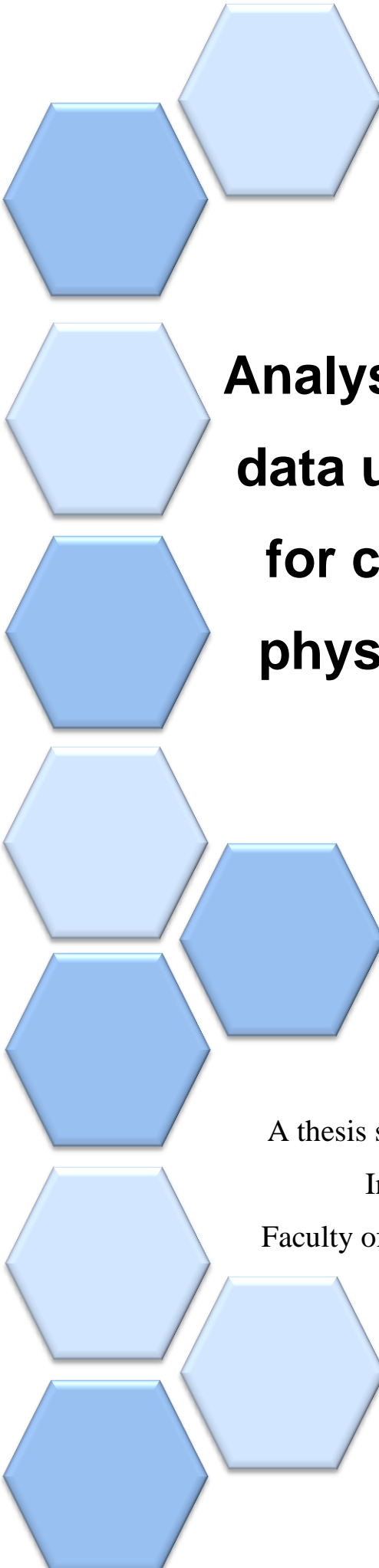




FCTUC

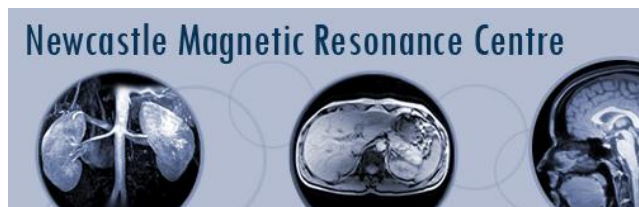
A decorative pattern of blue hexagons on the left side of the page. The hexagons are arranged in a vertical column, with some overlapping and some slightly offset to the left or right, creating a sense of depth and movement. The colors range from a light sky blue to a medium blue.

Analysis of non-invasive imaging data using self-organising maps for characterization of tumour physiological and biochemical properties

Mariana Simões Madeira

A thesis submitted for the degree of Master in Science (MSc)
Integrated in Biomedical Engineering Course
Faculty of Sciences and Technology of University of Coimbra
(FCTUC)

September 2012



Work developed in:

Newcastle Magnetic Resonance Centre
Newcastle upon Tyne
United Kingdom

In cooperation with:

Faculty of Medicine of the University of Coimbra – IBILI

Under supervision of:

Dr Ross Maxwell
Dr Francisco Caramelo

A thesis submitted for the degree of Master in Science (MSc)
Integrated in Biomedical Engineering Course
Faculty of Sciences and Technology of University of Coimbra (FCTUC)

September 2012



To my mum

Acknowledgments

I would like to acknowledge those that gave me great support during this year and made this thesis possible. Firstly I would like to thank my supervisor in Newcastle upon Tyne, Dr Ross Maxwell, who gave me the opportunity to come here to develop this work, who taught me so much. His huge knowledge helped me to deal and solve problems that arise along the development of this work. I would like to thank also my supervisor from Coimbra, Dr Francisco Caramelo for the important help and guidance, giving me complete freedom and supervision, even from such a long distance.

I am thankful to those that were closely involved in this work, Gilberto Almeida and Ian Wilson. Both taught me the most important principles of PET experiments, gave me their friendship, help and all necessary data to develop this work. Thank you!

I am also grateful to members of the clinical team, Dr Rachel Pearson, Dr Noor Haris and Elizabeth Howell with whom I have worked on patient data analysis.

Other important people for the development of this project were my office colleagues and now friends, Dr Eva Kenny and Dr Fiona Smith, with whom I improved my English and all the staff at the Magnetic Resonance Centre that received me there so well.

This thesis is the end of a cycle; as such, I have to thank all that accompany me on this journey, especially my friends Sofia Sousa, Zé Miguel, Rui Martins, Sónia, André Ventura, João, Sara, Vitor. Thank you for making this journey unforgettable.

To all my friends I ever lived with, who gave me so much, especially João Carvalho, Joana Carvalho, Sara Vale, Susana Pereira, Nica, Yuki and Eduardo Machado.

Thank you to all my new friends in Newcastle, in particular to Felipe, Ferran and Rafa that heard me speaking about my work and helped me so much.

To my parents, my sister, Telmo and my beautiful niece who is almost with us. Thank you all for believing and supporting me, even missing me as much as I do. To my uncles, cousins, and godparents my big thank you for always keeping in touch sending pictures and news to make this distance shorter. To my grandparents that I miss a lot and I know that always will be looking for me.

Finally a whole paragraph to thank the person that has always believed in me and that is responsible for this successful achievement: João. Thank you for teaching me so many things so useful for this project and life, for the patient, especially during these final months. What you gave me is much more than I could ever express by words. We did it and we will do much more. Thank you!

Abstract

Nowadays computer science applied to cancer research is of great importance. In this work, unsupervised and supervised learning techniques were applied to [^{18}F]Fluorothymidine Positron Emission Tomography ([^{18}F]FLT-PET) data.

As supervised learning we used an approach named Kinetic Spatial Filtering (KSF) developed at Imperial College London based on uptake of the radiopharmaceutical ([^{18}F]FLT) that differs from tissue to tissue. It is a filter shown to separate tumour and vertebra from other tissues in patients. However, when applied to images of mice implanted with HCT116 cancer cell lines, even when mouse time activity curves were used as template, it failed to achieve similar results to the ones obtained in analysis of patient images with pancreatic cancer with liver metastases where the filter was able to distinguish the tumour from the surrounding organs. This analysis intends to be used to detect early tumour response after one cycle of treatment.

As unsupervised learning we used Self-Organizing Maps (SOM). We applied this algorithm in [^{18}F]FLT-PET data from mice. First we applied SOM for HCT116 data analysis. The objective was successfully achieved since these maps could differentiate pre from post treatment images. In a second application, we used SOM with data from mice implanted with A2780, SJSA, SN40R2 and HT29 cancer cell lines, to see if we could distinguish and classify the tumours. We could only distinguish SJSA from other tumours. In two of the mice studies, data inputs for the SOM were voxels obtained from drawn regions of interest (ROI): tumour, muscle, bladder, heart and kidney. Characteristics of the derived maps were the clustering patterns of voxels from each tissue type. As final result, SOM was able to cluster the data as follows: tumour, muscle and bladder formed one cluster, heart and kidney formed another more diffuse cluster. When applying SOM just to tumour ROIs, the separation between different tumours was partly achieved. On a third approach we used this classification method on patient data, to see if it was able to distinguish between pre and post treatment data, so that we could evaluate if the patient was responding or not to treatment. We found that the algorithm was able to perform this task. Again, in this data, ROIs were drawn around tumour, vertebra, heart, liver and kidney. Here liver, tumour, vertebra and kidney performed one cluster and heart another.

Sumário

Nos dias de hoje a informática aplicada à investigação em cancro é de grande importância. Neste trabalho, técnicas de aprendizagem não supervisionada e supervisionada foram aplicadas a dados de [18F]Fluorotimidina – Tomografia por Emissão de Positrões ([¹⁸F]FLT-PET).

Como aprendizagem supervisionada usámos um filtro chamado Filtro Cinético Espacial (KSF) desenvolvido no Imperial College que é baseado na captação do radiofármaco ([¹⁸F]FLT) que difere de tecido para tecido. É um filtro capaz de separar tumor e vértebra dos restantes tecidos. Contudo, quando aplicado a imagens de ratinhos implantados com linhas celulares de cancro de HCT116, mesmo quando curvas de actividade ao longo do tempo do próprio ratinho foram usadas como modelo, este não conseguiu obter os mesmos resultados que obtivemos na análise de dados de um paciente com cancro no pâncreas com metastases no fígado onde o filtro foi capaz de distinguir o tumor dos órgãos circundantes. Esta análise pretende ser usada para detectar resposta precoce depois de um ciclo de tratamentos.

Como aprendizagem não supervisionada usámos Self-Organizing Maps (SOM). Aplicámos este algoritmo em dados de [¹⁸F]FLT-PET de ratinhos. Primeiro aplicámos SOM na análise de dados de HCT116. O objectivo foi alcançado com sucesso, uma vez que os mapas finais conseguiram diferenciar imagens de pré das de pós tratamento. Numa segunda aplicação, usámos SOM com dados de ratinhos implatados com linhas celulares de A2780, SJSA, SN40R2 e HT29 para ver se este conseguia distinguir e classificar os tumores. Conseguimos apenas distinguir SJSA das restantes linhas celulares. Em ambos os estudos, dados de input de SOM eram voxels obtidos de regiões de interesse: tumor, músculo, bexiga, coração e rim. Características dos mapas finais são os padrões de clustering dos voxels de cada tecido. No resultado final, o SOM conseguiu os seguintes clusters: tumor, músculo e bexiga num cluster e coração e rins noutra mais difuso. Quando aplicámos SOM apenas a ROIs sobre os tumores, a separação entre diferentes tumores foi apenas parcialmente conseguida.

Numa terceira aplicação usámos este método de classificação nos dados da paciente, para ver se este conseguia distinguir entre pré o pós tratamento, no sentido de poder avaliar se o paciente estaria a responder ou não ao tratamento. Descobrimos que o

algoritmo era capaz efectuar esta tarefa. Aqui o algoritmo agrupou fígado, tumor, vértebra e rins e colocou o coração noutra cluster.

Table of contents

ACKNOWLEDGMENTS	VI
ABSTRACT	VIII
SUMÁRIO	IX
TABLE OF CONTENTS	XI
LIST OF TABLES	XIII
LIST OF FIGURES	XIV
EQUATIONS	XVII
ABBREVIATIONS	XVIII
CHAPTER 1. INTRODUCTION	20
CHAPTER 2. GENERAL BACKGROUND	23
2.1 TUMOUR PHYSIOLOGICAL AND BIOCHEMICAL PROPERTIES	23
2.2 POSITRON EMISSION TOMOGRAPHY – PET	24
2.2.1 PROCEDURE	25
2.2.2 FUNDAMENTALS	25
2.2.3 DATA ACQUISITION	29
2.2.4 PET UNITS	29
2.2.5 PET/CT DATA ACQUISITION	30
2.3 RADIONUCLIDES AND RADIOPHARMACEUTICALS	30
2.3.1 [¹⁸ F] – FDG	33
2.3.2 [¹⁸ F] – FLT	34
2.4 KINETIC SPATIAL FILTER – BASED ON (GRAY <i>ET AL.</i> , 2010)	35
2.5 SELF-ORGANIZING MAPS	38
2.5.1 THE BASIC PROCESS	40

2.5.2 APPLICATIONS OF SOM	42
<u>CHAPTER 3. METHODS</u>	<u>43</u>
3.1 DATA ACQUISITION	43
3.1.1 PRE-CLINICAL (MOUSE STUDIES)	43
3.1.2 CLINICAL (HUMAN STUDY)	46
3.2 KINETIC SPATIAL FILTER (KSF)	47
3.2.1 PRE-CLINICAL (MOUSE)	47
3.2.2 CLINICAL (HUMAN)	48
3.3 SELF-ORGANIZING MAPS	48
3.3.1 PRE-CLINICAL (MOUSE)	49
3.3.2 CLINICAL (HUMAN)	50
<u>CHAPTER 4. RESULTS</u>	<u>51</u>
4.1 KINETIC SPATIAL FILTER (KSF)	51
4.1.1 PRE-CLINICAL (MOUSE)	51
4.1.2 CLINICAL (HUMAN)	53
4.2 SELF-ORGANIZING MAPS	58
4.2.1 PRE-CLINICAL (MOUSE)	58
4.2.2 CLINICAL (HUMAN)	65
<u>CHAPTER 5. DISCUSSION</u>	<u>67</u>
5.1 ANALYSIS WITH KSF	67
5.2 ANALYSIS WITH SOM	68
5.3 GENERAL DISCUSSION	70
<u>CHAPTER 6. CONCLUSION</u>	<u>71</u>
<u>APPENDIX</u>	<u>72</u>
<u>REFERENCES</u>	<u>74</u>

List of tables

TABLE 1 – PROPERTIES OF DIFFERENT SOLID SCINTILLATOR CRYSTALS FOR PET DETECTORS (ADAPTED FROM (CHERRY, 2006)	28
TABLE 2 – PROPERTIES OF THE MOST COMMON POSITRON EMITTERS.	31
TABLE 3 – MOST COMMONLY USED RADIOPHARMACEUTICALS. (ADAPTED FROM (HAUBNER, 2010))	32
TABLE 4 – TABLE WITH THE DESIGN OF MICE TREATMENT, IN HCT116 TUMOURS. (†) DEAD MOUSE.....	45

List of figures

FIGURE 1 – THE TWENTY MOST COMMONLY DIAGNOSED CANCERS EXCLUDING NON-MELANOMA SKIN CANCER IN UK 2009 (CANCERRESEARCH).....	20
FIGURE 2 – PET SCANNERS FOR PATIENTS (A) AND MICE (B) FROM NEWCASTLE UNIVERSITY.....	24
FIGURE 3 – THE PROCESS OF POSITRON EMISSION AND SUBSEQUENT POSITRON-ELECTRON ANNIHILATION WHICH PRODUCES TWO 511KeV PHOTONS EMITTED IN OPPOSITE DIRECTIONS (~180°). (ADAPTED FROM (SAHA, 2010))	26
FIGURE 4 – FALSE COINCIDENCES DETECTED. THE COMPTON INTERACTION WILL YIELD A SCATTERING EVENT (A) AND THE ACCIDENTAL INTERACTION WILL YIELD A RANDOM COINCIDENCE (B) (ADAPTED FROM(FERREIRA, 2009))	26
FIGURE 5 – COMPTON SCATTERING INTERACTION (ADAPTED FROM (CHERRY, 2006)).....	27
FIGURE 6 – SCINTILLATION DETECTOR OF PET SCANNER. (ADAPTED FROM (FERREIRA, 2009)).....	28
FIGURE 7 – [¹⁸ F]FDG UPTAKE INTO TUMOUR CELLS (HAUBNER, 2010).	33
FIGURE 8 - [¹⁸ F]FLT UPTAKE INTO TUMOUR CELLS(HAUBNER, 2010).	35
FIGURE 9 – KINETIC CURVES USED FOR THE CLASSIFICATION ALGORITHM IN PATIENTS. THESE CURVES ARE DERIVED FROM DE TEMPLATES USED IN THE KSF. THE ERROR BARS REPRESENT THE STANDARD ERROR OF THE MEAN (SEM) WHERE WE HAVE THE FLT UPTAKE FOR HEART (A), LIVER (B), LUNG (C), NORMAL BREAST (D), TUMOUR (E) AND VERTEBRA (F).	37
FIGURE 10 – KINETIC CURVES USED FOR THE CLASSIFICATION ALGORITHM IN MICE. THE ERROR BARS REPRESENT THE STANDARD ERROR OF THE MEAN (SEM) WHERE WE HAVE THE FLT UPTAKE FOR HEART (A), KIDNEY (B), BLADDER (C), MUSCLE (D) AND TUMOUR (E).....	38
FIGURE 11 – COMPETITION STEP OF SOM. ANY GIVEN INPUT NODE IS COMPARED TO THE WEIGHT VECTOR OF EACH MAP NODE (OR NEURON) AND THE CLOSEST MAP NODE IS THE WINNER.....	40
FIGURE 12 – COOPERATION STEP OF SOM. HERE THE DIAMETER BEGINS IN THE OUTER DASHED LINE AND IS GOING TO SHRINK TO THE INNER ONE WITH THE ITERATIONS(KOHONEN <i>ET AL.</i> , 1996A).....	41
FIGURE 13 – ADAPTATION STEP OF THE SOM (KOHONEN <i>ET AL.</i> , 1996A). THE WINNER MAP NODE (NEURON), REPRESENTED HERE BY SPHERES, AND ITS NEIGHBOURS ARE GOING TO ADAPT TO MAKE THEIR WEIGHT VECTORS IDENTICAL TO THE INPUT NODES.....	41
FIGURE 14 – IMAGE OF THE THREE MICE PLACED ON PET SCANNER.	44
FIGURE 15 – TRANSVERSE SLICES OF THE [¹⁸ F]FLT-PET SCAN OF MICE IMPLANTED WITH HCT116 DATA, BEFORE APPLYING THE KSF. THIS IMAGE IS FROM MOUSE 2, DAY 0. THE YELLOW ARROWS ARE INDICATING THE TUMOUR REGION.	51
FIGURE 16 – TRANSVERSE SLICES OF THE [¹⁸ F]FLT-PET SCAN OF MICE IMPLANTED WITH HCT116 DATA, AFTER APPLYING THE KSF. THE YELLOW ARROWS ARE INDICATING THE TUMOUR REGION.....	52
FIGURE 17 – IMAGE FROM THE FIRST [¹⁸ F]FLT-PET PATIENT SCAN NORMAL (A) AND FILTERED USING KSF (B). TUMOURS ARE INDICATED WITH YELLOW ARROWS. NOTE THE INTENSITY BAR AND THE Z-AXIS.	

