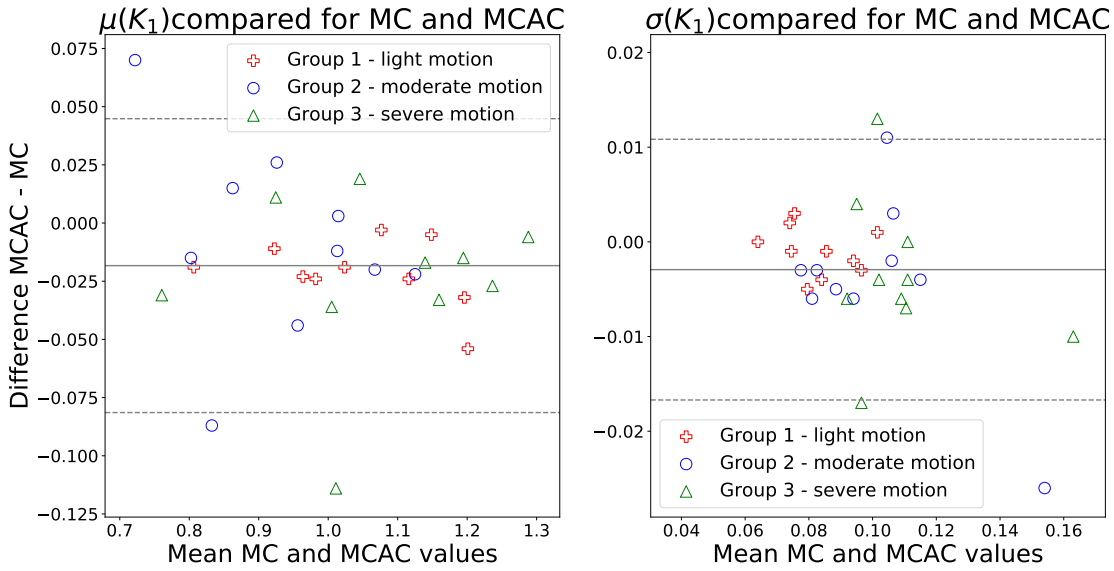


Figure 4.6: Bland-Altmann plot of σ_{K_1} and μ_{K_1} values after motion correction for MC and MCAC datasets for observer 1. No statistical differences were found between MC and MCAC correction for each motion group (p-value > 0.05)

Figure 4.4 can be done about the change of CV_{K_1} after motion correction: the uncertainty reduced most for high and moderate motion, while being more modest for light motion. The combined information from Figure 4.4 and 4.7 show that the uncertainty was reduced after motion correction (see Figure 4.4) and brought to a level similar to motion free datasets (see Figure 4.7), for every motion group.

4.4 Discussion

In this work, the reliability of motion correction was evaluated by comparing the uncertainty of the K_1 parameter (used here as a surrogate for MBF) extracted from the Variational Bayes framework.

Motion correction was shown to have significant impact on K_1 values (see Figure 4.5), so by extension on MBF values as well. The most affected coronary territory was the RCA, which is in agreement with previous studies [65, 67]. This is explained by the type of motion in dynamic cardiac PET, as observed in Figure 4.2. The motion is more pronounced and frequent in the vertical long axis direction, which creates misalignments of both the myocardium segmentation and the ROI

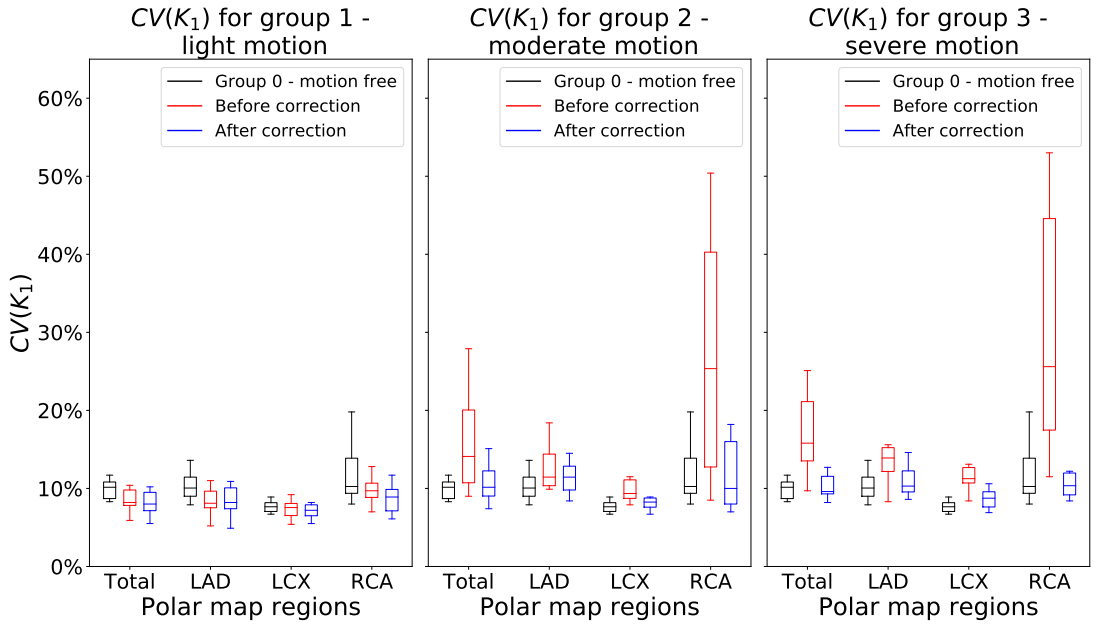


Figure 4.7: Comparison of CV_{K_1} values before and after motion correction for MCAC datasets for observer 1. The values of CV_{K_1}, across motion groups after motion correction, were similar to motion free values for each territory.

sampling the input function. That misalignment increases the spillover parameter S and the K_1 parameter.

In order to assess the effectiveness of manual motion correction carried out by expert clinical observers, the VB methods were applied to clinical data and an uncertainty metric was calculated before and after motion correction. This was done first by comparing the values of σ_{K_1} (Figure 4.4), before and after motion correction. A significant drop in σ_{K_1} was observed for moderate and high motion for both the total myocardial tissue and the RCA territory, with a trend toward reduction in the other territories. Values of CV_{K_1} after motion correction were also compared to CV_{K_1} before correction and those values found from motion-free datasets (Figure 4.7). The results showed that the level of uncertainties variations after motion correction were similar between each motion group and reduced to a comparable level to the data judged to be motion free. This tends to imply that for the data considered here, manual motion correction was effective for all levels of motion corruption seen and could produce parametric estimates with comparable accuracy to data not corrupted by substantial motion.

The proposed uncertainty measure was applied in the comparison of two different motion correction schemes, namely MCAC and MC. With regards to uncertainty in σ_{K_1}, no significant difference between the two strategies was seen. The fact that the two methods produced similar uncertainties could mean that the difference between the two schemes is less important than other factors that contribute to the uncertainty; for example, time sampling, intra-frame motion, scanner noise etc. While MCAC would be the preferred correction scheme as it takes into account the misalignment between the PET series and the attenuation map, our results tend to imply that the MC scheme is sufficient with regards to uncertainty measurement. However, in order to generalize to the wider problem of when correction should be applied, more experiments would be needed, for example exploring the use of deformable registration instead of rigid body translations for attenuation correction.

By measuring the uncertainties, before and after motion correction, we have obtained a quantitative metric that could measure the effectiveness of motion correction. This might be especially useful for high motion cases, where the change was observed to be significant and thus provided reassurance that motion correction had been effective. It is also worth noting that while individual observers corrected the datasets to obtain the best visual assessment of motion-free images, the final $\Delta\sigma_{K_1}$ was not the same for each observer. In this scenario the uncertainty measure could potentially be used to choose the most effective correction and hence used as an objective way to determine whether motion correction has been of value or to compare different motion correction algorithms. An example of the uncertainties before and after motion correction for the three different users is shown in Figure 4.8. This could be especially useful when automated motion correction is employed allowing the user to objectively determine whether it has been successful.

A potential limitation of this study is the choice of a white noise model for representing the corruption induced by motion. While it would be very difficult to describe mathematically the noise induced by motion, it is certainly more complex than a Gaussian distribution as used in our model. However, in a similar manner to the observation in the simulated data where a higher level of noise

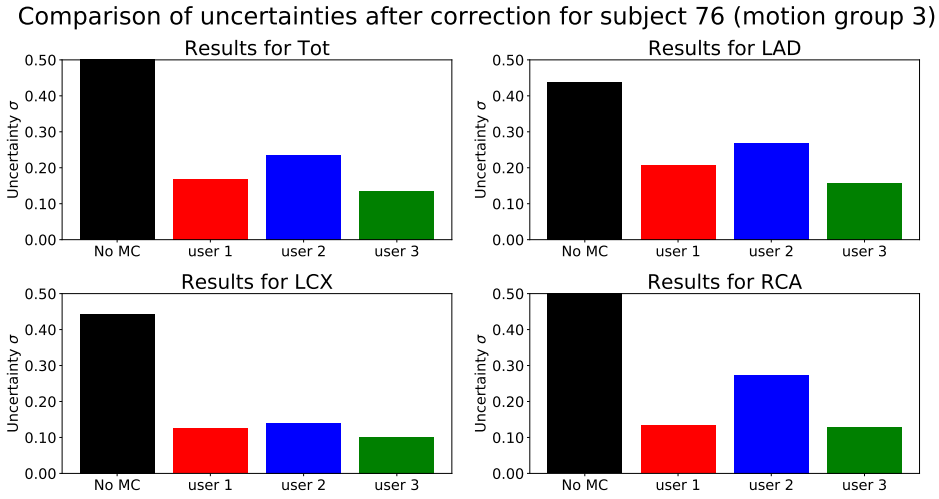


Figure 4.8: Comparison of uncertainties after correction of three manual realignments. Each graph shows the uncertainty of a group of TACs (either the whole myocardium or a coronary territory), before correction and after manual realignment of each user. While a significant reduction can be seen after correction of each user, a difference can also be seen among users, user 3 providing the smallest uncertainty.

correlated with higher uncertainty (see chapter 3), a similar correlation between higher level of motion and higher uncertainty was observed in the clinical data (Figures 4.4 and 4.7). The fact that the decrease of uncertainty correlates with the independent categorization of the data, implied that our assumption may be valid as far is required for this study.

This study focused primarily on the assessment of motion correction; there are numerous other factors that also have an influence on the reliability of estimated K_1 values. We observed, for example, that in regions with limited tracer signal in the late frame, which is characteristic of ischemic or infarcted patients, the uncertainty of K_1 was higher than in healthy tissues. This might influence the evaluation of motion correction using an uncertainty measure, because the change in uncertainty caused by motion may be smaller than for patients without the pathological regions of high uncertainty. The flow uncertainty in abnormal scans is further discussed in chapter 6. Those observations are corroborated by the values of CV_{K_1} for light motion datasets before and after motion correction (Figure 4.7). The values did not change significantly and were in the range of CV_{K_1} for motion-free images,

indicating that the remaining uncertainty was dominated by other factors and that the influence of motion correction becomes less noticeable. For this data it might thus be possible to give an absolute threshold of K_1 uncertainty on which we could base a quality control metric. In this chapter we have used uncertainty to compare different motion corrections in order to judge whether motion correction is required or has been successful. In principle it might be possible to establish a threshold on the K_1 uncertainty more generally on which to base decisions regarding motion correction. However, there are various sources of artifact and noise that contribute to uncertainty and these will vary from dataset to dataset, making it challenging to derive such a threshold from this study alone, and requires further studies with additional datasets.

Although not a limitation in this study, one issue with the proposed method in its current implementation is the inability to translate the K_1 uncertainty into MBF uncertainty. Whilst a reduction of uncertainty in K_1 will be translated in a reduction in MBF uncertainty (cf. equation 4.1), it is more challenging to translate K_1 uncertainty into that for MBF for any given image, since other parameters play a role in the translation of K_1 into MBF as shown in previous studies [61]. However quantitative MBF uncertainty is not a necessity for the purpose of assessing motion correction reliability; and therefore for this limited purpose K_1 can be used a surrogate for MBF.

4.5 Conclusion

In this chapter, we saw that the uncertainty metric on K_1, derived from the VB methods, could be used to assess the quality of motion correction. For high and moderate motion datasets, the uncertainty reduced significantly after motion correction for the whole myocardium and the RCA territory, indicating the success of the applied motion correction. Furthermore, while the reduction in uncertainty was similar for most datasets for the three different experts, it showed some differences as well, indicating that the metric was sensitive to differences between users.

5 Towards automated motion correction

5.1 Introduction

It was demonstrated in the previous chapter, as well as in the literature review, that patient motion plays a major role with regards to the accuracy and reliability of myocardial blood flow (MBF) estimates. A common pattern of patient motion is a vertical shifting along the anterior-posterior direction of the myocardium. Specifically, motion leads to a misalignment of regions of interest (ROI) that are used to model the myocardial blood flow (MBF) with dynamic PET, as shown in Figure 5.1. This misalignement creates significant inaccuracies in the MBF quantification, as well as higher uncertainties of the MBF estimates. Unlike technical factors affecting MBF estimates, which can be possibly controlled with training and standardization, patient motion is unavoidable. It is therefore necessary to provide a method that can correct the motion in order to improve the reliability of MBF studies with dynamic PET.

Motion correction of dynamic cardiac PET is a complex task for two main reasons. First, the spatial distribution of the tracer changes significantly over the PET series. The tracer is sequentially observed in the right ventricle (RV), the left ventricle (LV) and the myocardium tissue (MT), as shown in Figure 5.1. These changes of pattern make any kind of realignment between structures challenging

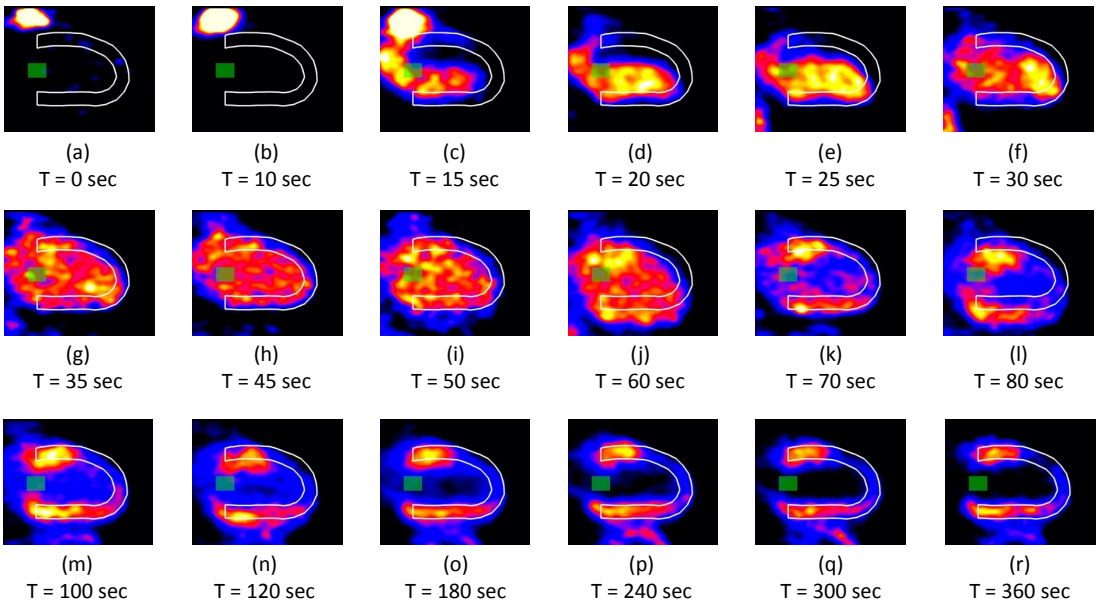


Figure 5.1: Temporal evolution of the tracer in the heart. Each image is a view of the horizontal long axis of the heart. The caption indicates the starting time of the frame acquisition relative to the beginning of the scan. A different windowing level was used for each frame in order to facilitate visualization. The tracer is observed initially in the RV ((a) and (b)) then in the LV cavity ((c) to (f)) then the LV and MT ((g) to (j)) and finally the MT only ((k) to (r)). The short acquisition frames of the early phase ((a) to (j)), as well as the pattern changes throughout the scan make any kind of motion correction a difficult task. The effect of motion can be seen by observing the misalignment of the two ROIs (white contour defining the MT and green rectangle sampling the arterial blood) in the frames (c) to (e), (j), (l) and (n).

as they might not be visible from one frame to another. Second, the frames have a low signal to noise ratio, due to short acquisition times in early frames. While this is necessary to capture the rapid evolution of the tracer, it makes any type of contouring and registration more difficult.

