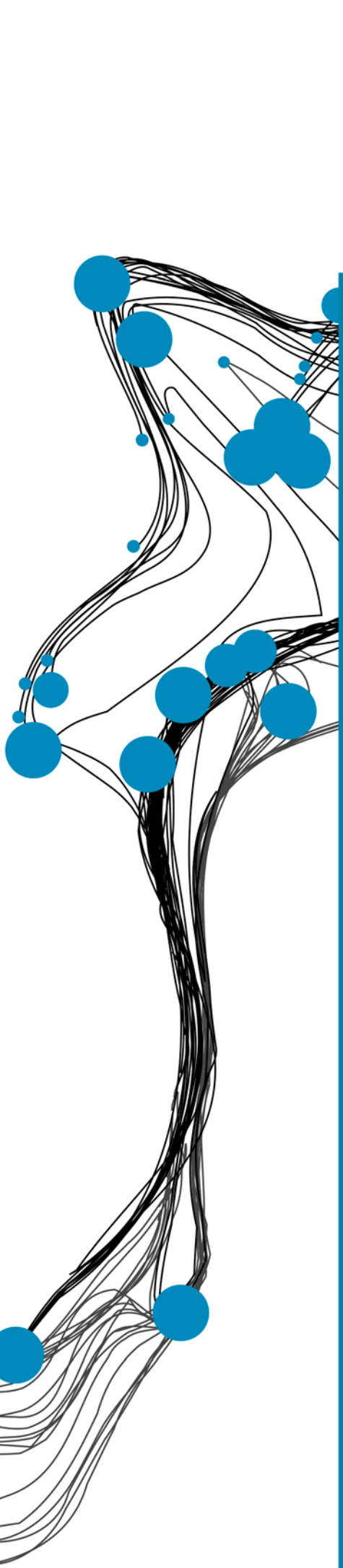


STATIC, PARAMETRIC AND DYNAMIC RADIOMICS FROM FDG-PET IN NON-SMALL CELL LUNG CANCER

W.A. Noortman

MASTER THESIS TECHNICAL MEDICINE
MEDICAL IMAGING AND INTERVENTIONS

EXAMINATION COMMITTEE
Chairman:
prof.dr. L.F. de Geus-Oei,
Medical supervisor:
dr. D. Vriens
Technical supervisor:
dr. F.H.P. van Velden
Technical supervisor:
prof.dr.ir. C.H. Slump
Professional behaviour supervisor:
drs. P.A. van Katwijk
External member:
M.E. Kamphuis, MSc

LUMC

UNIVERSITY OF TWENTE.

STATIC, PARAMETRIC AND DYNAMIC RADIOMICS FROM ^{18}F-FDG PET IN NON-SMALL CELL LUNG CANCER

W.A. Noortman

07-12-2018

Master thesis technical medicine - Medical imaging and interventions

Examination committee:

Chairman:
prof.dr. L.F. de Geus-Oei,
Nuclear medicine physician, department of radiology, Leiden University Medical Center
Biomedical Photonic Imaging, University of Twente

Medical supervisor:
dr. D. Vriens
Nuclear medicine physician, department of radiology, Leiden University Medical Center

Technical supervisor:
dr. F.H.P. van Velden
Medical physicist, department of radiology, Leiden University Medical Center

Technical supervisor:
prof.dr.ir. C.H. Slump
Robotics and Mechatronics, University of Twente

Professional behaviour supervisor:
drs. P.A. van Katwijk
Clinical internships and professional behaviour, University of Twente

External member:
M.E. Kamphuis, MSc
Robotics and Mechatronics, University of Twente

Leiden University Medical Center

UNIVERSITY OF TWENTE.

GLOSSARY

AC	adenocarcinoma
CT	computed tomography
DFS	disease-free survival
DSS	disease-specific survival
^{18}F-FDG	2-18F-fluoro-2-deoxy-D-glucose
GLCM	grey level cooccurrence matrix (group of radiomic features)
GLDM	grey level dependence matrix (group of radiomic features)
GLRLM	grey level run length matrix (group of radiomic features)
GLSZM	grey level size zone matrix (group of radiomic features)
MRI	magnetic resonance imaging
MTV	metabolic tumour volume (measure for quantitative PET)
NGTDM	neighbouring grey tone difference matrix (group of radiomic features)
NSCLC	non-small cell lung carcinoma
OS	overall survival
PCA	principal component analysis
PVE	partial volume effect
PET	positron emission tomography
TLG	total lesion glycolysis (measure for quantitative PET)
TOF	time of flight
SCC	squamous cell carcinoma
SUV	standardized uptake value (measure for quantitative PET)
VOI	volume of interest

PREFACE

In this master thesis, I present the research I have done during my graduation internship at the section of nuclear medicine at department of radiology at the Leiden University Medical Center. The past year I have conducted research to radiomics derived from positron emission tomography (PET) in non-small cell lung carcinoma. The field of radiomics studies the extraction of quantitative features from medical imaging with the goal to find stable and clinically relevant image-derived biomarkers or radiomic features that provide a non-invasive way of quantifying and monitoring tumour characteristics in clinical practice. The thesis starts with a general introduction about non-small cell lung carcinoma, PET imaging and radiomics. This section is followed by an article in which I present my research project. A general discussion follows, describing the potential and difficulties within the field of radiomics and also providing a future perspective of this interesting field. The last section of this document consists of the 'verantwoording', in which I reflect on my clinical- and personal development during this graduation internship and describe side project I did. I hope you will enjoy reading this thesis.

Wyanne Noortman,

november 2018

TABLE OF CONTENTS

GENERAL INTRODUCTION

Non-small cell lung carcinoma

Primary lung carcinomas have been the most common cancer type worldwide for several decades. It is the most common cause of cancer-related death with an estimated age-standardised mortality rate worldwide of 19.7 per 100.000 in 2012 (1). In the Netherlands, every year around 12.000 patients are diagnosed with lung cancer (2). The 5-year survival of lung cancer was only 19% between 2011 and 2015 in the Netherlands (3), but is highly dependent of tumour stage. The low 5-year survival is partly due to the fact that only 20% of patients is eligible for a primary resection, because most newly diagnosed patients already have metastases and are therefore inoperable or patients are inoperable due to comorbidities (4). Lung cancers arise in cigarette smokers in 85% to 90% of cases. Other risk factors include exposure to secondhand smoke, asbestos, radon gas or air pollution and pre-existing non-malignant lung diseases (5, 6). Also, lung carcinomas are associated with germline mutations (5).

Histology

Primary lung cancers can be divided into non-small cell lung carcinoma (NSCLC, 85%) and small cell lung carcinoma (SCLC, 15%), based on histopathological size and appearance of the malignant cells (7). Subtypes of NSCLC are adenocarcinomas (37.5%), squamous cell carcinomas (26.8%) large-cell carcinomas (5.7%), bronchioloalveolar carcinomas (3.5%) and other (26.5%) (7). Squamous cell carcinomas often arise after injury of the bronchial epithelium, for example as a result of smoking. These tumours are often centrally located near the major or segmental bronchi and regularly show central cavitation (necrosis). Adenocarcinomas often arise in the periphery of the lung (8). Histologic subtypes of adenocarcinoma are lepidic, acinar, papillary, micropapillary, solid and mixed subtypes (9). Classification is based on the predominant histological subtype. Metastases of NSCLC occur predominantly in regional lymph nodes, mostly hilar- and mediastinal nodes, but also in the adrenal glands, brain, bone and liver (8). Staging of NSCLC is performed according to the TNM classification (T: primary tumour, N: lymph node involvement, M: distant metastasis). Based on the TNM classification, cancer stages were determined, ranging from 0 to IV. Table 1 shows the TNM classification for lung cancer and table 2 shows the stages of NSCLC (10).

Diagnosis

The clinical presentation of lung cancer highly depends on the location and type of the tumour and the presence of metastases. Symptoms of NSCLC include cough, dyspnoea, haemoptysis, chest pain, obstructive pneumonia and pleural effusion (due to bronchial obstruction) (8). When suspicion for lung carcinoma arises a chest radiograph or a diagnostic computed tomography (CT) scan with intravenous contrast of the thorax and upper abdomen is acquired. Additionally, ^{18}F-fluorodeoxy glucose (18F-FDG) positron emission tomography (PET) can be acquired for evaluation of the primary tumour as well as for detection of regional and distant metastases. ^{18}F-FDG PET/CT imaging is indicated for any patient eligible for curative therapy (4).

The histological diagnosis is based on tissue sampling (5). A biopsy is taken during bronchoscopy or mediastinoscopy or a CT-guided transbronchial biopsy is performed. The choice of procedure depends on the location and stage of the tumour. Accurate lymph node staging is essential for treatment and prognosis. When suspicion for a lymph node metastasis arises based on medical imaging, lymph node sampling is performed. Depending on the location of the lymph node this is performed during endobronchial ultrasound transbronchial needle aspiration (EBUS-TBNA), transoesophageal endoscopic ultrasound-guided fine-needle aspiration (EUS-FNA) or surgically (mediastinoscopy) (4).

Table 1: TNM (tumour, lymph node involvement, distant metastasis) descriptors from the eight edition of TNM classification for lung cancer (10)

T: PRIMARY TUMOUR	
Tx	Primary tumour cannot be assessed or tumour proven by presence of malignant cells in sputum or bronchial washing, but not visualised by imaging or bronchoscopy
T0	No evidence of primary tumour
Tis	Carcinoma in situ
T1	Tumour ≤ 3 cm in greatest dimension surrounded by lung of visceral pleura without bronchoscopic evidence of invasion more proximal than the lobar bronchus (i.e., not in the main bronchus)
T1a(mi)	Minimally invasive adenocarcinoma
T1a	Tumour ≤ 1 cm in greatest dimension
T1b	Tumour > 1 cm but ≤ 2 cm in greatest dimension
T1c	Tumour > 2 cm but ≤ 3 cm in greatest dimension
T2	Tumour > 3 m in but ≤ 5 cm or tumour with any of the following features: - Involved main bronchus regardless of distance from the carina but without involvement of the carina - Invaded visceral pleura - Associated with atelectasis or obstructive pneumonitis that extends to the hilar region, involving part or all of the lung
T2a	Tumour > 3 cm but ≤ 4 cm in greatest dimension
T2b	Tumour > 4 cm but ≤ 5 cm in greatest dimension
T3	Tumour > 5 cm but ≤ 7 cm in greatest dimension or associated with separate tumour nodule(s) in the same lobe as the primary tumour or invades any of the following structures: chest wall (including the parietal pleura and superior sulcus tumours), phrenic nerve, parietal pericardium
T4	Tumour > 7 cm in greatest dimension or associated with separate tumour nodules(s) in a different ipsilateral lobe than that of the primary tumour or invades any of the following structures: diaphragm, mediastinum, heart, great vessels, traches, recurrent laryngeal nerve, oesophagus, vertebral body and carina

N: REGIONAL LYMPH NODE INVOLVEMENT	
Nx	Regional lymph nodes cannot be assessed
N0	No regional lymph node metastasis
N1	Metastasis in ipsilateral peribronchial and/or ipsilateral hilar lymph nodes ad intrapulmonary nodes, including involvement by direct extension
N2	Metastasis in ipsilateral mediastinal and/or subcarinal lymph node(s)
N3	Metastasis in contralateral mediastinal, contralateral hilar, ipsilateral or contralateral scalene, or supraclavicular lymph node(s)

M: DISTANT METASTASIS	
M0	No distant metastasis
M1	Distant metastasis present
M1a	Separate tumour nodules(s) in a contralateral love; tumour with pleural or pericardial nodule(s) or malignant pleural or pericardial effusion
M1b	Single extrathoracic metastasis
M1c	Multiple extrathoracic metastases in one or more organs

Table 2: Stages of non-small cell lung carcinomas from the eight edition of TNM classification for lung cancer (10).

	N0	N1	N2	N3
T1	IA	IIB	IIIA	IIIB
T2a	IB	IIB	IIIA	IIIB
T2b	IIA	IIB	IIIA	IIIB
T3	IIB	IIIA	IIIB	IIIC
T4	IIIA	IIIA	IIIB	IIIC
M1a	IVA	IVA	IVA	IVA
M1b	IVA	IVA	IVA	IVA
M1c	IVB	IVB	IVB	IVB

Treatment

Treatment is highly dependent on tumour stage. In early disease stages (I-IIIA), a radical resection, mostly in the form of a lobectomy, is the treatment of choice, in higher stages combined with chemotherapy or chemoradiotherapy. These are also treatment options when a patient is not eligible for surgery due to comorbidities. In advanced stage NSCLC (IIIB-IV), palliative treatment in the form of chemotherapy or chemoradiotherapy in combination with best supportive care is induced. Also, in recent years, targeted therapy in the form of biological therapeutic agents targeting specific molecular pathways have been assessed for anti-cancer therapy (11), most recently in the form of immune-checkpoint inhibitors such as nivolumab (PD-1 (programmed cell death protein) inhibitor) (12), pembrolizumab (PD-1 inhibitor) (13), possibly followed by atezolizumab (PD-L1 (programmed cell death ligand) inhibitor) in the near future (14).

Positron emission tomography

Positron emission tomography (PET) is major functional imaging technique in nuclear medicine (15). PET is used to image the distribution of a radionuclide administered in the body. When a radionuclide is rich in protons in relation to neutrons, it decays by the emission of a positron (anti-electron, β^+), further resulting in a neutron an electron-neutrino (ν_e). An example is the decay of fluorine-18 (^{18}F) to oxygen-18 (^{18}O): ${}^{18}_{9}F \rightarrow {}^{18}_{9}F + \beta^+ + v_e$ ($T_{1/2}$: 109.771(20) minutes). The positron loses kinetic energy by interaction with electrons (e^-). When the positron has lost enough energy, it combines with an electron. This process is called annihilation and leads to the production of two gamma photons of 511 keV, that are emitted in nearly opposite directions (~180°). The annihilation process is shown in figure 1. The emitted gamma photons are detected by the ring of high-energy detectors of the PET scanner. These detectors consist of a scintillator and a photo-detector. When a gamma photon interacts with the scintillator, visible scintillation light is emitted. The light is converted into an electrical signal by the photo-detector. When two gamma photons are detected by the ring of detectors within a time window of several nano-seconds, this is called a coincidence. If a coincidence pair is detected, it is expected that the annihilation event occurred somewhere on the 'line of response' (LOR) between both activated detector elements. Current state-of-the-art PET/CT scanners use the time difference between the detection of the photons of a coincidence pair to calculate the location on the LOR where the annihilation took place. This technique is called time-of-flight (TOF) PET.

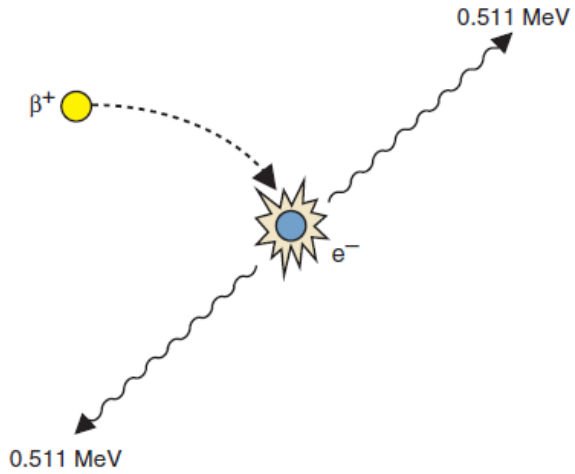


Figure 1: Annihilation of a positron (β^+) and an electron (e^-) with emission of a pair of 511 keV annihilation photons at ~180 degrees to each other (11).

All recorded coincidences together provide information about the quantity and location of positron emitting isotopes in the body. Coincidence data are stored in list mode. List mode data contain the x-y position signals from the camera, stored with periodic clock markers (15). The time data enable retrospective framing useful for data analysis. PET data are reconstructed into the spatial distribution of a tracer using different algorithms. Reconstruction algorithms are often based on iterative expectation maximization. This method aims to find the source distribution that would have created the observed projection data. Imaging data are processed assuming that the total number of coincidence events detected by the two detector elements is proportional to the total amount of tracer. In current clinical practice, PET imaging is combined with low dose computed tomography (CT) to provide anatomical information and for attenuation and scatter correction.

One of the major limitations of PET is the limited spatial resolution leading to uncertainties in the expectation of the tracer distribution (16). The spatial resolution of an imaging technique is defined as the minimum distance between two objects in an image, in which they can be distinguished as two separate points. The spatial resolution of a PET-scanner is predominantly limited by the size of the individual detector elements, the positron range and acollinearity of annihilation (15). The spatial resolution of a PET scanner is around 4-5 mm. Due to the limited spatial resolution, the

partial volume effect (PVE) occurs (17). This effect is characterised by a lower apparent activity in small objects with spill out to surrounding objects. Due to this effect small lesions seem larger, but the uptake is underestimated.