AT THE BOTTOM WE ZOOM THE LAST SLICE ($z = 37$) TO HELP VISUALIZATION OF KSF. THE LEGEND FOR LIVER, VERTEBRA AND KIDNEY ARE COMMON FOR (A) AND (B).	54
FIGURE 18 - IMAGE OF THE SECOND (POST TREATMENT) [^{18}F]FLT-PET SCAN NORMAL (A) AND FILTERED USING KSF (B). TUMOURS ARE INDICATED WITH YELLOW ARROWS. NOTE THE INTENSITY BAR AND THE Z-AXIS. AT THE BOTTOM WE ZOOM THE LAST SLICE ($z = 37$) TO HELP VISUALIZATION OF KSF. THE LEGEND FOR LIVER, VERTEBRA AND KIDNEY ARE COMMON FOR (A) AND (B).	56
FIGURE 19 – CT IMAGES OF THE PATIENT. THE YELLOW ARROW SHOWS THE TUMOUR REGION. THE BLACK DOT IN THE TUMOUR REGION IS THE CATHETER LEFT THERE ON PURPOSE TO HELP TO IDENTIFY THE TUMOUR IN THE IMAGE.	57
FIGURE 20 – [^{18}F]FLT UPTAKE CURVES IN THE REGION THOUGHT TO BE THE TUMOUR. THE GRAPHS HAVE SUV UPTAKE AGAINST TIME FRAMES.	57
FIGURE 21 – VISUALIZATION OF THE SOM APPLIED TO HCT116: DATA FROM MOUSE 4. U-MATRIX IS ON THE TOP LEFT, THEN WE HAVE COMPONENT PLANES FOR DAY0 AND DAY2 AND ON BOTTOM RIGHT WE HAVE MAP UNIT LABELS. IN THE BOTTOM RIGHT MAP THE LABELS MEAN: H IS FOR HEART, T FOR TUMOUR, B FOR BLADDER AND K FOR KIDNEY. ($QE=0.032$ AND $TE=0.097$).....	59
FIGURE 22 – VISUALIZATION OF THE SOM APPLIED TO HCT116: DATA FROM MOUSE 5. U-MATRIX IS ON THE TOP LEFT, THEN WE HAVE COMPONENT PLANES FOR DAY0 AND DAY2 AND ON BOTTOM RIGHT WE HAVE MAP UNIT LABELS. IN THE BOTTOM RIGHT MAP THE LABELS MEAN: H IS FOR HEART, T FOR TUMOUR, B FOR BLADDER AND K FOR KIDNEY. ($QE=0.090$ AND $TE=0.028$).....	60
FIGURE 23 – VISUALIZATION OF THE SOM APPLIED TO HCT116: DATA FROM MOUSE 6. U-MATRIX IS ON THE TOP LEFT, THEN WE HAVE COMPONENT PLANES FOR DAY0 AND DAY2 AND ON BOTTOM RIGHT WE HAVE MAP UNIT LABELS. IN THE BOTTOM RIGHT MAP THE LABELS MEAN: H IS FOR HEART, T FOR TUMOUR, B FOR BLADDER AND K FOR KIDNEY. ($QE=0.154$ AND $TE=0.107$).....	61
FIGURE 24 – VISUALIZATION OF THE SOM APPLIED TO HCT116: DATA FROM MOUSE 8. U-MATRIX IS ON THE TOP LEFT, THEN WE HAVE COMPONENT PLANES FOR DAY0 AND DAY2 AND ON BOTTOM RIGHT WE HAVE MAP UNIT LABELS. IN THE BOTTOM RIGHT MAP THE LABELS MEAN: H IS FOR HEART, T FOR TUMOUR, B FOR BLADDER AND K FOR KIDNEY. ($QE=0.060$ AND $TE=0.027$).....	62
FIGURE 25 – VISUALIZATION OF THE SOM APPLIED TO ALL TUMOUR CELL LINES: A2780, SJSA, SN40R2 AND HT29. U-MATRIX IS ON THE TOP LEFT, THEN WE HAVE COMPONENT PLANES FOR DAY0 AND DAY2 AND ON BOTTOM RIGHT WE HAVE MAP UNIT LABELS. IN THE BOTTOM RIGHT MAP THE LABELS MEAN: H IS FOR HEART, T FOR TUMOUR, B FOR BLADDER AND K FOR KIDNEY($QE=0.172$ AND $TE=0.103$).	63
FIGURE 26 – KINETIC UPTAKE CURVES OF ALL THE MICE TUMOUR CELL LINES: A2780, SJSA, SN40R2 AND HT29. CURVES WERE TAKEN FROM MOUSE 1 (M1), MOUSE 2 (M2) AND MOUSE 3 (M3) OF ALL THE TUMOUR TYPES.....	64
FIGURE 27 – VISUALIZATION OF THE SOM APPLIED TO ALL TUMOUR CELL LINES: A2780, SJSA, SN40R2 AND HT29. U-MATRIX IS ON THE TOP LEFT, THEN WE HAVE COMPONENT PLANES FOR $t=22$, $t=23$ AND $t=24$ AND ON BOTTOM RIGHT WE HAVE MAP UNIT LABELS. IN THE BOTTOM RIGHT MAP THE LABELS MEAN: H IS HT29, SN FOR SN40R2, SJ FOR SJSA AND A FOR A2780.....	65
FIGURE 28 – VISUALIZATION OF THE SOM APPLIED TO PATIENT DATA. WE HAVE DAY 0 AS DATA FROM FIRST SCAN AND DAY 2 AS DATA FROM THE SECOND SCAN. U-MATRIX IS ON THE TOP LEFT, THEN WE HAVE COMPONENT PLANES FOR DAY0 AND DAY2 AND ON BOTTOM RIGHT WE HAVE MAP UNIT	

LABELS. IN THE BOTTOM RIGHT MAP THE LABELS MEAN: H IS FOR HEART, T FOR TUMOUR, L FOR LIVER, V FOR VERTEBRA AND K FOR KIDNEY. (QE=0.040 AND TE=0.108) 66

Equations

EQUATION 1	25
EQUATION 2	36
EQUATION 3	40
EQUATION 4	44

Abbreviations

ANN	Artificial Neural Networks
BMU	Best Matching Unit
CT	Computed Tomography
DICOM	Digital Imaging and Communications in Medicine
DNA	Deoxyribonucleic Acid
DPK	(Nucleotide) Diphosphate Kinase
FDG	Fluorodeoxyglucose
FLT	Fluorothymidine
FOV	Field of View
GF	Growth Factors
GUI	Graphical User Interface
HK II	Hexokinase II
KSF	Kinetic Spatial Filter
LOR	Line of Response
LSO	Lutetium Oxyorthosilicate
MRI	Magnetic Resonance Imaging
NT	Nucleoside Transporters
PCA	Principal Component Analysis
PET	Positron Emission Tomography
PMT	Photomultiplier Tube
QE	Quantization Error
ROI	Region of Interest
SEM	Standard Error of the Mean
SOM	Self-Organizing Maps
SUV	Standardized Uptake Value
TAC	Time Activity Curve
TE	Topographic Error
TK1	Thymidine Kinase 1
TPK	(Nucleotide) Triphosphate Kinase

Chapter 1. Introduction

Throughout history there have been diseases concerning human civilization. In the bible we have evidence that show leprosy as the most fearful disease, then in the Middle Ages and Renaissance in Europe, the Plague haunted the human race and last century, it was tuberculosis that terrified the people (Pitot & Loeb, 2002). This infectious disease started to be less significant worldwide in the twentieth century, with the development of antimicrobial therapy.

Since the last century, the most frightening disease in developed countries is cancer (Tobias *et al.*, 2010).

Cancer has long been known as a human infirmity. However, records showing cancer as a cause of death first appeared in Europe in the eighteenth century. Since then, the mortality rate for cancer increased and accentuated after the industrialization of the nineteenth century (Glade, 1999).

In 2009, 320467 new cases of cancer were diagnosed in the United Kingdom and more than 1 in 3 people will develop some form of cancer during their lifetime. In the United Kingdom, the most common types of cancer are breast cancer, prostate cancer, lung cancer, bowel cancer, bladder cancer, uterine cancer (NHS, 2012).

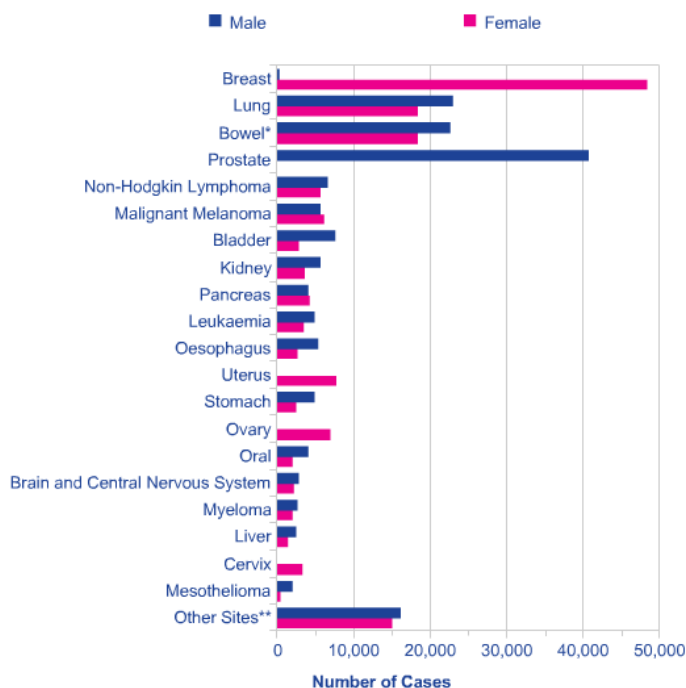


Figure 1 – The twenty most commonly diagnosed cancers excluding non-melanoma skin cancer in UK 2009 (CancerResearch)

In Portugal there are more than forty thousand new cancer patients diagnosed every year and more than 20 thousand die from this disease (LPCC, 2012).

World-wide, millions of people live with cancer diagnosis and it is thought that in 2020, the number of people diagnosed with cancer will rise from 10 million new cases every year to 16 million, with the deaths rising from 7 million to 10 million. Taking into account that the average life expectancy is growing every year, what we can take from here is that the number of cancer cases will rise, but more people will survive this condition (Nakajima, 2003; Tobias *et al.*, 2010).

For many years, the way to fight this pathology was based on surgery and radiotherapy. Nowadays, due to high level of research and developments in chemotherapy and radiotherapy, the planning in cancer treatment looks at each case as an individual. There are collaborations established on a national and international scale leading to an improvement in cancer treatment (Pitot & Loeb, 2002; Tobias *et al.*, 2010).

Cancer treatments are expensive and substantially toxic. To proceed with them, the doctors need to be completely sure of their efficacy in each individual patient. They need to be sure that they are increasing survival or improving quality of life somehow. For that reason, the collaboration between specialists is of massive importance: surgeons, specialists of the area where the tumour is located and radiologists need to work in cooperation (Pitot & Loeb, 2002; LPCC, 2012; NHS, 2012).

Before the treatment is started, the team need to be sure about the type of cancer that the patient has. Tumour pathology must be defined and the extent of local and systemic disease must be determined. It is in this last section that histopathology, cytology, haematology, chemical pathology and diagnostic imaging help (Pitot & Loeb, 2002; LPCC, 2012; NHS, 2012).

Accuracy in the staging of the tumour, prior to and after surgery, is achieved nowadays thanks to modern imaging techniques (e.g. Positron Emission Tomography (PET), Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Ultrasound). The techniques are expensive and the information that we get from there needs to be assembled with other clinical results to accomplish good answers (Tobias *et al.*, 2010).

The main problem with the imaging is due to human bias presented in all the analysis. That is the main reason why currently large investigation is being developed in the field of Oncology together with Pattern Recognition/Classification imaging methods. The aim of this coupling is to allow the detection of cancer as early as possible, in situations

where the human eye cannot distinguish anything as pathological and more important, non-invasively. Since screening is one of the main weapons to fight this sickness, if we can access software that can give us this answer, without the need of being invasive and with high reliability, we could alert the patient and start the treatment at a pre-symptomatic stage, where it is easier to ensure a successful response (Tobias *et al.*, 2010).

That is where my project gains relevance and it is now time to introduce it.

Data are available from a wide range of imaging techniques and can include dynamic, metabolic and spatial information. Improved methods are needed to extract the maximum information from these multi-parametric datasets and so aid tumour characterization.

This study will evaluate the use of advanced pattern recognition techniques, especially self-organizing maps (an unsupervised neural network method), for analysis of dynamic PET data to help characterize and/or classify animal and human tumour models. This involved data that has been obtained at Newcastle University over the last three years.

Due to the cooperation between Imperial College London and Newcastle University, we had access to software developed there, named Kinetic Spatial Filter (KSF), that helps to identify the tumour area in pancreatic tumours. Because our clinical team is trying to develop and improve new protocols for this malignancy, they gave us their software and we tried that on our data. That is the other part of the analysis developed in this project.

This thesis is divided in five chapters. This first chapter makes a contextualisation of the motivations and objectives of the thesis. In the second chapter we have a general background over all the subjects developed in this project. In the third chapter we have the methods and respective results on the fourth chapter. In the fifth, we have discussion and conclusion of the thesis. Finally, we have one appendix with parallel work developed to help with the achievement of our results.

Chapter 2. General Background

2.1 Tumour physiological and biochemical properties

Cancer is a problem of cellular growth and differentiation. The tumour cell has two major characteristics: autonomy and anaplasia. They are autonomous because they get complete independence from the normal internal control of the cells, continuously proliferating and we say that they are anaplastic because they lose differentiation capability. The normal process of programmed cell death no longer operates leading to complete uncontrollable growth of the tumour cells (Macdonald *et al.*, 2004; Lehninger *et al.*, 2008).

The tumour cells increase the size and density of the nucleus, which means, their DNA (Deoxyribonucleic Acid) replicates. They have abnormal chromosomes and abnormal mitosis. They develop modifications in their cellular surface: loss or modification of the glycolipids/glycoproteins; alterations in cellular response to external Growth Factors (GF), responding only to their own GFs (autocrin stimulation). They exhibit abnormal cellular proliferation, alteration in cell to cell recognition and less adhesion than normal cells (Macdonald *et al.*, 2004; Tobias *et al.*, 2010).

Tumours can be divided into benign and malignant. The benign ones are usually not fatal. They are well separated from the organs and grow inside a well-defined capsule that permits its elimination with no major difficulties. Malignant tumours can spread through all the body and generate further growth and metastasis, therefore becoming life-threatening (Pitot & Loeb, 2002; Macdonald *et al.*, 2004).

In this project we used image data from mice and patient, all with malignant types of tumour.

In human we just used data from one patient with pancreatic cancer with liver metastases.

The mice data collected is from two types of colorectal tumour (HCT116 and HT29), two from osteosarcoma (SJSa and SN40R2) and one from ovarian type (A2780). These tumour cell lines are of particular interest for Newcastle University either because of the

frequency of these tumour types in the population or because useful genetic variants are available.

2.2 Positron Emission Tomography – PET

Positron Emission Tomography (PET) is a non-invasive nuclear imaging technique that emerged during the 1980's and is based in the interaction between radioisotopes and matter (Brownell, October 15, 1999). Using this technique we can scan a subject (human or animal) and obtain a two or three-dimensional (2D or 3D) image based on cellular function and physiological processes. It is one of the techniques of Nuclear Imaging and it is very useful for research due to his wide application in diagnosis of different diseases in a vast number of areas: Cardiology, Respiratory Disease, Endocrinology, Neurology, Oncology, etc. Clinical applications for PET continue to increase, particularly in the field that we are mostly interested in this study: Oncology (Sharp *et al.*, 2005). For each type of study, a different range of radiopharmaceuticals is used, depending on the organ and function that we want to visualize (e.g. Fluorodeoxyglucose (FDG) and Fluorothymidine (FLT) among many others) (Sattler *et al.*, 2010).



Figure 2 – PET scanners for patients (a) and mice (b) from Newcastle University.

2.2.1 Procedure

The first procedure is the injection, inhalation or even ingestion of the radiopharmaceutical by the patient. The criteria to choose this radiopharmaceutical is usually related to the organ and disease that we intend to visualize. After the injection there will typically be a waiting period where we wait for the biodistribution of the radiopharmaceutical. This interval between the first step and the beginning of the scan is dependent on the organ that we want to visualize (Saha, 2010).

2.2.2 Fundamentals

‘Radioactivity is a property of atomic nuclei and may be defined as the spontaneous transformation of a structurally unstable nucleus to a structurally more stable nucleus, with the emission of energy in the form of ionizing radiation’ (Biersack, 2007). These transformations are known as radioactive decays and occur by emission of radiation such as α particles, β^- particles, β^+ particles, electron capture and isomeric transition(Saha, 2010). These radioactive isotopes are termed radionuclides.

Out of all the above radioactive decays, the one that we are interested in for PET studies is the positron (β^+) decay characteristic of radionuclides which are proton rich or neutron deficient. In essence, a proton in the nucleus is converted to a neutron in the process (Phelps, 2006; Saha, 2010):



Where p is the proton, n is the neutron, β^+ is the positron and ν is the neutrino.

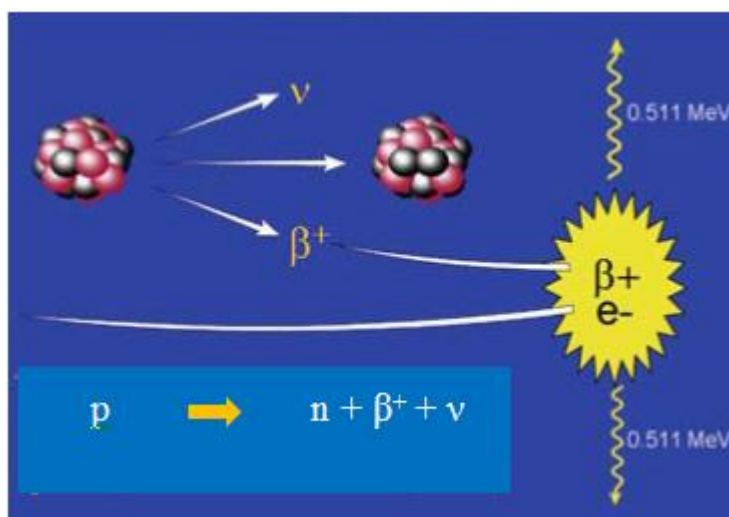


Figure 3 – The process of positron emission and subsequent positron-electron annihilation which produces two 511KeV photons emitted in opposite directions (~180°). (Adapted from (Saha, 2010))

In Figure 3, we have the decay process of a radionuclide, producing a positron that annihilates with an electron, emitting two γ photons (511KeV of Energy) in opposite directions (~180°). The site of annihilation is usually very close to the point of positron emission because the emitted positrons rapidly lose their energy in tissue (Cherry, 2006). If these two photons are detected by the detectors of the PET scanner within a short (~ns) timing window (the coincidence timing window), an event called true coincidence is recorded along a straight line that connects the centres of the two detectors – denominated the Line of Response (LOR). This is the basis of PET scanning (Saha, 2010).

There are some events that can yield false coincidences. These are present in Figure 4.

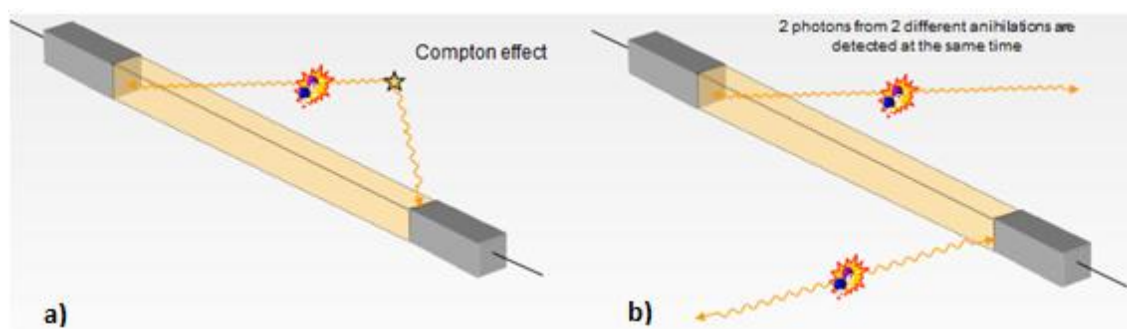


Figure 4 – False coincidences detected. The Compton interaction will yield a scattering event (a) and the accidental interaction will yield a random coincidence (b) (Adapted from(Ferreira, 2009))

In Figure 4 a) is present the Compton scattering interaction (Ollinger, 1997). This is a result of Compton Scattering of one or both photons with matter. The Compton Scattering interaction results from a partial energy transfer from γ photon to an outer shell electron, forcing it to change the direction and lose some of its energy (See Figure 5) (Ollinger, 1997; Hendee & Ritenour, 2002).