Various methods are available to correct for motion in dynamic PET. Some studies proposed to use external markers [66, 131], which are considered as reference landmarks to guide the correction of the overall scan. While those methods are certainly valid, they require additional preparation in the scanning protocol, which we did not have control over for the data available. Furthermore, in order to be widely available in clinical routine, data driven motion correction methods (i.e. that do not rely on external devices) are preferred. Other approaches automatically correct the motion in the late frames of the PET series, since the tracer distribution is stabilized

in the MT; using image registration based techniques [67, 98]. The drawback of these methods is their inability to realign the early frames. Realignment of early frames has been shown to play an important role in the accurate quantification of MBF with dynamic cardiac PET [65], and there is yet a solution for correcting these frames automatically to be found. Finally is is also possible to realign each frame manually [65]. This approach was chosen, for example, in the previous chapter. While it has the advantage of correcting the entire duration of the PET scan its reliability and accuracy depends on the user expertise and interpretation. Furthermore, the manual realignment of each individual frame is time-consuming. In the experiment conducted in the previous chapter, it could take up to 10 minutes for an expert user to realign the frames manually, which is not acceptable in a clinical setting.

To the best of our knowledge, only one study detailed an automated motion correction framework for the entire course of a dynamic cardiac PET scan [86]. This framework relies on the generation of two templates, one representing the MT and one representing the LV. These two templates are created from a segmentation of a single static image in the late phase of the dynamic scan. Each frame of the PET series is then matched and registered to one of the two templates, depending on the tracer spatial distribution of that given frame. While this study did not provide exhaustive validation, evaluating the methods on only three subjects, it developed an interesting concept with promising results. They showed that the model fitting residuals was smaller in one subject out of three, and that motion seemed to be removed when assessing visually the dynamic PET series in the three cases.

The work presented in this chapter is an extension and an improvement of the method in [86], previously described. Specifically we developed a framework to correct for motion over the whole duration of the dynamic PET scan, using image-based registration methods. The reference images used for correcting the motion is a series of templates that change progressively throughout the scan to match the spatial distribution of the tracer. The framework is then evaluated on 40 datasets with various magnitudes of motion. The results are assessed qualitatively by an expert user and quantitatively by comparing uncertainties in K_1 before and after correction.

5.2 Methods

5.2.1 Motion correction framework

The motion correction framework relies on generating template images adapted for each patient and each frame of the dynamic series. Taken sequentially, the template frames are motion free, hence by registering the dynamic PET series to the dynamic template, the motion would be corrected. This motion correction framework can be separated into two main steps:

1. The motion correction technique begins by detecting automatically the myocardium tissue (MT) from a frame with a long duration in order to maximize the signal to noise ratio. The frame used is the average of the last 3 frames of the dynamic series, which is 3 minutes long and taken 3 minutes after the beginning of acquisition. From that frame, the heart is segmented and reoriented into the cardiac planes (short axis, vertical and horizontal long axes). Each frame of the dynamic series is reoriented with the same orientation matrix. From the segmentation of the MT, a 3-dimensional template is generated with 3 different labels, corresponding to the main structures seen in the dynamic series: MT, left ventricle (LV) and right ventricle (RV).

2. Each frame of the dynamic PET series is pre-processed with a 3D spatial band-pass filter; and is matched to a corresponding template, derived from the main template previously generated, where only visible structures (i.e. LV, RV and MT) on the PET frame are kept.

Details of the motion correction framework are given in the following sub-sections.

5.2.1.1 Myocardium segmentation and template generation

The first step is to detect the myocardium within the dynamic PET series, in order to register the images using only a subregion of the whole scan. The myocardium is oriented and segmented by applying the automated heart detection and segmentation

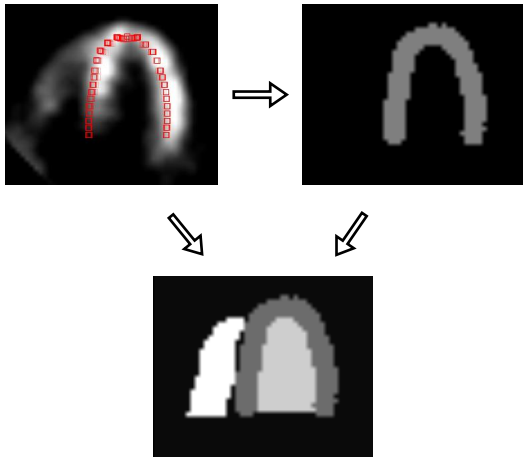


Figure 5.2: Patient specific myocardium template generation. The first step of the motion correction framework is to create a patient-specific myocardium template. This template is generated from the segmentation of the myocardium tissue of a late static image, as represented by the red squares of top left image. Details of the segmentation algorithm are provided in appendix A. A binary mask representing the myocardium tissue (dark gray, top right) is obtained. From this mask, the left ventricle cavity (light gray) can be defined with a simple closing operation. Finally, the right ventricle can be segmented by thresholding high uptake values on a sub-region of the image (e.g. the left of the myocardium for this view; bottom image).

algorithm of syngo.PET MBF [128]. Each segmentation was reviewed visually to ensure that artifacts did not occur.

The segmentation extracted from syngo.PET MBF is a list of 540 3-dimensional points from which TACs are extracted. A sphere of 5mm was generated around each of those 540 points, to create a 3D mask of the myocardium, as shown in Figure 5.2. Details of the segmentation algorithm for the myocardium are given in appendix A. The LV cavity was obtained by applying a 3D flood-fill algorithm, restricted to the bounding box of the myocardium segmentation, and using the center of the myocardium as a seed point. Finally the RV was obtained by applying a threshold of 50% of maximum uptake in a rectangular ROI outside the myocardium where the RV is located. Another closing operation was applied to fill the RV cavity. The different steps of that template generation are illustrated in Figure 5.2.

Once the main 3D template with 3 different labels (RV, LV and MT) is generated for a given dataset, multiple masks can be generated, according to the spatial distribution of the tracer for a specific frame. For example, if uptake was only

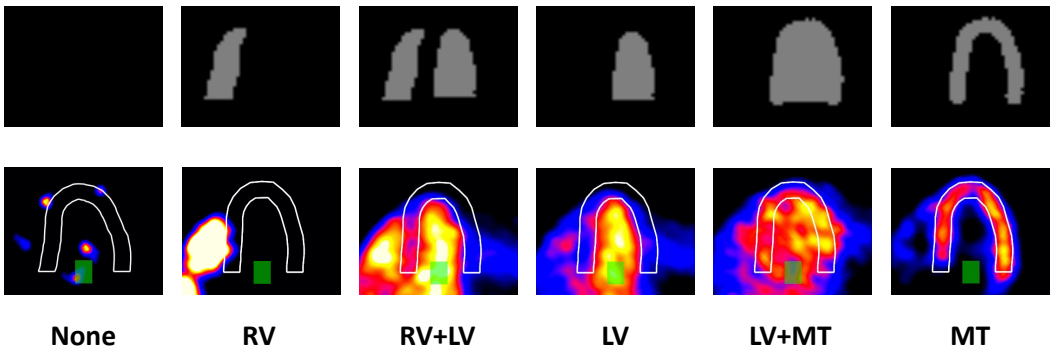


Figure 5.3: Frame specific mask generation. Depending on the spatial distribution of the tracer in the frame studied, a different template is created, generated by combining the different labels of the myocardium mask previously described (see Figure 5.2). Out of three different labels, 8 different templates could be generated theoretically. However, from the 40 datasets observed, only 6 combinations of tracer distribution were found, which are shown here.

visible in the myocardium for a given frame, only the MT label is kept for the template to which that frame is registered. From our observations, 6 different templates were generated: MT, MT+LV, LV, RV+LV, RV and None (no uptake). The different templates with a corresponding frame example are shown in Figure 5.3.

5.2.1.2 Frame pre-processing and template alignment

A band-pass filter was applied to each frame of the dynamic PET series, in order to isolate high uptake regions in the image. This band-pass filter consisted of a difference of 3D Gaussians smoothing filters. The kernel of each Gaussian filter was isotropic. The sigma value of the first Gaussian kernel was set to 12mm, while the second Gaussian kernel was set to 6mm. The negative values in the image were thresholded to 0. An example of a late frame on which a difference of Gaussians has been applied can be found in Figure 5.4

For each dataset, we annotated manually the different frames with a label describing the spatial distribution of the tracer, so that the corresponding template could be used for registration.

The 3D cross correlation was used as a similarity metric to register the frame, on which the band-pass filter was applied, to its corresponding template:

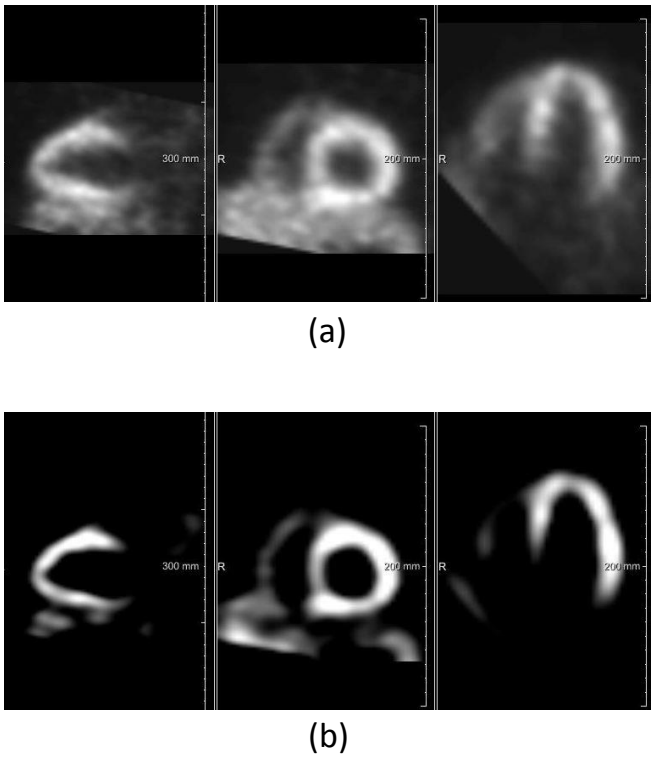


Figure 5.4: Comparison of original image with a difference of Gaussians filtered image. Images are shown in the three cardiac planes (from left to right: horizontal long axis, short axis and vertical long axis)
(a) Original last frame of a dynamic cardiac PET scan.
(b) The same frame on which a difference of Gaussians filter was applied. The sigma values of the two Gaussians functions were set to 12mm and 6mm.

$$CC(i,t) = \frac{\sum_{x\in\Omega}(i(x)-\bar{i})(t(x)-\bar{t})}{\sqrt{\sum_{x\in\Omega}(i(x)-\bar{i})^2}\sqrt{\sum_{x\in\Omega}(t(x)-\bar{t})^2}} \tag{5.1}$$

Where: i is the difference of Gaussians image that needs to be registered to the template t; $\bar{i}$ is the mean value of the image i; and Ω is the common domain definition of the image i and the template t.

For the MT templates, 3D translations were used to maximize the similarity metric. For the other templates (MT+LV, LV, RV+LV and RV), only 2D translations in the short axis plane were used. This is to avoid artifacts due to the definition of template. For example the RV and LV masks are defined from the late phase of the image, which could have a different shape in the earlier frames, especially in the long axis direction.

The outcome of this framework is a list of 3D translations corresponding to the motion extracted in each frame. The motion vectors were then applied to the original dataset. The resulting dataset is referred as the motion corrected series.

5.2.2 Data

Fourty datasets were used for this study, which are the datasets used in the previous chapter. Since manual realignment of these frames are available for these dynamic datasets, they can be used as a reference. As a reminder, datasets were categorized into four motion groups: no motion (group 0), light motion (group 1), moderate motion (group 2) and severe motion (group 3), according to [44]. Ten datasets were selected for each group.

5.2.3 Evaluation and comparison to other methods

The motion correction framework was run on each dataset with the same parameters, in order to verify that the motion correction framework would behave similarly for different motion magnitudes and that the algorithm would not introduce artifacts in motion- free datasets (group 0).

To evaluate the importance of applying motion correction to the early frames, a second motion corrected dataset was generated, where only the late frames were automatically corrected. The algorithm used for this work was the automated motion correction algorithm of syngo.PET MBF [128]. The concept of this algorithm is to use the last frame as a reference image, and use a cross-correlation based registration algorithm to register all late frames (i.e. frames where acquisition started between 3 and 6 minutes after tracer injection). The motion vectors extracted for the first frame of the late frames, were propagated to the earlier frames. The resulting data is referred to here as the late frame corrected series.

Finally, the results were also compared to the datasets realigned manually by the three users, as presented in the previous chapter.

5.2.4 Analysis

5.2.4.1 Qualitative assessment

Original and corrected datasets were assessed by a clinical expert. Datasets were blinded, so that the reader did not know which dataset was motion corrected and which dataset was the original. The motion of each dataset was independently rated with a 4 point scale in a similar manner to how the datasets were previously graded: 0 was considered as motion free, 1 was light motion, 2 was moderate motion and 3 was severe motion. However, the clinical expert who reviewed these datasets did not belong to the same institution that initially graded the images. Consequently, datasets could be rated differently, e.g. a dataset could be graded initially as severe motion (group 3) but only moderate motion (group 2) by our clinical reader (and vice versa), as this assessment was only qualitative. The assessment was based on the MBF maps.

Additionally, a score on a 5-point scale was given regarding the quality of motion correction algorithm as follows:

- -2: the motion corrected series was significantly worse than the original
- -1: the motion corrected series was moderately worse than the original dataset
- 0: both dynamic series were comparable
- 1: the motion corrected series was moderately better than the original dataset
- 2: the motion corrected series was significantly better than the original dataset.

Another score, using a 3-point scale, was given to the clinical impact of the motion correction:

- 0: The correction did not change the diagnosis.
- 1: The correction might have changed the diagnosis, depending on additional information (e.g. by reading the rest data).
- 2: The correction definitely changed the diagnosis.

Finally, a vote on the dataset that provided the most confidence to give a clinical interpretation was given, with the possibility to vote for either one series or both in case there was no difference.

5.2.4.2 Quantitative assessment

To evaluate the impact and the quality of the automated motion correction, the same methodology was used as in the previous chapter.

Each dataset was processed with the syngo.PET MBF [128] in order to extract time-activity curves (TACs). The default automatic motion correction was disabled for the processing for Original and MC series. For each dataset 540 TACs were extracted, and regrouped into 4 categories: the 3 coronary territories (LAD, LCX and RCA) and the whole myocardium (Tot). Each TAC was fitted individually with the VB method with a one-tissue compartment model.

For each TAC, the means μ_{K_1} and standard deviations σ_{K_1} of the parameter K_1 were averaged across the whole group of TACs.

The effectiveness and reliability of the motion correction framework was evaluated by computing the normalized difference of the estimated uncertainty of K_1:

$$\Delta\sigma_{K_1} = \frac{\sigma_{K_{1_C}} - \sigma_{K_{1_O}}}{\sigma_{K_{1_O}}} \tag{5.2}$$

where: $\sigma_{K_{1_C}}$ is the averaged uncertainty on K_1, for the group of TACs studied, after correction and $\sigma_{K_{1_O}}$ is the uncertainty on K_1 before correction. If $\Delta\sigma_{K_1} < 0$ the correction has reduced the uncertainty of K_1, i.e. motion correction resulted in a higher confidence in K_1. In other words, a negative change means that the motion corrected image provided a more reliable estimate of K_1 than the original series. A paired t-test was performed in order to evaluate the statistical difference between $\sigma_{K_{1_C}}$ and $\sigma_{K_{1_O}}$.