^{18}F-FDG

2-^{18}F-fluoro-2-deoxy-D-glucose (^{18}F-FDG) is the most common radiopharmaceutical used in clinical PET imaging. This non-metabolizable glucose analogue consists of a tracer (2-deoxy-D-glucose, DG) radiolabelled with a positron emitting radionuclide (^{18}F) (18). Figure 2 shows the structural formulas of glucose and ^{18}F-FDG (18). ^{18}F-FDG PET images in vivo whole body glucose metabolism. Since many pathological conditions cause regional alterations in glucose metabolism, ^{18}F-FDG PET is an important tool in detection and staging of cancer and active inflammations. Not only malignant lesions show higher glucose metabolism, but ^{18}F-FDG is accumulated in all cells using glucose as primary energy source. ^{18}F-FDG shows physiologic uptake in the brain, myocardium, bowel, liver, spleen and (active) muscles and is excreted by the kidneys to the bladder (19). ^{18}F-FDG is administered intravenously and the recommended interval between administration and the start of acquisition is 60 min (20).

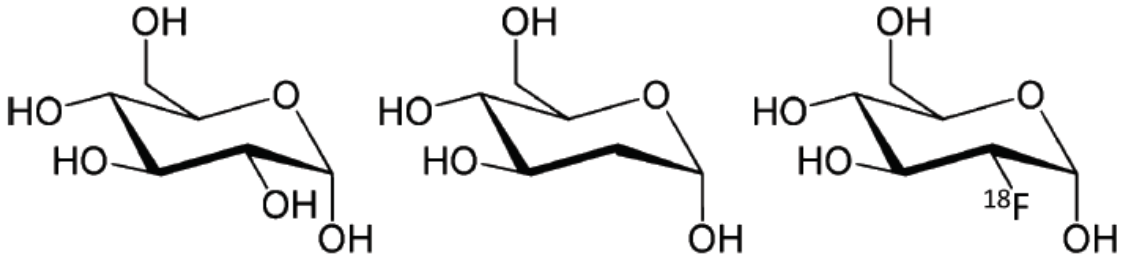


Figure 2: Structural formulas of D-glucose (left), 2-deoxy-D-glucose (middle) and 2-^{18}F-fluoro-2-deoxy-D-glucose (^{18}F-FDG) (right) (18). ^{18}F is synthesized in a cyclotron and is added to the 2-deoxy-D-glucose. In several intermediate reactions of precursor molecules, ^{18}F-FDG is formed.

Quantitative PET

Next to visual inspection of PET images for diagnosis, which is mostly used in clinical practice, (semi-)quantitative analysis allows an objective complement to visual interpretation (21). Quantitative measures are especially useful in response monitoring. Measures used in quantitative PET are the standardised uptake value (SUV), the metabolic tumour volume (MTV) and the total lesions glycolysis (TLG). The SUV expresses the ratio between the activity concentration at a single time point and the administered activity, taking into account a measure for distribution (e.g. body weight), the MTV expresses the functional tumour volume and the TLG is the product of the mean SUV and MTV.

Pharmacokinetic modelling

In PET, the spatial distribution of the radiotracer is measured. This distribution is varying in time and therefore the timing of imaging relative to administration has to be considered. The radiotracer concentration can also be measured as a function of time, providing quantitative measures of physiologic parameters and biochemical rates. Combined with knowledge of the biological behaviour of the natural molecule the radionuclide is bound to, pharmacokinetic analysis is possible (22). This pharmacokinetic analysis is possible when dynamic PET studies are acquired. A dynamic study protocol includes only one field of view (FOV) of 15-20 cm in which the PET signal ($C_{PET}(t)$) is measured over time. Also, the tracer concentration of the arterial blood plasma ($C_{plasma}(t)$) is measured (21).

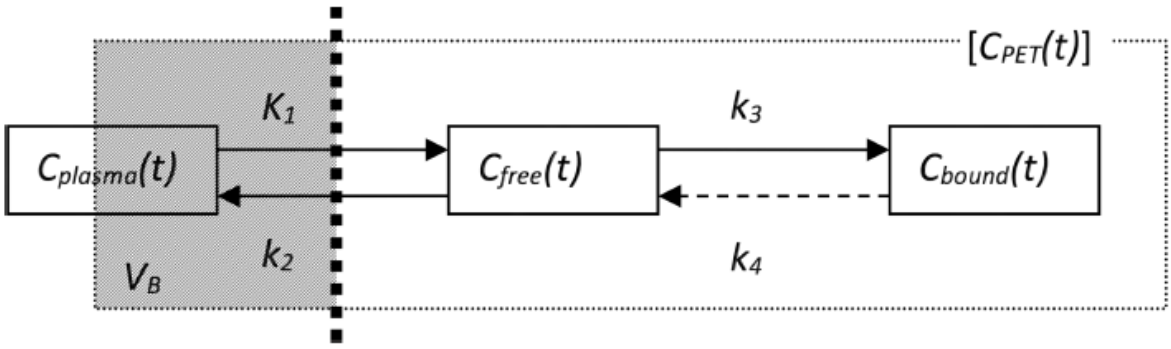


Figure 3: Two-compartment model of ^{18}F-FDG with four first-order rate constants ($\boldsymbol{K_1}$-k4) describing transport between the compartments. The vertical dotted line symbolises the cell membrane, $\boldsymbol{C_{plasma}(t)}$ is the activity concentration of ^{18}F-FDG in the arterial blood plasma, $\boldsymbol{C_{free}(t)}$ is the intracellular activity concentration of free ^{18}F-FDG and $\boldsymbol{C_{bound}(t)}$ is the intracellular activity concentration of ^{18}F-FDG-6-phosphate. $\boldsymbol{C_{PET}(t)}$ is the measured PET signal, which is a combination of $\boldsymbol{C_{free}(t)}$ and $\boldsymbol{C_{bound}(t)}$ and a fraction of $\boldsymbol{C_{plasma}(t)}$ (21).

Two compartment model

Pharmacokinetic modelling in ^{18}F-FDG PET is based on glucose metabolism. ^{18}F-FDG metabolism can be simplified in a two-compartment model, which is shown in figure 3. The vertical dotted line symbolises the cell membrane, $C_{plasma}(t)$ is the activity concentration of ^{18}F-FDG in the arterial blood plasma, $C_{free}(t)$ is the intracellular activity concentration of free ^{18}F-FDG and $C_{bound}(t)$ is the intracellular activity concentration of ^{18}F-FDG-6-phosphate. $C_{PET}(t)$ is the measured PET signal, which is a combination of $C_{free}(t)$ and $C_{bound}(t)$ and a fraction of $C_{plasma}(t)$. The arrows indicate the fluxes between the compartments, indicated with rate constants K_1-k_4. K_1 and k_2 indicate ^{18}F-FDG influx and outflux by membrane-bound sodium-dependent glucose transporter family (GLUT). k_3 indicates cytosolic phosphorylation by the hexokinase family. Different from D-glucose metabolites, ^{18}F-FDG metabolites cannot be catabolised further due to the replacement of the 2-hydroxyl group with hydrogen and cannot diffuse across cell membranes. Therefore ^{18}F-FDG-6-phosphate is trapped inside the cell and rate constant k_4 is zero (21) in most cancer cells. Rate constant K_1 is capitalized, because it takes into account blood flow and tracer extraction. Therefore, K_1 is expressed in mL blood per minute per gram of tissue, while the other rate constants have units of inverse minute (23, 24).

To translate the complex biological system of glycose metabolism into a simple compartmental model, some assumptions have to be made. In the first place it is assumed that all compartments are homogenous and well mixed. This means there are no concentration gradients within a compartment and every tracer molecule has equal probability to exchange into another compartment. Secondly, it is assumed that the underling physiological processes are in steady state, i.e. the rate constants of the systems do not change with time during the study. Therefore, the model can be expressed using linear differential equations. Also, it is assumed that the tracer behaves similarly to the non-radioactive natural biological substrate and the concentration of first is negligible compared to the concentration of the latter ($[S*] << [S]$) (23). Lastly, the assumption is made that the delivery of ^{18}F-FDG is independent of blood flow (21).

In enzyme kinetics, the Michaelis-Menten hypothesis describes the reaction of a substrate and an enzyme forming an intermediate complex, which is converted to a product with release of an enzyme. This reaction is shows in equation 1:

$$S + E \underset{k_2}{\overset{k_1}{\rightleftarrows}} SE \overset{k_3}{\rightarrow} P + E \qquad [1]$$

where S is the substrate, E the enzyme, P the product and k_1, k_2 and k_3 the rate constants for the steps of the reaction process.

The reaction rate R for the conversion of S to P is stated in the Michaelis-Menten equation 2:

$$R = \frac{V_{max} \cdot [S]}{[S] + K_m} \quad [2]$$

where V_{max} (mg/min) is the maximum rate of the reaction and K_m is the concentration of S that produces a reaction rate of one half the maximum value, i.e. $R = \frac{1}{2} V_{max}$.

When more than one substrate is competing for the enzyme E (S and S^*), but with a much lower concentration, the reaction rate of the competing substrate is stated in equation 3:

$$R^* \cong \frac{V^*_{max} \cdot \frac{K_m}{K^*_m}}{[S] + K_m} [S^*] \quad [3]$$

where R^* is the reaction rate of competing substrate, as long as the concentration of the competing substrate (S^*) is much lower than the concentration of the natural substrate (S), i.e. $[S*] << [S]$ (21).
The two-compartment model of FDG as shown in figure 3 can be expressed using differential equations. The net flux into a compartment can be expressed as the sum of all inflows minus the sum of all outflows. It is equal to the rate of change (d/dt) of the concentration in the compartment (dC/dt).

The rate of change of the tissue concentrations in the different compartments can be expressed by two differential equations:

$$\frac{dC_{free}(t)}{dt} = K_1 \cdot C_{plasma}(t) - (k_2 + k_3) \cdot C_{free}(t)$$
$$\Rightarrow C_{free}(t) = K_1 \cdot e^{-(k_2 + k_3) \times t} * C_{plasma}(t) \quad [4]$$

$$\frac{dC_{bound}(t)}{dt} = k_3 \cdot C_{free}(t)$$
$$\Rightarrow C_{bound}(t) = \frac{K_1 \cdot k_3}{k_2 + k_3} \cdot \left(1 - e^{-(k_2 + k_3) \cdot t}\right) * C_{plasma}(t) \quad [5]$$

where * stands for the operation of convolution, $C(t)$ is the concentration of the tracer in a certain compartment and K_1-k_3 are the rate constants of the different processes. Differential equations are solved with LaPlace transforms.

The activity concentration measured in the PET scan is expressed in equation 6:

$$C_{PET}(t) = (1 - V_B) \cdot \Big(C_{free}(t) + C_{bound}(t) \Big) + V_B \cdot C_{plasma}(t) \quad [6]$$

The ratio between the rate of phosphorylation of glucose (MR_{glc}) and ^{18}F-FDG (MR_{FDG}) is based on the Michaelis-Menten equations for the rate constants of the competing substrates (equation 2 and 3) and can also be expressed as the rate constant of phosphorylation. The ratio is expressed in equation 7:

$$\frac{MR_{FDG}}{MR_{glc}} = \frac{V_{max,FDG} \cdot K_{m,glc} \cdot C_{free,FDG}(t)}{V_{max,glc} \cdot K_{m,FDG} \cdot C_{free,glc}(t)} = \frac{k_{3,FDG} \cdot C_{free,FDG}(t)}{k_{3,glc} \cdot C_{free,glc}(t)} \quad [7]$$

where $V_{max}$and K_m are the Michaelis-Menten constants for the hexokinase mediated phosphorylation of both FDG and glucose.

Since it is assumed that the concentrations of glucose and FDG in the plasma are constant, equation 6 can also expressed as:

$$\frac{MR_{FDG}/C_{plasma,FDG}}{MR_{glc}/C_{plasma,glc}} = \frac{V_{max,FDG} \cdot K_{m,glc} \cdot C_{free,FDG}(t)/C_{plasma,FDG}}{V_{max,glc} \cdot K_{m,FDG} \cdot C_{free,glc}(t)/C_{plasma,glc}} \quad [8]$$

In this equation, the ratio between the intracellular activity concentration and the activity concentration in the arterial plasma of glucose and FDG ($C_{free,glc}/C_{plasma,glc}$ and $C_{free,FDG}/C_{plasma,FDG}$) is also known as the partition coefficient or the tissue-to-blood concentration ratios of glucose and FDG (λ_{glc} and λ_{FDG}). Also, the process is dependent of the blood flow F (22). This is expressed in equation 9:

$$\frac{MR_{FDG} \cdot (C_{plasma,FDG} \cdot F)^{-1}}{MR_{glc} \cdot (C_{plasma,glc} \cdot F)^{-1}} = \frac{V_{max,FDG} \cdot K_{m,glc} \cdot \lambda_{FDG}}{V_{max,glc} \cdot K_{m,FDG} \cdot \lambda_{glc}} = LC_{FDG} \quad [9]$$

with LC_{FDG} the lumped constant or the steady-state ratio of the net extraction of FDG to that of glucose at constant plasma levels of FDG and glucose. It illustrates competitive enzyme kinetics. It is used as a correction term that measures the difference in use of FDG and glucose in tissue (22).

The goal of pharmacokinetic modelling is to measure glucose metabolism, consequently equation 8 is rewritten in equation 9:

$$MR_{glc} = \frac{dC_{bound,glc}(t)}{dt} = \frac{MR_{FDG}}{C_{plasma,FDG(t)}} \cdot \frac{C_{plasma,glc}(t)}{LC_{FDG}} = \frac{K_1^* k_3^*}{k_2^* + k_3^*} \cdot \frac{C_{plasma,glc}}{LC_{FDG}} \quad [10]$$

$\frac{K_1^* k_3^*}{k_2^* + k_3^*}$ or K_i^* are the rate constants for FDG and are derived from the steady state equations of the two-compartment model. The notation is derived by Gambhir (24).

Patlak graphical method

Graphical analysis is applied to tracer kinetic data. This concept uses a mathematical transformation on the measured data in order to acquire a straight-line plot, where the slope and/or intercept have physiological meaning (23). Patlak et al. derived a graphical method that uses linear regression to analyse pharmacokinetics in a compartment model when there is an irreversible or nearly irreversible reaction in the model (i.e. $k_4=0$ or $k_4 << k_3$) (25). After combining equation 4, 5 and $\frac{K_1^* k_3^*}{k_2^* + k_3^*} = K_i$, we have:

$$\frac{C_{PET}(t)}{C_{plasma}(t)} = \left(K_i(1 - V_B)\right) \cdot \left(\frac{\int_0^t C_{plasma}(\tau) d\tau}{C_{plasma}(t)}\right) + \left(\frac{(1 - V_B) \cdot K_1 \cdot k_2}{(k_2 + k_3)^2} + V_B\right) \quad [11]$$

where $K_i(1 - V_B)$ is the slope of the linear regression line between $\frac{\int_0^t C_{plasma}(\tau) d\tau}{C_{plasma}(t)}$ and $\frac{C_{PET}(t)}{C_{plasma}(t)}$ after giving the system some time to stabilise (often the first 15 minutes of normalised Patlak-time

($\frac{\int_0^t C_{plasma}(\tau)d\tau}{C_{plasma}(t)}$) are left out of the fit). With an estimation of V_B, LC_{FDG} (unknown and often set at unity), a measured $C_{plasma,glc}$ and equation 10, it is possible to determine MR_{glc} (21).

The model assumes that all reversible compartments are in equilibrium with the plasma and that the bolus injection is a constant infusion (25). Figure 4 gives an overview of the Patlak graphical analysis.

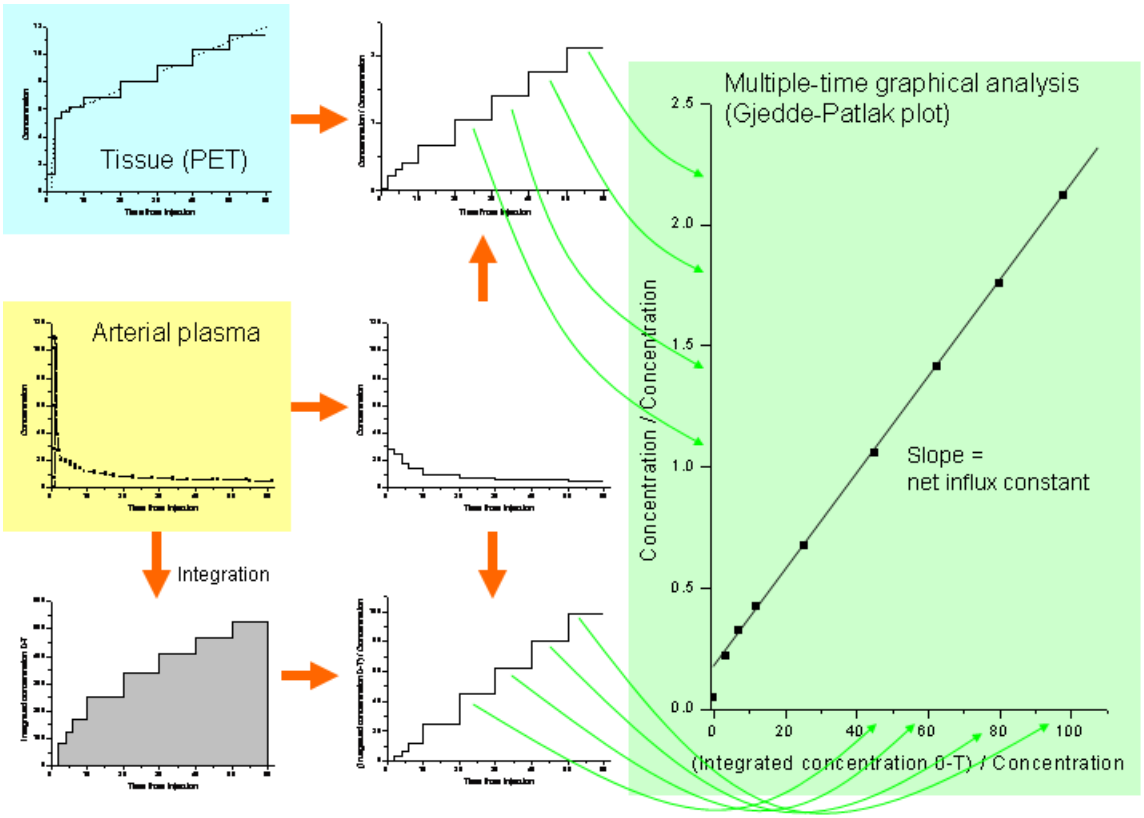


Figure 4: Schematic overview of Patlak graphical analysis. The plot becomes linear after the tracer concentrations in the reversible compartments and in plasma are in equilibrium. The slope of the linear phase of plot is the net uptake (influx) rate constant $\boldsymbol{K_i}$, taking into account the blood volume (26).