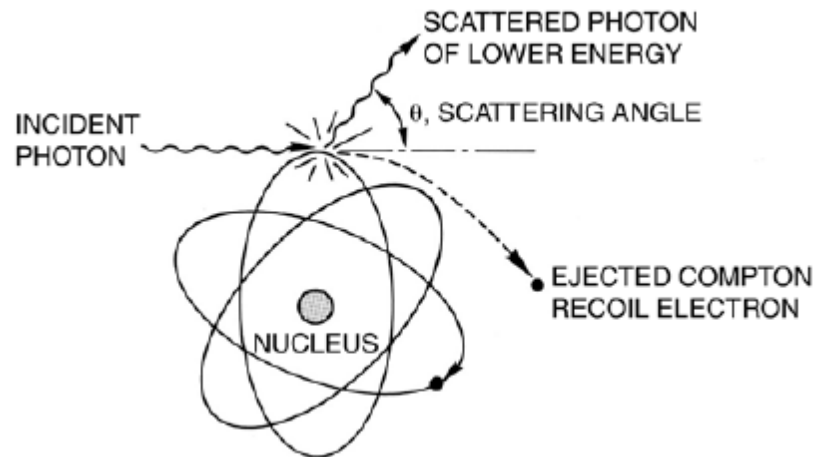


Figure 5 – Compton Scattering interaction (Adapted from (Cherry, 2006)).

It is important to point that whether this is counted as a coincidence event, depends on the final photon energy compared to the energy window set for the detectors.

To understand the accidental interaction (Figure 4 b)), we need to know that there are a large number of scattered photons. As a result, there will be single photons detected. If two single photons from different annihilations are detected, we have an accidental interaction. (Hendee & Ritenour, 2002; Cherry, 2006).

The forthcoming process after the annihilation is the detection of these γ photons by the detectors of the PET scanner.

These are solid scintillation detectors that after absorbing γ radiation emit photons of light that will be converted to an electrical signal by photomultiplier tubes (PMTs). This electrical current is then registered as a count (Cherry, 2006; Saha, 2010). The configuration of these detectors is given in Figure 6. As shown there, a block of detectors consists of an array of individual scintillation crystals viewed by a number of PMTs. All these components together generate a module that will be part of the PET scanner ring that surrounds the patient.

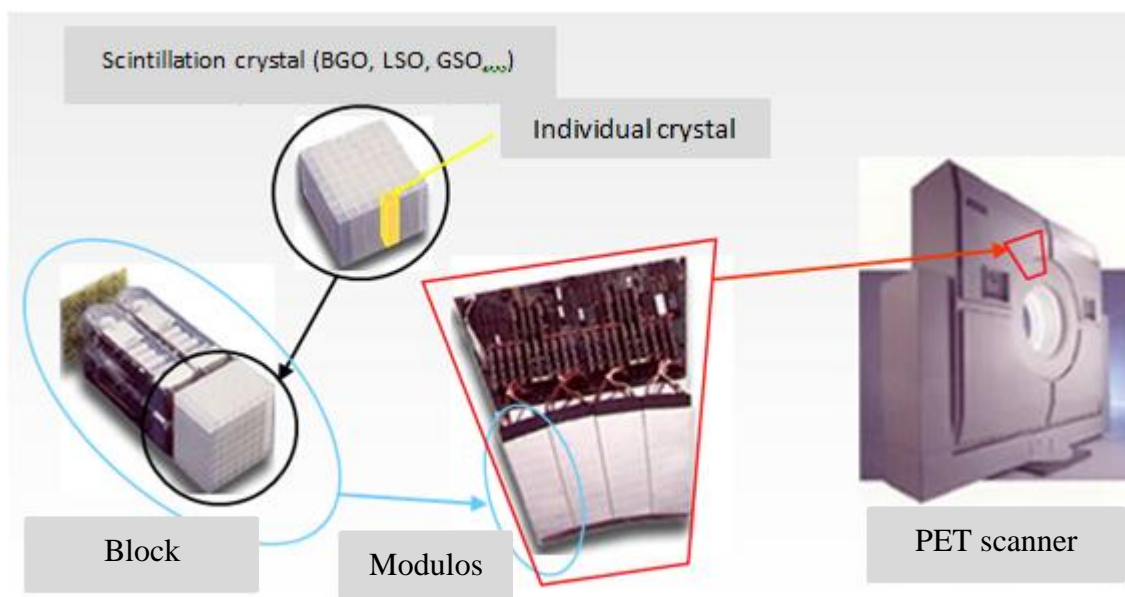


Figure 6 – Scintillation detector of PET scanner. (Adapted from (Ferreira, 2009))

These detectors need to have some characteristics that make them optimum for the detection of the γ photons. They should have good stopping power for 511KeV (i.e. high density), they should have quick reaction time, good energy resolution, low Compton scatter inside the detector crystal and matching of wavelength of fluorescence to response of the light detector (Sharp *et al.*, 2005). In Table 1 we have some of the most commonly used materials for solid scintillator crystals detectors and its respective characteristics.

Table 1 – Properties of different solid scintillator crystals for PET detectors (adapted from (Cherry, 2006))

Scintillation	Density (g/cc)	Light output (photons per 511Kev)	Decay time (ns)	Index of refraction	Linear attenuation at 511KeV(cm^{-1})	Ratio Between Photoelectric and Compton
Sodium Iodide [Nai(Tl)]	3.67	19400	230	1.85	0.34	0.22
Bismuth Germanate (BGO)	7.13	4200	300	2.15	0.96	0.78
Lutetium Oxyorthosilicate(LSO:Ce)	7.4	~13000	~47	1.82	0.88	0.52
Gadolinium Oxyorthosilicate(GSO:Ce)	6.71	~4600	~56	1.85	0.70	0.35
Barium Fluoride(BaF2)	4.89	700, 4900	0.6, 630	1.56	0.45	0.24
Yttium Aluminium Perovskite (YAP:Ce)	5.37	~9200	~27	1.95	0.46	0.05

With correct calibration, PET can yield very precise quantitative estimates of the concentration of the radiopharmaceutical, giving us precise quantitative measurements of specific physiological quantities such as blood flow, glucose metabolism, receptor binding characteristics, among many other physiological, biochemical and pharmacological processes (Ollinger, 1997). The biological pathway visualized, depends on the radiopharmaceutical chosen.

2.2.3 Data acquisition

The data acquisition in PET can be static or dynamic.

The most basic data acquisition is a static PET protocol that implies an acquisition of a single data set or static frame over a fixed length of time. The data obtained is useful to estimate the gross tracer uptake which means, the average tissue activity concentration of the tracer. Usually this acquisition mode is applied in studies where the concentration of the tracer is stable or the best time for imaging is already well defined (e.g. [¹⁸F]-FDG)(Phelps, 2006; Saha, 2010).

In dynamic PET, the data is acquired as a sequence of dynamic time frames. Here the information given is relevant to understand how the tracer biodistribution is done along the time of the acquisition in the body. A kinetic curve of the tracer can be achieved. The studies developed on this work are all based on dynamic data acquisition, in particular, the KSF is based on the kinetic curves obtained from this PET dynamic protocol (Phelps, 2006; Saha, 2010).

2.2.4 PET Units

Standard Uptake Value (SUV) is broadly used as the standard unit in PET studies, and this project was not exception. All the analyses were developed using SUV units.

The SUV is the concentration of activity at the region of interest (ROI), which is normalized for body weight and to injected dose activity.

$$SUV = \frac{\text{Concentration of the radionuclide (MBq/ml)}}{\text{Injected Activity (MBq) / Weight (g)}}$$

For mice analysis, we used Philips PET scanner. PET statistics in the Philips scanner are shown, by default, in counts. To visualize data in Standardized Uptake Value (SUV)

instead of counts, we need to calibrate the scanner to convert counts per pixel into millicuries per cubic cm (mCi/cc) (PhilipsManual, 2010). We can also convert the data from counts to SUV, after the acquisition, as described in methods.

For patient data, the scanner used was Siemens and the final data was already in SUV with no need for conversions.

2.2.5 PET/CT Data acquisition

A multimodality scanner can acquire structure and functional information in the same study. With the CT we can more accurately identify the anatomic structures and with PET we can determine where, for example, an abnormal function, such as a tumour, is present. When we have these images precisely coregistered, we can more precisely determine where in the body the tumour is (Townsend, 2008).

The introduction of combined PET/CT scanners to the market was in 2001 and nowadays it is fully adopted in all the Nuclear Medicine Services, but specially used for oncology studies (Muehlehner & Karp, 2006; Townsend, 2008).

In this project, for the patient study the PET/CT scanner was used. The objective of the KSF filter was to identify the tumour region of interest in the neighbourhood of high [¹⁸F]FLT signals from the liver in a PET study. CT images, especially making use of CT contrast agent injection, have been used to help identify the pancreatic tumour in this patient.

2.3 Radionuclides and radiopharmaceuticals

The majority of the radionuclides are produced artificially. They are atoms with unstable nuclei, due to excess or deficit of neutrons. Tending towards stability, they will spontaneously decay (radioactive decay), emitting different types of radiation: α particles, β^- particles, β^+ particles, electron capture and isomeric transition. The practical applications of these radionuclides are dependent on their properties and interaction with matter. As stated before, the radionuclides that we are interested in for PET study purposes are the ones that decay with positron (β^+) emission (deficit of neutrons) (Allisy-Roberts *et al.*, 2007).

The characteristics to look at when choosing the radionuclide for *in vivo* studies are the half-life – it needs to have a half-life big enough to permit transport from the production source until image acquisition but it cannot be too high, to avoid the unnecessary

exposure of the patient to the radiation. We need to be also aware of the radiotoxicity of the radionuclide, which is dependent on the nature and energy of the emitted radiation, on the tissues' uptake of the radionuclide and finally on the pharmacokinetic elimination of the radionuclide. Another important aspect is the relation between cost and benefit. There are some interesting radionuclides, but due to high production costs, they are generally not used (e.g. ^{67}Cu). Finally, the perfect radionuclide should be easy to attach to the pharmaceutical (Hendee & Ritenour, 2002).

In

Table 2 we have some of the more commonly used radionuclides with respective properties, where half-life is the time that the activity of the radionuclide will fall to one-half. The Maximum Energy (E_{max}) is the energy released in the positron after the decay (Martin *et al.*, 2003). The β^+ Branching Fraction is the probability that the radionuclide will decay with positron emission (β^+) (Hendee & Ritenour, 2002). Finally, the precursor columns show the isotopes that are used to generate the radioisotope by the reaction showed in parenthesis.

Table 2 – Properties of the most common positron emitters.

Radionuclide	Half-life	E_{max} (MeV)	β^+ Branching Fraction	Precursor
^{11}C	20.4 min	0.96	1.00	$^{14}\text{N}(\text{p},\alpha)^{11}\text{C}$
^{13}N	9.97 min	1.2	1.00	$^{16}\text{O}(\text{p},\alpha)^{13}\text{N}$
^{15}O	2.03min	1.73	1.00	$^{14}\text{N}(\text{d},\text{n})^{15}\text{O}$
^{18}F	109.8min	0.63	0.97	$^{18}\text{O}(\text{p},\text{n})^{18}\text{F}$

Radiopharmaceuticals are biochemical or drug molecules labelled with a radionuclide. It consists of two parts: a carrier molecule and the incorporated radionuclide. With the proper radiopharmaceutical we can image specific biochemical processes or we can image a certain organ that we are interested in (Schiepers & Baert, 2006; Biersack, 2007).

The criterion to choose this radiopharmaceutical is the ability to target a specific organ or disease, its solubility and capacity to cross cell membranes, along with all the criteria stated before to choose the radionuclide (Guy & Ffytche, 2005).

Some of the more commonly used radiopharmaceuticals are presented in Table 3.

Table 3 – Most commonly used radiopharmaceuticals. (adapted from (Haubner, 2010))

Molecular uptake Mechanism	Tracer	Radio nuclide	Organs of highest physiological uptake	Availability
Amino acid transport and protein synthesis	Methionine	^{11}C	Liver, salivary glands, lachrymal glands, bone marrow, pancreas, bowels, renal cortical, urinary bladder	In house production/ cyclotron
	Fluoroethyltyrosine	^{18}F	Pancreas, kidneys, liver, heart, brain, colon, muscle	In house production/ cyclotron
	FDOPA	^{18}F	Pancreas, liver, duodenum, kidneys, gallbladder, biliary duct	Commercially available
Glucose metabolism	FDG	^{18}F	Brain, myocardium, breast, liver, spleen, stomach, intestine, kidney, urinary bladder, skeletal muscle, lymphatic tissue, bone marrow, salivary glands, thymus, uterus, ovaries, testicle, brown fat	Commercially available
Proliferation	FLT	^{18}F	Bone marrow, intestine, kidneys, urinary bladder, liver	In house production/ cyclotron
Hypoxia	FMISO	^{18}F	Liver, urinary excretion	In house production/ cyclotron
	FAZA	^{18}F	Kidneys, gallbladder, liver, colon	In house production/ cyclotron
	Cu-ATSM	^{64}Cu	Liver, kidneys, spleen, gallbladder	In house production/ cyclotron
Lipid Metabolism	Choline	^{11}C	Liver, pancreas, spleen, salivary glands, lachrymal glands, renal excretion, bone marrow, spleen	In house production/ cyclotron
	Fluoroethylcholine	^{18}F	Liver, kidneys, salivary glands, urinary bladder, bone marrow, spleen	In house production/ cyclotron
	Acetate	^{11}C	Gastrointestinal tract, prostate, bone marrow, kidneys, liver, spleen, pancreas	In house production/ cyclotron
Angiogenesis/ integrin binding	Galacto-RGD	^{18}F	Bladder, kidneys, spleen, liver	In house production/ cyclotron
	AH111585	^{18}F	Bladder, liver, intestine, kidneys	In house production/ cyclotron
SSTR Binding	DOTATOC	^{68}Ga	Pituitary and adrenal glands, pancreas, spleen, urinary bladder, liver, thyroid	In house production/ cyclotron
	DOTATATE	^{68}Ga	Spleen, urinary bladder, liver	In house production/ cyclotron

Using PET with the radiopharmaceuticals presented on Table 3.

Table 3, we can image for tumour staging, we can predict tumour response to therapy, we can detect recurrence, and finally, evaluate modifications in organ function after treatment (Rohren *et al.*, 2004).

In cancer research, the most commonly used radiopharmaceutical is [^{18}F]-FDG. The FDG-PET combined with a structural imaging given by Computed Tomography (CT), provides the necessary information for a cancer stage definition.

Another radiopharmaceutical that recently started to be used is [^{18}F]-FLT. In our studies we used just this one.

In the following subsections we are going to write about both radiopharmaceuticals.

2.3.1 [^{18}F] – FDG

The [^{18}F]-FDG is an analogue of the glucose molecule. Fluorine is incorporated into deoxyglucose yielding FDG and this compound is very successfully used in oncology (see Figure 7) (Guy & Ffytche, 2005).

It is intravenously administrated and it will initially follow the same pathway as glucose. To ensure their fast growth, the tumour cells are avid for glucose and usually exhibit an increase in the membrane glucose receptors. The [^{18}F]-FDG will enter the cell using the glucose transporters and inside the cell, it will be phosphorylated by Hexokinase II (HK II). After this conversion from [^{18}F]-FDG to [^{18}F]-FDG-6-phosphate, it is not metabolized any further, so it will be trapped in the cell (Saha, 2010).

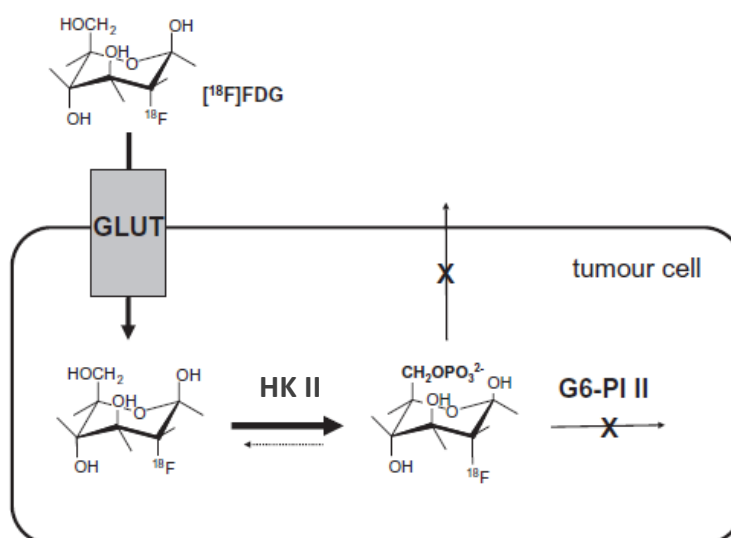


Figure 7 – [^{18}F]-FDG uptake into tumour cells (Haubner, 2010).

The major disadvantage of this radiopharmaceutical is that it images all the pathways that involve glucose. With this radiopharmaceutical we can image benign tumours, acute or chronic inflammatory tissue, normal tissue with a high physiological or metabolic activity (as bladder, kidney, heart, brain, muscle...) which yields a high background that will hamper the visualization and diagnosis. In other words, [^{18}F]-FDG is relatively unspecific. Nowadays it is the gold standard in oncology although, FLT has started to be used recently as a marker for proliferation, rather than a marker for more energetically and glycolytically active cells (Gray *et al.*, 2010).

2.3.2 [^{18}F] – FLT

Thymidine is the only nucleoside that is incorporated into DNA and not into RNA, becoming for that reason a good candidate to measure cell proliferation (Backes *et al.*, 2009).

The [^{18}F]FLT (Fluorothymidine) is a radiopharmaceutical analogue to thymidine where the hydroxyl function in position 3' is replaced by ^{18}F . It enters the cell using mainly nucleoside transporters (NT), but it can also enter by passive diffusion. After its entrance in the cell, it is phosphorylated by thymidine kinase 1 (TK1) to [^{18}F]FLT-monophosphate, which is further phosphorylated by nucleoside triphosphate kinase (TPK) and nucleoside diphosphate kinase (DPK) to [^{18}F]FLT-diphosphate and [^{18}F]FLT-triphosphate but the incorporation in DNA is restricted to minor amounts, due to the missing 3'-hydroxyl function (see Figure 8). Due to the negative charge of the phosphate group, it is not going to cross biological membranes getting accumulated inside the cell. This is the basis of [^{18}F]FLT – PET imaging (Backes *et al.*, 2009; Barwick *et al.*, 2009; Buck *et al.*, 2009; Haubner, 2010).

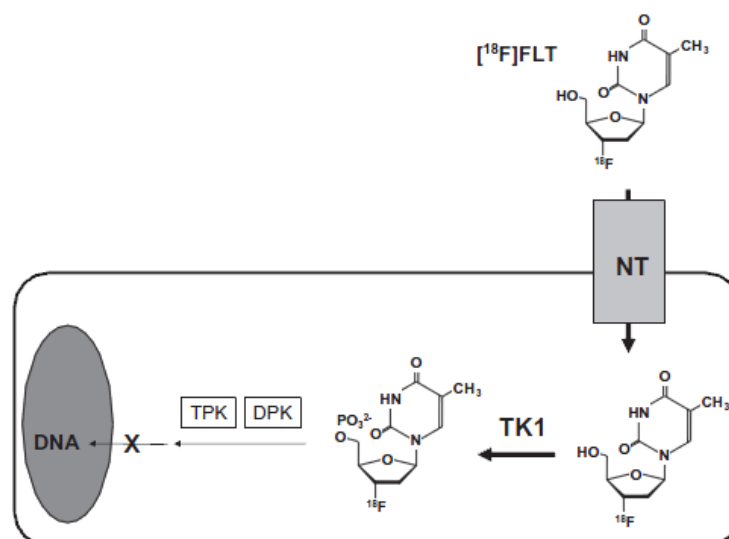


Figure 8 - [¹⁸F]FLT uptake into tumour cells(Haubner, 2010).

This radiopharmaceutical is nowadays used as a promising tracer for imaging cell proliferation. It goes mostly to hepatic metabolism as well as rapid proliferating tumours and some other tissues as the heart and vertebra (Gray *et al.*, 2010). And because the changes in proliferation in a tumour cell happen before changes in glucose metabolism or changes in the size of the lesion, it is described as a successful tracer in monitoring response to oncological therapy for some tumour types (e.g. pancreatic, colorectal, liver cancers) (Quon *et al.*, 2008; Contractor *et al.*, 2012).