5.3 Results

One dataset was removed from the result since the automated segmentation of the myocardium did not succeed, which prevented the patient-specific template generation. This dataset was categorized as light motion.

The clinical reader categorized the 39 remaining original datasets as follows: 4 datasets had severe motion (group 3), 19 had moderate motion (group 2), 10 had light motion and 6 were motion free (group 0), showing some disagreement with the initial grading.

Of these, 16 corrected datasets (41%) were rated as significantly better than the original series, 17 corrected datasets (44%) were rated as moderately better than the original series, 4 corrected datasets (10%) were rated as comparable with the original series and 2 datasets (5%) were rated as moderately worse than the original series. No datasets were rated as significantly worse than the original data. The 4 datasets that were categorized as comparable were rated with no motion (group 0), implying that no artifacts were introduced by implementing motion correction. The 2 datasets that were rated as moderately worse were shifted from motion free to light motion, implying that the motion correction framework might have introduced some artifacts. Out of the 33 datasets where motion correction had a positive impact, 12 datasets were categorized as motion free (group 0) after correction, while 16 corrected datasets were categorized as light motion and 5 were categorized as moderate motion (group 2). There were 5 corrected datasets rated as moderately better than the original series, even though both corrected and original series were rated with the same motion score (4 were with moderate motion and 1 with light motion). The reader commented on that observation by noting some improvement in the motion corrected series, but the motion could not be corrected to such extent that it would be graded with a lower scale. In addition, a definite change of diagnosis was observed for 12 datasets (31%). The reader stated there might be a change of diagnosis for 9 datasets (23%), while no change of diagnosis was found for 18 datasets (46%). Finally, the reader expressed a preference to interpret the results with the corrected dataset for 31 subjects

(79%). According to the reader there was the same confidence for the remaining 8 datasets (21%). The original series was never preferred.

Figure 5.5 compares the reduction in uncertainty of K_1 for the motion correction framework developed in this work with the results of the three users presented in the previous chapter, as well as with the correction framework that focuses only on late frames. Overall, the reduction of uncertainty for our motion correction framework was comparable with the results obtained from manual realignments, for every territory as well as every motion group. The correction framework presented here showed a significant reduction in uncertainty in the case of severe motion cases for the whole myocardium and the RCA territory, which was not the case for the late-frames motion correction framework. Additionally, for the RCA territory in severe motion cases, the framework presented provided higher reduction in uncertainty, compared to manual realignments of two users (user 2 and 3). The late-frames correction algorithm showed a reduction in uncertainty for each group and each territory, however it was only significant for moderate motion datasets. This finding correlates with the initial categorization [44], since one of the criteria of categorizing datasets in moderate (group 2) or severe (group 3) motion was the ability of the automated motion correction of syngo.PET MBF software to correct for it.

5.4 Discussion

We presented in this work a motion correction framework that aims to correct the entire timecourse of the dynamic PET series. The algorithm relied on a dynamic series of templates, adapted for each patient, and designed from the segmentation of a late static frame. These templates were matched to the spatial distribution of the tracer in each frame of the PET series. Consequently, by registering individually each frame to its corresponding template, the motion was extracted for each frame, and could then be corrected. A similar idea was detailed in [86], but we could not replicate their methods with satisfying results. Thus, modifications have been applied in order to improve the accuracy of the motion correction. In particular,

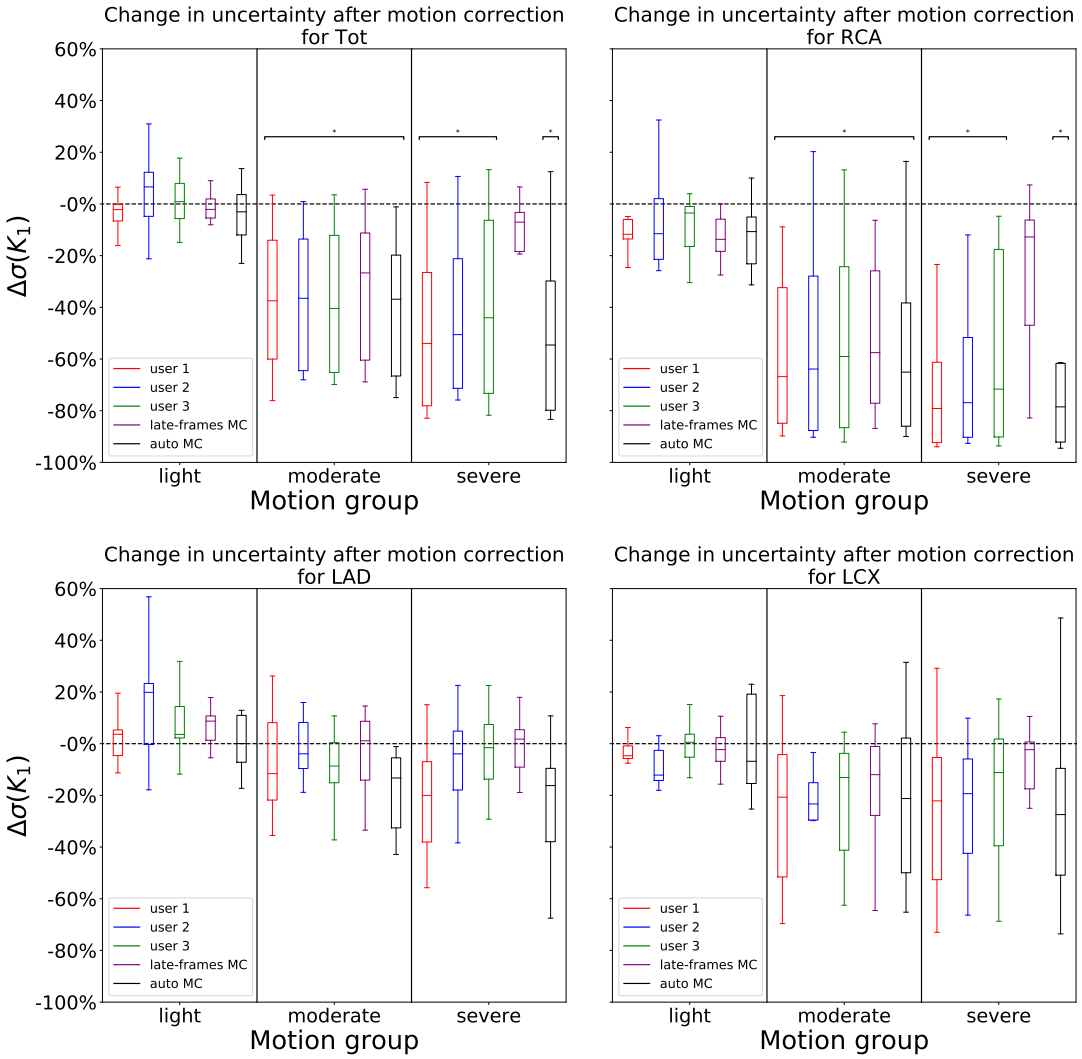


Figure 5.5: Change in uncertainty after motion correction for datasets corrected with manual realignment and automated motion correction strategies. Each graph shows the value $\Delta\sigma_{K_1}$ for the whole myocardium (Total) or a coronary territory (RCA, LAD and LCX). A global decrease can be seen for each motion group. The automated motion correction framework performed similarly to the three users for each motion group and in each territory. The difference in uncertainty between $\sigma_{K_{1_C}}$ and $\sigma_{K_{1_O}}$ was statistically significant for Total and RCA territory in groups 2 and 3 (p-value < 0.05). The reduction in uncertainty was comparable between users (median and quartiles are similar) and our correction framework. Unlike our motion correction framework, there was not a significant difference between $\sigma_{K_{1_C}}$ and $\sigma_{K_{1_O}}$ for the late-frames motion correction algorithm in severe motion (group 3).

we introduced the RV label in the mask generation, and used a band-pass filter on the PET series to improve the registration accuracy.

The method was first evaluated qualitatively through a blind review process. An

independent expert reader evaluated both original and corrected datasets without knowing which series our framework was applied to. The qualitative assessment showed the benefit of our algorithm in many aspects. First, an improvement was observed in the majority of cases, with 88% of the datasets which were classified as either significantly or moderately improved. The motion corrected dataset was preferred in 31 cases, the 8 remaining cases provided equally confident interpretation. The algorithm showed an improvement when comparing to the original series, and also managed to correct for motion in 12 datasets out of 33, and leaving 16 datasets with light motion. For 5 cases, the motion correction framework did not manage to correct a substantial fraction of the motion, since they were categorized as moderate motion after correction. The corrected datasets were, however, preferred for those 5 subjects, as well as rated moderately better than the original series, demonstrating some improvement from our method. One reason these datasets were categorized as moderate motion after correction is due to the fact that some frames had pronounced intra-frame motion artifacts. Since data were reconstructed with a fixed temporal sampling, the motion correction framework was not able to correct for intra-frame motion. An example of such a case is shown in Figure 5.6. In the context of our work, it might not be possible to significantly improve the results in these cases, as this would require reconstructing the PET data on a finer time grid. Reconstructing the PET signal on a finer time grid is challenging with the current scanners as the signal to noise ratio (SNR) at peak enhancement is between 5 and 10 for the blood region, but below 5 for the tissue region, for a 5 seconds frame duration. Therefore, while reconstructing the data with shorter frames may be beneficial for correcting motion more accurately, the quality of the images may affect the quality of the image registrations. Finally, the algorithm introduced some artifacts for 2 cases, since these 2 datasets were graded motion-free before correction and with light motion after correction. While this would not be desired in a general case of motion correction, we believe that in the context of clinical software where the motion correction algorithm could be an option to activate or not, the user would not necessarily activate this option after review of the original

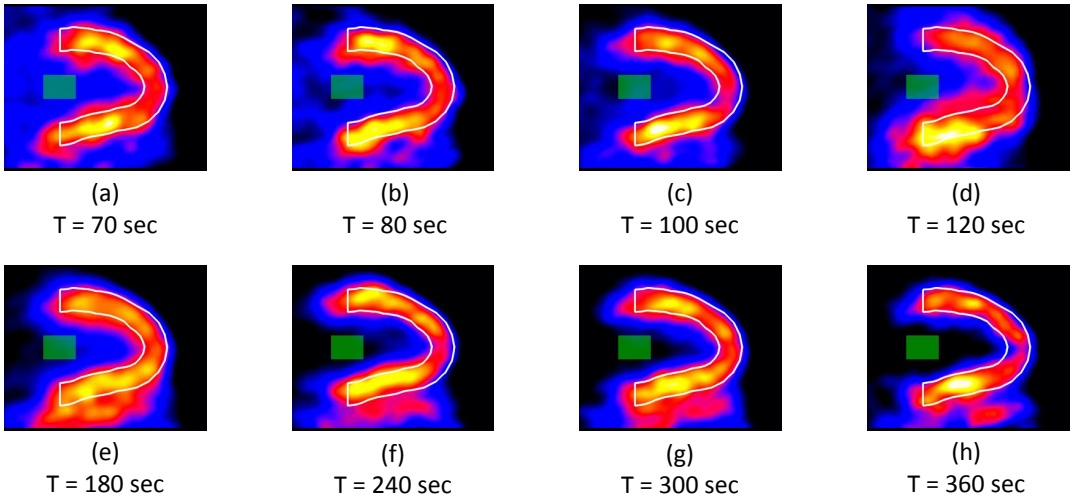


Figure 5.6: Example of intra frame motion artifacts. Each image is one of the last 8 frames of a dynamic scan. The caption is the starting time of the acquistion of that frame relative to the beginning of the scan. The 8 frames show uptake in the myocardium, and the MT segmentation is well aligned from frame (a) to frame (h). However for frames (d), (e) and (g), some artifacts can be seen in the inferior wall, due to intra-frame motion. In order to correct for such effects, the data would have to be reconstructed with a finer time scale.

dynamic series. Furthermore, the user could possibly return to the original series after running the algorithm in case the results were not satisfactory.

We also evaluated the motion correction framework by calculating the uncertainties in K_1 before and after motion correction. Uncertainty reduced significantly for severe motion cases and moderate motion cases, for both the Total myocardium tissue as well as the RCA territory. Reduction of uncertainty was also observed for the late-frames correction framework, but the difference between $\sigma_{K_{1_o}}$ and $\sigma_{K_{1_c}}$ was not significant for severe motion cases. This justifies the need for a method that is able to correct for motion in the entire timecourse of the dynamic PET scan. The method was also compared to the manual realignment performed by three expert readers, whose correction can be considered as a reference standard to aim for. Our method provided similar performance in terms of uncertainty reductions compared to the manual realignment of the users for every motion group and every territory, demonstrating the reliability of the framework. Finally, using the uncertainty measurement to assess the success of the motion correction framework

was advantageous for this work as it provided a quantitative metric that could be used to compare to different parameter settings. For example, the first study which aimed at correcting the early frames [86] did not apply any kind of pre-processing to the PET data. It was found in our case that applying a spatial band-pass filter helped to improve the registration of the late and early frames, as shown by assessing the uncertainty of K_1 after correction with and without a band-pass filter (data not shown). Other parameters such as the size of the Gaussian kernel for the difference of Gaussians, or the choice of the similarity metric between the PET data and the template were tested with different values and assessed using the uncertainty measure, in order to obtain the settings detailed above.

The framework developed in this chapter relies on the generation of a patient-specific template with 3 different labels corresponding to the three main structures visible through a dynamic PET scan: RV, LV and MT. Consequently for the proposed method to succeed, it is necessary that the automated segmentation of the myocardium runs successfully. In our work, the segmentation failed in only one case. This problem could be overcome in clinical software, where the user can manually define some key landmarks such as the base or the apex of the myocardium, so that the segmentation algorithm is facilitated. Additionally, the segmentation of the LV cavity and the RV are done on a late static image. This is necessary in order to create a consistent reference dynamic series. However it is not guaranteed that the shapes of these structures would be exactly identical at different time points, which could create some artifacts in the overall frame realignment. While not being essential in the overall correction framework, this issue remains to be solved.

A potential limitation of this study is the assignment of a frame to its corresponding template, since this step was done manually in this case. In order to fully automate the process, each frame would have to be associated with its label without any user input. A first solution to circumvent that problem would be to exhaustively generate every combination of template for the series and register them to each frame. The uncertainty metric could then be used to determine, which template provides the minimum uncertainty. While this is theoretically feasible,

this is not practical as the number of possibilities increases exponentially with the number of frames. It is however possible to refine the number of possibilities by using prior knowledge on the temporal evolution of the tracer. For example, the first frame of the sequence has either no uptake visible (template = None) or uptake in the RV. Similarly the last frame will have uptake only in the MT. Consequently, the number of possibilities could be reduced to a smaller subset, based on the temporal information of the frame studied. In addition, the frames follow a sequence, similarly to a state machine: if the current frame studied is LV, then the following frame will be either LV again or LV+MT (providing the scan is without any artifact, for example, there is no double pass of the tracer). These two approaches combined could reduce the number of possibilities to only a limited number of template sequences. Alternatively, other methods could be used in order to classify automatically the frames, such as calculating the center of mass of the frame to detect any change of spatial distribution of the tracer, or comparing the value of the similarity metric after optimization by the registration algorithm for each frame and keep the template with the highest similarity metric. Both of these methods have been tried but did not show any convincing results. It is challenging to observe a change in center of mass when switching from LV to LV+MT, or differentiate the correlation factor between an RV and LV phase as they have similar shapes. Finally, some machine learning techniques such as convolutional neural networks could be used to classify automatically the frames. Considering the technicality of such methods as well as the relative limited number of datasets, this work was considered outside of the scope of this chapter.

5.5 Conclusion

In this chapter we developed a framework which aims to correct for the motion in a dynamic PET scan, including its early phase. Through this framework, it has been shown that motion correction of early frames is not only necessary but also feasible, by registering the PET series to a series of various templates, adapted

for each subject. This work is a contribution towards fully automated motion correction in dynamic cardiac PET.