Radiomics

The term radiomics refers to a rapidly-emerging discipline within medical image processing and analysis with the goal to extract large amounts of quantitative data from medical images. Radiomics are mainly used in oncology for the characterization of specific aspects of patient health. Computer assisted interpretation is used to extract information from medical imaging studies like (PET/)CT and magnetic resonance imaging (MRI). Usually, these studies are used to assess intensity, size and shape of tumours. Some commonly used quantitative image-derived features in PET/CT are the standardized uptake value (SUV), the metabolic tumour volume (MTV) and the total lesion glycolysis (TLG) (21). The SUV expresses the ratio between the activity concentration at a single time point and the administered activity, taking into account a measure for distribution (e.g. body weight), the MTV expresses the functional tumour volume and the TLG is the product of the mean SUV and MTV.

Nevertheless, these features do not express the tracer uptake heterogeneity, which contains additional information about the biological behaviour of the tumour. Biologically, heterogeneity of the microenvironment of the tumour might be reflected in medical images and provide information about cellular density, proliferation, angiogenesis, hypoxia, necrosis and fibrosis (27). In PET, these biological processes are expressed in the spatial distribution of radiotracer uptake (28). Still, it has to be taken into account that the features do not bear a direct relationship to the underlying cellular biology on a microscopic scale, since, especially in PET, the features relate to a relatively macroscopic scale based on the used voxel size (29). For instance, with an estimation of 10^8 tumour cells in 1 cm³ of tumour tissue (30), a voxel of 2x2x2 mm³ contains already 8×10^5 tumour cells and a voxel of 4x4x4 mm³ contains even 6.4×10^6 cells.

The field of radiomics has the potential to improve knowledge in tumour biology and guide management of patients (31). Patient prognosis and treatment of choice vary between different cancer types and depend on tumour stage. Currently, the gold standard in tumour classification is histological tissue sampling (6). However, the biopsy techniques are invasive and since tumours do not represent a homogeneous entity, the biopsy represents only a small sample of the tumour as a whole (32). An important advantage of radiomics is that it is possible to sample the tumour as a whole in a non-invasive setting (29). When integrated and analysed with patient information like pathology, blood biomarkers and genomics, radiomics can play an important role in precision medicine and clinical decision making (33, 34). Thus, the field of radiomics aims to find stable and clinically relevant image-derived biomarkers or radiomic features that provide a non-invasive way of quantifying and monitoring tumour characteristics in clinical practice (35). PET-based radiomic features have been studied for prediction of treatment response (36, 37), overall survival (38, 39) and for identification of tumour phenotypes (40).

Radiomics consist of image acquisition and reconstruction, volume segmentation, feature extraction (radiomics as well as clinical and molecular) and storage and signature development and validation on one or several datasets (32). The full pipeline is shown in figure 5 (32). Feature extraction uses the segmented VOI for the creation of two masks: an intensity mask and a morphological mask. The intensity masks consists of different intensities in the VOI expressed in voxel values and the morphological mask describes the shape of the tumour (41). Radiomic features are classified into categories quantitatively describing intensity, shape and texture (42). Intensity based features are assessed by statistics, which characterize distribution of voxel values without considering spatial relationship based on the intensity mask. Shape features are based on the morphological mask and describe the surface or specific shape of the tumour. Texture features describe relationships between image voxels and are divided, among others, into fractals, grey-level co-occurrence matrices and grey-level run-length matrices within the VOI (41). Currently, over 5000 quantitative features are described in literature.

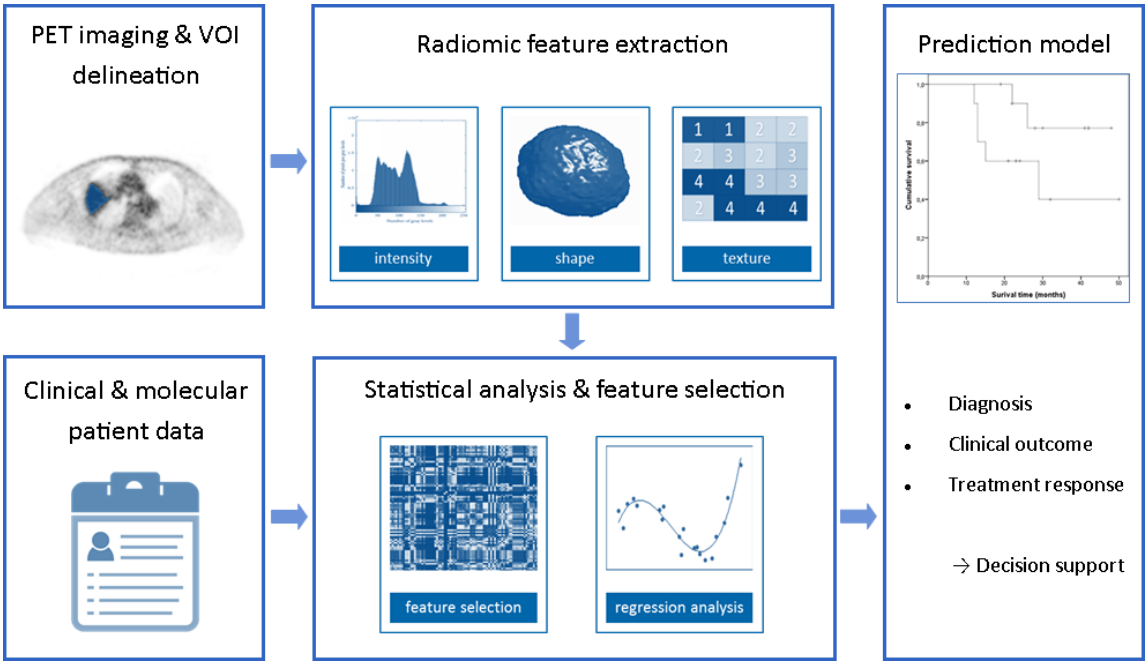


Figure 5: Overview of methodological processes within the field of radiomics: data collection, preparation, modelling and validation.

Difficulties within the field of radiomics are caused by inhomogeneity of data. Most data are extracted retrospectively from standard-of-care images, where acquisition parameters and reconstruction might vary between scans. To be able to attribute differences in radiomic features to tumour biology, it is important that data are homogeneous when it comes to acquisition and reconstruction (33). Therefore, cohorts are often small. Also, the number of patients in the cohort is small compared to the number of evaluated features. This introduces the 'curse-of-dimensionality', a problem that arises when analysing data in high-dimensional spaces (the hundreds of radiomic features). The data space increases exponentially with the number of variables, while the number of data points or samples stays the same. This leads to overfitting of the model. Thereby the generalisation performance of the model is negatively impacted, since the model is too specific (43).

Care has to be taken selecting a statistical method. Feature reduction or other adjustments for multiple testing are crucial to reduce the risk of overfitting in the field of radiomics (44). Dimension reduction can be performed using clustering approaches and principal component analysis (PCA) (45). Repeatability and variability of the radiomic features should be considered in feature selection (42, 46). Especially for response monitoring, it is critical to know whether an observed change in tracer uptake (heterogeneity) or tumour geometry is caused by a true response or by methodological variations, i.e. biological, technical or observer variability (42). Repeatability can be assessed with a test-retest analysis, where double baseline scans are acquired (42, 46). Inter-observer variability is a useful tool to assess repeatability (46). Also, validation of the model is important to test whether the model is predictive for the target patent population or just for a particular subset of samples analysed (45). This is done by splitting the dataset in a training- and a validation dataset or by using an external dataset for validation. In many studies validation of the model is not performed, since the number of patients is often limited.

^{18}F-FDG PET in non-small cell lung carcinoma

The ability to make suitable treatment decisions in patients with lung cancer is highly dependent on accurate disease staging. Accurate disease staging is important for the selection of patients eligible for surgery with curative intent and reduces the number of futile mediastinoscopies and thoracotomies in patients with inoperable advanced stage tumours (47). ^{18}F-FDG PET/CT imaging plays an important role in clinical decision making and is indicated for any patient eligible for curative therapy (4). In patients with (suspected) lung cancer, evaluation is focused on (1) the primary lung tumour(s), (2) intrathoracic lymph nodes and (3) regions where metastases of lung cancer predominantly occur. Also, for patients with unresectable disease treated with chemo(radiation) therapy, ^{18}F-FDG PET/CT imaging may be useful for treatment response evaluation (48), which includes shrinkage of disease burden (morphological as well as metabolic) and radiation-induced inflammation and fibrosis.

CT imaging used to be the gold standard in imaging for the detection and staging of lung tumours (49). However, while CT provided anatomical and morphological information about suspected long tumours, its ability to distinguish between benign and malignant lesions is limited (49). PET imaging provides complimentary metabolic information, which benefits a more accurate characterization of pulmonary lesions (50). Combined ^{18}F-FDG PET/CT imaging is preferred over diagnostic CT imaging alone for a high sensitivity for the detection of pulmonary lesions (90%), hilar and mediastinal lymph nodes (74-85%) and distant metastases (93%) (51). Note that the CT scan that is acquired as a part of ^{18}F-FDG PET/CT imaging is a low dose CT used for attenuation correction and anatomical matching and not a diagnostic CT scan. In 10-20% of patients with NSCLC, ^{18}F-FDG PET/CT detects unexpected distant metastasis (52). ^{18}F-FDG PET/CT imaging focuses on the investigation of abnormalities in the contralateral lung, liver, adrenal glands and bone. Due to the high background uptake in the brain, ^{18}F-FDG PET/CT imaging is not suitable for the detection of brain metastases. Therefore, magnetic resonance imaging (MRI) is the modality of choice (4). For the detection of bone metastases, ^{18}F-FDG PET/CT is the imaging method of choice over bone scintigraphy with a sensitivity, specificity, accuracy and prognostic value of >90% (4). The accuracy for the detection of adrenal gland metastases approximates 100% in lesions with a minimal diameter larger than 15 mm (52). From an imaging standpoint, detection of liver metastases is least challenging, since the liver is rarely the only site affected. Thus, for the detection of liver metastases, ^{18}F-FDG PET/CT imaging did not show added value over diagnostic CT and MRI (4).

It has to be noted that non-malignant inflammatory diseases can also show ^{18}F-FDG uptake. Also, ^{18}F-FDG uptake in lymph nodes is not only seen in metastatic lymph nodes, but also as a reaction of inflammation. To avoid false-positive results, in case of an enlarged or ^{18}F-FDG-positive lymph or other suspect ^{18}F-FDG uptake, tissue sampling is mandatory (4).

Limitations of ^{18}F-FDG PET/CT imaging in lung cancer are the spatial resolution of PET and respiratory motion artefacts. The limited spatial resolution of PET and the voxel grid result in the partial volume effect (17), which is explained in section 1.2. The PVE plays a role in smaller intrabronchial lesions and in lymph nodes with a diameter smaller than approximately 1 cm. The PVE leads to an underestimation of the uptake of the and an overestimation of the tumour volume. Also, respiratory motion artefacts are mostly reflected in smaller (and peripherally situated) intrabronchial lesions. Since the acquisition time of PET imaging is several minutes per bed position, during which the patient is instructed to breathe freely, images are averaged over several breathing cycles. This results in a substantial underestimation of tracer uptake in a lesion and an overestimation of the volume (53).

STATIC, PARAMETRIC AND DYNAMIC RADIOMICS FROM ^{18}F-FDG PET IN NON-SMALL CELL LUNG CANCER

W.A. Noortman, D. Vriens, C.H. Slump, J. Bussink, T.W.H. Meijer, L.F. de Geus-Oei, F.H.P. van Velden

Abstract

Purpose: The aim of this study was to assess prognostic and predictive abilities of static, parametric and, as a proof of concept, dynamic GLCM radiomic features derived from 2-^{18}F-fluoro-2-deoxy-D-glucose (^{18}F-FDG) positron emission tomography (PET) in comparison to traditional quantitative PET measures in non-small cell lung cancer (NSCLC).

Methods: Patients with newly diagnosed or suspected NSCLC of stage IB to stage IIIA underwent dynamic ^{18}F-FDG PET combined with computed tomography (CT) using the Biograph Duo or the Biograph 40 mCT (Siemens Healthineers, Erlangen, Germany). The final time frame (50–60 minutes after injection) was used as static ^{18}F-FDG PET scan. Parametric glucose metabolic rate (MR_{glc}) images were created based on tissue- and blood time-activity concentration curves using Patlak linearization, with data acquired between 15 and 60 minutes normalised Patlak-time. The dynamic images consisted of 16 frames of 150 s acquired between 10 and 50 minutes after injection. Volumes of interest (VOI) were drawn using a fuzzy locally adaptive Bayesian (FLAB) algorithm on the static and parametric images. Images and VOIs of both were interpolated using nearest neighbour interpolation with alignment of the grid centres to isotropic voxels of the maximum voxel dimension, leading to isotropic voxels with a dimension of $3.38 \times 3.38 \times 3.38\ mm^3$. Radiomic feature extraction of the static- and parametric images was performed using PyRadiomics 2.0 and for both scans 105 features were extracted. Radiomic features of the dynamic frames was performed PyRadiomics 1.3 and 22 grey level cooccurrence matrix features were extracted. Unique radiomic features were identified with correlation clustering and principal component analysis (PCA). Selected features and traditional quantitative PET features were compared with histopathology using an independent sample t-test. Univariate and multivariate Cox regression analyses were used to correlate selected radiomic features, traditional quantitative PET features and clinical characteristics with disease-free survival (DFS), disease-specific survival (DSS) and overall survival (OS). Survival curves were estimated using Kaplan-Meier analysis. Differences in feature set between scanners were assessed using logistic regression.

Results: Thirty-five lesions in 34 patients were included. PCA returned three radiomic features: the metabolic tumour volume (MTV), the GLCM maximum probability ($GLCM_{\max prob}$) and the GLCM sum of squares ($GLCM_{sum\ sqrs}$). $GLCM_{\max prob}$, $GLCM_{sum\ sqrs}$ and SUV_{max} showed significant differences between histological subtypes of NSCLC. Cox regression analysis did not show significant associations between selected radiomic features and survival outcome measures. Kaplan-Meier survival curves for features dichotomised at the median showed separations between the low and the high group, but log rank statistics were insignificant.

Conclusion: GLCM features contain limited additional information compared to static radiomic features, but were not selected in PCA. Parametric features did not contain additional information over static features. Selected static GLCM features implied that SCC show more heterogeneous uptake patterns than AC. Selected features showed clearer, but insignificant, separations in survival curves compared to traditional quantitative PET measures, indicating that image data contain more information about tumour biology than meets the eye. Also, a trend of higher heterogeneity in tracer uptake was seen in patients with a bad prognosis. Cox regression analysis showed that clinical characteristics were superior to radiomic features for the prediction of survival in patients with stage IB-IIIA NSCLC treated with a resection.

Introduction

In personalized medicine, medical decisions and interventions are tailored to the needs of the individual patient based on their predicted response or risk of disease (54). In oncology, the choice of therapy for a patient is mostly based upon molecular characterisation of the tissue, for which biopsy is the gold standard (6). However, biopsy comes with the risk of a sampling error, since only a small fraction of a heterogeneous tumour is sampled or the tumour is entirely missed (55). This might lead to misinterpretations. The problems related to biopsies might be addressed by less invasive medical imaging, which is routine clinical practice for diagnosis and staging in oncology.

Medical imaging can, unlike biopsies, provide information about the entire tumour phenotype, including intra-tumour heterogeneity (35). The extraction of these quantitative data from standard medical imaging is studied in the field of radiomics (33). This field aims to find stable and clinically relevant image-derived biomarkers or radiomic features that provide a non-invasive way of quantifying and monitoring tumour characteristics in clinical practice (35). Radiomic feature extraction is performed in scans from computer tomography (CT), magnetic resonance imaging (MRI) and positron emission tomography (PET). Usually, these studies are used to assess intensity, size and shape of tumours. However, imaging information is much richer. The goal of radiomics is to extract quantitative features, describing texture, intensity, morphological complexity and intratumour heterogeneity (33). Several studies showed the prognostic or predictive abilities of radiomic features derived from MRI (56, 57), CT (40, 58-60) and PET (36, 37, 39, 61) in different tumour types. They illustrated the discriminating capabilities of radiomic features for the stratification of histology, tumour grades or stages and clinical outcome (35).