2.4 Kinetic Spatial Filter – Based on (Gray *et al.*, 2010)

In 2010, a publication from Imperial College titled ‘Kinetic filtering of [¹⁸F]Fluorothymidine in positron emission tomography studies’, brought a new nonlinear kinetic filtering technique based on the kinetic properties of the FLT that are different depending on the tissues, but similar in all the patients (Gray *et al.*, 2010).

The basic idea behind this filter, was the isolation of cancerous tissue from healthy organs, developed and validated using scan data from 29 patients with locally advanced metastatic breast cancer, using [¹⁸F]FLT-PET.

The dose-normalized average time versus radioactivity curves (TACs) for different organs (lungs, liver, normal breast, vertebra), background (for voxels located outside the

body but inside the Field of View (FOV)) and tumour, were generated using scans from the 29 patients in this study. The resulting curves are presented in Figure 9. Each image voxel was then classified according to the tissue type it was more likely to represent, taking into account all the time frames for this voxel and comparing to the TACs obtained as a template. This comparison was made using Mahalanobis distance, given by Equation 2:

$$D_M = \sqrt{\sum_{t=1}^N \left(\frac{p_t - \mu_t}{\sigma_t} \right)^2} \quad \text{Equation 2}$$

where p_t is the activity of the voxel, μ_t and σ_t are the mean and standard deviation activities, respectively, of the class it is being compared against. N is the number of discrete time frames of the TACs.

In the end, after applying the filter, the idea was to reproduce an image only with tumour and vertebra tissues. Ideally they wanted to have just tumour, but because the signal in both vertebra/bones and tumour tissue is caused by rapid proliferation, the potential to remove the vertebra signal with a kinetic filtering approach was not considered further.

The technique belongs to class of supervised approaches and uses predefined kinetic classes recently introduced in PET for neurological applications (Turkheimer *et al.*, 2007).

In the Figure 9 we have the kinetic curves used for the classification algorithm in humans.

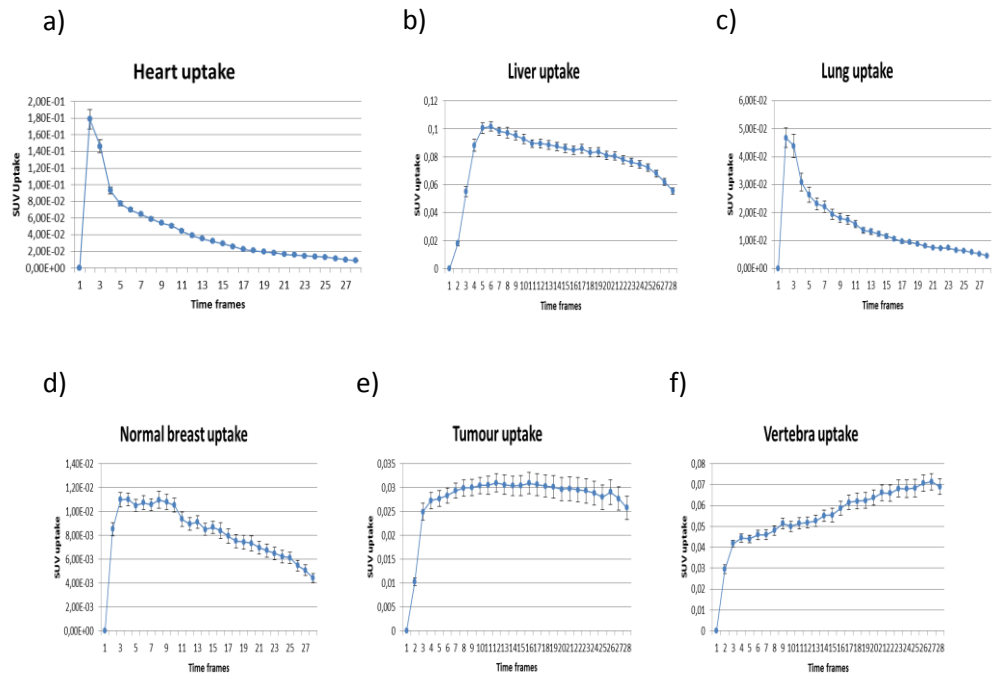


Figure 9 – Kinetic curves used for the classification algorithm in patients. These curves are derived from de templates used in the KSF. The error bars represent the standard error of the mean (SEM) where we have the FLT uptake for heart (a), liver (b), lung (c), normal breast (d), tumour (e) and vertebra (f).

They claim: ‘We have developed a novel kinetic filter that employs the time information in PET scans of FLT to remove high physiological background uptake signal and can be used to measure accurately changes in proliferation in liver tumours. The method was found to be successful in removing signal from the liver, heart, lungs and normal breast, whilst retaining that from tumour and vertebra tissue.’ Thus this technique appears to be very helpful for localization of the tumour lesion, reducing the background signal.

In discussion, they reveal that in one of the patients that had liver metastases, the filter was able to distinguish the tumour from the liver, what led to the study ‘Imaging of cellular proliferation in liver metastasis by [¹⁸F]Fluorothymidine PET: effect of therapy’ also from Imperial College (Contractor *et al.*, 2012), where they could distinguish between liver and metastases using what they now call kinetic spatial filter (KSF).

In mice, we can apply the same technique, but we need to take into account that the uptake curves are different in some tissues, as we can see in Figure 10

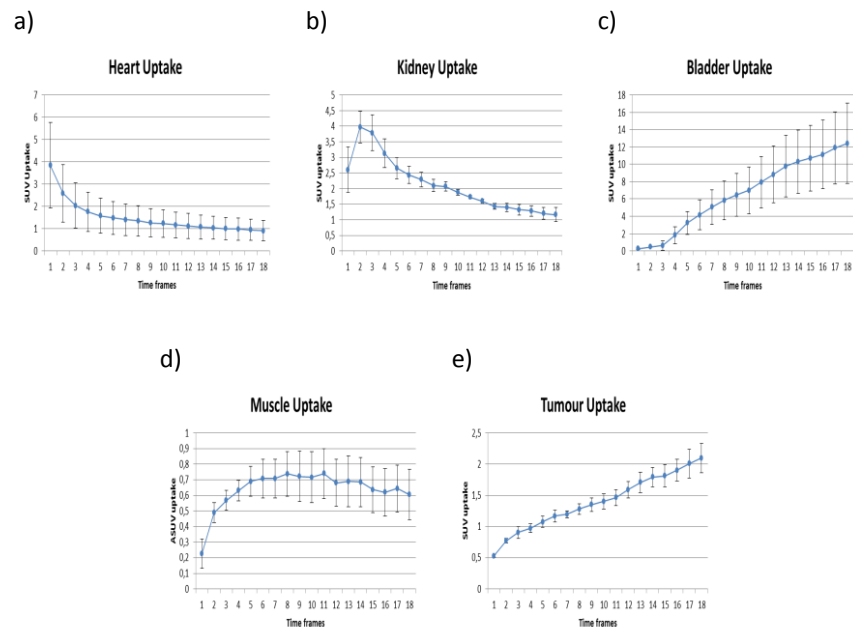


Figure 10 – Kinetic curves used for the classification algorithm in mice. The error bars represent the standard error of the mean (SEM) where we have the FLT uptake for heart (a), kidney (b), bladder (c), muscle (d) and tumour (e).

2.5 Self-Organizing Maps

To understand Self-Organizing Maps, we need to know its original basis. It is one of the most well-known Artificial Neural Networks (ANN). This ANN started in mid-1980s and was a result of the knowledge on how the brain works and the necessity to develop computer models that would be able to deal with complex tasks, as close as possible to brain ability to understand, learn, classify and cluster complex external information. Therefore, the basic idea was to mimic as perfectly as possible the brain cells and processes (Ritter *et al.*, 1992). Aiming to mimic the brain, ANN uses processing units (neurons) with connections between them to store the information needed to do a specific task (Cattinelli *et al.*, 2012). Results from (Tu, 1996) suggest that we can obtain better results using ANN than traditional statistical methods in medical applications, particularly when the data is complex and non-linear (Valkonen *et al.*, 2002).

We can say that ANN are models to learn from the environment. There are two types of learning processes inside the ANN: Supervised Learning and Unsupervised Learning.

Supervised Learning presupposes that the correct answer is already known, serving for training or simply for comparing the input data with the training, sometimes being the output just right or wrong (reinforcement learning)(Guthikonda, 2005).

The Unsupervised Learning implies that the correct answer is not known/ not given to the network. Therefore the aim of the network is to find by itself patterns in the input data. The advantage from this method is the independence from human interaction (Guthikonda, 2005).

The SOM is an Unsupervised Learning method. As unsupervised, it does not need any initial knowledge of the data and from this we just expect to collect some new information about the data. SOM is a non-hierarchical network with only lateral interconnections between computing units, which means that all neurons are connected to their nearest neighbours with respect to the structure of the network. All neurons have the same weighting or importance (Kohonen *et al.*, 1996a).

This algorithm was created by Kohonen *et al* as a model for visualization and interpretation of high-dimensional datasets (Kohonen & Oja, 1987). Also known as Kohonen's Neural Network, it transforms complex nonlinear statistical relationships in n-dimensional data into a low-dimensional grid map preserving the topological relations between the data(Chen *et al.*, 1999; Lobo *et al.*, 2007).

The package with this software was released in 1992 and since then it has had regular updates to improve its performance (Kohonen *et al.*, 1996a).

This software is very useful for visualization, dimension reduction, clustering, classification, sampling, vector quantization and data mining (Lobo *et al.*, 2004).

The advantages of the SOM are the capacity to cluster noisy and even incomplete data in an interpretable manner (Covell *et al.*, 2003; Hsu *et al.*, 2003). The ability to explore large data sets is another of the main advantages of this algorithm(Valkonen *et al.*, 2002). The fact that no a priori hypothesis needs to be provided by the user leads to an unbiased result, another of SOM's advantages (Cattinelli *et al.*, 2012). SOM is reported to be a really good technique able to achieve better dimensionality reduction and data visualization than other methods (Lee & Verleysen, 2007). After learning to interpret the final maps, they get easily-readable, what makes SOM a good method to use in data analysis (Cattinelli *et al.*, 2012). As described in Hsu *et al.*, SOM data has a good compression property, where nodes of SOM serve as prototypes , or mean values, for a

number of analogous input data, so that complete clustering can be achieved in a reasonable time (Hsu *et al.*, 2003).

As disadvantages we can name the necessity to define the number of clusters that we are going to visualize, or at least, the maximum number and the fact that it is not able to grow by itself, as we can see in GSOM(Hsu *et al.*, 2003).

2.5.1 The basic process

The Self-Organizing Map process is divided into three steps:

Competition

1. Initialization of weights for each node.
2. Randomly a vector is chosen from the set of training data and presented to the network (input).
3. All the nodes from the network are examined to determine which node has weights closest to the input vector, which means, the ones with minimal distance. In this step, the winning node is known as Best Matching Unit (BMU). The Euclidian distance is used to measure the closest input vector to a determined node:

$$\|x - m_c\| = \min_i \{\|x - m_i\|\} \quad \text{Equation 3}$$

where $\|\cdot\|$ is the distance measure, in this specific analysis, Euclidian distance.

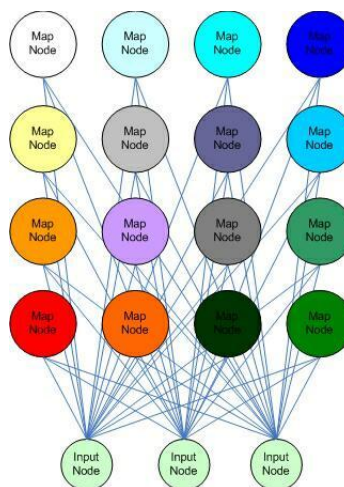


Figure 11 – Competition step of SOM. Any given input node is compared to the weight vector of each map node (or neuron) and the closest map node is the winner.

Cooperation

4. The radius of the neighborhood of the BMU is calculated. This value starts large. Typically it is set to be the radius of the network, diminishing at each new iteration.

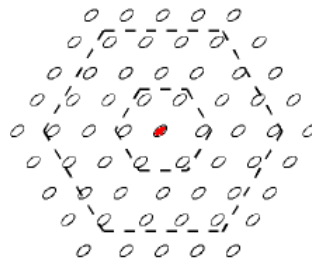


Figure 12 – Cooperation step of SOM. Here the diameter begins in the outer dashed line and is going to shrink to the inner one with the iterations(Kohonen *et al.*, 1996a).

Adaptation

5. Any nodes found within the radius of the BMU, calculated in step 4 are adjusted to make them more similar to the input vector. The closer a node is to the BMU the more its weights are transformed.

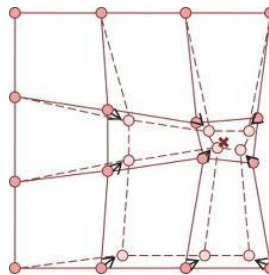


Figure 13 – Adaptation step of the SOM (Kohonen *et al.*, 1996a). The winner map node (neuron), represented here by spheres, and its neighbours are going to adapt to make their weight vectors identical to the input nodes.

6. Repeat from step 2. for N iterations.

It is important to note that the main idea of SOM after all these steps is not to cluster into a minimal amount, but to group the data into reduced and representative clusters (Valkonen *et al.*, 2002).

2.5.2 Applications of SOM

Self-Organizing Maps can be applied in a vast number of different areas; here I set down some of the ones found:

1. Colour reduction (Rasti *et al.*, 2007)
2. Computer Vision (Kohonen & Oja, 1987)
3. Classification of animals (Kohonen *et al.*, 1996b)
4. Image analysis(Rasti *et al.*, 2007)
5. Phoneme recognition(Arous & Ellouze, 2002)
6. Analysis of patients gene expression(Hsu *et al.*, 2003)
7. Analyse breast sonography to diagnose breast cancer (Chen *et al.*, 1999)
8. Modelling urban phenomena(Lobo *et al.*, 2007)
9. Tool for geodemographical data analysis(Lobo *et al.*, 2004)
10. Monitor the performances of dialysis clinic chains (Cattinelli *et al.*, 2012)

Chapter 3. Methods

This project was developed using data from mice and patients. The mice data was collected in Newcastle University (Northern Institute for Cancer Research) in the PET scanner provided there, and the mice treatment and care was also developed there.

The patient data was collected at Newcastle University PET Centre. It was acquired in a PET/CT scan and provided to me by the research team involved in the project.

The data used for KSF analysis and for SOM analysis is the same, so I will start with a description on common features of the acquired data and I will then specify what I have done differently in each analysis.

All the analyses were performed in a computer running Ubuntu 11.04, using Matlab 7.10 (Mathworks).

3.1 Data acquisition

3.1.1 Pre-clinical (mouse studies)

All animal experiments were conducted in compliance with the Animals (Scientific Procedures) Act 1986 and were approved by the local ethical review committee.

All imaging was carried out on a Philips Mosaic HP small animal PET scanner (Philips). The scanner has a gantry designed with an 18 cm port diameter, a transverse FOV of 12.8 cm and an axial extent of 11.9 cm. The scanner operates exclusively in 3D mode. The coincidence timing window is 12 ns and the standard energy window lies between 410 and 665 KeV.

Animals were selected for scanning when their tumour was approximately 5 x 5 mm (volume: $\sim 62.5 \text{ mm}^3$) in diameter, the three tumours closest to the median diameter are chosen for scanning. They were anaesthetised with a mixture of ketamine/medetomidine (Ketamine 50-75 mg/kg IP and medetomidine 0.5-1.0 mg/kg IP in a volume of 0.1 ml/10 g).

The animals then had their lateral tail vein cannulated and placed on the PET bed three at a time in the prone position (See Figure 14). The [^{18}F]FLT (produced in PETNET,

Nottingham) is injected (~10 MBq) at the start of the scan and a dynamic data set is collected over a period of 1 hour for the HCT116 study with FLT or 2 hours for FLT with all other tumour cell lines.



Figure 14 – Image of the three mice placed on PET scanner.

After the acquisition, a reconstruction for each mouse was conducted using 3D-RAMLA algorithm with no attenuation correction. It means that we need to do three reconstructions per scan, adjusting for mouse weight and radiation dose. In the end, for each mouse we get a Digital Imaging and Communications in Medicine (DICOM) file with correct values just for one of the mice presented in the image.

We had a small problem because we identified that the PET scanner was giving us the data in counts, and we believed that it was already converted to SUV. In the end we solved this problem by converting each pixel of the DICOM image, into SUV, applying the following equation:

$$SUV\ Value = ((SV \times m) + b) \times f \quad \text{Equation 4}$$

Where SV means original pixel value, m means rescale slope, b means rescale intercept and f means SUV scale Factor, all these values available in DICOM file header.

After this conversion, the mice data was ready to be analysed with both software tools.

The mice were scanned and analysed differently depending on the study: HCT116 study was different from other tumour cell lines study, so I have separated them to enable the correct description of both pre-clinical works.

The rest of the process is dependent on the software applied, so this will be split from now on.

3.1.1.1 HCT116

This study consisted of 9 mice in the beginning, but one died one day after the first scan, therefore we remained with 8 mice. The drugs used were both inhibitors of kinases (MEK or PI3K) used to target signalling pathways in tumour cells.

Mice were scanned on day 0, before any treatment and on day 2, after two days from the beginning of the study. The control mice were scanned on both days but remained without any treatment during the study. The others were scanned on both days too, but two hours after the first scan, on day 0, they started to be treated with 1mg/Kg PD0325901 (MEK inhibitor) and 100mg/Kg GDG-0941 (PI3K inhibitor). That treatment was repeated for the next two days. On the last day, two hours after the last treatment, they were scanned (See Table 4).

Table 4 – Table with the design of mice treatment, in HCT116 tumours. (†) dead mouse.

Mice	HCT116	
	Day 0	Day 2
Mouse 1	Not treated	†
Mouse 2	Not treated	Not treated
Mouse 3	Not treated	Not treated
Mouse 4	Not treated	Treated
Mouse 5	Not treated	Treated
Mouse 6	Not treated	Treated
Mouse 7	Not treated	Not treated
Mouse 8	Not treated	Treated
Mouse 9	Not treated	Not treated

These mice went on an 1 hour dynamic [¹⁸F]FLT-PET scan, where the sequence of the time frames was (10 x 1 min, 6 x 5 min, 2 x 10 min).

To analyse this data we applied the KSF to find out if it was able to separate tumour from other tissues, using just data from controls. We also applied SOM to try to identify patterns in the image data, comparing controls against treated, aiming to identify any differences between pre and post treatment data.

3.1.1.2 Other Tumour Cell Lines

This group includes mice implanted with the following tumour cell lines: A2780, SJSA, SN40R2 and HT29.

The mice were scanned on day 0, before any treatment and on day 2, after two days. The control mice were scanned on both days but remained without any treatment during the study. The others were scanned on both days also, but 2 hours after the first scan, in day 0, they started to be treated with specific cancer drugs, depending on the tumour cell line. That treatment was repeated for the next two days. On the last day, two hours after the last treatment, they went scanned again.

As mentioned before, these mice went on a 2 hour dynamic [¹⁸F]FLT-PET scan, where the sequence of the time frames was (10 x 1 min, 6 x 5 min, 8 x 10 min).

These mice were analysed just with SOM, aiming to distinguish between different types of tumour.

3.1.2 Clinical (human study)

This data is from a study of our clinical group, entitled ‘[¹⁸F]FLT-PET for assessment of treatment response in Exocrine Carcinoma of the Pancreas’.