6 Uncertainty and spillover effect

6.1 Introduction

The initial motivation of this work was to derive a quality metric, which could be used to assess the quality of a motion correction technique in dynamic cardiac PET. In Bayesian frameworks, model parameters are represented as probability distributions, from which an uncertainty metric can be drawn. Therefore a quality metric based on the uncertainty of the kinetic parameter K_1 was chosen. It has been demonstrated that patient motion plays a major role in the magnitude of uncertainty, as shown in chapter 4. However throughout chapter 3 and 4, it was observed that other factors would influence the uncertainty on K_1. One of them is the partial volume effect, also called spillover effect. The spillover effect arises due to the segmentation of the myocardial tissue, which is based on the later phase of the dynamic PET series. In the early frames, the arterial blood is located only in the left ventricle cavity. However due to the resolution of the PET image, as well as cardiac and respiratory motion, the arterial blood of the left ventricle cavity overlaps with the myocardium segmentation, creating a partial volume effect, as shown in Figure 6.1.

This spillover is taken into account in the equation that models the activity measured in the myocardium $C_m(t)$, through the S parameter:

$$C_m(t) = (1-S)C_{tiss}(t)+SC_a(t) \qquad \text{with:} \quad C_{tiss}(t) = K_1 \exp(-k_2 t) \circledast C_a(t) \quad (6.1)$$

where: $C_a(t)$ is the arterial blood activity and K_1, k_2 are the kinetic parameters representing the exchange rate of the tracer between the blood and the tissue, as explained in chapter 2, section 2.2.3.

The spillover effect has an impact on the accuracy of the kinetic parameters that are fitted. A high S implies less signal to extract information from $C_{tiss}(t)$, leading to higher uncertainties on the kinetic parameters. Figure 6.2 shows three polar maps representing the K_1 parameter, the K_1 parameter uncertainty and the S parameter. It can be noticed that regions of high S correlate with regions of high uncertainties. A pattern that has been observed throughout multiple datasets is that the LAD territory has higher uncertainties on the external segments of the polar map. This is explained by the shape of the myocardium, and the way it is segmented, as shown in Figure 6.3. The segmentation of the myocardium is defined symmetrically, while the shape of the myocardium is more complex and not symmetrical especially in the base area. Consequently some segments of the segmentation may have a high spillover

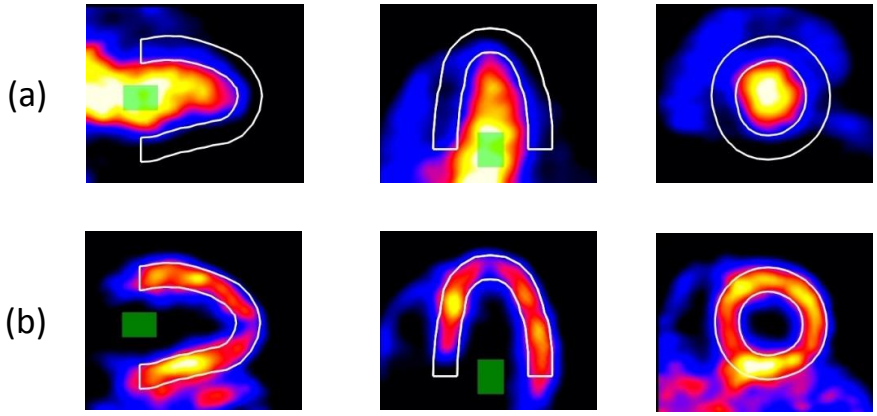


Figure 6.1: Spillover effect explanation. The top row (a) shows an early frame of a cardiac scan, reoriented in the three cardiac planes. In this frame, the tracer is only present in the arterial blood, as it has not yet perfused the myocardium. The bottom row (b) shows a late frame of the same dataset. The segmentation of the myocardial tissue, defined from a late frame, is shown with the white countour. The same contour is replicated in the early frames. It can be seen that the arterial blood signal, overlaps the white contour, creating a partial volume effect, also called spillover.

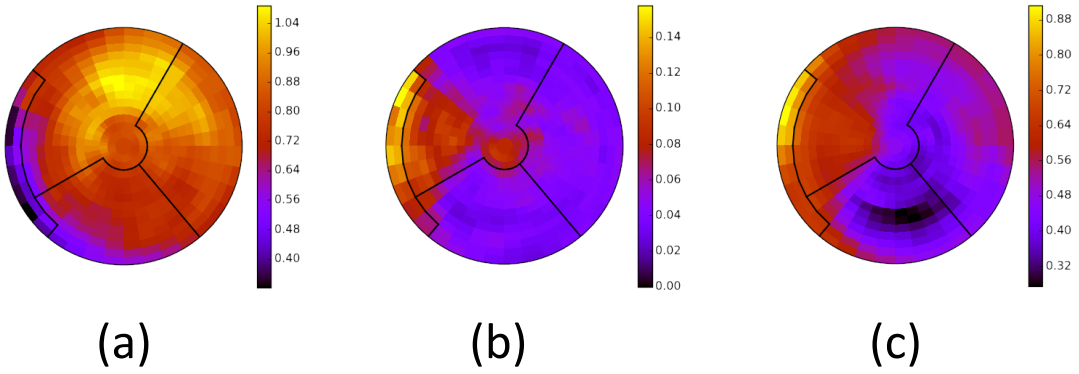


Figure 6.2: Relationship between uncertainty and spillover. Three polar maps are represented, showing the distribution of (a) the K_1 parameter, (b) the uncertainty of K_1 and (c) the S parameter. Regions of high spillover effect (high S) tend to have higher uncertainties.

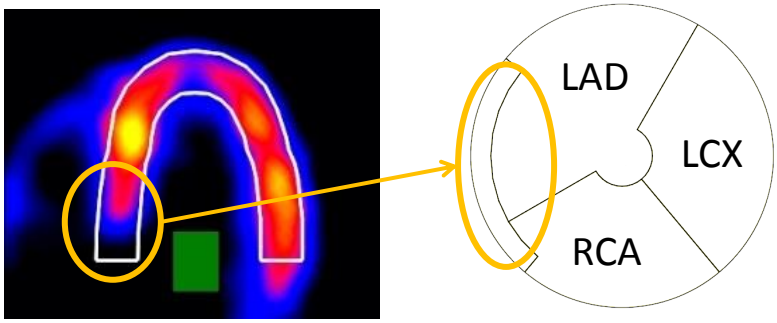


Figure 6.3: Asymmetry of the myocardium. The myocardium does not have a symmetrical shape, while it is the case for the segmentation (white contour). Consequently, when the 3D segmentation is converted into a polar map, some segments in the LAD may have high spillover.

effect due to a weak or no signal in the late frames. More information is given in Appendix A about the myocardium segmentation process of syngo.PET MBF.

Spillover effect may have an impact in the way flow values of coronary territories are calculated. To calculate the blood flow in each coronary territory, multiple segments are defined on the surface of the myocardium tissue. Taken temporally these segments define the TACs that are fitted to calculate individual kinetic parameters. For example, in syngo.PET MBF, 540 TACs are extracted, grouped into 4 categories: the 3 coronary territories (LAD, LCX and RCA) and the whole myocardium (Tot). The kinetic parameters are then averaged accross the coronary

territory. For example the K_1^T value for a given territory is calculated as follows:

$$K_1^T = \frac{1}{m} \sum_{i=1}^{m} K_1^i \tag{6.2}$$

where: T is the coronary territory studied (e.g. LAD); m is the number of segments in the coronary territory T; and K_1^i is the value of K_1 calculated from the individual TAC fitting process, in the i^{th} segment of the territory.

With this method all estimated K_1^i contribute equally to the territory average. This may create some issues with regions of high spillover, since the estimated K_1 parameters are more variable and less accurate. One potential solution to address this issue is to filter out segments in which spillover factor S is higher than a certain threshold for the calculation of the regional flow value. Another approach would be to use the uncertainty on K_1 to weight each segment estimate. This approach makes a better use of the data available, as it is possible that uncertainty of K_1 ends up being low, even with relatively high spillover factor, if the signal extracted to fit $C_{tiss}(t)$ is high enough.

Besides the accuracy in calculating regional flow values, due to inaccurate segmentation of the myocardium, the spillover effect may have also an impact in the detectability of abnormal scans. Classically, abnormal scans are associated with low flow values. Studies report that MBF values below 2.0 mL/min/g are considered as abnormal while values above 2.5 mL/min/g are considered as normal scans [132, 133]. MBF values between 2.0 and 2.5 mL/min/g represent an ambiguous interval where additional information would be needed to help establishing a clear diagnosis. In chapter 4, it was observed that patients with low perfusion, synonymous with an abnormal scan, have higher uncertainties. This is a combined effect of spillover and low perfusion, leading to low tracer concentration. While the spillover is manageable in the "normal" case (i.e. when the perfusion is high) it may not be possible to estimate the K_1 parameter in tissues with low tracer concentration, as shown in Figure 6.4. In that case the signal to be fit by $C_{tiss}(t)$ is dominated by noise. Therefore, the resulting K_1 may be less accurate or potentially wrong, which could induce doubt in a clinical reader as flow values may disagree with the visual

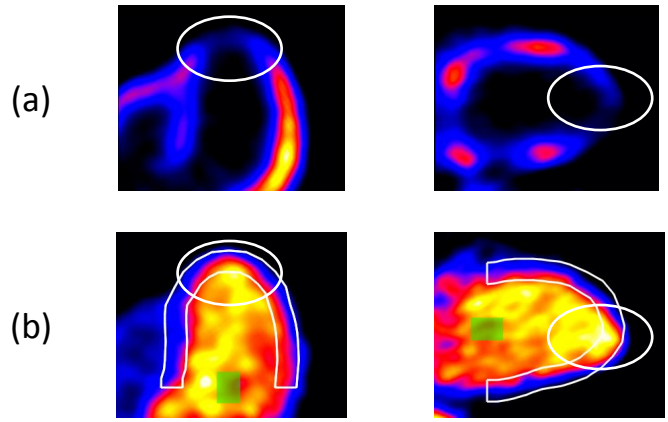


Figure 6.4: Effect of spillover in low perfusion area. The top row (a) shows a late frame of a cardiac scan, reoriented in the vertical and horizontal long axis cardiac planes. In this frame, the tracer is not present in the apical area of the myocardium, which is synonymous with blood perfusion defect. The bottom row (b) shows an early frame of the same dataset. The segmentation of the myocardium tissue, defined from a late frame, is shown with the white countour. The spillover from the early frame, combined with the low signal in the late frame make any kind of kinetic parameter inference challenging if not impossible.

assessment of the scan. In such cases, providing uncertainty of K_1 may give an indication to the reader that the flow values should not be trusted.

The work presented in this chapter describes situations in which the uncertainty measurement may convey additional information, other than an indicator of patient motion artifacts. Specifically, two experiments are designed to look at the influence of the spillover effect. The first experiment aims to evaluate the influence of uncertainty in the average territory parameter estimate. The second experiment looks at the detectability of abnormal territories, by looking simultaneously at K_1 estimates and the uncertainty of K_1 in territories of low perfusion, compared to normal scans.

6.2 Methods

6.2.1 Experiment 1: Weighted methods to calculate flow values

The objective of this experiment is to investigate alternative methods to calculate the territory flow values from the TACs extracted. Specifically these methods try to filter out TACs with high spillover factors, by either looking at the spillover

directly or with a more indirect approach, that is by weighting the TACs by the uncertainty of the kinetic parameter.

6.2.1.1 Coronary territory calculation: review of methods

In the case of a VB method, kinetic parameters are modeled as a Normal distribution, hence the K_1 parameter has a mean μ_{K_1} and standard deviation σ_{K_1}. The mean is taken as the best estimate of K_1, and the value of K_1^T is calculated with:

$$K_1^T = \frac{1}{m}\sum_{i=1}^{m}\mu_{K_1}^i \tag{6.3}$$

where: T is the coronary territory studied (e.g. LAD); m is the number of segments in the coronary territory T; and K_1^i is the value of K_1 calculated from the individual TAC fitting process, in the i^{th} segment of the territory. This approach is referred as the uniform average method.

Another approach is to weight values of $\mu_{K_1}^i$:

$$K_1^T = \sum_{i=1}^{m} w_i \mu_{K_1}^i \tag{6.4}$$

Three different types of weighting methods are investigated:

A. Filtering high S values:

In this method, only values of $\mu_{K_1}^i$ with low corresponding spillover S^i are considered:

$$w_i = \frac{\delta^i}{\sum_{i=1}^{m}\delta^i} \qquad \text{where:} \quad \delta^i = \begin{cases} 1 & \text{if } S^i \leq 0.7 \\ 0 & \text{if } S^i > 0.7 \end{cases} \quad ; \tag{6.5}$$

In this method, it is considered that with high spillover values, the resulting $\mu_{K_1}^i$ is too variable to be considered. The value of 0.7 was chosen as typical values of spillover range between 0.4 and 0.6, and it is reasonable to assume that above 0.7 the signal is dominated by the arterial blood component $C_a(t)$. This approach is referred as the capped S average.

B. Weighting by inverse uncertainty

In this method, values of $\mu^i_{K_1}$ are weighted by the corresponding uncertainties.

$$w_i = \frac{(\sigma^i_{K_1})^{-1}}{\sum_{i=1}^{m}(\sigma^i_{K_1})^{-1}} \tag{6.6}$$

The weights are defined such that the highest weight is given to the parameter with the lowest uncertainty. This approach is referred as the inverse uncertainty average.

C. Weighting by opposite uncertainty

This method is similar to the previous method, except that a different weighting scheme is considered:

$$w_i = \frac{\max_{i\in\{1;m\}}(\sigma^i_{K_1}) - \sigma^i_{K_1}}{\sum_{i=1}^{m}\left(\max_{i\in\{1;m\}}(\sigma^i_{K_1}) - \sigma^i_{K_1}\right)} \tag{6.7}$$

This approach is referred as the opposite uncertainty average.

6.2.1.2 Data

The same patient cohort was used as in chapter 3, where we compared the Bayesian inference schemes Monte-Carlo Markov Chain (MCMC) and Variational Bayes (VB). This cohort consists of 12 pairs of stress / rest datasets. The list of datasets used was as follows: 14_S0_R0; 17_S0_R0; 21_S0_R0; 22_S0_R0; 26_S0_R0; 34_S0_R0; 38_S0_R0; 39_S0_R0; 51_S0_R0; 58_S0_R0; 64_S0_R0; 66_S0_R0;

Details about image acquisition and reconstruction can be found in appendix A.

6.2.1.3 Analysis

Each dataset was processed in syngo.PET MBF. For each dataset 540 TACs were extracted, and regrouped into the 3 coronary territories. Each TAC was fitted individually with the VB method. For each TAC, the mean $\mu^i_{K_1}$ and standard deviation $\sigma^i_{K_1}$ of the parameter K_1 were obtained, as well at the spillover factor S^i. The values of K_1^T was calculated for each territory with the four different averaging methods. The mean relative difference between the weighted averages

and the uniform average were calculated. In addition a one-way ANOVA test was performed to look at the statistical differences between the distribution of K_1^T for each of the four methods across the patient cohort.

6.2.2 Experiment 2: uncertainty in abnormal scans

The objective of this experiment was to look at how the uncertainty extracted from Bayesian method can provide additional clinical information in a dynamic cardiac PET study. Specifically it looks at uncertainties in abnormal scans where the perfusion of the tracer is limited.

6.2.2.1 Data

A cohort of 18 Rubidium stress scans have been examined, equally split between patients with visually normal and abnormal myocardial perfusion. MBF values of the normal patient cohort were reviewed by an independent expert and considered as plausible representation of blood flow. Within the 9 abnormal datasets, 17 territories with reduced perfusion were considered (5 for LAD, 5 for LCX, 7 for RCA). To ensure that any change of uncertainty was not due to motion, datasets used were those categorized as motion free. Only stress datasets were considered since abnormal perfusion patterns may not be seen in rest scans, in case of stress induced ischemia for example.