In PET imaging, some quantitative image-derived features are used: the standard uptake value (SUV), metabolic tumour volume (MTV) and the total lesion glycolysis (TLG) (21). However, these features do not contain information related to the tracer uptake heterogeneity. Biologically, heterogeneity of the microenvironment of the tumour might be reflected in medical images and provide information about cellular density, proliferation, angiogenesis, hypoxia, receptor expression, necrosis and fibrosis (27). These biological processes are expressed in the spatial distribution of radiotracer uptake (28).

Most state-of-the-art radiomic features describe tracer uptake heterogeneity in a static image, but do not take into account tracer uptake heterogeneity over time, while this might also contain information about tumour biology. Meijer et al. found a difference in maximum standard uptake value ($SUVmax$) between non-small cell lung cancer (NSCLC) adenocarcinomas and squamous cell carcinomas (62), which might indicate differences in perfusion and uptake between histological subtypes. These differences might be reflected in temporal tracer uptake heterogeneity.

Research into radiomics in the temporal domain is limited. There are some studies that apply texture analysis on parametric images in MRI (63) and PET (64), but these are based upon 3D images created with pharmacokinetic modelling and do not assess time frames as the fourth dimension. Woods et al. investigated the use of 4D texture analysis, with time as the fourth dimension, to distinguish between non-malignant and malignant tissues in dynamic contrast-enhanced (DCE) MRI (65). This study showed promising results, but textures were calculated within a small window (66), instead of the lesion as a whole. Also, interchangeability of spatial and temporal dimensions was assumed. However, causality is a condition in the temporal dimension, while it is not in the spatial dimensions (67).

Another approach was found within proteomics, the field that studies proteins. Hu et al. studied the application of temporal texture features for the analysis of subcellular locations in time series fluorescence microscope images (68). They investigated the original 13 Haralick grey level co-occurrence (GLCM) features in the temporal domain. Originally, the Haralick GLCM features were developed for object identification in 2D, expressing combinations of grey levels of neighbouring pixels (69). In the dynamic approach of Hu et al., GLCMs were calculated for adjacent voxels in

time. This approach might be useful for the quantification of tracer uptake heterogeneity over time in PET imaging. To the best of our knowledge, there are no papers describing radiomics in the temporal dimension. In this study, we have evaluated the extension of radiomic feature set to the temporal domain, since tracer uptake heterogeneity in the time dimension might express other aspects of tumour biology than radiomics extracted from a static image. In this study, the added value of the dynamic GLCM features in the radiomic feature set was evaluated.

The aim of this study was to assess prognostic and predictive abilities of static, parametric and, as a proof of concept, dynamic GLCM radiomic features derived from 2-^{18}F-fluoro-2-deoxy-D-glucose (^{18}F-FDG) positron emission tomography (PET) in comparison to traditional quantitative PET measures in non-small cell lung cancer (NSCLC).

Materials and methods

Patient data

In this study, data of patients with newly diagnosed or suspected NSCLC of stage IB to stage IIIA (according to the TNM 7th edition), who underwent primary resection in the Radboud University Medical Center between 2009 and 2014 (62), were reanalysed. Routine staging was performed using contrast enhanced CT of the chest and/or upper abdomen and ^{18}F-FDG PET/CT with additional histologic staging of the mediastinum or other sites suspicious for cancer when necessary. Only tumours with a diameter larger than 30 mm were included, to minimize the partial volume effect. Thirty-five lesions in 34 patients were included (one patient had two synchronous primary NSCLCs (AC and SCC)). One patient from the original publication was excluded, because no surgery was performed for stage IV findings perioperatively (62).

Patients underwent a dynamic ^{18}F-FDG PET/CT scan within 7 days of surgery using either the Biograph Duo (n = 21) or Biograph 40 mCT with TrueV z-axis gantry extension (n = 17) (Siemens Healthineers, Erlangen, Germany). Patients were instructed to fast for at least 6 hours before imaging, were orally hydrated with 500 mL of water and euglycemia was validated. Patients underwent scanning in the supine position, with the tumour centrally located in the field of view (FOV), the axial length ranging from 162 to 216 mm, for the Biograph Duo and the Biograph 40 mCT, respectively. Prior to the PET scan, a low-dose breath-hold CT scan (Biograph Duo: 40 mAs, 130 kVp, Biograph 40 mCT: 50 mAs, 100 kVp) was performed for PET attenuation correction and anatomic matching. Free-breathing PET images were acquired at a single bed position for 60 minutes. A standardized infusion of approximately 3.45 MBq of ^{18}F-FDG per kilogram of body weight was administered intravenously in an intracubital vein contralateral to the tumour, directly after the start of infusion.

PET data from both scanners were reconstructed into 70 time frames for pharmacokinetic modelling: 20 × 5 seconds, 5 × 10 seconds, 10 × 15 seconds, 10 × 30 seconds, 16 × 75 seconds, 8 × 150 seconds and 1 × 600 seconds. The final time frame (50–60 minutes after injection) was used as static ^{18}F-FDG PET scan. Images acquired with the Biograph Duo PET/CT scanner were iteratively reconstructed with 256×256 matrices by means of ordered-subsets expectation maximization with four iterations and 16 subsets, followed by post smoothing using a 5-mm full-width-at-half-maximum (FWHM) Gaussian filter. The voxel size was 2.56×2.56×3.38 mm^3. Reconstruction parameters for the Biograph 40 mCT scans were three iterations and 21 subsets, 512×512 matrices, taking to account the time of flight information, with resolution modelling (i.e. point spread function–based reconstruction), followed by post smoothing using a 3-mm FWHM Gaussian filter, and. The voxel size was 1.59×1.59×2.03 mm^3. The acquisition and reconstruction of scans from both scanners were performed according to the first edition of the European Association of Nuclear Medicine (EANM) guidelines for tumour imaging using ^{18}F-FDG PET/CT (70). Acquisition and reconstruction parameters also met the criteria of the second edition, except for the reconstructed voxel size, which should be within 3.0-4.0 mm in any direction (20).

Parametric glucose metabolic rate (MR_{glc}) images were created based on tissue- and blood time-activity concentration curves using Patlak linearization, with data acquired between 15 and 60 minutes normalised Patlak-time. The image-derived input function (IDIF) was based on a 10 mL VOI of the descending aorta on which endothelial wall and calcifications were excluded to identify only blood, drawn on the images obtained during the first 60 seconds. Glucose metabolic rate was calculated as $K_i = \frac{C_{plasma,glc}}{LC_{FDG}}$ assuming $V_b = 0$, with a lumped constant (LC_{FDG}) of 1 and K_i the ^{18}F-FDG influx constant, or the slope of the Patlak plot, and $C_{plasma,glc}$, the plasma glucose concentration (62).

For the dynamic radiomic feature extraction the 16 × 75 s and the 8 × 150 s frames obtained between 10 and 50 minutes after injection of ^{18}F-FDG, are used. These frames were chosen, since the equilibrium between perfusion and uptake of ^{18}F-FDG starts after 10-15 minutes of Patlak time, which corresponds with 10-15 minutes in real time (21). The 16 75 seconds frames were combined per two frames, so that 16 frames of equal acquisition length (150 s) were assessed.

Image analysis

Image acquisition, pre-processing and radiomic feature extraction was reported according to the Image Biomarker Standardisation Initiative reporting guidelines. These guidelines can be found in appendix 1, which also contains figure 7, illustrating the pipeline for radiomic feature extraction. An overview of the procedure can be found in appendix 2 for the static and parametric images and in appendix 3 for the dynamic images.

VOIs

Volumes of interest (VOI) of the tumour in both the static and parametric PET scan were drawn in the source study using a fuzzy locally adaptive Bayesian (FLAB) algorithm (62), excluding ^{18}F-FDG-avid nontumour tissue by drawing an oversized container around the tumour and surrounding tissue by a radiation therapist under supervision of an experienced nuclear medicine physician. It was found that the FLAB-based segmentation volumes corresponded better with pathology volumes compared to uptake-based relative threshold delineation (ICC of 0.72 vs ICC of 0.29-0.45). The FLAB algorithm is based on the probability that a voxel belongs to the tumour tissue or background tissue based on the intensity of the voxels in various regions of the image, as well as the spatial correlation with neighbouring voxels (71). The FLAB segmentations of the static PET were used as VOIs for all dynamic frames.

Interpolation

Interpolation of the image and the VOI was performed to compare images from the different scanners. Isotropic voxels were preferred, so that texture features were calculated in a rotationally invariant volume (41). Within radiomics, there are no clear indications whether up-sampling or down-sampling schemes are preferable (41). A down-sampling approach was chosen, so that no artificial information was created. The slice thickness of the Biograph Duo had the largest voxel dimension (3.38 mm). Therefore, all images were scaled to this voxel dimension, i.e. 3.38 × 3.38 × 3.38 mm^3. The original grid and the interpolation grid were aligned by centre (41) and trilinear interpolation was performed (41). Image interpolation was performed using MATLAB 2017b (Mathworks, Natick, Massachusetts). The interpolation script for the static and parametric scans can be found in appendix 2. Also, the frames of the dynamic scans were interpolated and afterwards combined in order to create a 4D volume. The MATLAB script for this procedure can be found in appendix 3. The maximum standardized uptake value ($SUVmax$) values within the VOI of the interpolated scan were compared to the original study (62).

Static and parametric radiomic features

Radiomic feature extraction of the static and parametric images was performed using PyRadiomics version 2.0 (72) in Python 3.6 (Python Software Foundation, Wilmington, Delaware). For every VOI, 104 intensity -, shape – and texture features were calculated. Additionally, the total lesion glycolysis (TLG) was added to the feature set, since this classic feature was not included in the PyRadiomics package. The TLG was calculated by multiplying the mean standardized uptake

value (SUV_{mean}) and the metabolic tumour volume (MTV). Image normalization and distance weighting were not applied (41). Grey level co-occurrence matrix features (GLCM) were assessed in two directions per angle, taking into account rotational invariance (41). Features were calculated for the GLCM for the thirteen different angles combined to a 3D volume (41).

For the extraction of texture features, grey value discretization was performed using a fixed bin width. Leijenaar et al. compared resampling with a fixed number of bins and resampling with a fixed bin width and found that quantification of radiomic features were more robust to a change in bin size than to a change in the number of bins (46). They concluded that grey value discretization using a fixed bin width may be more appropriate for clinical studies, because a fixed number of bins implicitly assumes that the images of all patients have the same range of standardized uptake values (SUV). For SUV-based images, a bin width of 0.5 g/mL has been described in literature (46). To the best of our knowledge, optimal bin widths for parametric PET images are not known. Therefore, the bin width was determined according to the Freedman-Diaconis rule (73), stated in equation 12:

$$bin\ size = 2 \cdot IQR \cdot N^{-1/3} \quad [12]$$

with IQR the mean interquartile range within the VOI and N the number of voxels. The values for IQR and number of voxels were the mean values of all included tumours. The use of this formula resulted in a bin size of 0.55 g/mL for the static images and a bin width of 1.8E-08 mol/ml/min for the parametric images.

Dynamic radiomic features

Dynamic grey level co-occurrence matrix (GLCM) extraction was performed using in house software based on PyRadiomics version 1.3 (72) in Python 3.6 (Python Software Foundation, Wilmington, Delaware). In this version of PyRadiomics the grey level cooccurrence matrix (GLCM) was calculated in Python instead of in C, like in more recent versions of PyRadiomics. A GLCM expresses how combinations of discretised grey levels of neighbouring voxels are distributed along one of the image directions (41). In 3D, these image directions correspond to the 26 directly neighbouring voxels, calculated with equation 13:

$$N_{vox} = 3^{N_{dim}} - 1 \quad [13]$$

with N_{vox} the number of directly neighbouring voxels and N_{dim} the number of dimensions. The GLCM is calculated for the 13 unique direction vectors corresponding to the 26 neighbouring within a neighbourhood volume for distance 1 i.e. (0; 0; 1), (0; 1; 0), (1; 0; 0), (0; 1; 1), (0; 1;-1), (1; 0; 1), (1; 0;-1), (1; 1; 0), (1;-1; 0), (1; 1; 1), (1; 1;-1), (1;-1; 1) and (1;-1;-1). Since a 3D volume should be rotationally invariant, co-occurrences are calculated in two directions per angle. Mathematically, this is done by creating a symmetrical matrix.

The dynamic GLCM is calculated differently, since spatial and temporal dimensions are not interchangeable (67). The GLCM was not calculated for the 40 unique direction vectors of the 80 directly neighbouring voxels calculated with equation 13 (i.e. (0; 0; 0; 1), (0; 0; 1; 0), etc), but only in the temporal direction, (0; 0; 0; 1). In this approach, neighbouring voxels are subsequent voxels in time and the GLCM indicates variation of voxel values over time. This approach is illustrated in figure 5. Since causality plays a role in the temporal domain (67), the value of a voxel is dependent of the value of the preceding voxel. Therefore, co-occurrences were calculated in only one direction per angle instead of two. For that reason, the matrix was not made symmetrical. Twenty-two dynamic GLCM features were calculated for the different angles combined to a 3D volume (41). The Python script for the dynamic GLCM extraction can be found in appendix 5.

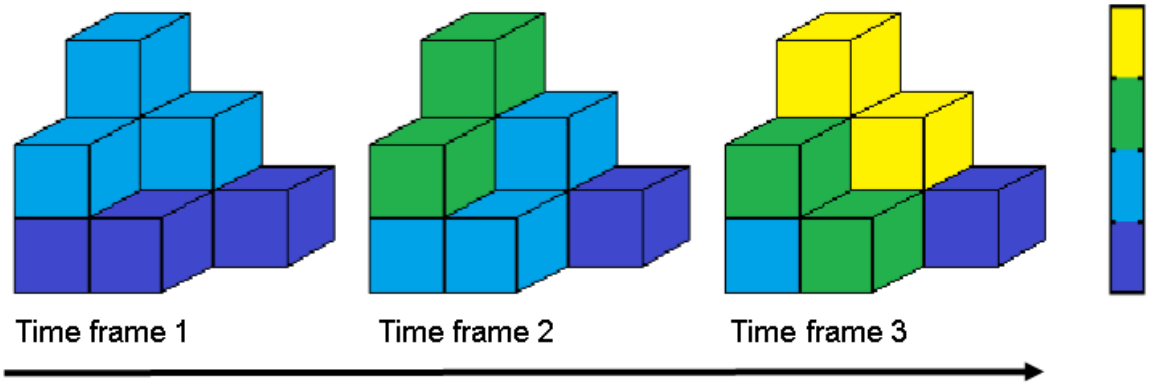


Figure 6: The dynamic grey level cooccurrence matrix (GLCM) are calculated in the temporal dimension (the different time frames). Values of subsequent voxels in time are compared: an increase of the value of the upper left voxel over the three time frames can be seen, while the value of the right voxel does not change.

Grey value discretization of the dynamic images was performed using a fixed bin width for the individual frames. For all frames the mean IQR in the population was calculated and the bin width was calculated with Freedman-Diaconis rule in equation 12.

Statistical analysis

Unsupervised feature selection

The radiomic features extracted from the static and parametric PET scan both account 105 features. There are 22 dynamic GLCM features. This leads to a total of 232 unique radiomic features. To decrease high dimensionality, which occurs when too many (multicollinear) features are introduced in a regression model, dimension reduction was performed. The same strategy for statistical analysis was applied as detailed in Collarino et al. (39). Groups of radiomic features with high common variance ($r^2 > 0.56, r > 0.75$) were created using MATLAB 2017b (Mathworks, Natick, Massachusetts). From each group one representative feature was selected, based on common features from literature. When both static features and parametric or dynamic GLCM features were present in a group, static PET features were preferred, since a static PET scans are routine in clinical practice, while dynamic scans are not. Also, common variance between corresponding static- and parametric features and between static features and dynamic features were assessed.

Principal component analysis (PCA) was performed to further reduce the dimensionality of the selected features in SPSS 23 (IBM Statistics, Chicago, IL). PCA with an orthogonal rotation (varimax with Kaiser normalisation) was performed so that the first principal component explained the largest possible variance in the dataset; succeeding components explained the highest variance in orthogonal directions. The sampling adequacy for each feature and for the complete model had to be higher than 0.5, which was determined by the Kaiser-Meier-Olkin (KMO) measure (74). Only components with eigenvalues above Kaiser's criterion of 1 were included for further analysis (75), so that a component does not explain less variance than an individual variable.

Correlation with histopathology and clinical data

Radiomic features selected with PCA and traditional quantitative PET features (SUV_{max} and TLG) were correlated to histopathology (squamous cell carcinoma or adenocarcinoma), tumour differentiation and TNM stage (7th edition). The features were tested for (log)normality using skewness and kurtosis parameters and standard errors. Statistical difference between histology were analysed using Mann-Whitney U test or independent samples t-test depending on (log)normality. p-values less than 0.05 were considered statistically significant. Correlations were assessed in SPSS 23 (IBM Statistics, Chicago, IL).