The target group in this study are patients with locally advanced or metastatic pancreatic carcinoma. The clinical group aim to find collaboration from 28 patients. By the end of this report, they just had data from one patient. This first patient has a pancreatic cancer with liver metastases and was attending the Northern Centre for Cancer Care, Newcastle upon Tyne, United Kingdom.

The patient experiments were conducted in compliance with the Ethics Committee (10/H0707/50, West London REC expiry – 31/07/2012).

All imaging was carried out on a Siemens Biograph – 40 PET/CT scanner. The scanner has an axial FOV of 21.6 cm. It has four rings of LSO (Lutetium Oxyorthosilicate) detectors giving longer FOV and higher sensitivity. The coincidence timing window is 4.5 ns.

This patient went in this PET/CT scanner to perform the baseline, before any treatment (day 0). Five days after she started her first chemotherapy cycle with Standard

Gemcitabine-based chemotherapy (IV). The second multimodality scan was performed on the fifteenth day after the first chemotherapy treatment (day 2).

Simultaneously with PET scan, the patient went on a CT scan on both days to help with the tumour localization. The data that we had access was properly anonymised to protect the patient identity.

These PET scans were one hour long list mode and the images acquired were reconstructed into 28 time frames with 168x168x74 voxels.

3.2 Kinetic Spatial Filter (KSF)

3.2.1 Pre-clinical (mouse)

The pre-clinical data used for KSF analysis was just the one from HCT116 cancer cell lines.

The analysis of this data aims to identify and separate tumour from other tissues.

For this study, we did not use any of the treated mice presented in Table 4. They were analysed afterwards with SOM. With KSF we only analysed all the data from day 0 (Mouse 2 to Mouse 9).

In this analysis we adapted the Software assigned by Imperial College to our data. We tried first with the kinetic curves already provided with the filter that is based on human kinetic curves. On a second approach, we developed some kinetic curves based on the data from the controls, doing the average of all controls and calculating the standard deviation to create the .struct file needed in this software, but the final results were not good.

We changed the filter to be able to analyse each mouse individually, applying a mask to look only to one mouse at a time.

The animal data has a different number of volumes compared to patient data, so we needed to change this number in the filter before applying it.

The biggest issue on this analysis was the fact that the software was prepared to receive the data in ANALYZE format. In a first approach, we converted data using ImageJ and Amide. But because we found that there were some scaling issues with the data, we

decided to change the filter to accept data straight from DICOM, avoiding problems either from scaling or related with the inversion of the data time frames.

3.2.2 Clinical (human)

For this analysis we had data in DICOM format. All the software was almost complete for human analysis, we just needed to do minor changes. We applied a Gaussian filter in the original data to smooth the images. Then we corrected the time points because the original software was prepared for data with more time frames than the ones we had.

To avoid again the problems with conversion from DICOM to ANALYZE using Amide or ImageJ, we changed the software to receive DICOM files.

3.3 Self-Organizing Maps

The SOM analysis was carried out in using the SOM Toolbox for Matlab freely available from the internet (CIS, Accessed in September 2011).

The way that the weight vectors were initialized was randomly, the type of algorithm used for training was batch and the neighbourhood function employed in weight updates was Gaussian.

We choose one of the visualizations presented in the SOM Toolbox that we found more relevant and easier to interpret, according to our data. The analysis is based on pixel values of the ROIs defined on different tissues.

To measure the quality of our SOM, it gives us a final quantization error and final topographic error. The first error states how well the prototypes represent the input data; the second one gives the degree of preservation of neighbourhood relations. The best result is achieved when both errors are minimized.

When applying this algorithm, we tried different approaches: we tried different types of analysis using Principal Component Analysis (PCA) with SOM, using different approaches to add or remove values from our data, we tried different maps. In the end we opted for applying just SOM with no previous application of PCA. We removed values randomly from the matrices to make sure that they had the same size and we opted for showing the final result using the U-Matrix map with respective labels.

Due to specificities of the application of this software in these different studies, from now on we are going to divide the information, depending on the data analysed.

3.3.1 Pre-Clinical (mouse)

In pre-clinical SOM analysis, the first steps in HCT116 analysis and the analysis of all the other tumour cell lines are the same.

The analysis was based on pixel values of the ROIs defined on tumour, bladder, muscle, heart and kidney of the mice. We chose these tissues because they are easy to define in mice [^{18}F]FLT-PET image data.

To apply the SOM to this data we converted the data from counts to SUVs, we scaled the initial data for the acquisition time frames and to smooth the image we applied a Gaussian filter (Gonzalez & Woods, 2008). After that, we created a mask to make sure that we were looking only for the mouse that we were interested in and after this, the image was ready to be analysed with SOM.

The next steps are different, depending on the study, so again, we are going to separate the studies in order to specify each ones procedures.

3.3.1.1 HCT116

The analysis of this data with SOM aims to find a pattern in the scan from day 0 against day 2 scan.

For this study, we did not use any of the controls. They were analysed just to make sure that the conditions before and after treatment were the same. We just picked the ones that went on treatment to compare the differences on day 0 against day 2. According to Table 4, we analysed mice 4, 5, 6 and 8.

We did one mouse at a time, with pre and post treated data from this same mouse, and we applied the SOM to that group. Next steps were repeated for mouse 4, 5, 6 and 8.

Initially we draw the ROIs described above, in day 0 and day 2 from the same mouse. After getting the values from these ROIs of all the above tissues we create a matrix with this data and we proceed with the SOM analysis to find the final maps.

From here we got four SOM maps.

3.3.1.2 Other Tumour Cell Lines

The analysis involving mice with tumours of A2780, SJSA, SN40R2 and HT29 and SOM aims to find if this algorithm is able to distinguish between different types of tumour and/or find if the ones with the same origin have some common pattern, different from the other tumour types. Thus for this study we just used the controls.

The analysis is again based on pixel values of the ROIs drawn around tumour, bladder, muscle, heart and kidney of the mice in a first approach. To help with the analysis of the resulting maps we got the kinetic curves that show the uptake of the radiopharmaceutical in each tumour cell line. Finally we developed a very similar mapping to the first one, using SOM, but drawing ROIs just around the tumours, to check if the algorithm could distinguish tumour alone, ignoring other tissues.

We draw the ROIs described above, in each mouse. After getting the values from these ROIs of all the above tissues of all the mice, we created a matrix with this data and we proceed with the SOM analysis.

3.3.2 Clinical (human)

The aim of this study was to find if there was any difference when analysing the patient data comparing between the first scan (before treatment) and the second scan (after treatment).

We choose the same visualization method as before.

This data came already in SUV, with no need of conversions. The analysis is based on pixel values of the ROIs defined on tumour, kidney, heart, liver and vertebra of the patient. We choose these tissues because they are easy to define in human [¹⁸F]FLT-PET data.

This data was scaled for the acquisition time frames and we applied a Gaussian filter to smooth the image. Then we draw the above ROIs in patient data on day 0 and on day 2 (corresponding to the second scan, which is fifteen days after starting treatment) images and saved the values resulting from here in a matrix, ready to be analysed with SOM.

Chapter 4. Results

In this chapter we are going to present the results from all the described studies.

We are going to start from the KFC applied to HCT116 and to the patient data.

Afterwards we are going to report the results from SOM applied to HCT116, to the comparison between different tumour cell lines and finally, to our patient data.

4.1 Kinetic Spatial Filter (KSF)

4.1.1 Pre-clinical (mouse)

In Figure 15 is present the final image that we get before applying the KSF. This image is from Mouse 2 Day 0. It is reconstructed taking into account weight and injected dose from the mouse in the middle. Thus, for the KSF analysis we applied a mask to be able to visualize just the mouse that we are interested in, the one with the right reconstruction in that image.

In Figure 16 we have the images resulting from the application of the KSF. The mice missing are the ones that did not achieve good results.

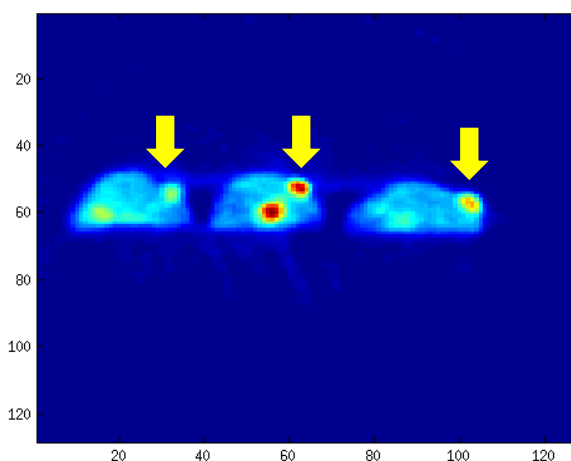


Figure 15 – Transverse slices of the $[^{18}\text{F}]\text{FLT}$ -PET scan of mice implanted with HCT116 data, before applying the KSF. This image is from Mouse 2, day 0. The yellow arrows are indicating the tumour region.

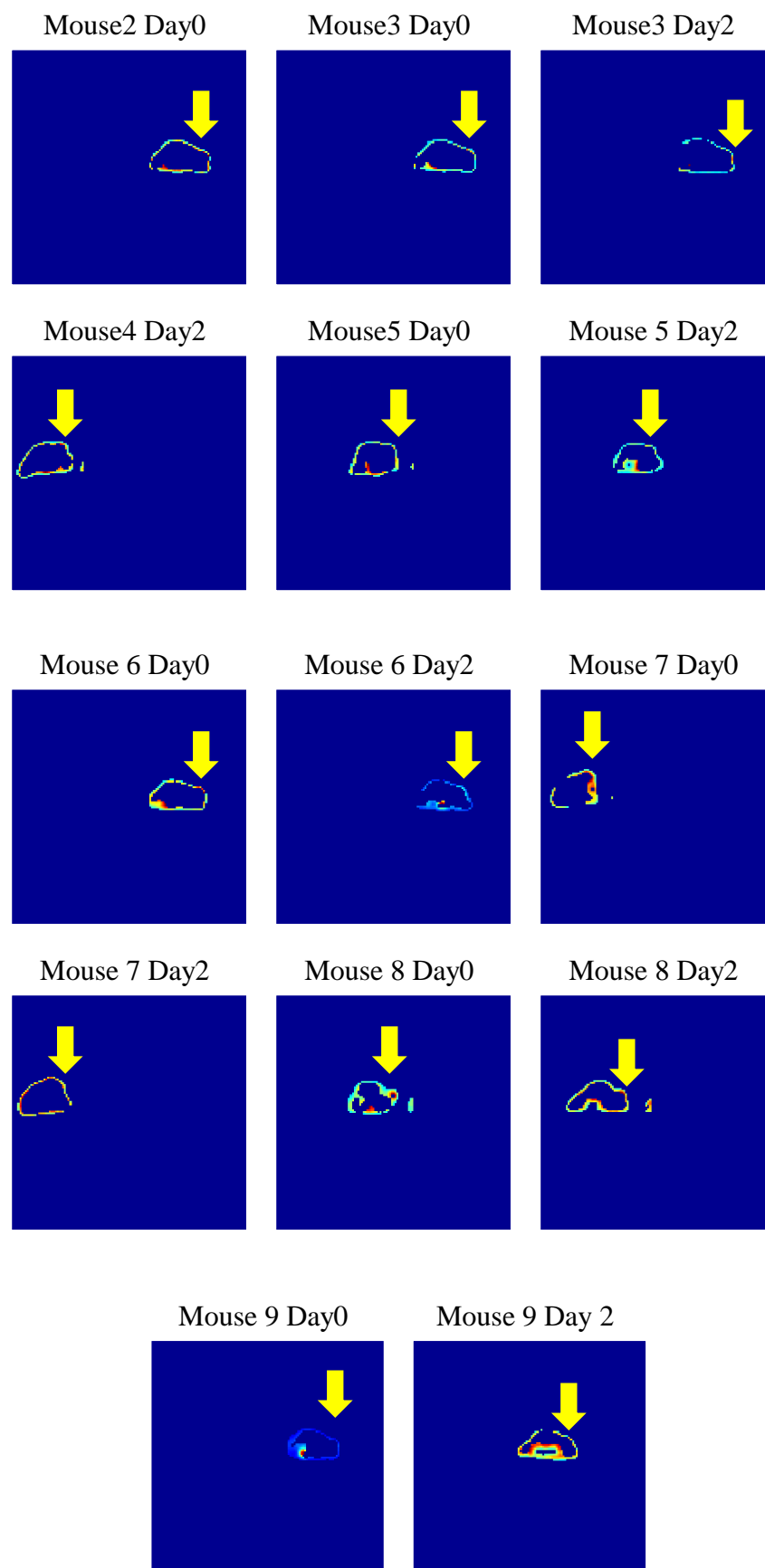


Figure 16 – Transverse slices of the $[^{18}\text{F}]\text{FLT}$ -PET scan of mice implanted with HCT116 data, after applying the KSF. The yellow arrows are indicating the tumour region.

In these results we show only the slices where we know that the tumour is present because the major aim for this study was to see if the filter was able to retain just the tumour signal. As we can see, this filter removed almost all the signal, keeping just the borders of the mice. The image of ‘Mouse 8 Day 0’ clearly reveals the tumour and a small part of the bladder. In other images, as Mouse 5 Day2 and Mouse 9 days 0 and 2 we can see that the signal of the bladder was kept too.

4.1.2 Clinical (human)

The results of applying KFC to data from the [^{18}F]FLT-PET scans of our patient, are presented in Figure 17 and Figure 18.

In Figure 17, in the left (A), we have a sequence of some representative slices where we can see the liver and vertebra (in the first three images) and in the following ones we can see the kidneys appearing on the right bottom of the image. Notice the colour bar presented on the top of the images, where 100% means the maximum uptake of the [^{18}F]FLT for that image, and 0% means no uptake. The z-axis is pointing in the direction of low to high z value of transversal slices of patient data.

In Figure 17 (B) we can see the same slices presented in Figure 17 (A) but after applying the filter. In all slices we can clearly distinguish the vertebra and the outline of the body. In the last image, all tissues were removed, except a small signal of the kidney and the tumour (note the yellow arrow).

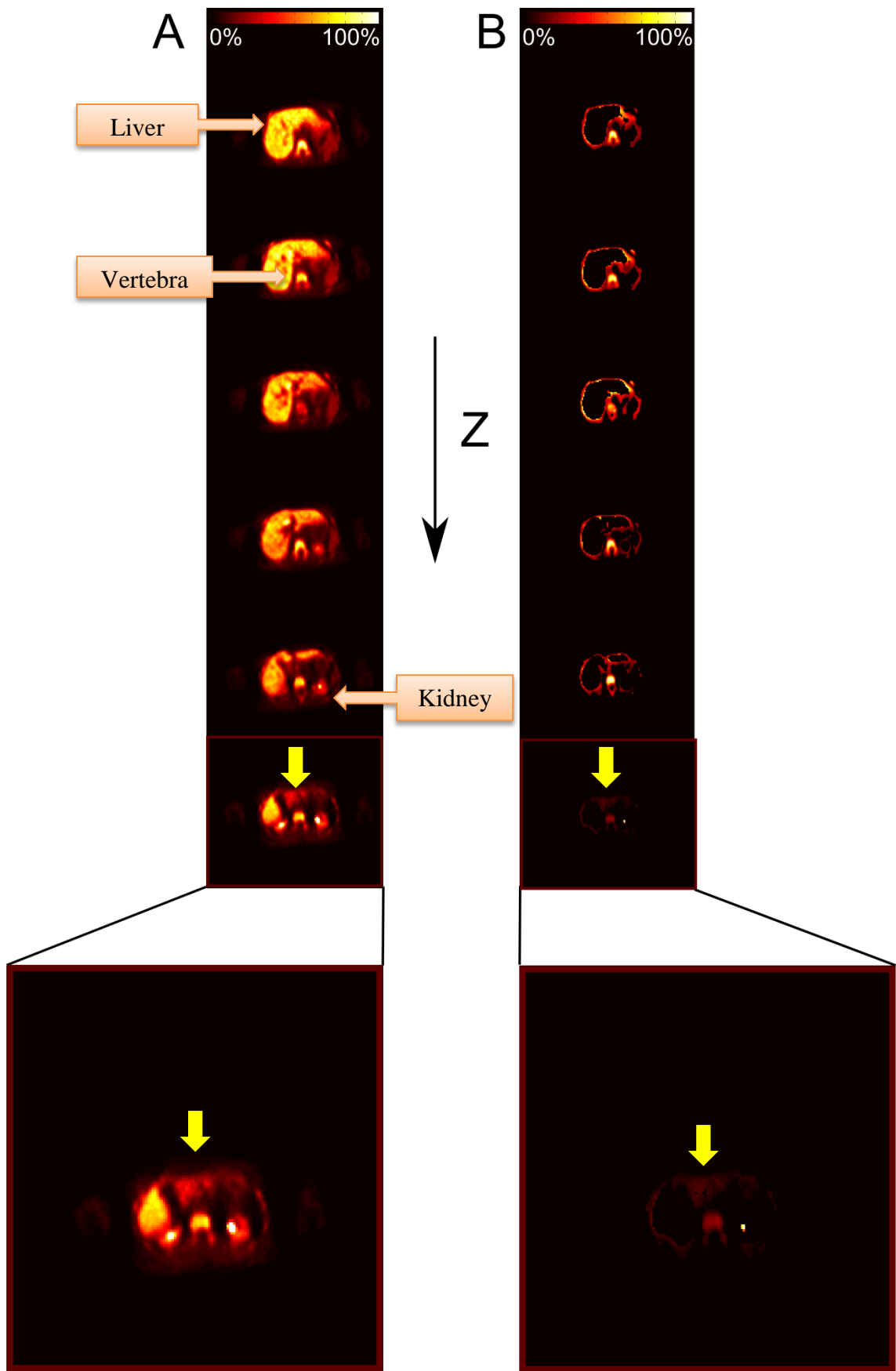


Figure 17 – Image from the first [^{18}F]FLT-PET patient scan normal (A) and filtered using KSF (B). Tumours are indicated with yellow arrows. Note the intensity bar and the z-axis. At the bottom we zoom the last slice ($z = 37$) to help visualization of KSF. The legend for Liver, Vertebra and Kidney are common for (A) and (B).

In Figure 18 we have the resulting filtered data of the second [^{18}F]FLT-PET scan of the patient, after treatment. Again we have the same colour bar and z axis with the same meaning described above. In Figure 18 (A) we have the images before applying the filter. The slices are the same as the ones in Figure 17. Again we can clearly see vertebra and liver in the first slices, but in the last two ones, we start to distinguish the tumour (yellow arrow) and kidneys.

Figure 18 (B) shows the filtered image. The vertebra is also present as it is all the outline of the patient body. In the last two images we can notice the tumour, pointed with yellow arrows.