The list of datasets considered in this study was:

- For the normal perfusion datasets: 17_S0_R0; 21_S0_R0; 22_S0_R0; 34_S0_R0; 38_S0_R0; 39_S0_R0; 50_S0_R0; 51_S0_R0; 64_S0_R0

- For the abnormal perfusion datasets: 16_S0_R0; 23_S0_R0; 25_S0_R0; 30_S0_R0; 32_S0_R0; 36_S0_R0; 46_S0_R0; 69_S0_R0; 70_S0_R0

Details about image acquisition and reconstruction can be found in appendix A.

6.2.2.2 Analysis

Each dataset was processed in syngo.PET MBF. For each dataset 540 TACs were extracted, and regrouped into the 3 coronary territories. Each TAC was fitted individually with the VB method through a one-tissue compartment model. For each TAC, the means μ_{K_1} and standard deviations σ_{K_1} of the parameter K_1 were obtained. The coefficient of variation $CV_{K_1} = \frac{\sigma_{K_1}}{\mu_{K_1}}$ was calculated and used as a unitless measure of the reliability of the flow estimate. The values of μ_{K_1}, σ_{K_1} and CV_{K_1} were averaged per territory for each dataset.

The reference MBF values that are reported in literature [132, 133] were converted into K_1 reference values, using the Renkin-Crone equation:

$$K_1 = \mathrm{MBF}\Big[1 - A\exp(\frac{-B}{\mathrm{MBF}})\Big] \tag{6.8}$$

The Lortie model [57], which sets $A = 0.77$ and $B = 0.63$ was used to calculate K_1 values. Studies report that MBF values below 2.0 mL/min/g were considered as abnormal while values above 2.5 mL/min/g were considered as normal scans. Consequently:

- Values of μ_{K_1} strictly below 0.87 mL/min/g were considered abnormal.
- Values of μ_{K_1} strictly above 1.00 mL/min/g were considered normal.

The mean μ_{CV} and standard deviation σ_{CV} of CV_{K_1} of the normal perfusion patient cohort were calculated. A threshold $\mathcal{T}$ above which the uncertainty would be considered as "unreliable" was defined as follows:

$$\mathcal{T} = \mu_{CV} + 1.96\sigma_{CV} \tag{6.9}$$

This threshold is defined as the upper limit of the 95% confidence interval of CV_{K_1} for normal scans.

6.3 Results

6.3.1 Experiment 1

Figure 6.5 shows the Bland-Altman plots between each weighted average method and the uniform average method, taken as reference. In each territory the difference between a weighted averaging method and the uniform averaging method is very small, with less than 1% across the patient cohort for each method. In addition the ANOVA test showed no statistical differences between the 4 averaging methods (p-value $>$ 0.05).

Among the 24 datasets, only 3 datasets had more than 5% segments with spillover S above 0.7. This indicates that the spillover was very limited in most cases and that the observed changes are probably not noticeable other than for these datasets.

6.3.2 Experiment 2

Table 6.1 shows the K_1 values for each territory and each patient. Out of 27 territories (9 patients with 3 territories each), 17 were annotated as low perfusion: patient 1, 6, 8 and 9 for all territories and patient 2, 5 and 7 for RCA only, patient 3 for LAD only, and patient 4 for LCX only. Out of these 17 territories, 14 had K_1 values strictly below the threshold defined for abnormality (0.87 mL/min/g). One territory (patient 3 for LAD) was annotated as low perfusion but was above the threshold that defines abnormal scan. Finally, patient 2, 4 and 6 for LCX, and patient 2 and 7 for LAD had K_1 values within a similar range, between 0.78 and 0.83 mL/min/g, close to the abnormality threshold, but only two of them were annotated as low perfusion, while three of them were considered normal. These 5 territories show that relying solely on flow values is challenging to tell apart between normal or abnormal cases.

Figure 6.6 shows the CV_{K_1} values of the abnormal perfusion patient cohort for each territory. The values were compared to the threshold $\mathcal{T}$ computed from the normal perfusion patient cohort. Every territory which uncertainty was above the threshold was annotated as low perfusion by the clinical reader, indicating

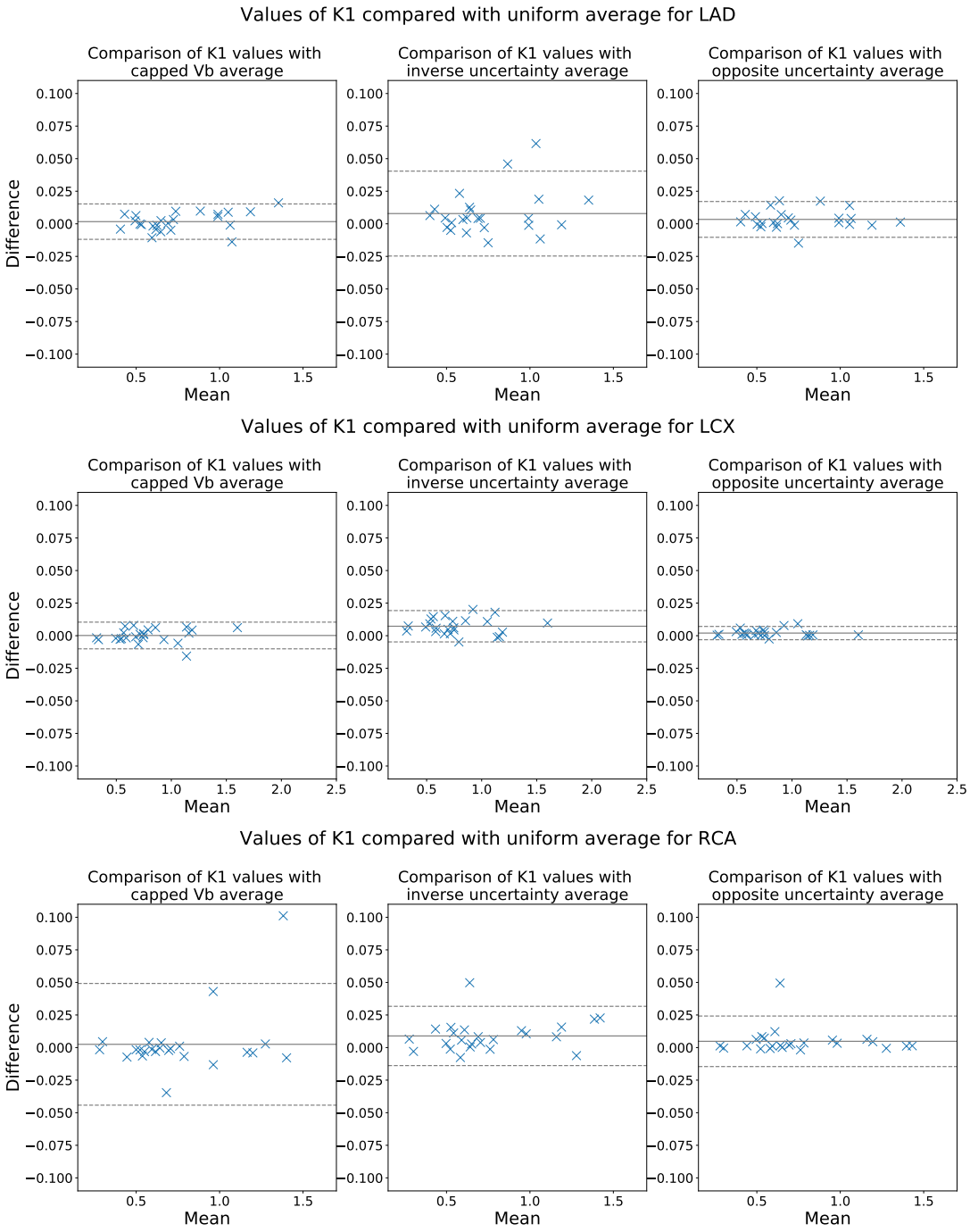


Figure 6.5: Bland Altman for average method comparison. For each territory, the difference between a proposed weighted average and the uniform average was limited, with less than 1% difference on average across the patient cohort. The difference was the most limited for the LCX territory which is the territory that is the least sensitive to spillover.

that flow values in low perfusion regions tend to be of higher variability. Some territories (e.g. patient 9) were annotated as low perfusion but were below the threshold, indicating that flow values can be estimated robustly even in low perfusion territories in some cases.

Combined information from Table 6.1 and Figure 6.6 allows us to refine the analysis for territories that were close to the threshold of abnormality (0.87 mL/min). For example, the K_1 values of patient 3 for LAD and patient 4 and 6 for LCX had high uncertainties, which indicate that the values were not as accurate as for normal cases, and might not be reliable to use in order to assess abnormality. On the other hand, the territories LAD and LCX for patients 2 and 7 have been found to have a relatively low uncertainty, indicating that those values may be reliable at least from a model fitting perspective.

Patient	1	2	3	4	5	6	7	8	9
LAD	**0.53**	0.79	**0.89**	1.02	1.15	**0.61**	0.83	**0.60**	**0.52**
LCX	**0.72**	0.78	1.11	**0.83**	1.03	**0.80**	0.80	**0.57**	**0.72**
RCA	**0.61**	**0.38**	1.04	1.05	**0.61**	**0.65**	**0.65**	**0.52**	**0.68**

Table 6.1: K_1 values for patients with abnormal perfusion. K_1 values are reported for each territory. Territories that were manually annotated as low perfusion are reported with a bold font. From the literature, K_1 values below 0.87 mL/min indicate an abnormal scan. While most of the territories annotated with low perfusion were strictly below 0.87, one territory (LAD patient 3) was above, and two territories were close to 0.87 (LCX for patient 4 and 6). On the contrary, some territories have K_1 values below 0.87 mL/min (patient 2 and 7 for both LAD and LCX) but were not annotated as abnormal. These 5 territories show that flow values are challenging to tell apart between normal or abnormal cases.

6.4 Discussion

In this chapter we have studied the clinical impact of uncertainty metrics. This was done through two experiments.

First, new methods to calculate regional K_1 parameter were proposed, by weighting the individual kinetic parameters. This framework was evaluated by doing a cross-comparison between the regional values obtained with the different

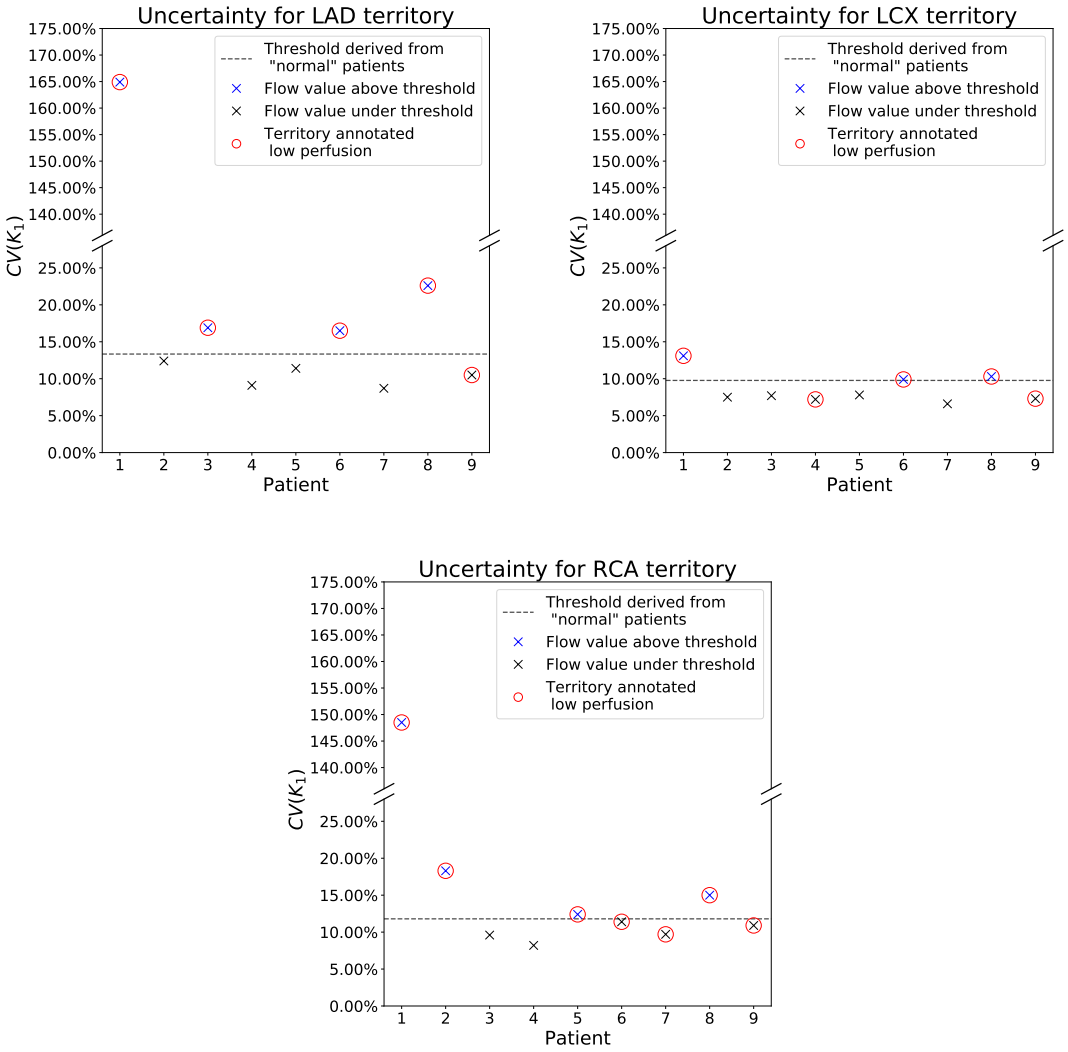


Figure 6.6: Uncertainty for abnormal scans. The uncertainty CV_{K_1} of the scans that were annotated with at least one territory with low perfusion are plotted, gathered by coronary territory. All the uncertainties CV_{K_1} that are above the threshold defining an "acceptable" reliability indicate that territories with low perfusion have reduced confidence and that flow values should be read with attention.

weighting schemes. It was expected that some differences would be seen in territories (LAD and RCA) where the spillover is in general higher due to segmentation issues, as shown in Figure 6.2 and Figure 6.3. The observation was somehow confirmed with the fact that the territory with the smallest change was the LCX. However no significant differences were found between the three newly proposed method and the uniform average for the data used. Furthermore the difference between each newly proposed method and the uniform average was at most 5%, which is

probably not significant enough to change any clinical decision based on flow values considering that more than 10% difference between a coronary territory K_1 value in experiment 2 was not significant to annotate confidently whether the scan was normal or abnormal. This may signify that the uniform average method is robust enough to calculate reliable K_1 territory values, or simply that the spillover was not important enough to make any significant differences. The latter hypothesis is reinforced by the fact that significant spillover was only found in 3 datasets, with 5% of the segments with a value above 0.7. The first method is somewhat easier to implement in a non-Bayesian model inference scheme where the uncertainty would not be available. However the two uncertainty-based weighted averages do not require a fixed threshold for TACs with high S and make a better use of the data, as the uncertainty and the spillover S capture different phenomenon, even if they are linked. The fact that no differences were seen with the uniform average also indicates that the segmentation issue detailed in the introduction and in Figure 6.3 does not impact significantly the coronary territory.

A potential limitation of these weighted methods is that in case of a high variability of uncertainty within a segment, as it is the case with motion or in low perfusion scans (as shown in the second experiment), it may create some artifacts by artificially favouring a small number of TACs within the territory. This could be overcome by leaving the choice to a user and displaying both polar maps (parameter and uncertainty) to have an understanding of how the method proceeded. Finally, this study was limited by the data available. In our case there was no significant difference, however the weighted-average methods might be justified for other studies and requires further evaluation. Including additional cases might showcase examples where a significant spillover was induced by a wrong segmentation, which could be corrected by refining the segmentation or simply filtering the unreliable segments with uncertainty. Furthermore, the weights defined are based on random variables (since noise is a random process), and this aspect needs to be studied more in depth as the noise in a reconstructed dynamic PET scan is a complex phenomenon.