Correlation with clinical outcome

Correlation of selected radiomic features, traditional quantitative PET features, clinical characteristics and histology with disease-free survival (DFS), disease-specific survival (DSS) and overall survival (OS) was performed with univariate Cox regression analysis. DFS was defined as the time between surgery and recurrence of the disease. DSS was defined as the time between surgery and patient death as a result of the disease. OS was defined as the time between surgery and patient death. For all survival measures, censoring was taking into account. Candidate covariates for multivariate regression were identified from univariate Cox regression analysis based on significance level ($p \leq 0.2$). Multivariate Cox regression analyses were performed with iterative forward and backward selection based of the candidate covariates based on the likelihood-ratio. A maximum of 3 features was included in the final model, since for linear models a minimum of 10 to 15 observations per predictor variable will result in a well-fitting model (76, 77). Also, survival curves of DFS, DSS and OS were estimated using Kaplan-Meier analysis for the selected static/parametric and dynamic radiomic features dichotomized at the median. Survival curves were compared using Log-Rank statistics. Correlations were assessed in SPSS 23 (IBM Statistics, Chicago, IL).

Differences between scanners

To test the dependence of the results on the type of scanner (Biograph Duo or Biograph 40 mCT), the association between the used scanner and the static feature set (N = 105), the parametric feature set (N = 105) the dynamic GLCM feature set (N = 22) and the complete feature set (N = 232) was assessed. This association was tested using a logistic regression in which the response variable was modelled as a function of the covariates. This analysis was performed in the GlobalTest package (78) using R Statistical Software version 3.3.3 (R Foundation for Statistical Computing, Vienna, Austria).

Results

Clinical characteristics of 35 included NSCLC lesions in 34 patients can be found in table 3. The maximum standardized uptake value ($SUVmax$) within the volume of interest (VOI) of the interpolated scans were compared to values in the original study. The average $SUVmax$ value in the VOI was 3.87% lower than in the original dataset, which can be explained by down-sampling interpolation.

Feature selection

Correlation clustering on the complete feature set of 232 features was performed. The common variance matrix is shown in figure 7. Parametric features showed to be highly similar to static radiomic features: only 6 out of 105 corresponding features showed a common variance below 0.56. Five out of 22 dynamic GLCM features showed a maximum common variance with any static feature below 0.56, indicating additional information over the static features. These were the correlation, the informational measure of correlation 1 and 2, the inverse difference moment normalized and the inverse difference normalized.

Correlation clustering of this set of 232 features with a Pearson correlation above 0.75 led to 36 groups of features. From each groups, one feature was selected. Principal component analysis was conducted on the 36 selected features with an orthogonal rotation (varimax with Kaiser normalization). The Kaiser-Meyer-Olkin sampling adequacy measure was ‘mediocre’ with a value of 0.697 (74). Eighteen features were rejected for individual sampling adequacy measures below 0.5 (75), leading to individual sampling adequacy measures greater than 0.532 In the final model. Three components had eigenvalues above Kaiser’s criterion of 1. In combination, the three components explained 81.4% of the variance. The features that corresponded best to the components were the metabolic tumour volume (MTV), the static PET GLCM maximum probability ($GLCM_{\max prob}$)and the static PET GLCM sum of squares $GLCM_{sum\ sqrs}$. Note that the MTV is also a traditional quantitative PET feature in clinical practice. Additional information about the correlation clustering and PCA can be found in appendix 6.

Table 3: Clinical characteristics of 35 NSCLC lesions in 34 patients.

Characteristic	Value
Age (years), mean (range)	64 (44-80)
Gender (M/F)	24/11
Scanner	
Siemens Biograph Duo	20
Siemens Biograph mCT	15
$SUVmax$**, mean (range)**	14.96 (5.62-30.98)
Histology	
Squamous cell carcinoma	20
Adenocarcinoma	12
Other	3
Differentiation	
Well differentiated	1
Moderately differentiated	13
Poorly differentiated	21
TNM stage (7th edition)	
Stage I	8
Stage II	20
Stage III	7
Surgical margins	
Free	32
Not free	2
Rx	1
Pleural invasion	
No	21
Yes	14
Adjuvant chemotherapy	
Yes	17
No	17
Unknown	1
Adjuvant radiation therapy	
Yes	4
No	31

Correlation with histopathology

Features selected with PCA and traditional quantitative PET features were correlated with histopathological characteristics. Radiomic features were tested for normality based on skewness and kurtosis. All features had a normal distribution after log-transformation. Statistical differences were analysed using the independent samples t-test. Static $GLCM_{\max prob}$ and static $GLCM_{sum\ sqrs}$ showed a statistically significant difference ($p = 0.008$ and $p = 0.001$, respectively) between histological subtypes. Also, the traditional features maximum standardized uptake value (SUV_{max}) showed a significant difference between histological subtypes ($p = 0.001$). The MTV and the traditional feature total lesion glycolysis (TLG) both showed significant differences between TNM stage ($p = 0.010$ and $p = 0.013$, respectively). Associations can be found in table 4.

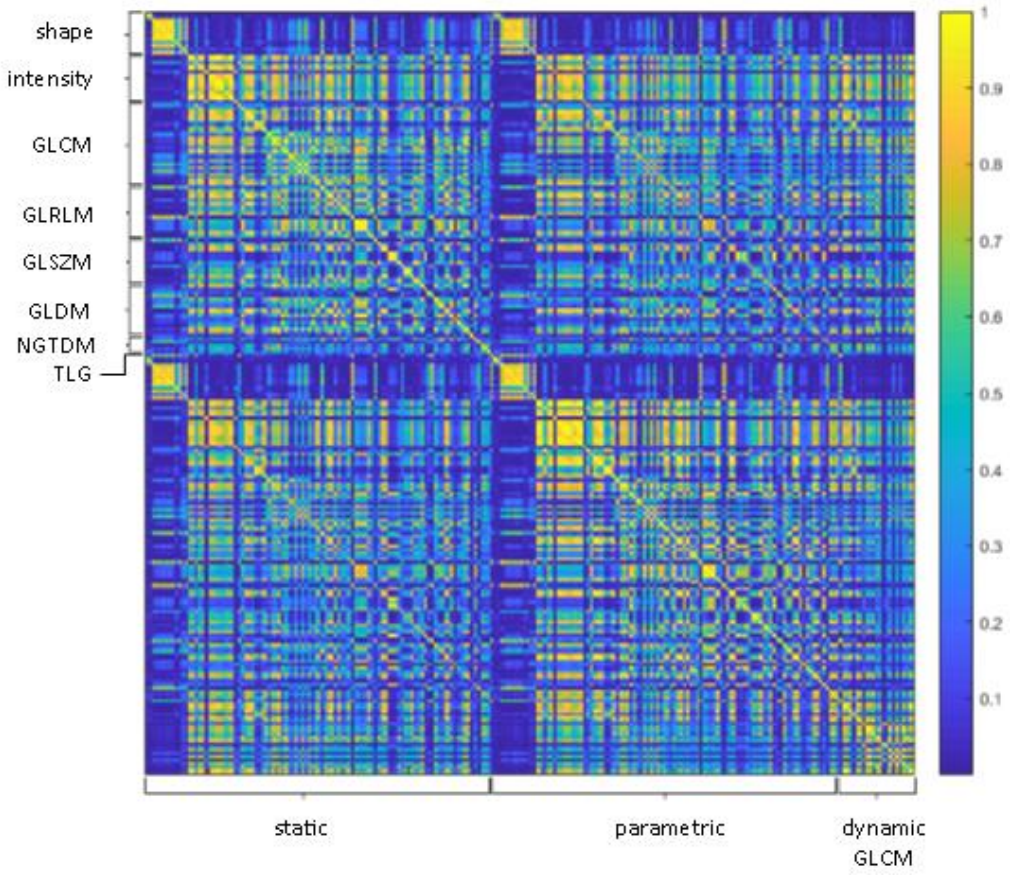


Figure 7: Variance matrix (r^2) of the complete feature set. On the x-axis the sets from the different scans (static, parametric and dynamic GLCM) are indicated. On the y-axis, the different groups are indicated for the static features. The same features were extracted from the parametric images and are stated in the same order. GLCM: grey level cooccurrence matrix, GLRLM: grey level run length matrix, GLSZM: grey level size zone matrix, GLDM: grey level dependence matrix, NGTDM: neighbouring grey tone difference matrix, TLG: total lesion glycolysis.

Correlation with clinical outcome

No significant univariate correlation between the selected radiomic features and traditional quantitative PET features and OS, DSS and DFS was found. In multivariate Cox regression, age of diagnosis was found to be an independent prognostic factor of OS and DFS and pleural ingrowth was found to be an independent prognostic factor of DSS. Hazard ratios and p-values of univariate and multivariate Cox regression of clinical characteristics and selected radiomic features for OS, DSS and DFS are shown in table 5. Also, the estimated Kaplan-Meier survival curves for survival outcome measure and the selected radiomic features dichotomized at the median were not significantly different. Although non-significant, survival curves of the features $GLCM_{\max prob}$ and $GLCM_{sum\ sqrs}$ showed a clear separation between the low and the high group. A high $GLCM_{\max prob}$ and a low $GLCM_{sum\ sqrs}$ were associated with a better prognosis. Kaplan-Meier survival curves are shown in figure 8. Sensitivity and specificity of the models were not calculated, since follow-up was too short to reach the median disease-free survival.

Differences between scanners

No significant difference between the two scanners (Biograph Duo and Biograph 40 mCT) were found for the static features ($p = 0.074$), parametric features ($p = 0.156$), dynamic GLCM features ($p = 0.061$) and all features together ($p = 0.093$).

Table 4: Association between histopathological characteristics and selected radiomic features and traditional quantitative pet features calculated with the independent samples t-test after testing for (log-)normality. Table shows means and ranges and p-values. MTV: metabolic active tumour volume, $GLCM_{\mathbf{max}\,prob}$: grey level cooccurrence matrix maximum probability, $GLCM_{sum\ sqrs}$: grey level cooccurrence matrix sum of squares, $SUVmax$: maximum standardized uptake value, TLG: total lesion glycolysis.

	Histopathological subtype			Tumour differentiation			TNM stage (7th edition)		
	AC	SCC	p-value	Well/ moderate	Poor	p-value	I-IIa	IIb-III	p-value
Selected radiomic features									
$\boldsymbol{MTV}$ $(\mathbf{cm}^3)$	38.3 (7.9 – 153.5)	35.3 (11.5 – 180.9)	0.777	36.6 (7.9 – 153.5)	38.1 (11.5 – 180.9)	0.891	26.8 (7.9 – 100.5)	53.6 (13.5 – 180.9)	**0.010**
$\boldsymbol{GLCM}_{\mathbf{max}\,prob}$	0.037 (0.011 – 0.249)	0.017 (0.006 – 0.080)	**0.008**	0.026 (0.008 – 0.249)	0.022 (0.006 – 0.080)	0.531	0.024 (0.006 – 0.249)	0.023 (0.008 – 0.108)	0.810
$\boldsymbol{GLCM}_{\boldsymbol{sum\ sqrs}}$	7.15 (2.06 – 45.21)	17.99 (6.67 – 76.25)	**0.001**	11.06 (2.06 – 76.25)	13.32 (4.46 – 57.46)	0.512	12.27 (2.31 – 57.46)	12.47 (2.06 – 76.25)	0.955
Traditional quantitative PET features									
$\boldsymbol{SUVmax}$ $(\mathbf{g/mL})$	10.41 (5.62 – 27.62)	16.75 (9.05 – 30.98)	**0.001**	12.95 (5.62 – 29.77)	14.34 (8.09 – 30.98)	0.487	13.71 (5.94 – 30.98)	13.81 (5.62 – 29.77)	0.963
$\boldsymbol{TLG}$ $(\boldsymbol{g})$	222.5 (60.9 – 630.5)	322.8 (100.9 – 1669.3)	0.196	260.8 (60.9 – 1233.3)	302134 (93.0 – 1669.3)	0.613	205.3 (60.9 – 886.0)	403.0 (97.9 – 1669.3)	**0.013**

Table 5: Univariate and multivariate Cox regression analysis of clinical characteristics, traditional quantitative PET features and selected radiomic features for overall survival (OS), disease-free survival (DFS) and disease-specific survival (DSS). Characteristics and features with a p-value < 0.20 in univariate analysis, were selected for multivariate analysis. MTV: metabolic active tumour volume, $GLCM_{\mathbf{max}\,prob}$: grey level cooccurrence matrix maximum probability, $GLCM_{sum\ sqrs}$: grey level cooccurrence matrix sum of squares, $SUVmax$: maximum standardized uptake value, TLG: total lesion glycolysis.

	OS		**DSS**		**DFS**	
Parameter	Hazard ratio (95% confidence interval)	p-value	Hazard ratio (95% confidence interval)	p-value	Hazard ratio (95% confidence interval)	p-value
Univariate Cox regression analysis						
Gender	2.000 (0.563 – 7.109)	0.284	1.575 (0.334 – 7.429)	0.566	2.549 (0.550 - 11.814)	0.232
Age of diagnosis	1.085 (1.011 – 1.164)	**0.023**	1.094 (1.006 – 1.190)	**0.037**	1.082 (1.001 – 1.171)	**0.048**

Table continues on the next page.

	OS		DSS		DFS	
Stadium						
IB	1	0.632	1	0.659	1	0.689
IIA	0.715 (0.159 – 3.204)		2.169 (0.242 – 19.438)		1.055 (0.193 – 5.771)	
IIB	1.486 (0.368 – 5.996)		3.449 (0.385 – 30.896)		1.600 (0.292 – 8.752)	
IIIA	0.661 (0.110 – 3.969)		1.142 (0.087 – 22.792)		0.428 (0.039 – 4.729)	
Pleural ingrowth	1.979 (0.712 – 5.502)	**0.190**	4.947 (1.259 – 19.436)	**0.022**	3.091 (0.901 – 10.608)	**0.073**
Negative resection margin	1.615 (0.202 – 12.925)	0.651	2.821 (0.332 – 23.957)	0.342	1.595 (0.201 – 12.671)	0.659
Adjuvant chemotherapy	0.395 (0.132 – 1.184)	**0.097**	0.549 (0.154 – 1.953)	0.354	0.484 (0.141 – 1.655)	0.247
Post-operative radiotherapy	1.145 (0.255 – 5.134)	0.859	1.682 (0.210 – 13.461)	0.624	0.865 (0.111 – 6.769)	0.890
Radiomic features						
MTV	1.000 (1.000 – 1.000)	0.665	1.000 (1.000 – 1.000)	0.819	1.000 (1.000 – 1.000)	0.727
$GLCM_{\max prob}$	0.039 (0 - 5018928189)	0.804	0.009 (0 - 4292555)	0.664	0.005 (0 – 914224)	0.580
$GLCM_{sum\ sqrs}$	1.003 (0.977 – 1.031)	0.807	1.012 (0.984 – 1.041)	0.395	1.016 (0.987 – 1.045)	0.281
Traditional quantitative PET features						
$SUVmax$	1.025 (0.951 – 1.104)	0.518	1.050 (0.966 – 1.142)	0.247	1.061 (0.970 – 1.159)	**0.194**
TLG	1.000 (1.000 – 1.000)	0.589	1.000 (1.000 – 1.000)	0.856	1.000 (1.000 – 1.000)	0.816

Multivariate Cox regression analysis

Iterative forward selection

Age of diagnosis	1.082 (1.007 – 1.163)	0.032			1.082 (1.001 – 1.171)	0.048
Pleural ingrowth			4.947 (1.259 – 19.436)	0.022		

Iterative backward selection

Age of diagnosis	1.082 (1.007 – 1.163)	0.032			1.082 (1.001 – 1.171)	0.048
Pleural ingrowth			4.947 (1.259 – 19.436)	0.022		