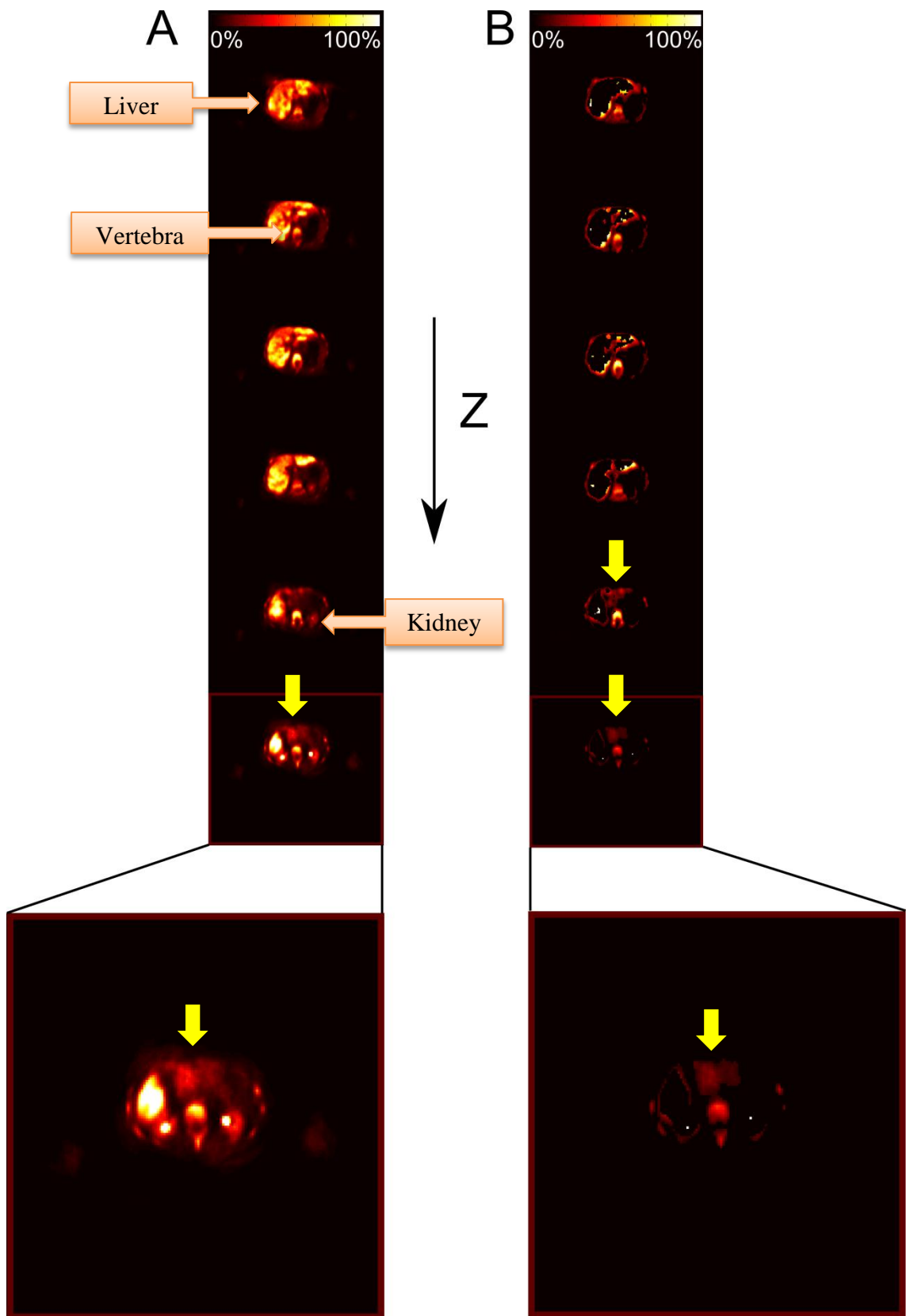


Figure 18 - Image of the second (post treatment) [^{18}F]FLT-PET scan normal (A) and filtered using KSF (B). Tumours are indicated with yellow arrows. Note the intensity bar and the z-axis. At the bottom we zoom the last slice ($z = 37$) to help visualization of KSF. The legend for Liver, Vertebra and Kidney are common for (A) and (B).

A radiologist that works closely with our team came to help on localization of the tumour. To support that task we had the [18F]FLT-PET data (Figure 17 and Figure 18) and the CT images (Figure 19). Nevertheless, to ensure that we were looking to the tumour we decided to draw ROIs around what we thought to be the tumour in pre and post treatment data and we could get the images presented in Figure 20. They have the curves of the tumour uptake and with that we could confirm that we were looking to the right place. In Figure 20(A) we have the tumour kinetic curve of the first scan (PET1) and in Figure 20(B) we have the respective curve for the second scan (PET2). These graphs of Figure 20 are in accordance with tumour kinetic curves presented in Figure 9.

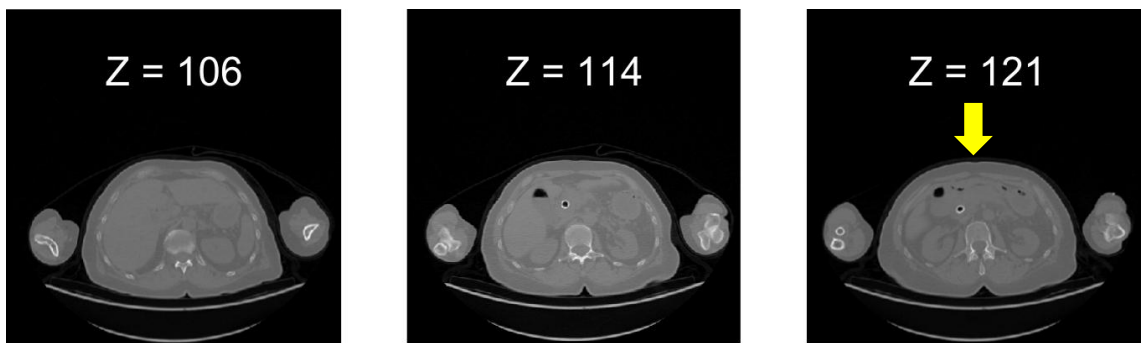


Figure 19 – CT images of the patient. The yellow arrow shows the tumour region. The black dot in the tumour region is the catheter left there on purpose to help to identify the tumour in the image.

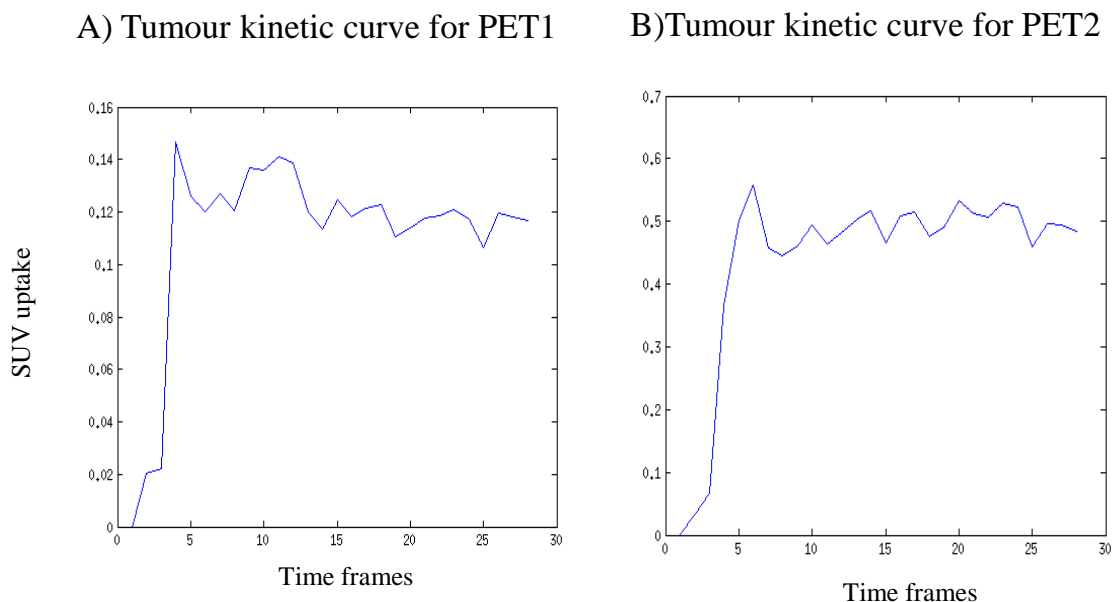


Figure 20 – [18F]FLT uptake curves in the region thought to be the tumour. The graphs have SUV uptake against time frames.

4.2 Self-Organizing Maps

In this chapter we present the resulting data from all the SOM studies developed on this project.

We begin showing the data obtained from the application of SOM to HCT116, aiming to find if SOM is able to distinguish between day 0 and day 2 in the same data.

Then we show the results of SOM applied to the other tumour cell lines (A2780, SJSA, SN40R2 and HT29) study, where we intend to find if this algorithm is able to distinguish between different types of cancer.

Finally we present the data of SOM applied to our patient data, aiming to see if there are differences, again, between the first and the second scans.

It should be noted that the colour bars of all the maps in this chapter have the same meaning. The hot colours with maximum value for red correspond to high distance between the vectors. In contrast, the cold (blue) colours correspond to low distance between values.

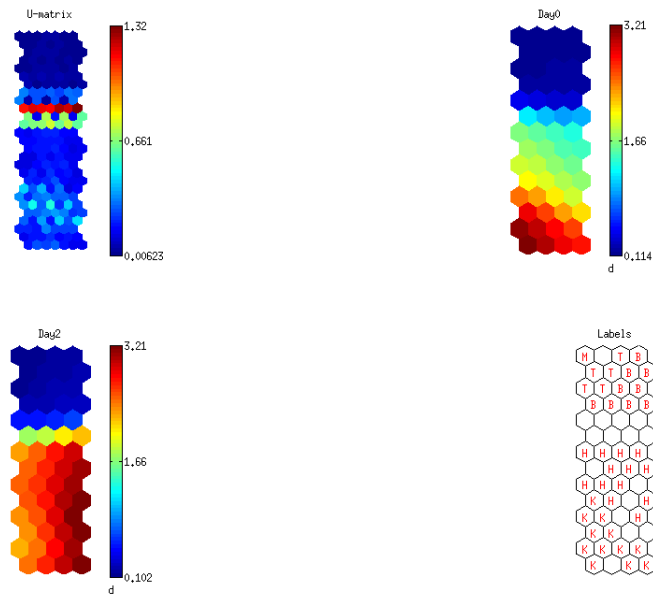
To measure the quality of the SOMs, we have the quantization error (QE) and topographic error (TE) given in the legend of all the figures.

4.2.1 Pre-clinical (mouse)

4.2.1.1 HCT116

In this subsection we have the analysis performed in mice implanted with HCT116 tumour cell lines. We need to remember here that Day 0 refers to the first scan of the mice and Day 2 refers to the scan that happened two days after.

In Figure 21 we have the visualization of the SOM applied to Mouse 4 with HCT116 tumour. As we can see here, the top left U-Matrix shows a well differentiated red line that in conjunction with the bottom right map, gives us the information that the algorithm clearly separates bladder, tumour and muscle from heart and kidney. Inside these major groups, we can still notice some distinction between different tissues. From the other maps we can notice that patterns from Day 0 and Day 2 are similar, but in Day 0 we have the hottest colours more to the bottom left of the map and on Day 2, more to the bottom right.



SOM 15-Aug-2012

Figure 21 – Visualization of the SOM applied to HCT116: data from Mouse 4. U-Matrix is on the top left, then we have component planes for Day0 and Day2 and on bottom right we have map unit labels. In the bottom right map the labels mean: H is for heart, T for Tumour, B for Bladder and K for Kidney. (QE=0.032 and TE=0.097)

In Figure 22 we have the visualization of the SOM applied to Mouse 5 with HCT116 tumour. The top left U-Matrix shows a less differentiated map. This map read in conjunction with the bottom right map, shows that the algorithm was able to separate bladder, tumour and muscle from heart and kidney. Again, inside these major groups, we can still notice some distinction between different tissues. Looking to the other two maps left, we can notice that patterns from Day 0 and Day 2 are similar, but in Day 0 we have the hottest colours more to the bottom left of the map and on Day 2, more to the bottom right.

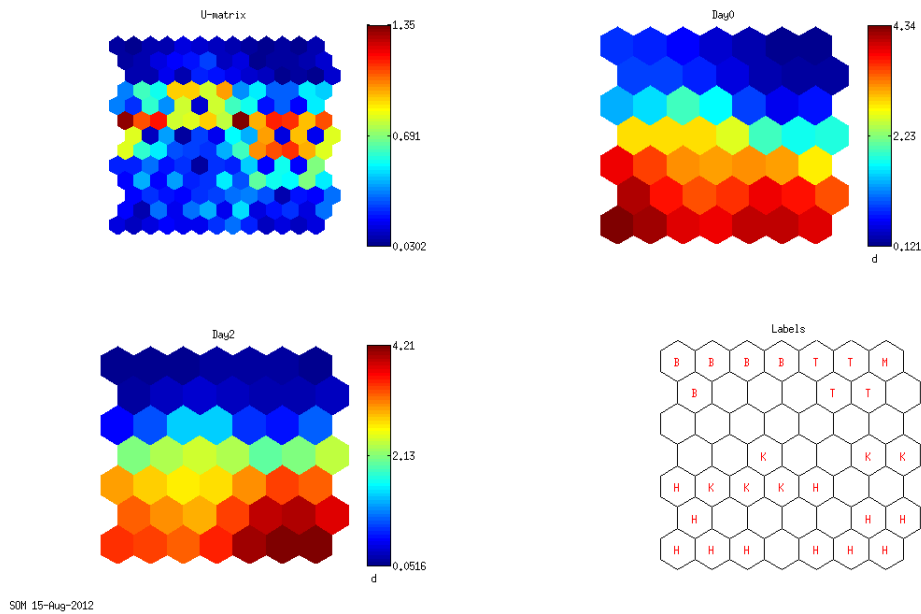
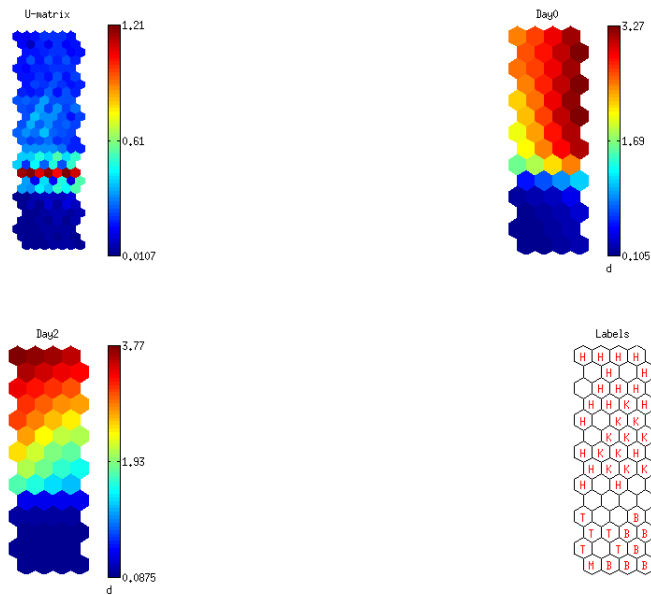


Figure 22 – Visualization of the SOM applied to HCT116: data from Mouse 5. U-Matrix is on the top left, then we have component planes for Day0 and Day2 and on bottom right we have map unit labels. In the bottom right map the labels mean: H is for heart, T for Tumour, B for Bladder and K for Kidney. (QE=0.090 and TE=0.028)

In Figure 23 we have the visualization of the SOM applied to Mouse 6. The top left U-Matrix shows a well differentiated map, more similar to Figure 21 than the former map. It shows that the algorithm was able to separate bladder, tumour and muscle from heart and kidney. In these two major groups the algorithm was able to separate between different tissues. Day 0 and Day 2 maps reveals an inversion of the map vectors distribution relative to the previous maps, but we have similar pattern between both days, just differing in the aspect that Day 0 has the red part localized on the top right of the map and Day 2 has the red part more to the top left.

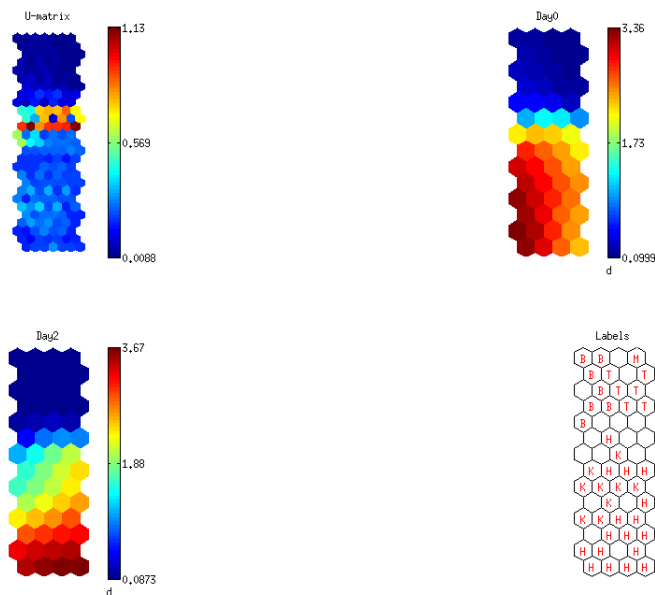


SOM 15-Aug-2012

Figure 23 – Visualization of the SOM applied to HCT116: data from Mouse 6. U-Matrix is on the top left, then we have component planes for Day0 and Day2 and on bottom right we have map unit labels. In the bottom right map the labels mean: H is for heart, T for Tumour, B for Bladder and K for Kidney. (QE=0.154 and TE=0.107)

Figure 24 shows the visualization of the SOM applied to Mouse 8. The top left U-Matrix shows some degree of differentiation similar to almost all the previous maps. Here the algorithm was able to separate bladder, tumour and muscle from heart and kidney, as it did for all of the above analysis. Inside these two major groups we have separation between different tissues.

The remaining maps reveal a pattern similar to Figure 21 and Figure 22. Again, Day 0 has its hottest colours on the bottom left and Day 2 has its hottest colours on bottom right.



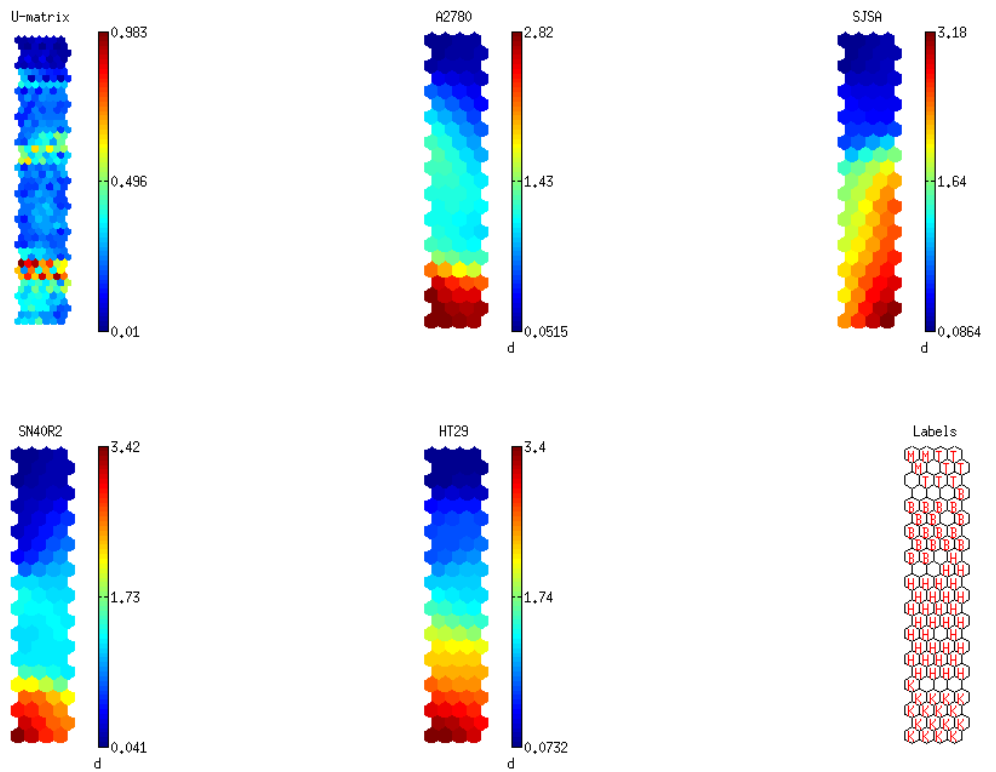
SOM 15-Aug-2012

Figure 24 – Visualization of the SOM applied to HCT116: data from Mouse 8. U-Matrix is on the top left, then we have component planes for Day0 and Day2 and on bottom right we have map unit labels. In the bottom right map the labels mean: H is for heart, T for Tumour, B for Bladder and K for Kidney. (QE=0.060 and TE=0.027).