The second experiment looked at the uncertainty in abnormal scans. In this study, uncertainties between abnormal and normal scans were compared, by calculating a threshold of "acceptable" uncertainty, above which the kinetic parameter could be considered unreliable. It appeared that values that were above that threshold corresponded to coronary territories of low perfusion patterns. For most of these territories the K_1 values calculated were considered as abnormal according to reference values found in the literature. However for some datasets, the information of the K_1 parameter alone was ambiguous and not sufficient to decide whether the scan was normal or abnormal. In this situation, the information provided by the uncertainty measurement would show that the variability of the K_1 parameter makes any kind of decision based on the K_1 parameter less reliable. For example, in the case of the LAD territory of patient 3, while the average K_1 value was above the threshold defined in the literature, the uncertainty CV_{K_1} tended to indicate that K_1 is more variable, and that interpretation made from K_1 values should be done with caution or simply avoided. The opposite observation can be made for LAD and LCX territory of patients 2 and 7, where a visual assessment considered that those territories were not abnormal, but the combined information of K_1 and CV_{K_1} for these territories indicate that the values may be analyzed confidently, from a model inference perspective. Furthermore, the discrepancy between the visual assessment of the myocardium and the flow values extracted, may induce some doubt in a clinical reader. The information given by the uncertainty may be helpful to clear these questions. As stated in the previous experiment discussion, displaying polar maps of parameters together with their uncertainties to a clinical reader could be beneficial. It could either reinforce the visual assessment of the coronary territories, or help to highlight unreliable values in other cases.

The correlation between S and the uncertainty on K_1 raises the question if it is necessary to calculate the uncertainty on K_1. Although the uncertainty is more appropriate by definition, one could wonder whether by simply observing the S parameter the same conclusions would be obtained. The answer to that question is multiple. First, while those metrics are linked, they have different meaning. The

uncertainty captures a more complex phenomenon than S does. For example, for a given S parameter, the uncertainty is higher with with a higher level of noise, hence looking at the uncertainty could help in understanding and analyzing different acquisition, reconstruction or post-processing protocols, the same way it was done for analyzing the impact of patient motion. Second, it might also be possible that the uncertainty of K_1 might be low, even with relatively high spillover factor, if the signal extracted to fit $C_{tiss}(t)$ is high enough. This was shown in the second experiment, where all the normal cases showed relatively low uncertainty, even if some data had similar spillover factor compared to the abnormal cases (data not shown). Finally, the uncertainty provides a direct quantitative measurement of the variability on K_1, while S would provide only a qualitative assessment on the K_1 variability.

6.5 Conclusion

In this chapter, we developed two experiments that looked at the influence of spillover effect on the quality of K_1 estimates. It was found that the spillover effect has an impact on the magnitude of the uncertainty measurement. A new method to calculate the regional flow values was introduced, in order to compensate for high spillover. The clinical benefit of reading uncertainty values in cases of abnormal scans was shown. Ambiguous flow values could be detected and provide some indication to a reader as to whether to trust the values obtained or not. To conclude, displaying the uncertainty polar map is likely to be beneficial to a clinical reader, and the work undertaken in this chapter contributes to this argument. The results exhibited in the second experiment were promising, and the work undertaken in this chapter could benefit from further investigations, with additional datasets or different use cases.

7 Conclusions and future work

7.1 Conclusions

The initial aim of this work was to apply a Bayesian model inference framework to estimate kinetic parameter uncertainty in dynamic PET. The underlying hypothesis was that the uncertainty of the kinetic parameters could be used as a metric to assess motion correction. A Variational Bayes framework was chosen for its relatively inexpensive computations.

Throughout this thesis, several studies have been conducted, in order to validate our hypothesis and explore the relationship between kinetic modeling and patient motion in dynamic PET. We have assessed the reliability of kinetic parameters estimates and uncertainty estimates using a Variational Bayes inference framework. The uncertainty of the K_1 parameter, which was used as a surrogate for the flow measurement, was calculated and compared before and after a manual realignment of the PET frames. We have designed a motion correction algorithm that includes the realignment of early frames. Finally we have assessed the benefit of using the uncertainty measurement besides assessing motion correction. Multiple observations and conclusions have been made, they are summarized in the following sections.

7.1.1 Applying a VB framework to calculate parameter uncertainty in kinetic PET

In chapter 3, it was found that a critical aspect of implementing a VB algorithm was to obtain reliable kinetic parameters estimates and uncertainty, for different levels of noise. Two different implementations of the one tissue compartment model have been compared. The first model, which was a straightforward implementation, has been found to be not robust for high level of noise, which may be problematic in case of high motion datasets, where the noise induced by mis-alignment of PET frames was significant. A reparametrization of the parameters has been proposed, so that kinetic parameters are constrained to their domain of definition, and showed to be robust to different level of noises.

The VB method was also compared to another Bayesian model inference, the Markov Chain Monte Carlo (MCMC), considered as a "gold standard" in the Bayes theory. The two methods were evaluated on simulated data as well as on clinical data.It was found that both methods estimated parameter uncertainty with a strong agreement, both on simulated and clinical data. VB tended to underestimate the K_1 parameter compared to MCMC on clinical data, but this difference was not found to be significant.

Overall, this chapter showed that VB method could be used for parameters inference of a one tissue compartment model with a close agreement to the MCMC, which had not been shown previously, as far as we are aware of, in the context of PET kinetic modeling.

7.1.2 Assessing PET frames realignment and motion correction

The uncertainty metric computed with the VB method was then used to evaluate motion correction techniques. Fourty datasets categorized in four motion categories were manually realigned by three experienced observers. Two different types of motion correction were used: Motion Correction with Attenuation Correction (MCAC), where the correction was applied before the reconstruction and by shifting

the mu-map accordingly, and Motion Correction (MC), where the correction was applied post-reconstruction. The uncertainty was calculated for the original dataset as well as each corrected dataset (MC and MCAC). Multiple findings have been found in this study. First, the uncertainty metric was able to assess the effectiveness of motion correction. Up to 60% of uncertainty reduction was observed for high motion cases. Moreover, the level of uncertainty after motion correction was brought to similar level than motion free datasets, demonstrating the success of motion correction with regards to uncertainty metric. Furthermore, no statistical difference was found between the uncertainty extracted from MC and MCAC data. This finding shows that these two methods could be used interchangeably, regarding the uncertainty of the parameter K_1. Finally, the uncertainty metric was sensitive enough to detect differences between users, validating the initial hypothesis of this work, that is using uncertainty to assess motion correction.

To conclude, this work demonstrated that uncertainty metric could be used for assessing motion correction in dynamic PET. We believe this finding contributes to a better understanding of the relationship between patient motion, kinetic parameter estimates and parameters uncertainty for dynamic cardiac PET.

7.1.3 Correcting motion on the entire dynamic PET scan timecourse

In chapter 5, we developed a framework which aimed to correct for motion of cardiac dynamic PET images, including early frames. The proposed framework was evaluated qualitatively by an experienced observer, through a blind review process, and quantitatively by the uncertainty metric developed in the previous chapters. We showed that an improvement was observed in the majority of cases, with 88% datasets that were either significantly or moderately improved. The cases where the algorithm did not improve, were in majority due to intra-frame motion which could not be corrected, in the context of this project. The motion correction framework was also assessed with the uncertainty metric and showed comparable results to manual realignment provided by experienced users. Finally, we showed that motion

correction of the early frames was necessary, as the automated motion correction algorithm correcting only for the second phase (late frames) of the PET series, did not succeed in reducing the uncertainty to a similar level than with manual realignment.

The impact of this work in the field of cardiac dynamic PET is multiple. We showed that motion correction of the early frames was necessary and feasible, thanks to the generation of a synthetic image, adapted to each patient.

7.1.4 Clinical benefit of using uncertainty

Chapter 6 explored alternative usages of the uncertainty metric developed in chapter 3. The uncertainty metric was found to be influenced by spillover factor, and we investigated different cases where it could be problematic. First, we investigated different methods to reduce the influence of spillover by providing a new method to calculate regional flow values. It was found that the methods did not change significantly the results. This may be due to the limited number of datasets available, or that the current method is robust enough to spillover. A second experiment was undertaken, where we showed that displaying uncertainty metric could be beneficial to assess the reliability of flow estimates, especially in abnormal scans where the absence of signal in the late frames combined to the spillover effect may lead to incorrect flow estimates, and wrong assessment of the scan, if not looked at with an uncertainty polar map.

7.2 Future work

7.2.1 Fully automate motion correction

In chapter 5, we manually annotated frames in order to define which synthetic template should be register to the frame. It would be beneficial if this labeling could be done automatically, in order to fully automate the motion correction algorithm. Different options are available. First, considering the limited number of different templates, as well as the prior knowledge that we have on each frame (e.g. intensity, time since beginning of acquisition), and between the frames (there will not be any uptake in LV if there was none in RV in the previous frames) a

set of rules could be defined in order to characterize each frame and assign it a label automatically, or at least narrow down the possibilities. A second option could be to apply machine learning classification techniques, such as convolutional neural network, to label the different phases automatically.

7.2.2 Defining what is an "acceptable" uncertainty

Another avenue that could be investigated is to define an "acceptable" uncertainty. This would benefit for prospective assessment of PET images in multiple ways. For example, if the uncertainty is below the threshold, then the dataset could be considered trustworthy, at least with regards to kinetic modeling. However if the uncertainty is above the acceptable threshold, it may mean that motion correction need to be performed, or that the parameter estimate is not reliable enough to be read, in a similar way than in chapter 6, where it was found that abnormal datasets could potentially be found with unreliable estimate.

To achieve such work, an extended experiment based off the second experiment in chapter 6 could be performed. By assessing reliable and motion free datasets by a clinician, sufficient statistics could be derived to calculate an "acceptable" threshold.

7.2.3 Assessing other factors that influence MBF estimation

Multiple factors are influencing the MBF modeling estimation, among which the reconstruction parameters and the temporal sampling of the PET series. We could potentially investigate the use of time of flight reconstruction, compared with OSEM or FBP reconstructions. Another key parameter in kinetic modeling, that is often overlooked, is the choice of the temporal sampling, which needs to balance between high signal to noise ratio and high number of points. The uncertainty metric could help define an optimal temporal framing, with regards to kinetic modeling.

7.2.4 Extending to further applications

The concept of using uncertainty estimates to assess motion correction could be applied for other organs where dynamic PET is playing a role, as well as other tracers. For example ^{13}N ammonia follows a two tissue irreversible compartment model, and the same experiments that have been conducted in this thesis would be needed to be repeated. Finally it could also be extended to other imaging modalities such dynamic contrast MRI or dynamic CT, where the concept of modeling physiological processes with a dynamic scan is relatively similar.

7.3 Summary

The work presented in this thesis is a contribution to the field of dynamic imaging in general, and cardiac dynamic PET in particular. The uncertainty of kinetic parameters can be calculated in a fast and reliable way with a Variational Bayes framework, and could be used to assess many factors influencing kinetic modeling in dynamic imaging. In this work we exhibited the relationship between kinetic parameter uncertainty and patient motion for cardiac PET studies. We used the uncertainty metric to evaluate a new motion correction framework which incorporates the correction of early frames. Finally, we showed that uncertainty could provide clinical information in cases where the signal obtained with dynamic images is such that the kinetic parameters obtained are unreliable. This work could be extended in the short term term by fully automating the motion correction framework, and in the longer term by looking at other applications of the uncertainty metric both in dynamic PET imaging and other imaging modalities.

Appendices

A

Clinical data: acquisition and processing

This appendix provides some details about the clinical data available throughout this work. Datasets were provided by the department of Nuclear Medicine of the Central Manchester University Hospital (CMUH) NHS Foundation Trust in the United-Kingdom. Those datasets are not publicly available.

A.1 Data collection

86 datasets from a patient cohort, referred for clinically indicated Rb-82, PET/CT were collected. Data was completely anonymised before inclusion, collected as part of routine clinical management and retrospectively analysed. Advice from the local ethics committee deemed that the use of retrospective anonymised patient data did not require formal ethical approval. Each dataset consists in a pair of stress/rest dynamic PET acquisition. Details about the acquisition and reconstruction protocol can be found in the next section.

Stress and rest scans were categorized in 4 different groups based on the amount of motion, according to the method in [44]. Group 0 was characterized by no motion. Group 1 contained light motion (less than half the width of the left ventricular (LV) myocardial wall). Group 2 contained moderate motion, i.e. greater than half the width of the LV myocardial wall but could be corrected partially by automated

	Statistics
Gender (Female / Male)	50 / 36
Weight (kg)	93.7 ± 21.5 [56.0-155.0]
Body Mass Index (kg/m^2)	32.9 ± 6.7 [21.2-55.6]
Age (years)	59.5 ± 12.7 [25-86]
Perfusion pattern (Normal / Abnormal)	48 / 38
Motion 3 (Stress / Rest)	14 / 4
Motion 2 (Stress / Rest)	27 / 5
Motion 1 (Stress / Rest)	24 / 12
Motion 0 (Stress / Rest)	21 / 65

Table A.1: Demographics of the 86 patient cohort that was analysed. For some entries (e.g. weight), the first figure represents the mean value of the cohort, the second figure its standard deviation, and the figures in bracket the minimum and maximum of the cohort.

motion correction algorithm of syngo.PET MBF [128]. Group 3 contained severe motion cases, same as 2 but could not be corrected by automated motion correction algorithm of syngo.PET MBF, visually assessed by review of the dynamic PET series post-correction. A naming strategy was used to identify easily the type of motion for a given dataset. Every dataset was renamed with the following pattern: ID_S{0-3}_R{0-3}, where ID is the patient number, S{0-3} is the motion score for the stress data and R{0-3} for the rest data. For example, 13_S3_R1 is the patient 13 whose stress scan was categorized with severe motion (group 3) and whose rest scan was categorized with light motion (group 1). The demographics of the patients can be found in Table A.1.

A.2 Data acquisition and reconstruction

A.2.1 Acquisition protocol

Images were acquired on a Biograph mCT (Siemens Medical Solutions USA, Inc., Knoxville, TN, USA), with extended 21.6-cm axial FOV and 64 slice CT. The acquisition protocol was done in two phases, called the rest and stress acquisition, corresponding to the physiological state of the myocardium. Prior to the rest scan, a low-dose non-diagnostic CT was acquired to generate the mu-map for reconstructing the datasets with attenuation correction. For both acquisition

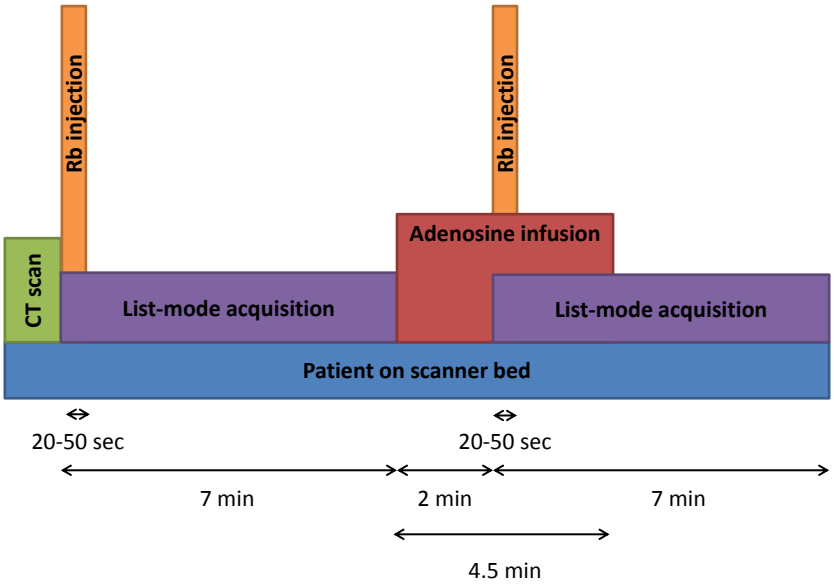


Figure A.1: Acquisition protocol. The protocol starts with the acquisition of a low dose CT, which is used for generating the mu-map for attenuation correction. Then a first injection of Rubidium-82 is done. Data is continuously collected for 7 minutes, from the beginning of the tracer injection. This first acquisition is the rest scan. A stress agent is then infused for 4.5 minutes. The stress scan is acquired with the same settings than the rest scan.

patients were administered with 1110 MBq (30 mCi) of Rb-82, generated from a Cardiogen 82-Rb generator (Bracco Diagnostics, Milan, Italy). List-mode data were collected for 7 minutes, from the beginning of the injection of the tracer. After the rest acquisition, a pharmacological stress agent was injected to the patient to induce the stress state of the myocardium. In this work the stress agent used was adenosine, with a rate of 140 μg/kg/min, for a total duration of 4.5min. The injection of the Rubidium tracer was done after 2 minutes of the beginning of the adenosine infusion. The acquisition protocol is summarized in Figure A.1.