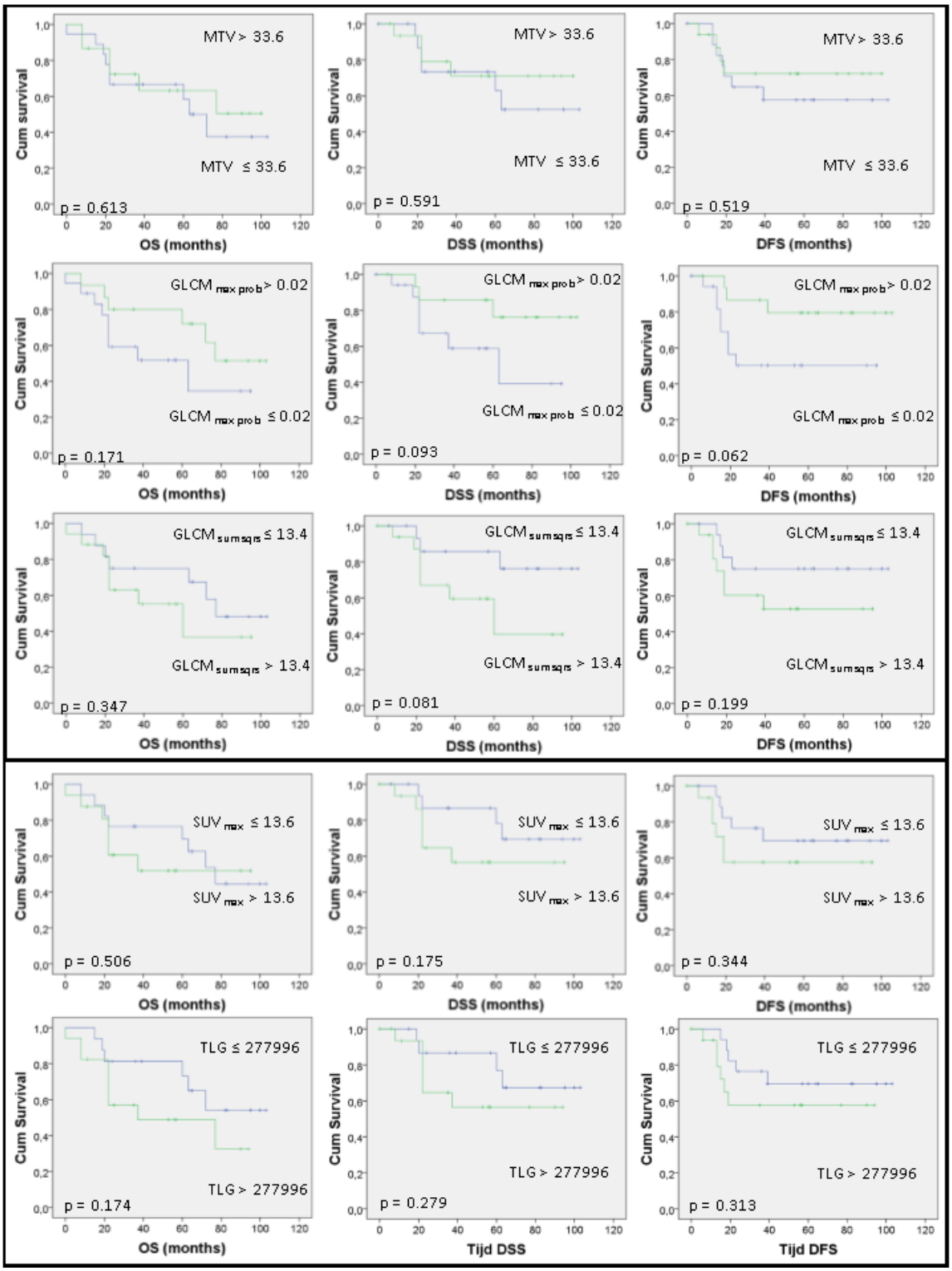


Figure 8: Kaplan Meier survival curves for overall survival (OS), disease-specific survival (DSS) and disease-free survival (DFS) for selected radiomic features (MTV: metabolic active tumour volume, $GLCM_{\mathbf{max}\,prob}$: grey level cooccurrence matrix maximum probability, $GLCM_{sum\ sqrs}$: grey level cooccurrence matrix sum of squares) and classic quantitative features ($SUVmax$: maximum standardized uptake value, TLG: total lesion glycolysis), dichotomised at the median, compared using Log-Rank statistics.

Discussion

In this study, prognostic and predictive abilities of static, parametric and, as a proof of concept, dynamic GLCM radiomic features derived from ^{18}F-FDG PET were assessed in comparison to traditional quantitative PET measures and clinical characteristics in non-small cell lung cancer (NSCLC). In the field of radiomics, the use of features describing tracer uptake heterogeneity for the non-invasive assessment of tumour biology is explored. However, to the best of our knowledge, there are no publications describing temporal changes in tracer uptake heterogeneity in PET, while this might express heterogeneity of ^{18}F-FDG over time. Therefore, this study investigated the added value of dynamic GLCM features and parametric radiomic features in addition to static radiomic features for the prediction of histological subtype and survival in patients with stage IB to IIIA NSCLC treated with surgical resection.

Unsupervised data reduction using principal component analysis returned the metabolic tumour volume (MTV), the static grey level cooccurrence matrix (GLCM) maximum probability ($GLCM_{\max prob}$) and the static GLCM sum of squares ($GLCM_{sum\ sqrs}$). Kaplan Meier survival curves show that a high $GLCM_{\max prob}$ and a low $GLCM_{sum\ sqrs}$ are associated with better survival, despite the difference between curves being insignificant. The $GLCM_{\max prob}$, also known as the joint maximum in literature, expresses the occurrences of the most predominant pair of neighbouring intensity values (41, 72). A high maximum probability in the GLCM indicates that there are many the same neighbouring intensity values, which suggests that the volume is relatively homogeneous. The $GLCM_{sum\ sqrs}$, also known as the joint variance, expresses the variance of the distribution of neighbouring intensity level values around the mean intensity value in the GLCM (72). A higher variance in the GLCM indicates more spread in the neighbouring intensity levels and therefore a more heterogeneous lesion. Both features indicate that a small difference in grey levels is associated with a better survival and more heterogeneity within the tumour is associated with bad prognosis. This corresponds with findings in literature that state that a higher heterogeneity is associated with a poorer survival in NSCLC (79-81). Kaplan-Meier survival analysis was also performed for the traditional quantitative PET features maximum standardized uptake value (SUV_{max}) and total lesion glycolysis (TLG), showing no significant differences between high and low values. Some separations between high and low SUV_{max} and TLG can be observed, but separations were less clear than for the selected radiomic features. This suggests that the static radiomic features $GLCM_{\max prob}$ and $GLCM_{sum\ sqrs}$ might be better predictors of survival than traditional quantitative PET measures. This indicates that the image data contain more information about tumour biology than meets the eye.

Also, the mean $GLCM_{\max prob}$ and mean $GLCM_{sum\ sqrs}$ were significantly different between squamous cell carcinomas (SCC) and adenocarcinomas (AC) ($p = 0.008$ and $p = 0.001$, respectively). A higher mean $GLCM_{\max prob}$ and lower mean $GLCM_{sum\ sqrs}$ were found in AC, which indicates a more homogeneous uptake pattern. This can be explained by a better vascularisation and perfusion of AC compared to SCC (82), resulting in a more homogeneous uptake. The differences in mean GLCM values for both AC and SCC cannot be used for the stratification of the individual patient, since overlap of the ranges of the values between the different tumour types can be observed. The mean of the traditional quantitative PET measure SUV_{max} is significantly higher in SCC than in AC ($p = 0.001$), which was already found by Meijer et al. (62). This corresponds with findings of Ha et al., showing that several radiomic features, among which the SUV_{max} and some GLCM features, are able to discriminate between AC and SCC (83). The selected radiomic feature MTV and the TLG both show significantly different means in TNM-stage (I-IIa vs IIb – III) ($p = 0.010$ and $p = 0.013$, respectively). Tumour diameter is a variable determining the T-classification in TNM-classification (10) and is therefore of influence in the stage of the tumour. The MTV is dependent on the tumour diameter and the TLG is derived from the MTV, which explains this difference.

Cox regression analysis showed no significant univariate correlation between the selected radiomic features and traditional quantitative PET features with OS, DSS and DFS. Significant correlations with survival were found for the clinical characteristics age of diagnosis, pleural ingrowth and adjuvant chemotherapy. This shows that in this dataset clinical characteristics are better predictors of patient outcome than radiomic features, selected using unsupervised data reduction, and traditional quantitative PET features in patients with stage IB to IIIA NSCLC treated with surgical resection. Some studies assess the predictive value of ^{18}F-FDG radiomics for survival in NSCLC, indicating the added value of the combination of radiomic features and clinical characteristics (84). However, comparison of these studies with our study is difficult, since most studies were conducted in other disease stages or treatment was different (36, 81, 85-87). Apostalova et al. showed that asphericity of the tumour volume and the clinical characteristic "primary surgical treatment' were significant independent predictors of disease-free survival and overall survival in 60 patients with newly diagnosed NSCLC of stage I-III in multivariate Cox regression, while traditional quantitative PET features were not (88). Asphericity was not included in PyRadiomics feature set, but it highly correlates to sphericity (72), which was deprecated in PCA for an insufficient individual sampling adequacy measure.

While static as well as parametric and dynamic radiomic features were included in principal component analysis, the three principal components turned out to be only static features. Parametric features showed to be highly similar to static radiomic features: only six out of 105 corresponding features showed a common variance below 0.56. In correlation clustering, parametric and dynamic features were exclusively selected when no static features were present in a group. Four groups with only parametric features were returned leading to four parametric features in PCA. Three out of four features were deprecated in PCA for an individual sampling adequacy measure below 0.5. The last feature did not correlate best to the three principal components. This agrees with findings in literature, where high correlations between nine radiomic features extracted from static and parametric images of 20 patients were found (64). This indicates that the additional information of radiomic features extracted from parametric PET images is limited.

Five out of 22 dynamic GLCM features showed a maximum common variance below 0.56 with any static feature in the dataset. Two out of these five features were selected in correlation clustering, since there were no static features present in their groups. Both features were deprecated in PCA for an individual sampling adequacy measure below 0.5. This suggests that the additional information of dynamic GLCM features is limited. This amplifies that there is no additional benefit from dynamic scanning over static scanning for ^{18}F-FDG, which is strengthened by the fact that dynamic ^{18}F-FDG PET imaging is not standard-of-care in patients with NSCLC.

However, the GLCM might not express all temporal heterogeneity expressed on dynamic PET. The GLCM was chosen to asses, as a proof on concept, the added value of radiomics in the fourth dimension, since it expresses combinations of grey levels of neighbouring pixels (69), in dynamic setting subsequent voxels in time. Other types of radiomic features might also contain additional information about tracer uptake heterogeneity. Intensity features over the different time frames express differences in intensity values. Shape features over time might express whether lesions show early uptake in the lesion as a whole or in subvolumes. Intensity- and shape feature values over the different time frames would lead to a curve from which ideally one value should be derived indicating the trend of the data points. Other dynamic texture features would also express changes in grey levels in subsequent voxels. The grey level run length matrix (GLRLM) features are somewhat similar to the GLCM features and count the number of the same grey levels in a row (72); in dynamic setting mapping rows of subsequent voxels. Grey level size zone matrices (GLSZM) quantify regions of pixels with the same voxel value (72). This features might be less useful in dynamic setting, since it would connect a 3D volume over the different time frames.

The methods used in PCA could have played a role in the deprecation of parametric and dynamic features. PCA tries to reduce the high-dimensional feature space into a meaningful representation (75). Multicollinearity of the features in PCA negatively influences the final result and therefore the input of PCA highly influences the outcome. Also, the order of removal of features with an insufficient individual sampling accuracy measure affects the final results. In this study, the general KMO value was 'mediocre' with a value of 0.697, which indicates a small sample size and a relatively large number of multicollinear features in the dataset (74). This might indicate that the final result of the PCA was influenced by multicollinearity (75). Multicollinearity is difficult to quantify, since in PCA four features with a correlation of 0.6 could lead to a worse analysis than two features with a correlation of 0.8 (75). Multicollinearity results in an overrepresentation of a certain feature type in the dataset, which increases the common variance in a certain direction. This effect could also have played a role in the deprecation of the dynamic radiomic features from PCA, since the dynamic features might have been underrepresented compared to the static features. Following the approach of Collarino et al., unsupervised feature reduction with PCA was preceded with correlation clustering (39), attempting to reduce multicollinearity by preselecting features based on correlation clustering of features with a common variance larger than 0.56. One feature from each cluster was selected, based on the experience and preferences of the authors. It has to be noted that while PCA has its limitations, it is currently considered one of the better options when it comes to feature reduction (45, 89).

Many studies have been performed applying radiomics in NSCLC (84); methodological considerations as well as publications reporting diagnostic, prognostic and predictive ability. However, it is difficult to draw conclusion from these results, since methods for image acquisition, reconstruction, pre-processing, feature extraction and statistical analysis vary highly or documentation is incomplete (84), which makes reproduction of the study impossible. This leads to rather preliminary results, only applicable for the specific population of a study, making it difficult to draw general conclusions. To improve the value of this study, methodological considerations were substantiating with findings in literature and using the guidelines of the Image Biomarker Standardisation Initiative (IBSI) (41). Also, validation of the model is important to investigate general applicability of the model (45). A limitation of this study is that validation was not performed, since the number of patients in this dataset was limited.

No significant differences between the features from the different scanners were found. Data acquisition took place using two PET scanners, the Siemens Biograph Duo and the Siemens Biograph 40 mCT. While scans from both scanners met the EANM guidelines for tumour PET imaging (20), differences in hardware and software of the scanners still lead to (small) differences in PET images. For the scans acquired with the Biograph mCT, time of flight (TOF) reconstruction was used, which improves the signal to noise ratio (SNR) and detectability of small lesions. Also, the resolution of the mCT was better than the resolution of the Duo, resulting in higher noise per voxel. Interpolation of the images of both scans to the largest voxel dimension was performed to have the same resolution, but noise levels still varied between both scans due to the difference in system sensitivities (90, 91). This indicates that the used interpolation is useful for harmonization of images acquired with different scanners.

Conclusion

This study shows that dynamic GLCM features contain limited additional information compared to static radiomic features in patients with stage IB-IIIA non-small cell lung carcinoma, but dynamic features were not selected in principal component analysis. Parametric features did not contain additional information over static radiomic features. Feature selection based on unsupervised feature reduction returned MTV, $GLCM_{\max prob}$ and $GLCM_{sum\ sqrs}$. Both GLCM features showed differences in values between the histological subtypes of NSCLC, implying that SCC show more heterogeneous uptake patterns. Selected features showed clearer, but insignificant, separations between high and low values in Kaplan-Meier analysis, compared to traditional quantitative PET measures, indicating that image data contain more information about tumour biology than meets

the eye. The features showing this trend suggest higher heterogeneity in tracer uptake of patients with a bad prognosis, which corresponds with findings in literature. Cox regression analysis did not show significant correlations between selected radiomic features and traditional PET features and survival, while clinical parameters did. This implies that clinical characteristics are superior to radiomic features for the prediction of survival in patients with stage IB-IIIA NSCLC treated with a resection. No significant differences between the features from the different scanners were found, indicating that interpolation is useful for the comparison of images from different scanners.

GENERAL DISCUSSION

The field of radiomics studies the extraction of quantitative features from medical imaging like CT, MRI and PET, based on the hypothesis that medical images contain quantifiable information about underlying tumour biology. The goal is to find stable and clinically relevant image-derived biomarkers or radiomic features that provide a non-invasive way of quantifying and monitoring tumour characteristics in clinical practice. Radiomics have the potential to improve knowledge in tumour biology and guide patient management, but the field faces several difficulties within different steps of the pipeline (figure 5). Some of these difficulties are pointed out in this discussion.

Images and feature extraction

Many studies have been performed applying radiomics derived from CT, MRI and PET in several pathological conditions (33). In PET imaging, heterogeneity of tracer uptake in the microenvironment of the tumour might be reflected in the images and provide information about cellular density, proliferation, angiogenesis, hypoxia, necrosis and fibrosis (27). However, it is not known how the tumour biology reflects in this heterogeneity. The features do no bear a direct relationship to the underlying cellular biology on a microscopic scale, since, especially in PET, the features relate to a relatively macroscopic scale based on the used voxel size (29). For instance, with an estimation of 10^8 tumour cells in 1 cm³ of tumour tissue (30), a voxel of 2x2x2 mm³ contains already 8×10^5 tumour cells and a voxel of 4x4x4 mm³ contains even 6.4×10^6 cells. To deepen understanding of the relation between tumour biology and radiomic features, it is encouraged to discuss the underlying biological process that causes differences in feature values (45). This further emphasizes the explorative character of radiomics.

Another difficulty within radiomics is that methods for image acquisition, reconstruction, pre-processing, feature extraction and used definitions vary highly or documentation is incomplete (84), which impedes drawing conclusions from these results. To attribute differences in radiomic features to tumour biology, it is important that data are homogeneous when it comes to acquisition and reconstruction (33). In PET imaging, the EANM guidelines are useful for standardisation (20). Also, reproducibility of features in quantitative image analysis is a major challenge (31). The Image Biomarker Standardisation Initiative (IBSI) aims to provide a common nomenclature and definitions for image biomarkers, benchmarks for image processing and feature extraction and reporting guidelines (41). Several pre-processing steps influence the feature values. In the first place, radiomic features are sensitive to changes in voxel size (92), which might be a reason to perform voxel interpolation, when comparing images from different scanners. ComBat, a post-reconstruction harmonization method that was recently proposed (93), might be useful for the this comparison, also facilitating multicentre radiomic studies. Secondly, the segmentation method influences radiomic features (94). In manual segmentation inter- and intra-observer variability play a role. Semi-automatic segmentation methods will show better repeatability, but feature values vary between segmentation methods: differences up to 51% between fixed threshold segmentation and fuzzy locally adaptive Bayesian (FLAB) were observed (94). Thirdly, the grey level discretization used for texture feature calculation influences feature values (46). Grey level discretization can be performed using a fixed number of bins or a fixed bin width, from which the second led to more robust features. The reporting guidelines can be used for documentation of the different steps concerning the feature extraction.