4.2.1.2 Other Tumour Cell Lines

In this subsection we have the analysis performed in mice implanted with A2780 (ovarian carcinoma cancer cell line), SJSA (osteosarcoma cell line), SN40R2 (osteosarcoma cell line selected for drug resistance) and HT29 (colorectal adenocarcinoma cell line).

Figure 25 shows the visualization of SOM applied to all of the above tumour cell lines. Looking to the U-Matrix we can notice some segregation and completing this information with the one given in the bottom right map, we notice perfect dissociation between different tissues, from top to bottom, muscle, tumour, bladder, heart and kidney. Regarding the four maps left, they reveal a very similar pattern between them, SJSA being the most disparate from the others.



SOM 07-Sep-2012

Figure 25 – Visualization of the SOM applied to all tumour cell lines: A2780, SJSA, SN40R2 and HT29. U-Matrix is on the top left, then we have component planes for Day0 and Day2 and on bottom right we have map unit labels. In the bottom right map the labels mean: H is for heart, T for Tumour, B for Bladder and K for Kidney(QE=0.172 and TE=0.103).

In Figure 26, we have the kinetic curves from all of the above tumour cell lines. In general there are no obvious differences between FLT uptake in any of these tumours. Mouse 1 (M1) refers to the mouse on the left, Mouse 2 (M2) to the one in the middle and Mouse 3 (M3) to the one on the right.

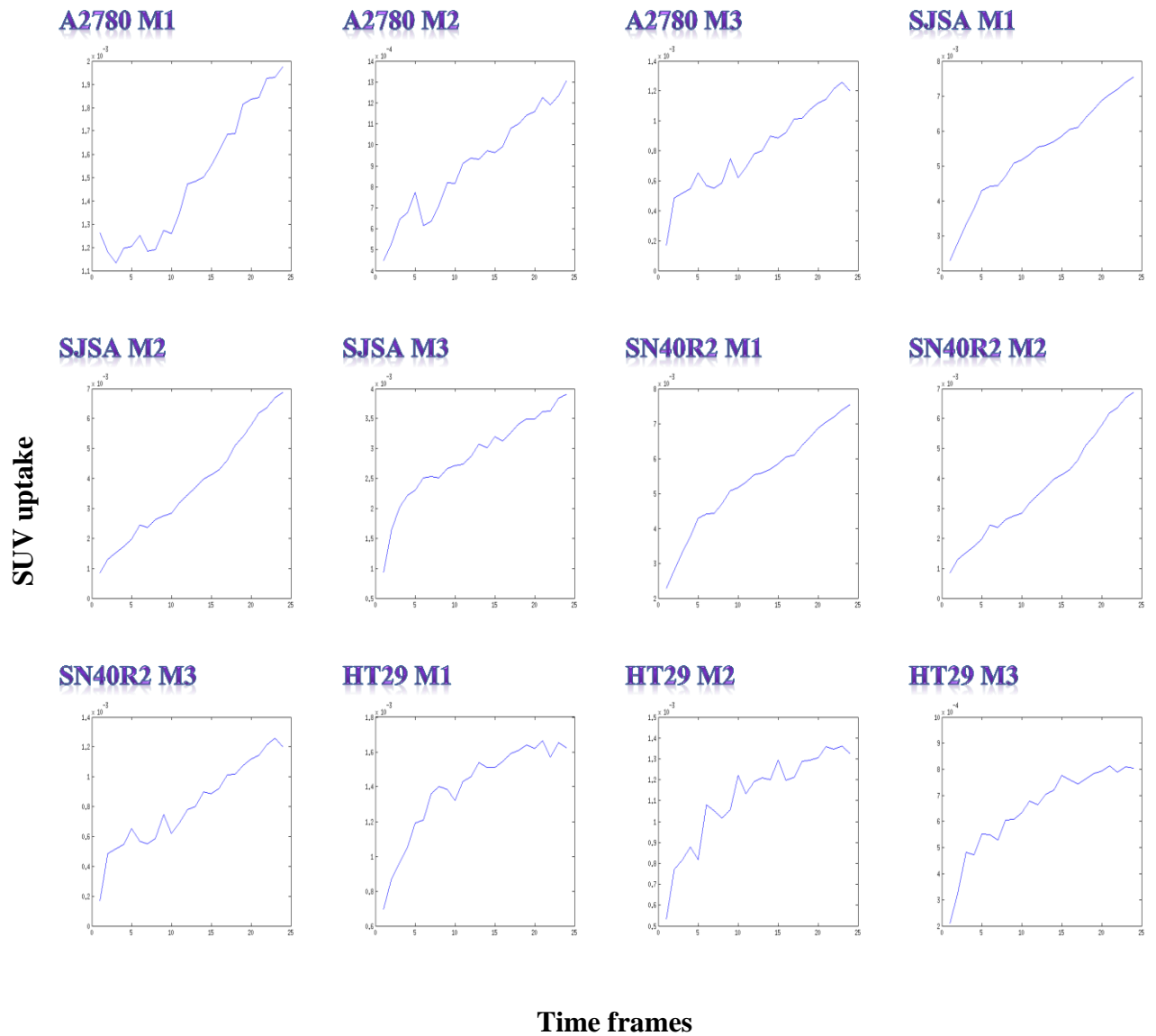


Figure 26 – Kinetic uptake curves of all the mice tumour cell lines: A2780, SJSA, SN40R2 and HT29. Curves were taken from mouse 1 (M1), mouse 2 (M2) and mouse 3 (M3) of all the tumour types.

As a final confirmation, we used SOM again to analyse just the tumour ROIs along the time frames. What we got was the result presented in Figure 27. We notice that between the time frames 22 to 24 there are no differences and in U-matrix we have a map a bit diffuse. Analysing it carefully, HT29 is grouped on the top left of the map, SJSA is grouped more to the bottom left of the map, SN40R2 more on the bottom right and A2780 is in the centre, partly mixed with the other tumour cell lines.

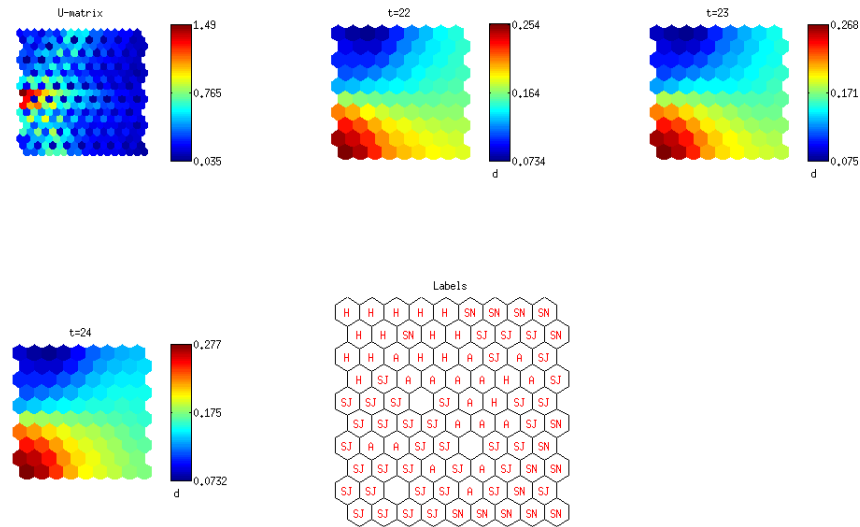


Figure 27 – Visualization of the SOM applied to all tumour cell lines: A2780, SJSA, SN40R2 and HT29. U-Matrix is on the top left, then we have component planes for t=22, t=23 and t=24 and on bottom right we have map unit labels. In the bottom right map the labels mean: H is HT29, SN for SN40R2, SJ for SJSA and A for A2780.

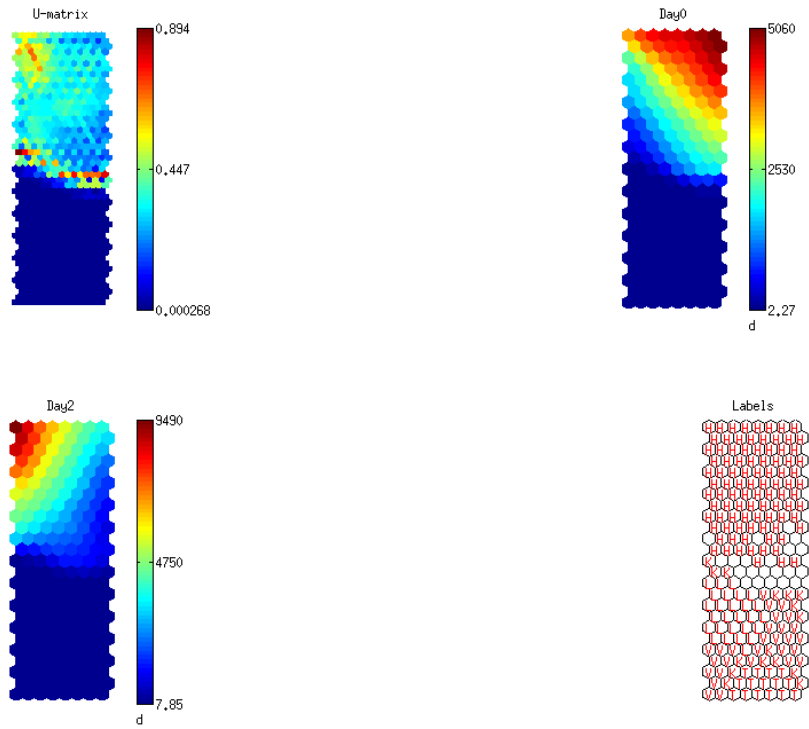
4.2.2 Clinical (human)

Here we had developed the analysis of patient data with SOM algorithm.

Presented in Figure 28 we have the resultant data from this study. Comparing Day 0 and Day 2 maps, we notice some pattern but we can distinguish both of the maps because on Day 0 we have the hot colours on the top right and on Day 2 we have them on top left. This characteristic is very similar to the previous HCT116 study.

The top left U-Matrix shows some degree of differentiation. Accomplishing this analysis with the labels map, we can notice that the heart is completely separated from other organs. Inside this group where all the other organs are together, we can find again some separation between tissues (tumour, kidney, liver and vertebra). Vertebra is the only tissue more mixed with the others.

The maps presented have a quantization error of 0.04 and a topographic error of 0.108.



SOM 31-Aug-2012

Figure 28 – Visualization of the SOM applied to patient data. We have Day 0 as data from first scan and Day 2 as data from the second scan. U-Matrix is on the top left, then we have component planes for Day0 and Day2 and on bottom right we have map unit labels. In the bottom right map the labels mean: H is for heart, T for Tumour, L for Liver, V for Vertebra and K for Kidney. (QE=0.040 and TE=0.108)

Chapter 5. Discussion

After the described work, we can discuss the application of KSF and SOM to mice and patient data.

The objectives of both in image analysis are different thus, to illustrate better the conclusions for each of the above applications, this chapter will have three subsections.

5.1 Analysis with KSF

When applying this filter to our mice data, we found that the results were not in accordance with the analysis in patients. After trying this filter with kinetic curves created by us, based on HCT116 data without success, we decided to apply patient curves that came with this software and we could achieve the results presented in Figure 16. In Figure 9 and Figure 10, we have access to kinetic curves from patients and mice and as we can see, there are differences within the same tissue. Heart uptake is similar in both humans and mice, but human tumour uptake is more similar to mice muscle uptake and human vertebra is more similar to mice tumour uptake. In Figure 10, curves for neither liver nor vertebra are shown because in mice images, these tissues are not as evident as they are in the human data. What happened when we filtered mice data using human kinetic curves was that bladder and some parts of the tumour were kept, because they have similar pattern to vertebra and tumour. Nevertheless, the results are not good and the best solution would be to acquire more data of mice implanted with HCT116 and create new kinetic curves. To draw the TACs, the study from Imperial College was using data from 29 different patients. We used data just from 9 mice which explains the size of our error bars compared to theirs and maybe our lack of success when applying them to our images.

Our findings regarding the application of this filter to patient data is consistent with previous reports (Kenny *et al.*, 2007; Gray *et al.*, 2010). As we can see in Figure 17 and Figure 18 this filter is working almost perfectly on removing signal from all tissues, leaving just the outline, vertebra and tumour, as expected. We can say that it is efficient when analysing data from a patient with pancreatic cancer with liver metastases. The only problem in final results of our data was that this filter kept the kidneys, as we can see in the bottom images of both filtered figures (Figure 17(B) and Figure 18(B)). Initially this filter was developed to detect cancerous tissues in cases of breast tumours

that are far above kidneys, so it did not take into account the kidney uptake. To improve its performance, I would suggest adding kidneys TACs to the software. Another TAC that could be added was the pancreas one, since we are applying this filter specifically to pancreas studies.

It is important to highlight the relevance of CT data that gave us some help for tumour localization (see Figure 20). If still in doubt we can always confirm if what is thought to be tumour tissue is really tumour, drawing a ROI around the area and performing an average of this ROI values across the volume. In the end we get graphs similar to the ones presented on Figure 20, and from this kinetic curves we can ensure what the type of tissue presented.

This KSF is a really good tool for data analysis. However, in heterogeneous tumours (e.g. necrotic cores) it will keep high proliferating tissues, removing the necrotic part from the tumour in the images. In that situation is not so easy to distinguish between tumour and other organs. It is important to note also that for bone or bladder tumours, this filter is not so efficient because proliferation is already high in these tissues. Maybe the next step could be using this principle to apply in PET scans with other radiopharmaceuticals (Gray *et al.*, 2010). In addition, it has other limitations such as not making any corrections for artefacts, not considering the hypothesis that different tissues can be mixed in the same voxel, having reduced sensitivity to small lesions and having reduced number of subjects included to create the template curves.

5.2 Analysis with SOM

The analysis of HCT116 with SOM presented in Figure 21 to Figure 24 shows us a clear pattern between the data from Day 0 against Day 2 for all the mice. In all the images we can see that SOM was able to distinguish data from Day 0 to Day 2 since they are revealing similar pattern when looking at all the figures, Day 0 and when we do the same for Day 2, we can see that the map is revealing a pattern from these group of days, in all mice data analysed.

Regarding the U-Matrix and respective label maps, there was always a separation between all tissues, but especially tumour, bladder and muscle were grouped on one side and kidney and heart on the other. If we compare these result with the graphs presented in Figure 10, we notice that the uptake curves of kidney and heart, are very

similar between them and the shapes of the curves from tumour, bladder and muscle have also a high degree of similarity.

From here we can conclude that this unsupervised learning method is able to distinguish between all the tissues and between Day 0 and Day 2. It is important to notice the small error values acquired (quantization (QE) and topographic (TE) errors).

Applying the SOM to different tumour cell lines to find out if it was able to make some distinction between them, was not as clear as expected, because A2780, SJSA and SN40R2 appear with the same pattern and just SJSA has a different one. Having in mind that A2780 is an ovarian cancer, SJSA and SN40R2 are osteosarcomas and HT29 is a colorectal type of tumour, we should expect some grouping of SJSA and SN40R2, differentiated from the others.

For this study we just analysed data from 4 scans, 12 mice. To take any conclusions it would be better to add more data to this study. Considering just the present results, we can conclude that SOM is partly able to distinguish between these cell lines. Looking to Figure 26 and Figure 27 we find that uptake curves are not very different between different tumours and as we can see in the last SOM analysis, when having ROIs just around tumours, SOM could not differentiate so well between different tumour cell lines. We can conclude that this task for the SOM is difficult since it is trying to distinguish FLT uptake patterns between different tumours with similar curves. In general, most tumours are expected to have high uptake compared to most normal tissues. After these analyses we can infer that in some respects it is maybe easier to use uptake information from other mouse tissues which takes into account toxicology between tumour and host.

In patient analysis we have a similar behaviour as in the analysis of HCT116. We can distinguish data from Day 0 and Day 2. This is helpful for future analysis because it can reflect the response to treatment. For the next patient, the same analysis can be developed and if the pattern of the maps are the same, we can say that the patient is responding to the drug, if the pattern between Day 0 and Day 2 are similar, we can conclude that this treatment is not having any, or at least no significant effect on the patient. With this information the treatment can be conducted in another direction.

In all the above SOM studies, the resulting errors are low. As explained in methods, we tried different types of analysis using PCA and SOM, using different approaches to

make sure that the size of each group of tissues had the same amount of values and the best results – results with smaller errors – were achieved using SOM, removing random values from the groups of tissues instead of using other approaches to add values to these groups and using the correct number of time frames from mice and patient data.

The choice of hexagonal lattice in SOM maps is optional, we could have chosen rectangular lattice, but because as there is evidence that the hexagonal is better, we kept this one (Kohonen, 1997; Hsu *et al.*, 2003)

5.3 General Discussion

Comparing the supervised classification technique used in this project (KSF) and the unsupervised one (SOM) we could suggest that KSF is more objective. KSF does not have the same uncertainties that SOM has. With SOM we should analyse more data aiming to confirm what each map pattern means so that, in future analysis, we would not need to do any ROIs delineations. It would automatically give us the answer that we would be able to interpret with no need of any basic knowledge of what was the type of tumour presented.

The advantage and curious thing from SOM is that it can reveal unknown relations between data. So it can find useful relations to apply and test with supervised learning. Nowadays there are several studies that suggest different approaches for tumour classification such as Growing Self-Organizing Maps (GSOMs) (Hsu *et al.*, 2003), algorithmic methods for multiclass tumour classifications based on gene expression (Covell *et al.*, 2003), application of vector machines to analyse carcinomas dataset (Su *et al.*, 2001).

I believe that the analysis of image texture and dynamic data all together, could yield good results within the field of oncological medical imaging. There is a study where benign and malignant tumours could be classified with SOM using interpixel texture (Chen *et al.*, 1999).

Chapter 6. Conclusion

The main goal of this thesis was to use SOM for analysis of non-invasive imaging data, aiming to achieve characterization of tumour physiological and biochemical properties. Collaboration between Imperial College and Newcastle University allowed us to use the KSF that we applied to pre-clinical and clinical data. Thus, we not only studied tumour characterization but also validated the KSF in our data and tested the importance of SOM in another type of classification, as visualizing if it was able to distinguish between pre and post treated data.

The R Software was introduced to me and was also supposed to be used for this analysis. We started with it, using Kohonen Package Toolbox for R. Although, the Matlab analysis was complex so we decided to focus in that analysis and leave R analysis as suggestion for future work. We know that R is a really good tool for statistics and visualization so in the future we can compare the results obtained with Matlab with the ones obtained with R (Ron Wehrens, 2007).

As future work it would be helpful to adapt KSF to other radiopharmaceuticals, as [^{18}F]FDG. There was already a study where this was tried (Janssen *et al.*, 2009), but since the KSF is giving good results, it could be adapted. Along with this adaptation, we could try to improve the KSF limitations described in the last chapter.

At the end of this work we are aware of the importance of cancer research, not only in biology and genetics, but in computer sciences applied to this field. Nowadays it is believed that early screening together with good algorithms for image analysis can be the key for a non-invasive and early stage detection of cancer, yielding a best survival chance for the patient.