A.2.2 Reconstruction protocol

From the 7 minutes list mode data acquired, only the first 6 minutes were used to reconstruct the dynamic PET series. The last minute of list mode data was used by the site for the static reconstruction protocol (not used in this thesis). The time sampling used for the reconstruction was: 1x10 seconds, 8x5 seconds, 3x10 seconds, 2x20 seconds and 4x60 seconds, for a total of 18 PET frames. Each frame

was reconstructed with a 3D Ordered Subset Expectation Maximization (OSEM) [129] using 2 iterations, 24 subsets and a 6.5mm full-width half maximum Gaussian post-filter. In addition, each frame was decay corrected. The matrix size used was 128x128x109 and the voxel spacing was 3.18x3.18x2.03mm. During the stress protocol, some frames are still being acquired after the effect of the adenosine washes out (the adenosine half-life is about 10 seconds). This implies that the myocardial tissue is not at stress for the last 3 minutes of the acquisition. These late frames are necessary to define the different regions of interest (myocardial tissue and blood pool regions), since the tracer is still visible in the myocardial tissue with a good contrast, which facilitates the task for automatic segmentation algorithms. The consequences of using non stressed frames is minimal, since the MBF estimates are primarily defined in the first two minutes of the acquisition, during which the adenosine is still being injected.

A.3 Data post-processing

Each pair of stress / rest dynamic PET series was processed with syngo.PET MBF. The software automatically locates and segments the myocardium from a summed image of the last 3 minutes of the dynamic PET series.

The heart localization algorithm defines three landmarks which are the center, the base and the apex of the myocardium. These three points define the long axis of the myocardium. If needed these 3 landmarks can be manually adjusted, and the long axis can also be reangulated. From the long axis definition, the myocardium is segmented into 15 rings from apex to base with 36 angular segments in each ring. The segments are placed at the maximum intensity along orthogonal lines to the long axis. In total 540 points are defined. An example can be seen in Figure A.2. Activity concentration in each segment is measured using the mean of pixel values within a region of interest of 1-cm radial thickness centered on the middle of the myocardial surface. The 540 segments are then replicated in each frame of the dynamic series, which defines 540 Time Activity Curves (TACs). These curves are grouped in 4 different categories: the whole myocardium (Total) and the 3 coronary territories

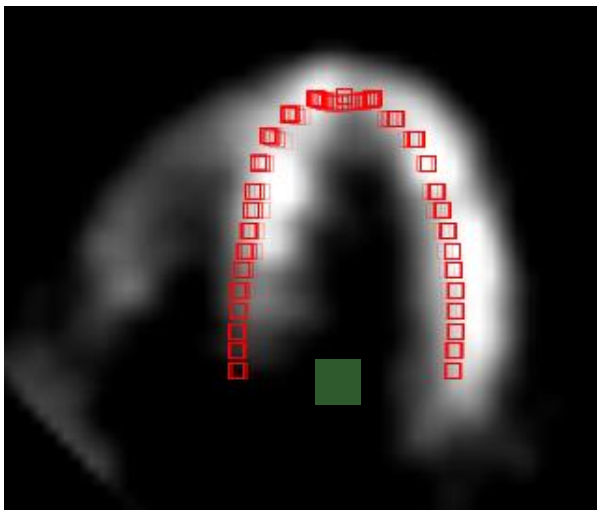


Figure A.2: Myocardium segmentation generated with syngo.PET MBF. The segmentation consists in 540 segments defined in 15 rings of 36 points located at the maximum of orthogonal lines to the long axis definition. These 540 segments define subsequently the 540 TACs used for kinetic modeling of the myocardial blood flow.

(LAD, LCX and RCA). In addition a cylindrical volume of interest, of 2-cm diameter and 1.5-cm long, is placed at the base position to define the arterial blood Time Activity Curve. This ROI was defined relatively small to avoid any spillover effect from the cardiac tissue. To limit the error related to the size of the ROI, the long axis of the myocardium was reangulated manually in order to maximize the signal to noise ratio from the blood pool region. Additionally, the position of the ROI is close to the mitral valve. While this could be a source of error for frames with very short duration (less than 1 or 2 cardiac cycles), the consequences on the datasets used in this project should be minimal (the shortest frames were 5 seconds long).

A.4 Implementation details

Various algorithms have been used and implemented throughout this thesis. All results were obtained with an Intel Xeon CPU E5-1620 of 3.70 GHz and with 16.0 GB of RAM. Our implementation of the Variational Bayes (VB) algorithm was based on C++, and wrapped with a Python 3.5 script, interfacing with the Numpy library. An implementation of the VB algorithm can be found in the FMRIB Software library (www.fmrib.ox.ac.uk/fsl) [120], which is freely available

for academic use. The VB algorithm processed 1000 TACs (see experiments in chapter 3) in 6 seconds. Our implementation of the Metropolis Hastings algorithm was also based on C++ and wrapped with a Python 3.5 script. It processed the same 1000 TACs in 42 minutes and 20 seconds.

The motion correction algorithm developed in chapter 5, was based on the Elastix software suite [134, 135] (C++). The motion correction algorithm processed one dynamic image in 1min15sec (including the initial segmentation of the myocardium).

Finally, the reconstruction of the list-mode acquisition data to DICOM files was done with a C++ core and wrapped in a Matlab routine. The reconstruction procedure, including the DICOM file creation, for an 18 frames dataset lasted about 35min.

References

[1] World Health Organization. *World Heatlh Statistics - A wealth of information on global public health*. URL: apps.who.int/iris/bitstream/10665/82058/1/WHO_HIS_HSI_13.1_eng.pdf.

[2] Cutler David. "Rising medical costs mean more rough times ahead". In: *Journal of American Medical Association* 318.6 (2017), pp. 508–509.

[3] Sunil Mathur and Joseph Sutton. "Personalized medicine could transform healthcare (Review)". In: *Biomedical reports* 7.6 (2017), pp. 3–5.

[4] National Research Council. *Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease*. National Academies Press, 2011.

[5] *The 100,000 genomes project*. URL: https://www.genomicsengland.co.uk/the-100000-genomes-project/.

[6] *The precision medicine initiative*. URL: https://obamawhitehouse.archives.gov/node/333101.

[7] Angela Giardino et al. "Role of Imaging in the Era of Precision Medicine". In: *Academic Radiology* 24 (5 2015), pp. 639–649.

[8] Simon Cherry, James Sorenson, and Michael E. Phelps. *Physics in nuclear medicine*. Philadelphia: Elsevier/Saunders, 2012.

[9] Richard Carson. "Tracer Kinetic Modeling in PET". In: *Positron Emission Tomography: Basic Science*. Ed. by Peter Valk et al. 2003. Chap. 6, pp. 127–158.

[10] Rathan Subramaniam. "Precision Medicine and PET/Computed Tomography: Challenges and Implementation". In: *PET clinics* 12 (2017), 1—5.

[11] Juhani Knuuti et al. "Quantification of myocardial blood flow will reform the detection of CAD". In: *Journal of Nuclear Cardiology* (2009), pp. 497–506.

[12] British Heart Foundation. *BHF CVD Statistics compendium 2017*. 2017.

[13] P Camici et al. "The impact of myocardial blood flow quantitation with PET on the understanding of cardiac diseases". In: *European Heart Journal* (1996), pp. 25–34.

[14] Halia Al-Shehri, Gary Small, and Benjamin J.W. Chow. "Cardiac CT, MR, SPECT, ECHO, and PET: What test, when?" In: *Applied Radiology* 40 (5 2011).

[15] Adam Ghotbi, Andreas Kjaer, and Philip Hasbak. "Review: comparison of PET rubidium-82 with conventional SPECT myocardial perfusion imaging". In: *Clinical Physiology and Functional Imaging* (2014), pp. 163–170.

[16] Ratika Parkash et al. “Potential utility of rubidium 82 PET quantification in patients with 3-vessel coronary artery disease”. In: *Journal of Nuclear Cardiology* 11 (4 2004), pp. 440–449.

[17] Maria Ziadi et al. “Does quantification of myocardial flow reserve using rubidium-82 positron emission tomography facilitate detection of multivessel coronary artery disease?” In: *Journal of Nuclear Cardiology* (2011), pp. 671–679.

[18] Martin Omlin and Peter Reichert. “A comparison of techniques for the estimation of model prediction uncertainty”. In: *Ecological Modelling* (1999), pp. 45–59.

[19] Giulio D’agostini. “Bayesian inference in processing experimental data: principles and basic applications”. In: *Reports on Progress in Physics* (2003), pp. 1383–1419.

[20] Vasken Dilsizian et al. “ASNC imaging guidelines/SNMMI procedure standard for positron emission tomography (PET) nuclear cardiology procedures”. In: *Journal of Nuclear Cardiology* (2016), pp. 1187–1226.

[21] Sten Carlsson. “A glance at the history of nuclear medicine”. In: *Acta Oncologica* 34 (8 1995), pp. 1095–1102.

[22] John Lawrence. “Nuclear Physics and Therapy: Preliminary Report on a New Method for the Treatment of Leukemia and Polycythemia”. In: *Radiology* 35.1 (1940), pp. 51–60.

[23] Néstor Herran. “Isotope networks: Training, sales and publications, 1946-1965”. In: *Dynamis* 29 (2009), pp. 285–306.

[24] S. M. Seidlin, L. D. Marinelli, and Eleanor Oshry. “Radioactive Iodine Therapy - Effect on Functioning Metastases of Adenocarcinoma of the Thyroid”. In: *Journal of the American Medical Association* (1946).

[25] Brian Hutton. “The origins of SPECT and SPECT/CT”. In: *European Journal of Nuclear Medicine and Molecular Imaging* 41 (2014), S3–S16.

[26] Dayton Rich. “A Brief History of Positron Emission Tomography”. In: *Journal of Nuclear Medicine Technology* 25 (1997), pp. 4–11.

[27] Simon Cherry, James Sorenson, and Michael Phelps. *Physics in Nuclear Medicine*. Saunders, 2003.

[28] Michael Phelps. *PET Molecular Imaging and Its Biological Applications*. Springer, 2004.

[29] A. Shukla and Utham Kumar. “Positron emission tomography: An overview”. In: *Journal of Medical Physics / Association of Medical Physicists of India* 31 (1 2006), pp. 13–21.

[30] Gopal Saha. *Basics of PET imaging*. Springer, 2010.

[31] Craig Levin and Edward Hoffman. “Calculation of positron range and its effect on the fundamental limit of positron emission tomography system spatial resolution”. In: *Phys. Med. Biol.* (44 1999), 781–799.

[32] Frederic Fahey. “Data Acquisition in PET Imaging”. In: *Journal of Nuclear Medicine technology* 30 (2002), pp. 39–50.

[33] Geoffrey Currie, Johnathan Hewis, and Stewart Bushong. “Tomographic Reconstruction: A Nonmathematical Overview”. In: *Journal of Medical Imaging and Radiation Sciences* 46 (4 2015), pp. 403–412.

[34] Oleh Tretiak and Charles Metz. “The Exponential Radon Transform”. In: *SIAM Journal on Applied Mathematics* 39 (1980), pp. 341–354.

[35] Sandro Bellini, Maurizio Piacentini, and Ciro Cafforio. “Compensation of Tissue Absorption in Emission Tomography”. In: *IEEE Transaction on acoustics, speech, and signal processing* 27 (1979), pp. 213–218.

[36] Roger Gunn, Steve Gunn, and Vincent Cunningham. “Positron Emission Tomography Compartmental Models”. In: *Journal of Cerebral Blood Flow and Metabolism* (2001), pp. 635–652.

[37] Evan Morris et al. “Kinetic Modeling in Positron Emission Tomography”. In: *Emission Tomography*. Ed. by Elsevier. 2004. Chap. 23, pp. 499–540.

[38] Ran Klein. “Kinetic Model Based Factor Analysis of Cardiac 82-Rb PET Images for Improved Accuracy of Quantitative Myocardial Blood Flow Measurement”. PhD thesis. University of Ottawa - School of Information Technology and Engineering, 2010.

[39] Gary Heller et al. “ASNC Model Coverage Policy: Cardiac positron emission tomographic imaging”. In: *Journal of Nuclear Cardiology* 20 (5 2013), 916–947.

[40] Bailing Hsu. “PET tracers and techniques for measuring myocardial blood flow in patients with coronary artery disease”. In: 27 (6 2013), pp. 452–459.

[41] John Prior et al. “Quantification of myocardial blood flow with 82 Rb positron emission tomography : clinical validation with 15 O-water”. In: *European Journal of Nuclear Medicine* (39 2012), pp. 1037–1047.

[42] Hsiao ming Wu et al. “Quantification of Myocardial Blood Flow Using Dynamic Nitrogen-13-Ammonia PET Studies and Factor Analysis of Dynamic Structures”. In: *Journal of Nuclear Medicine* (36 1995), pp. 2087–2093.

[43] Jamshid Maddah and Rene Packard. “Cardiac PET Perfusion Tracers: Current Status and Future Directions”. In: *Seminars in Nuclear Medicine* 44 (5 2015).

[44] Matthew Memmott et al. “Impact of pharmacological stress agent on patient motion during rubidium-82 myocardial perfusion PET/CT”. In: *Journal of Nuclear Cardiology* (2017).

[45] Elise Vleeming et al. “Cardiac displacement during 13 N-Ammonia myocardial perfusion PET/CT: comparison between adenosine and regadenoson induced stress”. In: *Journal of Nuclear Medicine Technology* (2017).

[46] Christoffer Hagemann et al. “Quantitative myocardial blood flow with Rubidium-82 PET: a clinical perspective”. In: *American Journal of Nuclear Medicine and Molecular Imaging* (2015).

[47] Keiichiro Yoshinaga et al. “Incremental Diagnostic Value of Regional Myocardial Blood Flow Quantification Over Relative Perfusion Imaging With Generator-Produced Rubidium-82 PET”. In: *Circulation Journal* 75 (2011).

[48] Maria Ziadi et al. “Impaired Myocardial Flow Reserve on Rubidium-82 Positron Emission Tomography Imaging Predicts Adverse Outcomes in Patients Assessed for Myocardial Ischemia”. In: *Journal of the American College of Cardiology* 58 (7 2011), pp. 740–748.

[49] Jeffrey Kolthammer and Raymond Muzic. "Optimized dynamic framing for PET-based myocardial blood flow estimation". In: *Physics in Medicine and Biology* 58 (2013).

[50] M'hamed Bentourkia and Habib Zaidi. "Tracer Kinetic Modeling in PET". In: *PET Clinics* 2 (2 2007), pp. 267–277.

[51] Manuel Desco et al. "Automatic TAC extraction from dynamic cardiac PET imaging using iterative correlation from a population". In: *Computer Methods and Programs in Biomedicine* 111 (2013), pp. 308–314.

[52] Pilar Herrero et al. "Implementation and evaluation of a two-compartment model for quantification of myocardial perfusion with rubidium-82 and positron emission tomography." In: *Circulation Research* 3 (70 1992), pp. 496–507.

[53] Pamela Coxson, Ronald Huesman, and Lisa Borland. "Consequences of using a simplified kinetic model for dynamic PET data." In: *Journal of Nuclear Medicine* (38 1997), pp. 660–667.