Feature reduction

So far, over 5000 radiomic features have been described in literature and this number is still increasing (33). In this study we investigated the added value of tracer uptake heterogeneity over time, even increasing the total amount of radiomic features. However, this large number of features compared to the number of subjects introduces the 'curse-of-dimensionality', a problem that arises

when analysing data in high-dimensional spaces (the hundreds of radiomic features). Thereby the generalisation performance of the model is negatively impacted; overfitting occurs, since the model is too specific for the training dataset (43).

To reduce the curse of dimensionality, a selection of features has to be made. Feature selection methods from statistics and machine learning are evaluated for their use in radiomics. A selection of features can be made using supervised and unsupervised methods (89, 95). Supervised approaches take into account outcome measures for feature selection and are based on their discriminative value of outcomes, using classification and regression. Supervised feature selection is prone to overfitting and ignores the effects of interaction of features among themselves (multicollinearity) (89). Unsupervised selection does not use outcome measures, but creates clusters of features showing similar patterns, thereby maintaining interactions between features (89). This is also called dimension reduction. Examples of dimension reduction are singular value decomposition (SVD), principal component analysis (PCA), self-organising map (SOM) and independent component analysis (ICA) (96). A difficulty of dimension reduction is that most machine learning algorithms generally need 80 to 560 observations to achieve a root-mean-square error below 0.01 (97). Radiomic datasets, however, are often small, concerning homogenous acquisition and reconstruction of images. Optimal dimension reduction methods in small datasets vary between outcome measures and require subsampling (the creation of artificial data points), a dimension reduction technique like PCA and afterwards classification to find the association between the selected features and outcome measures (89).

Validation of the prediction model is an important step towards the clinical use of image-derived biomarkers like radiomics. In many radiomic studies, the dataset is divided in a training- and a validation dataset. However, external validation of the model on at least one dataset is preferred, ideally on datasets from other institutes (45).

Future perspective

In recent years, the number imaging-based procedures in clinical practice has increased considerably and is still growing. This is caused by a growing number of indications for radiologic imaging, due to the need for more rapid, accurate, cost-effective, and less invasive treatment and technological advancements (98). The increased amount of medical imaging has led to an increased workload for the radiologists. This is where automated image analysis, like image segmentation, registration, computer-aided diagnosis and detection play a role. Radiomics have the potential to improve computer-aided diagnosis and detection. Especially in combination with clinical characteristics, radiomics might be used in prognostic and predictive models used for clinical decision support. However, we are not there yet and many steps have to be taken in order to reach the goal of radiomics being part of routine clinical practice. Currently, the field of radiomics would benefit most when more insight was gained in robust dimension reduction methods. Within feature extraction, methodological problems are also faced, but standardisation of image acquisition and feature extraction is already fostered by the Image Biomarker Standardisation Initiative (IBSI).

REFERENCES

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. International journal of cancer. 2015;136(5):E359-86.
2. Integraal kankercentrum Nederland. Cijfers over kanker 2017 [Available from: www.cijfersoverkanker.nl].
3. KWF kanker bestrijding. Longkanker: KWF kanker bestrijding; 2018 [Available from: www.kwf.nl/kanker/longkanker].
4. Landelijke Werkgroep Longtumoren. Niet-kleincellig longcarcinoom, Landelijke richtlijn, Versie: 2.0. Oncoline. 2011.
5. Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. Mayo Clinic proceedings. 2008;83(5):584-94.
6. Collins LG, Haines C, Perkel R, Enck RE. Lung cancer: diagnosis and management. American family physician. 2007;75(1):56-63.
7. Morgensztern D, Ng SH, Gao F, Govindan R. Trends in stage distribution for patients with non-small cell lung cancer: a National Cancer Database survey. Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer. 2010;5(1):29-33.
8. Rubin E, Reisner HM. Essentials of Rubin's pathology. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2014.
9. Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger KR, Yatabe Y, et al. International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer. 2011;6(2):244-85.
10. Goldstraw P, Chansky K, Crowley J, Rami-Porta R, Asamura H, Eberhardt WEE, et al. The IASLC Lung Cancer Staging Project: Proposals for Revision of the TNM Stage Groupings in the Forthcoming (Eighth) Edition of the TNM Classification for Lung Cancer. Journal of Thoracic Oncology. 2016;11(1):39-51.
11. Zarogoulidis K, Zarogoulidis P, Darwiche K, Boutsikou E, Machairiotis N, Tsakiridis K, et al. Treatment of non-small cell lung cancer (NSCLC). Journal of Thoracic Disease. 2013;5(Suppl 4):S389-S96.
12. Brahmer J, Reckamp KL, Baas P, Crinò L, Eberhardt WEE, Poddubskaya E, et al. Nivolumab versus Docetaxel in Advanced Squamous-Cell Non–Small-Cell Lung Cancer. New England Journal of Medicine. 2015;373(2):123-35.
13. Herbst RS, Baas P, Kim D-W, Felip E, Pérez-Gracia JL, Han J-Y, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. The Lancet. 2016;387(10027):1540-50.
14. Fehrenbacher L, Spira A, Ballinger M, Kowanetz M, Vansteenkiste J, Mazieres J, et al. Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial. The Lancet.387(10030):1837-46.
15. Cherry SR, Sorenson JA, Phelps ME. chapter 18 - Positron Emission Tomography. Physics in Nuclear Medicine (Fourth Edition). Philadelphia: W.B. Saunders; 2012. p. 307-43.
16. Tong S, Alessio AM, Kinahan PE. Image reconstruction for PET/CT scanners: past achievements and future challenges. Imaging in medicine. 2010;2(5):529-45.
17. Soret M, Bacharach SL, Buvat I. Partial-volume effect in PET tumor imaging. J Nucl Med. 2007;48(6):932-45.
18. Vriens D. Quantitative characterisation of solid tumours by 18F-FDG PET. What's in a number? Nijmegen: Radboud Universiteit Nijmegen; 2015.
19. Dutch Society of Nuclear Medicine. Procedure Guidelines Nuclear Medicine: Dutch Society of Nuclear Medicine; 2016.

20. Boellaard R, Delgado-Bolton R, Oyen WJG, Giammarile F, Tatsch K, Eschner W, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. European journal of nuclear medicine and molecular imaging. 2015;42:328-54.
21. Vriens D, Visser EP, de Geus-Oei LF, Oyen WJ. Methodological considerations in quantification of oncological FDG PET studies. European journal of nuclear medicine and molecular imaging. 2010;37(7):1408-25.
22. Cherry SR, Sorenson JA, Phelps ME. chapter 21 - Tracer Kinetic Modeling. Physics in Nuclear Medicine (Fourth Edition). Philadelphia: W.B. Saunders; 2012. p. 379-405.
23. Carson RE. Chapter 6 - Tracer Kinetic Modeling in PET. Positron Emission Tomography: Basic Science and Clinical Practice. London: Springer-Verlag; 2003. p. 147-79.
24. Gambhir SS. Quantitative Assay Development for PET. In: Phelps ME, editor. PET: Molecular Imaging and Its Biological Applications. New York, NY: Springer New York; 2004. p. 125-216.
25. Patlak CS, Blasberg RG, Fenstermacher JD. Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism. 1983;3(1):1-7.
26. Oikonen V. Multiple Time Graphical Analysis (MTGA) Turku: Turku PET Centre; 2008 [Available from: http://www.turkupetcentre.net/petanalysis/model_mtga.html].
27. Junttila MR, de Sauvage FJ. Influence of tumour micro-environment heterogeneity on therapeutic response. Nature. 2013;501:346.
28. Hatt M, Tixier F, Pierce L, Kinahan PE, Le Rest CC, Visvikis D. Characterization of PET/CT images using texture analysis: the past, the present… any future? European journal of nuclear medicine and molecular imaging. 2017;44(1):151-65.
29. Cook GJR, Azad G, Owczarczyk K, Siddique M, Goh V. Challenges and Promises of PET Radiomics. International journal of radiation oncology, biology, physics. 2018.
30. Del Monte U. Does the cell number 109 still really fit one gram of tumor tissue? Cell Cycle. 2009;8(3):505-6.
31. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology. 2016;278(2):563-77.
32. Scrivener M, de Jong EEC, van Timmeren JE, Pieters T, Ghaye B, Geets X. Radiomics applied to lung cancer: a review. Translational Cancer Research. 2016;5(4):398-409.
33. Limkin EJ, Sun R, Dercle L, Zacharaki EI, Robert C, Reuze S, et al. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. Ann Oncol. 2017;28(6):1191-206.
34. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer. 2012;48(4):441-6.
35. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJWL. Machine Learning methods for Quantitative Radiomic Biomarkers. Scientific Reports. 2015;5:13087.
36. Cook GJ, O'Brien ME, Siddique M, Chicklore S, Loi HY, Sharma B, et al. Non-Small Cell Lung Cancer Treated with Erlotinib: Heterogeneity of (18)F-FDG Uptake at PET-Association with Treatment Response and Prognosis. Radiology. 2015;276(3):883-93.
37. Antunes J, Viswanath S, Rusu M, Valls L, Hoimes C, Avril N, et al. Radiomics Analysis on FLT-PET/MRI for Characterization of Early Treatment Response in Renal Cell Carcinoma: A Proof-of-Concept Study. Transl Oncol. 2016;9(2):155-62.
38. Tixier F, Hatt M, Valla C, Fleury V, Lamour C, Ezzouhri S, et al. Visual Versus Quantitative Assessment of Intratumor 18F-FDG PET Uptake Heterogeneity: Prognostic Value in Non–Small Cell Lung Cancer. Journal of Nuclear Medicine. 2014;55(8):1235-41.
39. Collarino A, Garganese G, Fragomeni SM, Pereira Arias-Bouda LM, Ieria FP, Boellaard R, et al. Radiomics in vulvar cancer: first clinical experience using (18)F-FDG PET/CT images. J Nucl Med. 2018.
40. Parmar C, Leijenaar RT, Grossmann P, Rios Velazquez E, Bussink J, Rietveld D, et al. Radiomic feature clusters and prognostic signatures specific for Lung and Head & Neck cancer. Sci Rep. 2015;5:11044.
41. Zwanenburg A, Leger S, Vallieres M, Lock S. Image biomarker standardisation initiative - feature definitions v1.7. CoRR. 2016;1612.07003.

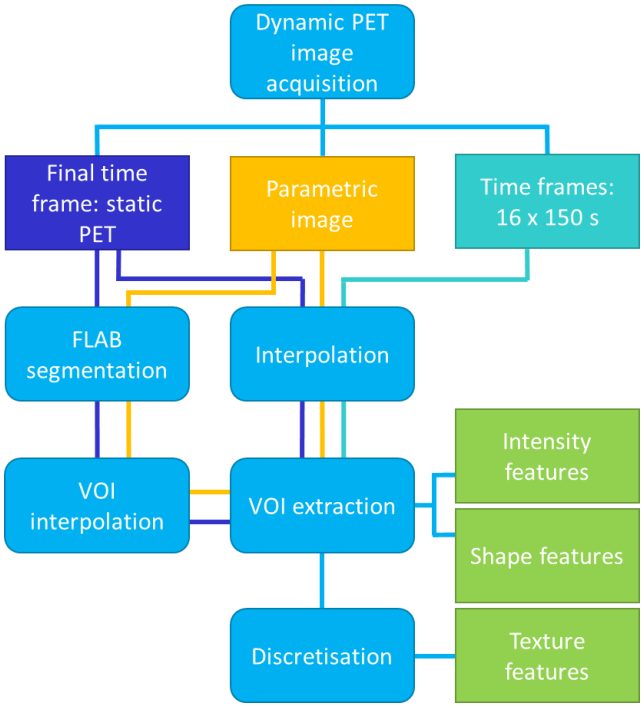
42. van Velden FH, Kramer GM, Frings V, Nissen IA, Mulder ER, de Langen AJ, et al. Repeatability of Radiomic Features in Non-Small-Cell Lung Cancer [(18)F]FDG-PET/CT Studies: Impact of Reconstruction and Delineation. Mol Imaging Biol. 2016;18(5):788-95.
43. Clarke R, Ressom HW, Wang A, Xuan J, Liu MC, Gehan EA, et al. The properties of high-dimensional data spaces: implications for exploring gene and protein expression data. Nature reviews Cancer. 2008;8(1):37-49.
44. Sanduleanu S, Woodruff HC, de Jong EEC, van Timmeren JE, Jochems A, Dubois L, et al. Tracking tumor biology with radiomics: A systematic review utilizing a radiomics quality score. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology. 2018;127(3):349-60.
45. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nature reviews Clinical oncology. 2017;14(12):749-62.
46. Leijenaar RT, Nalbantov G, Carvalho S, van Elmpt WJ, Troost EG, Boellaard R, et al. The effect of SUV discretization in quantitative FDG-PET Radiomics: the need for standardized methodology in tumor texture analysis. Sci Rep. 2015;5:11075.
47. van Tinteren H, Hoekstra OS, Smit EF, van den Bergh JH, Schreurs AJ, Stallaert RA, et al. Effectiveness of positron emission tomography in the preoperative assessment of patients with suspected non-small-cell lung cancer: the PLUS multicentre randomised trial. Lancet (London, England). 2002;359(9315):1388-93.
48. Cuaron J, Dunphy M, Rimner A. Role of FDG-PET scans in staging, response assessment, and follow-up care for non-small cell lung cancer. Frontiers in oncology. 2013;2:208-.
49. Subedi N, Scarsbrook A, Darby M, Korde K, Mc Shane P, Muers MF. The clinical impact of integrated FDG PET–CT on management decisions in patients with lung cancer. Lung Cancer. 2009;64(3):301-7.
50. Gregory DL, Hicks RJ, Hogg A, Binns DS, Shum PL, Milner A, et al. Effect of PET/CT on Management of Patients with Non–Small Cell Lung Cancer: Results of a Prospective Study with 5-Year Survival Data. Journal of Nuclear Medicine. 2012;53(7):1007-15.
51. Hellwig D, Baum RP, Kirsch C. FDG-PET, PET/CT and conventional nuclear medicine procedures in the evaluation of lung cancer: a systematic review. Nuklearmedizin Nuclear medicine. 2009;48(2):59-69, quiz N8-9.
52. Silvestri GA, Gould MK, Margolis ML, Tanoue LT, McCrory D, Toloza E, et al. Noninvasive Staging of Non-small Cell Lung Cancer: ACCP Evidenced-Based Clinical Practice Guidelines (2nd Edition). Chest. 2007;132(3, Supplement):178S-201S.
53. Grootjans W, de Geus-Oei L-F, Troost EGC, Visser EP, Oyen WJG, Bussink J. PET in the management of locally advanced and metastatic NSCLC. Nature Reviews Clinical Oncology. 2015;12:395.
54. Ashley EA. Towards precision medicine. Nature Reviews Genetics. 2016;17:507.
55. Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger K, Yatabe Y, et al. Diagnosis of lung cancer in small biopsies and cytology: implications of the 2011 International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society classification. Archives of pathology & laboratory medicine. 2013;137(5):668-84.
56. Gevaert O, Mitchell LA, Achrol AS, Xu J, Echegaray S, Steinberg GK, et al. Glioblastoma Multiforme: Exploratory Radiogenomic Analysis by Using Quantitative Image Features. Radiology. 2014;273(1):168-74.
57. Khalvati F, Wong A, Haider MA. Automated prostate cancer detection via comprehensive multi-parametric magnetic resonance imaging texture feature models. BMC Medical Imaging. 2015;15(1):27.
58. Coroller TP, Grossmann P, Hou Y, Rios Velazquez E, Leijenaar RTH, Hermann G, et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. Radiotherapy and Oncology. 2015;114(3):345-50.
59. Ganeshan B, Goh V, Mandeville HC, Ng QS, Hoskin PJ, Miles KA. Non–Small Cell Lung Cancer: Histopathologic Correlates for Texture Parameters at CT. Radiology. 2013;266(1):326-36.
60. Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nature Communications. 2014;5:4006.