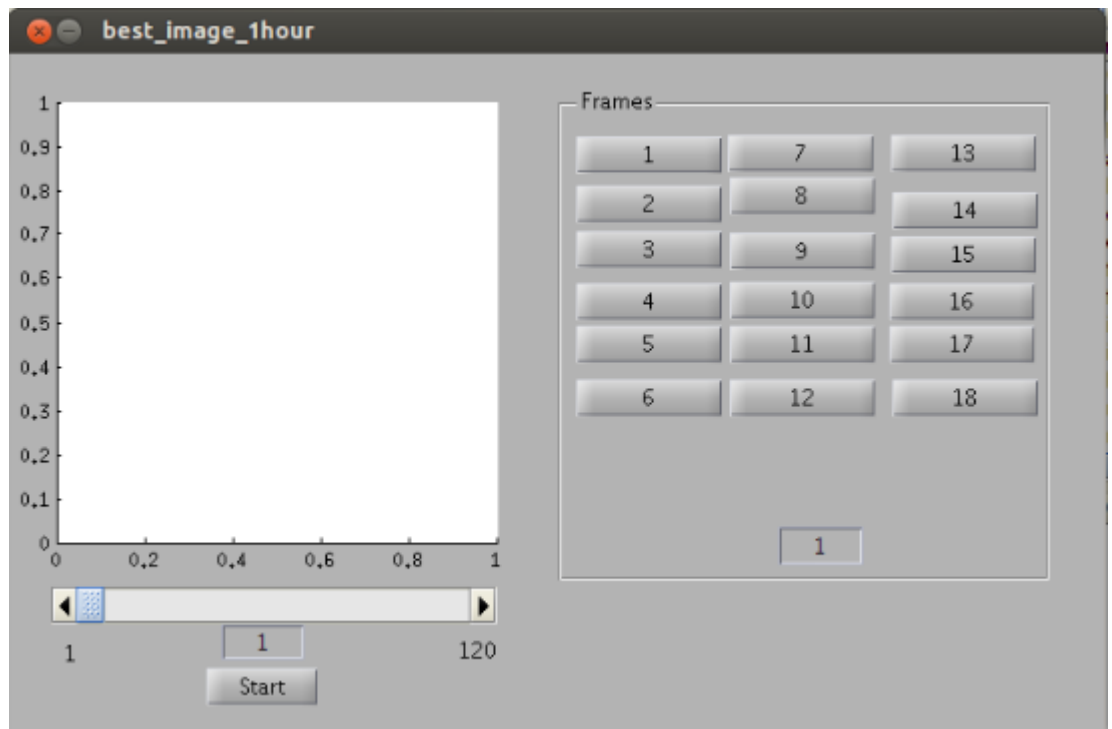
After working with pre-clinical and clinical data, I also found important to remember that all this software should be adapted to enable and improve pre-clinical image analysis. Before applying any treatment to patients, there is great investigation in mice, therefore it is important to provide the best tools to help with interpretation of these results.

Appendix

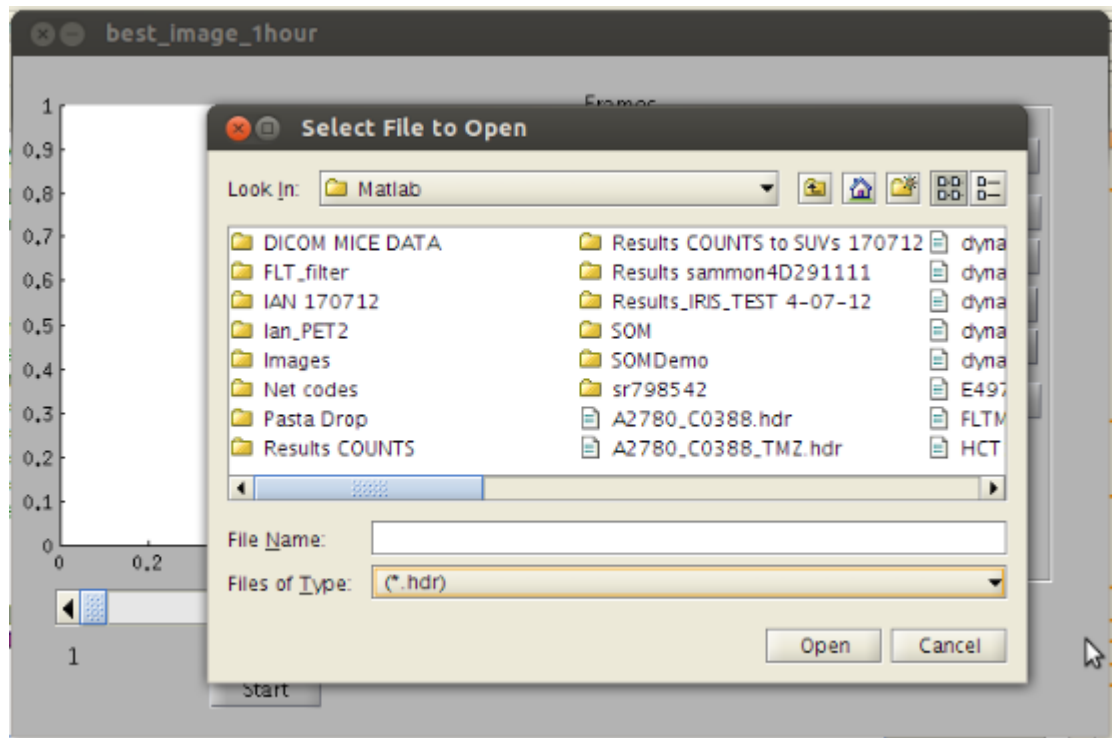
For the analysis of the data, I have created a Graphical User Interface (GUI).

The usage is very simple, handy and instinctive.

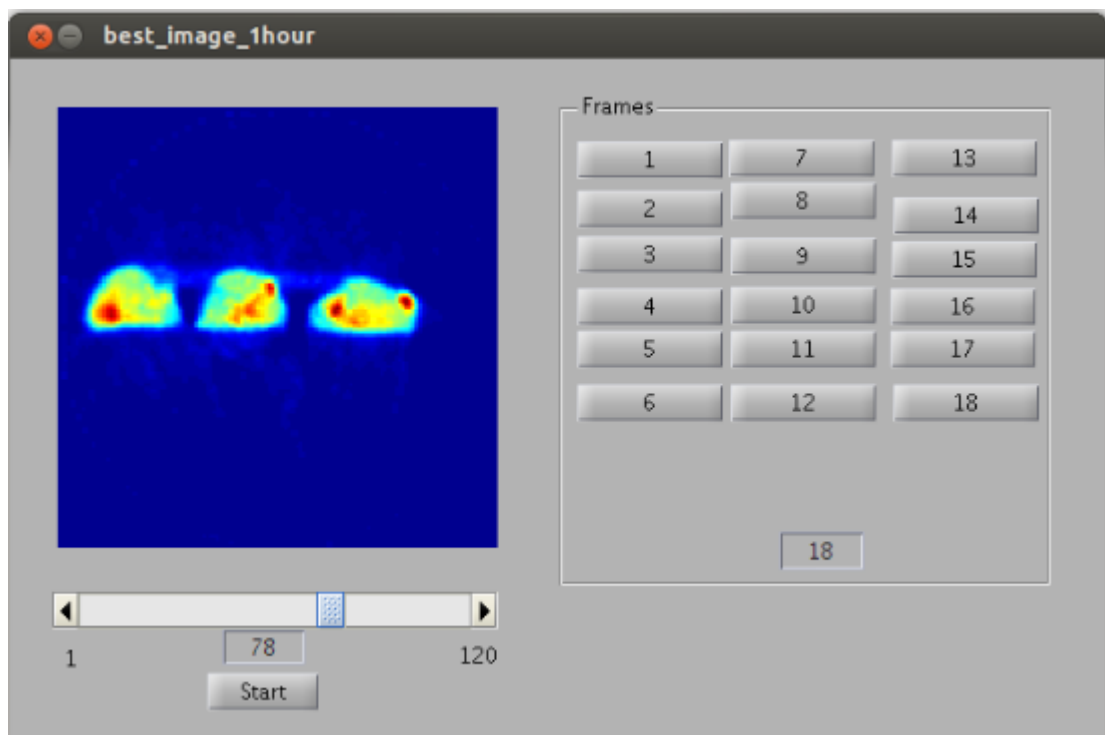
- 1) First we just need to call for the GUI on Matlab command line and we will get the following GUI.



- 2) Click on the Start button and select the file that we want to analyse with the following extensions: .mat or .hdr.



- 3) Then the image will automatically open and we just need to click on the time frame that we want to visualize and we can use the slider to select the z value that we want to visualize. In this example we have all the mice together, but with simple changes we can visualize just the one that we are interested in.



References

- Allisy-Roberts, P., Williams, J.R. & Farr, R.F. (2007) *Farr's physics for medical imaging*. Saunders, Edinburgh ; New York.
- Arous, N. & Ellouze, N. (2002) Cooperative supervised and unsupervised learning algorithm for phoneme recognition in continuous speech and speaker-independent context. *Neurocomputing*, **51**, 225-235.
- Backes, H., Ullrich, R., Neumaier, B., Kracht, L., Wienhard, K. & Jacobs, A.H. (2009) Noninvasive quantification of 18F-FLT human brain PET for the assessment of tumour proliferation in patients with high-grade glioma. *Eur J Nucl Med Mol Imaging*, **36**, 1960-1967.
- Barwick, T., Bencherif, B., Mountz, J.M. & Avril, N. (2009) Molecular PET and PET/CT imaging of tumour cell proliferation using F-18 fluoro-L-thymidine: a comprehensive evaluation. *Nucl Med Commun*, **30**, 908-917.
- Biersack, H.J. (2007) *Clinical nuclear medicine*. Springer, Berlin ; New York.
- Brownell, G.L. (October 15, 1999) A History of Positron Imaging. *Presentation prepared in celebration of the 50th year of services by the author to the Massachusetts General Hospital*, 1-11.
- Buck, A.K., Herrmann, K., Shen, C., Dechow, T., Schwaiger, M. & Wester, H.J. (2009) Molecular imaging of proliferation in vivo: positron emission tomography with [18F]fluorothymidine. *Methods*, **48**, 205-215.
- CancerResearch
<http://info.cancerresearchuk.org/cancerstats/incidence/commoncancers/#Twenty>
(accessed 21st August 2012).
- Cattinelli, I., Bolzoni, E., Barbieri, C., Mari, F., Martin-Guerrero, J.D., Soria-Olivas, E., Martinez-Martinez, J.M., Gomez-Sanchis, J., Amato, C., Stopper, A. & Gatti, E. (2012) Use of Self-Organizing Maps for Balanced Scorecard analysis to monitor the performance of dialysis clinic chains. *Health Care Manag Sci*, **15**, 79-90.
- Chen, D.-R., Chang, R.-F. & Huang, Y.-L. (1999) Breast cancer diagnosis using self-organizing map for sonography. *Ultrasound in Med & Biol*, **26**, 405-411.
- Cherry, S.R. (2006) *PET: Physics, Instrumentation, and Scanners*. Springer New York, New York.

CIS (Accessed in September 2011) Laboratory of Computer and Information Science,
<http://www.cis.hut.fi/somtoolbox/>

Contractor, K., Challapalli, A., Tomasi, G., Rosso, L., Wasan, H., Stebbing, J., Kenny, L., Mangar, S., Riddle, P., Palmieri, C., Al-Nahhas, A., Sharma, R., Turkheimer, F., Coombes, R.C. & Aboagye, E. (2012) Imaging of cellular proliferation in liver metastasis by [18F]fluorothymidine positron emission tomography: effect of therapy. *Phys Med Biol*, **57**, 3419-3433.

Covell, D.G., Wallqvist, A., Rabow, A.A. & Thanki, N. (2003) Molecular classification of cancer: unsupervised self-organizing map analysis of gene expression microarray data. *Mol Cancer Ther*, **2**, 317-332.

Ferreira, N. (2009) Notes from Instrumentation for Medical Imagiology.

Glade, M.J. (1999) Food, nutrition, and the prevention of cancer: a global perspective. American Institute for Cancer Research/World Cancer Research Fund, American Institute for Cancer Research, 1997. *Nutrition*, **15**, 523-526.

Gonzalez, R.C. & Woods, R.E. (2008) *Digital image processing*. Prentice Hall, Upper Saddle River, N.J.

Gray, K.R., Contractor, K.B., Kenny, L.M., Al-Nahhas, A., Shousha, S., Stebbing, J., Wasan, H.S., Coombes, R.C., Aboagye, E.O., Turkheimer, F.E. & Rosso, L. (2010) Kinetic filtering of [(18)F]Fluorothymidine in positron emission tomography studies. *Phys Med Biol*, **55**, 695-709.

Guthikonda, S.M. (2005) Kohonen Self-Organizing Maps *Wittenberg University*. Wittenberg University, Wittenberg, pp. 19.

Guy, C. & Ffytche, D. (2005) *An introduction to the principles of medical imaging*. Imperial College Press ;

Distributed by World Scientific Pub., London

Singapore ; Hackensack, NJ.

Haubner, R. (2010) PET radiopharmaceuticals in radiation treatment planning - synthesis and biological characteristics. *Radiother Oncol*, **96**, 280-287.

Hendee, W.R. & Ritenour, E.R. (2002) *Medical imaging physics*. Wiley-Liss, New York.

Hsu, A.L., Tang, S.L. & Halgamuge, S.K. (2003) An unsupervised hierarchical dynamic self-organizing approach to cancer class discovery and marker gene identification in microarray data. *Bioinformatics*, **19**, 2131-2140.

- Janssen, M.H., Aerts, H.J., Ollers, M.C., Bosmans, G., Lee, J.A., Buijsen, J., De Ruyscher, D., Lambin, P., Lammering, G. & Dekker, A.L. (2009) Tumor delineation based on time-activity curve differences assessed with dynamic fluorodeoxyglucose positron emission tomography-computed tomography in rectal cancer patients. *Int J Radiat Oncol Biol Phys*, **73**, 456-465.
- Kenny, L., Coombes, R.C., Vigushin, D.M., Al-Nahhas, A., Shousha, S. & Aboagye, E.O. (2007) Imaging early changes in proliferation at 1 week post chemotherapy: a pilot study in breast cancer patients with 3'-deoxy-3'-[18F]fluorothymidine positron emission tomography. *Eur J Nucl Med Mol Imaging*, **34**, 1339-1347.
- Kohonen, T. (1997) Self-organizing maps. Springer, Berlin ; London.
- Kohonen, T., Hynninen, J., Kangas, J. & Laaksonen, J. (1996a) SOM_PAK: The Self-Organizing Map Program Package. *Helsinki University of Technology*.
- Kohonen, T. & Oja, E. (1987) Computing with neural networks. *Science*, **235**, 1227a.
- Kohonen, T., Oja, E., Visa, A. & Kangas, J. (1996b) Engineering applications of the self-organizing map. *Proc IEEE*, **84**, 1358-1384.
- Lee, J.A. & Verleysen, M. (2007) *Nonlinear dimensionality reduction*. Springer, New York ; London.
- Lehninger, A.L., Nelson, D.L. & Cox, M.M. (2008) *Lehninger principles of biochemistry*. W.H. Freeman, New York.
- Lobo, V., Bação, F. & Pinheiro, M. (2004) The Self-Organizing Map and its variants as tools for geodemographical data analysis; the case of Lisbon's Metropolitan Area.
- Lobo, V., Cabral, P. & Bação, F. (2007) Self Organizing Maps for Urban Modelling.
- LPCC (2012) Liga Portuguesa Contra o Cancro.
- Macdonald, F., Ford, C.H.J. & Casson, A.G. (2004) *Molecular biology of cancer*. BIOS Scientific Publishers, London ; New York, N.Y.
- Martin, C.J., Dendy, P.P. & Corbett, R.H. (2003) *Medical Imaging and Radiation Protection for Medical Students and Clinical Staff*. The British Institute of Radiology, London.
- Muehllehner, G. & Karp, J.S. (2006) Positron emission tomography. *Phys Med Biol*, **51**, R117-137.

- Nakajima, H. (2003) The delicate balance of world health. Emerging plagues & resurrected pestilence. A report from the World Health Forum 1998. *Caring*, **22**, 6-10, 12.
- NHS (2012) NHS - Information on cancer and useful links;
<http://www.nhs.uk/conditions/Cancer/>.
- Ollinger, J.M. (1997) Positron Emission Tomography. *IEEE Signal Processing Magazine* **14**, 43-55.
- Phelps, M.E. (2006) *PET : physics, instrumentation, and scanners*. Springer, New York.
- PhilipsManual (2010) DICOM Conformance Statement, GEMINI PET/CT Systems with v3.5 or 3.6.
- Pitot, H.C. & Loeb, D.D. (2002) *Fundamentals of oncology*. M. Dekker, New York.
- Quon, A., Chang, S.T., Chin, F., Kamaya, A., Dick, D.W., Loo, B.W., Jr., Gambhir, S.S. & Koong, A.C. (2008) Initial evaluation of 18F-fluorothymidine (FLT) PET/CT scanning for primary pancreatic cancer. *Eur J Nucl Med Mol Imaging*, **35**, 527-531.
- Rasti, J., Monadjemi, A. & Vafei, A. (2007) Color reduction using a multi-stage Kohonen Self-Organizing Map with redundant features. *Expert Systems with Applications*, **38**, 13188-13197.
- Ritter, H., Martinetz, T. & Schulten, K. (1992) *Neural computation and self-organizing maps : an introduction*. Addison-Wesley, Reading, Mass.
- Rohren, E.M., Turkington, T.G. & Coleman, R.E. (2004) Clinical applications of PET in oncology. *Radiology*, **231**, 305-332.
- Ron Wehrens, L.M.C.B. (2007) Self- and Super-organizing Maps in R: the Kohonen Package. *Journal of Statistical Software*, **21**.
- Saha, G.B. (2010) *Basics of PET imaging : physics, chemistry, and regulations*. Springer, New York.
- Sattler, B., Lee, J.A., Lonsdale, M. & Coche, E. (2010) PET/CT (and CT) instrumentation, image reconstruction and data transfer for radiotherapy planning. *Radiother Oncol*, **96**, 288-297.
- Schiepers, C. & Baert, A.L. (2006) *Diagnostic Nuclear Medicine*, New York.

- Sharp, P.F., Gemmell, H.G. & Murray, A.D. (2005) *Practical nuclear medicine*. Springer, London ; New York.
- Su, A.I., Welsh, J.B., Sapinoso, L.M., Kern, S.G., Dimitrov, P., Lapp, H., Schultz, P.G., Powell, S.M., Moskaluk, C.A., Frierson, H.F., Jr. & Hampton, G.M. (2001) Molecular classification of human carcinomas by use of gene expression signatures. *Cancer Res*, **61**, 7388-7393.
- Tobias, J.S., Hochhauser, D. & Souhami, R.L. (2010) *Cancer and its management*. Wiley-Blackwell, Chichester, West Sussex, UK ; Hoboken, NJ.
- Townsend, D.W. (2008) Multimodality imaging of structure and function. *Phys Med Biol*, **53**, R1-R39.
- Tu, J.V. (1996) Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *J Clin Epidemiol*, **49**, 1225-1231.
- Turkheimer, F.E., Edison, P., Pavese, N., Roncaroli, F., Anderson, A.N., Hammers, A., Gerhard, A., Hinz, R., Tai, Y.F. & Brooks, D.J. (2007) Reference and target region modeling of [11C]-(R)-PK11195 brain studies. *J Nucl Med*, **48**, 158-167.
- Valkonen, V.P., Kolehmainen, M., Lakka, H.M. & Salonen, J.T. (2002) Insulin resistance syndrome revisited: application of self-organizing maps. *Int J Epidemiol*, **31**, 864-871.