[54] Jou-Wei Lin et al. "Quantification of myocardial perfusion in human subjects using 82Rb and wavelet-based noise reduction." In: *Journal of Nuclear Medicine* 42 (2001), pp. 201–208.

[55] Karin Knesaurek, Josef Machac, and Zhuangyu Zhang. "Repeatability of regional myocardial blood flow calculation in 82Rb PET imaging". In: *BMC Medical Physics* 2 (9 2009).

[56] Christian M. Wacker et al. "Determination of Regional Blood Volume and Intra-Extracapillary Water Exchange in Human Myocardium Using Feruglose: First Clinical Results in Patients With Coronary Artery Disease". In: *Magnetic Resonance in Medicine* (47 2002), pp. 1013–1016.

[57] Mireille Lortie et al. "Quantification of myocardial blood flow with 82Rb dynamic PET imaging". In: *European Journal of Nuclear Medicine and Molecular Imaging* 34 (2007), pp. 1765–1774.

[58] K. Yoshida, N. Mullani, and K. L. Gould. "Coronary flow and flow reserve by PET simplified for clinical applications using rubidium-82 or nitrogen-13-ammonia". In: *Journal of Nuclear Medicine* 37 (10 1996), 1701–1712.

[59] Adrian Ocneanu et al. "Optimally Repeatable Kinetic Model Variant for Myocardial Blood Flow Measurements with 82Rb PET". In: *Computational and Mathematical Methods in Medicine* 2017 (2017).

[60] Jonathan Moody et al. "Variance Estimation for Myocardial Blood Flow by Dynamic PET". In: *IEEE Transactions on Medical Imaging* (2015), pp. 2343–2353.

[61] Jonathan Moody et al. "Precision and accuracy of clinical quantification of myocardial blood flow by dynamic PET: A technical perspective". In: *Journal of Nuclear Cardiology* (2015), pp. 935–951.

[62] Jonathan Moody, Benjamin Lee, and Edward Ficaro. "Error estimation for dynamic PET myocardial blood flow (Abstract)". In: p. 323.

[63] Georges El Fakhri et al. "Reproducibility and Accuracy of Quantitative Myocardial Blood Flow Assessment with 82Rb PET: Comparison with 13N-Ammonia PET". In: *Journal of Nuclear Medicine* (2009), pp. 1062–1071.

[64] Robert deKemp, Ran Klein, and Rob Beanlands. "82-Rb PET imaging of myocardial blood flow—have we achieved the 4 "R"sto support routine use?" In: *EJNMMI Research* 6 (69 2016).

[65] Benjamin Lee et al. "Blood pool and tissue phase patient motion effects on 82-Rubidium PET myocardial blood flow quantification". In: *Journal of Nuclear Cardiology* (2018).

[66] Kazuhiro Koshino et al. "Effects of patient movement on measurements of myocardial blood flow and viability in resting 15O-water PET studies". In: *Journal of Nuclear Cardiology* (2012).

[67] Chad Hunter et al. "Patient motion effects on the quantification of regional myocardial blood flow with dynamic PET imaging". In: *Medical Physics* (2016), pp. 1829–1840.

[68] Marina Piccinelli, John Votaw, and Ernest Garcia. "Motion Correction and Its Impact on Absolute Myocardial Blood Flow Measures with PET". In: *Current Cardiology Reports* (2018).

[69] Axel Martinez-Moller et al. "Artifacts from Misaligned CT in Cardiac Perfusion PET/CT Studies: Frequency,Effects, and Potential Solutions". In: *Journal Of Nuclear Medicine* (2007), pp. 188–193.

[70] Hendrik Pretorius, Karen Johnson, and Michael King. "Evaluation of Rigid-Body Motion Compensation in Cardiac Perfusion SPECT Employing Polar-Map Quantification". In: *IEEE Transactions on Nuclear Science* 63 (3 2016).

[71] Arman Rahmim, Olivier Rousset, and Habib Zaidi. "Strategies for Motion Tracking and Correction in PET". In: *PET Clinics* (2007), pp. 251–266.

[72] S. A. Nehmeh et al. "Effect of respiratory gating on reducing lung motion artifacts in PET imaging of lung cancer". In: *Medical physics* 29 (3 2002), pp. 366–372.

[73] Wenjia Bai. "Respiratory Motion Correction in Positron Emission Tomography". PhD thesis. University of Oxford - Department of Engineering Science, 2010.

[74] Audrey Pepin et al. "Management of respiratory motion in PET/computed tomography: the state of the art". In: *Nuclear Medicine Communications* 35 (2014), pp. 113–122.

[75] Oline Vinter. "Markerless 3D Head Tracking for Motion Correction in High Resolution PET Brain Imaging". PhD thesis. University of Copenhagen - Faculty of Health Sciences, 2011.

[76] Philip Noonan et al. "Accurate Markerless Respiratory Tracking For Gated Whole Body PET Using the Microsoft Kinect". In: 2012.

[77] Yani Picard and Christopher Thompson. "Motion Correction of PET Images Using Multiple Acquisition Frames". In: *IEEE Transactions on Medical Imaging* (1997).

[78] M. Menke, M. S. Atkins, and K. R. Buckley. "Compensation methods for head motion detected during PET imaging". In: *IEEE Transactions on Nuclear Science* 43 (1 1996), pp. 310–317.

[79] Ashley Gillman et al. "PET motion correction in context of integrated PET/MR: Current techniques, limitations, and future projections". In: *Medical Physics* 44 (12 2017), pp. 430–445.

[80] Sebastian Furst et al. “Motion Correction Strategies for Integrated PET/MR”. In: *Journal of Nuclear Medicine* (2015).

[81] Yoann Petibon et al. “Impact of motion and partial volume effects correction on PET myocardial perfusion imaging using simultaneous PET-MR”. In: *Physics in Medicine and Biology* 62 (2017).

[82] Antoine Maintz and Max Viergever. “A survey of medical image registration”. In: *Medical Image Analysis* 2 (1 1998).

[83] Barbara Zitova and Jan Flusser. “Image registration methods: a survey”. In: *Image and Vision Computing* (2003).

[84] Jesper Andersson. “How to obtain high-accuracy image registration: application to movement correction of dynamic positron emission tomography data”. In: *European journal of Nuclear Medicine and molecular imaging* 25 (6 1998).

[85] Kang ping Lin et al. “Automated image registration for FDOPA PET studies Automated image registration for F-DOPA PET studies”. In: *Physics in Medicine and Biology* 41 (1996).

[86] Timothy Turkington et al. “Alignment of Dynamic Cardiac PET Images for Correction of Motion”. In: *IEEE Transactions on Nuclear Science* 44 (2 1997).

[87] Thomas Lehmann, Claudia Gonner, and Klaus Spitzer. “Survey: Interpolation Methods in Medical Image Processing”. In: *IEE Transactions on Medical Imaging* 18 (11 1999), pp. 1049–1075.

[88] Aristeidis Sotiras and Nikos Paragios. *Deformable image registration: a survey.* Research Centre Saclay – Ile-De-France, 2012.

[89] Paul Besl and Neil McKay. “A method for registration of 3D shapes”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14 (2 1992), pp. 239–256.

[90] “Using mutual information and cross correlation as metrics for registration of images”. In: *Journal of Theoretical and Applid Information Technology* (2008).

[91] N. Raghunath et al. “Motion correction of PET brain images through deconvolution: II. Practical implementation and algorithm optimization”. In: *Physics in Medicine and Biology* (2009).

[92] Norbert Lang et al. “Organ Movement Reduction in PET/CT using Dual-Gated Listmode Acquisition”. In: *Zeitschrift für Medizinische Physik* (2006).

[93] D. Visvikis et al. “Respiratory motion in positron emission tomography for oncology applications: Problems and solutions”. In: *Nuclear Instruments and Methods in Physics Research* 569 (2006), pp. 453–457.

[94] Jurgen Mourik et al. “Off-line motion correction methods for multi-frame PET data”. In: *European Journal of Nuclear Medicine and Molecular Imaging* (2009).

[95] Nicolas Costes et al. “Motion correction of multi-frame PET data in neuroreceptor mapping: Simulation based validation”. In: *NeuroImage* 47 (4 2009), pp. 1496–1505.

[96] Jieqing Jiao et al. “Spatio-temporal pharmacokinetic model based registration of 4D PET neuroimaging data”. In: *NeuroImage* 84 (2014), pp. 225–235.

[97] Josien Pluim, J. B. Antoine Maintz, and Max A. Viergever. “Interpolation Artefacts in Mutual Information-Based Image Registration”. In: *Computer Vision and Image Understanding* 77 (2000), pp. 211–232.

[98] Jonghye Woo et al. “Automatic 3D registration of dynamic stress and rest 82Rb and flurpiridaz F 18myocardial perfusion PET data for patient motion detection and correction”. In: *Medical Physics* 38 (11 2011), pp. 6313–6326.

[99] Hui Xue et al. “Motion Correction for Myocardial T1 Mapping Using Image Registration with Synthetic Image Estimation”. In: *Magnetic Resonance in Medicine* (67 2012), pp. 1644–1655.

[100] Tautz L., Chitiboi T., and Hennemuth A. “Automatic Perfusion Analysis Using Phase-Based Registration and Object-Based Image Analysis.” In: *Statistical Atlases and Computational Models of the Heart - Imaging and Modelling Challenges. STACOM 2014. Lecture Notes in Computer Science* 8896 (2015).

[101] Devavrat Likhite, Ganesh Adluru, and Edward DiBella. “Deformable and Rigid Model-Based Image Registration for Quantitative Cardiac Perfusion”. In: *Statistical Atlases and Computational Models of the Heart - Imaging and Modelling Challenges. STACOM 2014. Lecture Notes in Computer Science* 8896 (2015).

[102] Mohamed Ahmed, Peng Xiao, and Qingguo Xie. “New approach for simultaneous respiratory and cardiac motion correction in cardiac PET (NAMC-CPET)”. In: *Physics in Medicine and Biology* (2015), pp. 7779–7804.

[103] Ian Armstrong, Christine Tonge, and Parthiban Arumugam. “Assessing time-of-flight signal-to-noise ratio gains within the myocardium and subsequent reductions in administered activity in cardiac PET studies”. In: *Journal of Nuclear Cardiology* (2017).

[104] Hadi Fayad et al. “Reconstruction-Incorporated Respiratory Motion Correction in Clinical Simultaneous PET/MR Imaging for Oncology Applications”. In: *Journal of Nuclear Medicine* (2015), pp. 884–889.

[105] Robin Willink and Rod White. *Disentangling Classical and Bayesian Approaches to Uncertainty Analysis*. 2011, pp. 1–19.

[106] Y. Vardi, L. A. Shepp, and L. Kaufman. “A statistical model for Positron Emission Tomography”. In: *Journal of the American Statistical Association* 80 (389 1985), pp. 8–20.

[107] Maqsood Yaqub et al. “Optimization algorithms and weighting factors for analysis of dynamic PET studies”. In: *Physics in Medicine and Biology* 51 (2006), pp. 4217–4232.

[108] Kewei Chen, Sung-Cheng Huang, and Dan-Chu Yu. “The effects of measurement errors in the plasma radioactivity curve on parameter estimation in positron emission tomography”. In: *Physics in Medicine and Biology* 36 (9 1991), pp. 1183–1200.

[109] Dagan Feng and Xinmin Wang. “A computer simulation study on the effects of input function measurement noise in tracer kinetic modeling with positron emission tomography (PET).” In: *Computers in biology and medicine* 23 (1 1993), pp. 58–68.

[110] A. Teymurazyan et al. “Properties of Noise in Positron Emission Tomography Images Reconstructed with Filtered-Backprojection and Row-Action Maximum Likelihood Algorithm”. In: *Journal of digital imaging* 26 (2013), pp. 447–456.

[111] Raymond Muzic and Bradley Christian. “Evaluation of objective functions for estimation of kinetic parameters”. In: *Medical Physics* (2006).

[112] Vincent Cunningham and Terry Jones. “Spectral Analysis of Dynamic PET Studies”. In: *Journal of cerebral blood flow and metabolism* 13 (1993), pp. 15–23.

[113] Roger Gunn et al. “Positron Emission Tomography Compartmental Models : A Basis Pursuit Strategy for Kinetic Modeling”. In: *Journal of cerebral blood flow and metabolism* 22 (2002), pp. 1425–1439.

[114] Mattia Veronese et al. “Spectral Analysis of Dynamic PET Studies : A Review of 20 Years of Method Developments and Applications”. In: (2016).

[115] William Vetterling et al. *Numerical Recipes in C The Art of Scientific Computing*. Cambridge University Press, 1992, pp. 656–706.

[116] Mustafa Kamasak. “Computation of variance in compartment model parameter estimates from dynamic PET data”. In: *Medical Physics* 39 (5 2012).

[117] F Hampel. “The influence curve and its role in robust estimation.” In: *Journal of American Statistician Association* 69 (1974), 383–93.

[118] William DuMouchel and Fanny O’Brien. “Integrating a Robust Option into a Multiple Regression Computing Environment”. In: *Computing and graphics in statistics*. Ed. by VA American Statistical Association: Alexandria. 1989, pp. 41 –48.

[119] John Geweke, Gary Koop, and Herman Van Dijk. *The Oxford Handbook of Bayesian Econometrics*. Oxford University Press, 2013.

[120] Michael Chappell et al. “Variational Bayesian Inference for a Nonlinear Forward Model”. In: *IEEE Transactions on Signal Processing* (2009), pp. 223–236.

[121] Edward Chib Siddhartha; Greenberg. “Understanding the Metropolis-Hastings Algorithm”. In: *The American Statistician* (1995).

[122] M Castellaro et al. “A Variational Bayesian inference method for parametric imaging of PET data”. In: *NeuroImage* (2017), pp. 136–149.

[123] Hagai Attias. “A Variational Bayesian Framework for Graphical Models”. In: *In Advances in Neural Information Processing Systems* (1999), pp. 209–215.

[124] Matthew Beal. “Variational algorithms for approximate bayesian inference”. PhD thesis. University College London - The Gatsby Computational Neuroscience Unit, 2003.

[125] Michael Chappell, Adrian Groves, and Mark Woolrich. *The FMRIB Variational Bayes tutorial*. 2008.

[126] Adrian Groves, Michael Chappell, and Mark Woolrich. “Combined spatial and non-spatial prior for inference on MRI time-series”. In: *NeuroImage* (2009), pp. 795–809.

[127] Moss Zhao et al. “A systematic study of the sensitivity of partial volume correction methods for the quantification of perfusion from pseudo-continuous arterial spin labeling MRI”. In: *NeuroImage* (162 2017), pp. 384–397.

[128] Bernard Bendriem et al. “The continual innovation of commercial PET/CT solutions in nuclear cardiology: Siemens Healthineers”. In: *J. Nucl. Cardiol.* (2018), pp. 1–12.

[129] H. Malcolm Hudson and Richard S. Larkin. “Accelerated image reconstruction using Ordered Subsets of projection data”. In: *IEEE Trans. Med. Imag.* 13 (4 1994), pp. 601–609.

[130] Sung-Hyuk Cha. “Comprehensive Survey on Distance/Similarity Measures between Probability Density Functions”. In: *International Journal of Mathematical models and methods in applied sciences* 1 (4 2007).

[131] Alexandru Naum et al. “Motion detection and correction for dynamic 15O-water myocardial perfusion PET studies”. In: *Eur. J. Nucl. Med.* (32 2005), 1378–1383.

[132] Roel Driessen et al. In: *Int. J. of Cardiovasc. Imaging* (33 2017), pp. 1021–1031.

[133] Schindler et al. “Cardiac PET imaging for the detection and monitoring of coronary artery disease and microvascular health”. In: *Journal of American College of Cardiology* (3 2010), 623–640.

[134] S. Klein et al. “Elastix: a toolbox for intensity based medical image registration"”. In: *IEEE Transactions on Medical Imaging* 29 (1 2010), pp. 196–205.

[135] D.P. Shamonin et al. “Fast Parallel Image Registration on CPU and GPU for Diagnostic Classification of Alzheimer’s Disease”. In: *Frontiers in Neuroinformatics* 7 (50 2014), pp. 1–15.