61. Cook GJR, Yip C, Siddique M, Goh V, Chicklore S, Roy A, et al. Are Pretreatment 18F-FDG PET Tumor Textural Features in Non–Small Cell Lung Cancer Associated with Response and Survival After Chemoradiotherapy? Journal of Nuclear Medicine. 2013;54(1):19-26.
62. Meijer TWH, de Geus-Oei LF, Visser EP, Oyen WJG, Looijen-Salamon MG, Visvikis D, et al. Tumor Delineation and Quantitative Assessment of Glucose Metabolic Rate within Histologic Subtypes of Non-Small Cell Lung Cancer by Using Dynamic (18)F Fluorodeoxyglucose PET. Radiology. 2017;283(2):547-59.
63. Kassner A, Thornhill RE. Texture Analysis: A Review of Neurologic MR Imaging Applications. American Journal of Neuroradiology. 2010;31(5):809-16.
64. Tixier F, Vriens D, Cheze-Le Rest C, Hatt M, Disselhorst JA, Oyen WJ, et al. Comparison of Tumor Uptake Heterogeneity Characterization Between Static and Parametric 18F-FDG PET Images in Non-Small Cell Lung Cancer. J Nucl Med. 2016;57(7):1033-9.
65. Woods BJ, Clymer BD, Kurc T, Heverhagen JT, Stevens R, Orsdemir A, et al. Malignant-lesion segmentation using 4D co-occurrence texture analysis applied to dynamic contrast-enhanced magnetic resonance breast image data. Journal of magnetic resonance imaging : JMRI. 2007;25(3):495-501.
66. Woods B, Clymer B, Saltz J, Kurc T, editors. A Parallel Implementation of 4-Dimensional Haralick Texture Analysis for Disk-Resident Image Datasets. SC '04: Proceedings of the 2004 ACM/IEEE Conference on Supercomputing; 2004 6-12 Nov. 2004.
67. Bianco-Martinez E, Baptista MS. Space-time nature of causality. Chaos: An Interdisciplinary Journal of Nonlinear Science. 2018;28(7):075509.
68. Hu Y, Carmona J, Murphy RF, editors. Application of temporal texture features to automated analysis of protein subcellular locations in time series fluorescence microscope images. Biomedical Imaging: Nano to Macro, 2006 3rd IEEE International Symposium on; 2006: IEEE.
69. Haralick RM, Shanmugam K, Dinstein I. Textural Features for Image Classification. IEEE Transactions on Systems, Man, and Cybernetics. 1973;SMC-3(6):610-21.
70. Boellaard R, O'Doherty MJ, Weber WA, Mottaghy FM, Lonsdale MN, Stroobants SG, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. European journal of nuclear medicine and molecular imaging. 2010;37(1):181-200.
71. Hatt M, Cheze le Rest C, Turzo A, Roux C, Visvikis D. A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET. IEEE transactions on medical imaging. 2009;28(6):881-93.
72. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. Cancer research. 2017;77(21):e104-e7.
73. Yip SS, Aerts HJ. Applications and limitations of radiomics. Physics in medicine and biology. 2016;61(13):R150-66.
74. Hutcheson GD, Sofroniou N. The Multivariate Social Scientist: Introductory Statistics Using Generalized Linear Models: SAGE Publications; 1999.
75. Field A. Discovering Statistics using IBM SPSS Statistics: Sage Publications Ltd.; 2013. 952 p.
76. Babyak MA. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. Psychosomatic medicine. 2004;66(3):411-21.
77. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. Journal of Clinical Epidemiology. 1996;49(12):1373-9.
78. Goeman JJ, van de Geer SA, de Kort F, van Houwelingen HC. A global test for groups of genes: testing association with a clinical outcome. Bioinformatics (Oxford, England). 2004;20(1):93-9.
79. Desseroit M-C, Visvikis D, Tixier F, Majdoub M, Perdrisot R, Guillevin R, et al. Development of a nomogram combining clinical staging with 18F-FDG PET/CT image features in non-small-cell lung cancer stage I–III. European journal of nuclear medicine and molecular imaging. 2016;43(8):1477-85.
80. Tixier F, Hatt M, Valla C, Fleury V, Lamour C, Ezzouhri S, et al. Visual versus quantitative assessment of intratumor 18F-FDG PET uptake heterogeneity: prognostic value in non-small cell lung cancer. J Nucl Med. 2014;55(8):1235-41.
81. Win T, Miles KA, Janes SM, Ganeshan B, Shastry M, Endozo R, et al. Tumor Heterogeneity and Permeability as Measured on the CT Component of PET/CT Predict Survival in Patients with Non–Small Cell Lung Cancer. Clinical Cancer Research. 2013;19(13):3591-9.

82. Zhang J, Chen L, Chen Y, Wang W, Cheng L, Zhou X, et al. Tumor Vascularity and Glucose Metabolism Correlated in Adenocarcinoma, but Not in Squamous Cell Carcinoma of the Lung. PLOS ONE. 2014;9(3):e91649.
83. Ha S, Choi H, Cheon GJ, Kang KW, Chung J-K, Kim EE, et al. Autoclustering of Non-small Cell Lung Carcinoma Subtypes on (18)F-FDG PET Using Texture Analysis: A Preliminary Result. Nuclear medicine and molecular imaging. 2014;48(4):278-86.
84. Sollini M, Cozzi L, Antunovic L, Chiti A, Kirienko M. PET Radiomics in NSCLC: state of the art and a proposal for harmonization of methodology. Sci Rep. 2017;7(1):358.
85. Fried DV, Mawlawi O, Zhang L, Fave X, Zhou S, Ibbott G, et al. Stage III Non–Small Cell Lung Cancer: Prognostic Value of FDG PET Quantitative Imaging Features Combined with Clinical Prognostic Factors. Radiology. 2016;278(1):214-22.
86. Carvalho S, Leijenaar RT, Velazquez ER, Oberije C, Parmar C, van Elmpt W, et al. Prognostic value of metabolic metrics extracted from baseline positron emission tomography images in non-small cell lung cancer. Acta oncologica (Stockholm, Sweden). 2013;52(7):1398-404.
87. Ohri N, Duan F, Snyder BS, Wei B, Machtay M, Alavi A, et al. Pretreatment 18F-FDG PET Textural Features in Locally Advanced Non–Small Cell Lung Cancer: Secondary Analysis of ACRIN 6668/RTOG 0235. Journal of Nuclear Medicine. 2016;57(6):842-8.
88. Apostolova I, Rogasch J, Buchert R, Wertzel H, Achenbach HJ, Schreiber J, et al. Quantitative assessment of the asphericity of pretherapeutic FDG uptake as an independent predictor of outcome in NSCLC. BMC Cancer. 2014;14(1):896.
89. Zhang Y, Oikonomou A, Wong A, Haider MA, Khalvati F. Radiomics-based Prognosis Analysis for Non-Small Cell Lung Cancer. Scientific reports. 2017;7:46349-.
90. Marco Brambilla M, Chiara Secco M, Marco Dominietto M, Roberta Matheoud M. Performance characteristics obtained for a new 3-dimensional lutetium oxyorthosilicate–based whole-body PET/CT scanner with the National Electrical Manufacturers Association NU 2-2001 Standard. J Nucl Med. 2005;46:2083-91.
91. Jakoby BW, Bercier Y, Conti M, Casey ME, Bendriem B, Townsend DW. Physical and clinical performance of the mCT time-of-flight PET/CT scanner. Physics in medicine and biology. 2011;56(8):2375-89.
92. Shafiq-Ul-Hassan M, Zhang GG, Latifi K, Ullah G, Hunt DC, Balagurunathan Y, et al. Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels. Medical physics. 2017;44(3):1050-62.
93. Orlhac F, Boughdad S, Philippe C, Stalla-Bourdillon H, Nioche C, Champion L, et al. A Postreconstruction Harmonization Method for Multicenter Radiomic Studies in PET. J Nucl Med. 2018;59(8):1321-8.
94. Larue RTHM, Defraene G, Ruysscher DD, Lambin P, Elmpt Wv. Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures. The British Journal of Radiology. 2017;90(1070):20160665.
95. Nie F, Xu D, Tsang IW, Zhang C. Flexible Manifold Embedding: A Framework for Semi-Supervised and Unsupervised Dimension Reduction. IEEE Transactions on Image Processing. 2010;19(7):1921-32.
96. Sembiring RW, Zain JM, Embong A. Dimension Reduction of health data clustering. arXiv preprint arXiv:11103569. 2011.
97. Figueroa RL, Zeng-Treitler Q, Kandula S, Ngo LH. Predicting sample size required for classification performance. BMC medical informatics and decision making. 2012;12:8.
98. Wang S, Summers RM. Machine learning and radiology. Medical Image Analysis. 2012;16(5):933-51.

Appendix 1: Image Biomarker Standardisation Initiative reporting guidelines

Written according Image Biomarker Standardisation Initiative reporting guidelines (41).

General			
Aim	Prediction of overall survival (OS), disease-specific survival (DSS), disease-free survival (DFS) and histopathology in patients with non-small cell lung cancer of stage IB to stage IIIA (according to the TNM 7th edition), who underwent primary resection		
Acquisition		Siemens Biograph Duo	Siemens Biograph mCT
	18F-FDG activity	±3.45 MBq/kg bodyweight	±3.45 MBq/kg bodyweight
	Reconstruction	ordered-subsets expectation maximization with 4 iterations per 16 subsets	TrueX, 3 iterations per 21 subsets, point-spread function-based time of flight
	Filter	5 mm Gaussian	3 mm Gaussian
	Matrix	256x256	512x512
	Resolution (mm²)	2.653x2.653	1.591x1.591
	Slice thickness (mm)	3.375	2.027
	Slices	47	109
	Time frames (1 bedposition)	20 x 5 s 5 x 10 s 10 x 15 s 10 x 30 s 16 x 75 s 8 x 150 s 1 x 600 s (static)	20 x 5 s 5 x 10 s 10 x 15 s 10 x 30 s 16 x 75 s 8 x 150 s 1 x 600 s (static)
Approach	The images were analysed as a volume		
Structure	 Figure 9: Pipeline for radiomic feature extraction.		

Software	Interpolation: MATLAB 2017b (Mathworks, Natick, Massachusetts) Static/parametric radiomics: PyRadiomics 2.0 (72) Dynamic GLCM radiomics: PyRadiomics 1.3 (72)		
Data conversion			
Procedure	Static images: Voxel values were converted from Bq/mL to SUV, based on the injected activity with decay correction and the body weight of the patient. The MATLAB script for this procedure can be found in appendix 2. Dynamic images: Dynamic images were created by combining the 16 75 s frames to 8 150 s frames and combining them with the 8 150 s frames to a 4D volume. Voxel values were converted from Bq/mL to SUV, based on the injected activity with decay correction and the body weight of the patient. The MATLAB script can be found in appendix 3.		
Image post-processing			
Procedure	None		
Segmentation			
Procedure	Segmentation of the static and parametric tumour volumes was done using a fuzzy locally adaptive Bayesian (FLAB) algorithm (62), excluding ^{18}F-FDG-avid nontumour tissue by drawing an oversized container around the tumour and surrounding tissue. The volume of interest (VOI) drawn on the static scan was also used for all frames of the dynamic scan.		
Voxel interpolation			
Isotropic voxel dimension		Siemens Biograph Duo	Siemens Biograph mCT
	Original voxel dimensions (mm³)	2.653x2.653x3.375	1.591x1.591x2.027
	Interpolated voxel dimensions (mm³)	3.375x3.375x3.375	3.375x3.375x3.375
Interpolation procedure	Trilinear interpolation, grids aligned by centre.		
Grey level rounding	Not applicable		
Grey level cut-off	Not applicable		
ROI mask interpolation procedure	Trilinear interpolation, grids aligned by centre.		
ROI partial volume	Integer rounding.		
Re-segmentation			
Inclusion/exclusion criteria	Not applicable.		
Volume resection			
Bounding box	Not applicable.		
Discretisation			
Discretisation algorithm	Fixed bin width calculated with the Freedman Diaconis rule: $bin\ size = 2 * IQR \cdot N^{-1/3}$ (73), where IQR is the interquartile range in the VOI and N is the number of voxels in the VOI. The values for IQR and number of voxels are the mean values of all included tumours. For the dynamic scans discretisation was performed for the 16 individual time frames.		
Discretisation parameters	Static PET: bin width = 0.55 Parametric PET: bin width = 0.000000018 Dynamic GLCM, per frame: bin width: 0.26, 0.29, 0.31, 0.34, 0.36, 0.36, 0.39, 0.40, 0.42, 0.44, 0.48, 0.49, 0.50, 0.51, 0.53, 0.54		

Feature calculation	
Feature set	Static and parametric scans: - First Order Statistics (18 features): Energy, Total Energy, Entropy, Minimum, 10th Percentile, 90th Percentile, Maximum, Mean, Median, Interquartile Range, Range, Mean Absolute Deviation, Robust Mean Absolute Deviation, Root Mean Squared, Skewness, Kurtosis, Variance, Uniformity - Shape based (13 features): Volume, Surface Area, Surface Area to Volume Ratio, Sphericity, Maximum 3D Diameter, Maximum 2D Diameter Slice, Maximum 2D Diameter Column, Maximum 2D Diameter Row, Major Axis, Minor Axis, Least Axis, Elongation, Flatness - Gray Level Cooccurrence Matrix (22 features): Autocorrelation, Joint Average, Cluster Prominence, Cluster Shade, Cluster Tendency, Contrast, Correlation, Difference Average, Difference Entropy, Difference Variance, Joint Energy (= Angular Second Moment), Joint Entropy, Informational Measure of Correlation 1, Informational Measure of Correlation 2, Inverse Difference Moment, Inverse Difference Moment Normalized, Inverse Difference, Inverse Difference Normalized, Inverse Variance, Maximum Probability (= Joint Maximum), Sum Entropy, Sum of Squares (=Joint Variance) - Gray Level Run Length Matrix (16 features): Short Run Emphasis, Long Run Emphasis, Gray Level Non-Uniformity, Gray Level Non-Uniformity Normalized, Run Length Non-Uniformity, Run Length Non-Uniformity Normalized, Run Percentage, Gray Level Variance, Run Variance, Run Entropy, Low Gray Level Run Emphasis, High Gray Level Run Emphasis, Short Run Low Gray Level Emphasis, Short Run High Gray Level Emphasis, Long Run Low Grey Level Emphasis, Long Run High Gray Level Emphasis - Gray Level Size Zone Matrix (16 features): Small Area Emphasis, Large Area Emphasis, Grey Level Non-Uniformity, Gray Level Non-Uniformity Normalized, Size-Zone Non-Uniformity, Size-Zone Non-Uniformity Normalized, Zone Percentage, Gray Level Variance, Zone Variance, Zone Entropy, Low Grey Level Zone Emphasis, High Grey Level Zone Emphasis, Small Area Low Gray Level Emphasis, Small Area High Gray Level Emphasis, Large Area Low Gray Area Emphasis, Large Area High Gray Level Emphasis - Neighbouring Gray Tone Difference Matrix (5 features): Coarseness, Contrast, Busyness, Complexity, Strength - Gray Level Dependence Matrix (14 features): Small Dependence Emphasis, Large Dependence Emphasis, Gray Level Non-Uniformity, Dependence Non-Uniformity, Dependence Non-Uniformity Normalized, Gray Level Variance, Dependence Variance, Dependence Entropy, Low Grey Level Emphasis, High Gray Level Emphasis, Small Dependence Low Gray Level Emphasis, Small Dependence High Gray Level Emphasis, Large Dependence Low Gray Level Emphasis, Large Dependence High Gray Level Emphasis - Total Lesion Glycolysis Dynamic scans: - Gray Level Cooccurrence Matrix (22 features): Autocorrelation, Joint Average, Cluster Prominence, Cluster Shade, Cluster Tendency, Contrast, Correlation, Difference Average, Difference

	Entropy, Difference Variance, Joint Energy (= Angular Second Moment), Joint Entropy, Informational Measure of Correlation 1, Informational Measure of Correlation 2, Inverse Difference Moment, Inverse Difference Moment Normalized, Inverse Difference, Inverse Difference Normalized, Inverse Variance, Maximum Probability (= Joint Maximum), Sum Average, Sum Entropy
Feature settings	Static and parametric: Image normalization and distance weighting were not applied. GLCMs were calculated in 13 directions, made symmetrical and combined to one 3D GLCM on which the features were calculated. Dynamic GLCM features: Image normalization and distance weighting were not applied. Features were calculated in only the temporal direction (0;0;0;1), comparing grey levels of subsequent voxels in time. GLCMs for the individual angles were not made symmetrical to take into account causality in the time dimension and were combined to one 3D GLCM.
Standardisation	Feature extraction was tested using the digital phantom with a bin width of 1. Results corresponded to the standards values provided.

Appendix 2: MATLAB script pre-processing and interpolation static FDG-PET radiomics

A similar code was used to pre-process and interpolate the parametric PET images, but these images did not require conversion of voxel unit.

The function ‘interpolation’ used in this script can be found in appendix 4.

```
clear all
close all
clc
```

Load Nearly raw raster data read and write package

```
addpath('\\vf-d2-home\d2home$\wanoortman\MyDocs\MATLAB\Packages\nrrdread');
addpath('\\vf-d2-home\d2home$\wanoortman\MyDocs\MATLAB\Packages\nrrdWriter');
```

Load PET information: patient weight, injected activity

```
PET_info = xlsread('Z:\PLUTARCh\Databases\PET gegevens.xlsx');
```

Select patients

```
path.PET = 'Z:\PLUTARCh\FLAB\Plutarch_FLAB_definitief';
cd(path.PET);
subjects = dir('*_*');

for p = 1:length(subjects)
% Create a list with all subjects in the subject directory
Se{p} = subjects(p).name;
end

[pat_select, ok] = listdlg('PromptString', 'Select subjects', 'ListString', Se,
'SelectionMode', 'multiple'); % select one or more subjects
```

Loop through patients

```
for i = pat_select
```

Read DICOM file

```
% Read dicom files in directory
path.subj = [path.PET '\' subjects(i).name '\' subjects(i).name '_StatischePET'];
cd(path.subj)
files = dir('*.dcm');

for j = 1:length(files)
    % Read dicom info for an individual slice
    info = dicominfo(files(j).name);
```

```
    slopes(j) = info.RescaleSlope;
    % Read dicom file, correcting for rescale slope and set class to double
    % Rescale intercept was 0 for all patients and slices
    I = info.RescaleSlope .* double(dicomread(files(j).name));
    % Create volume from slices, order based on image index
    vol(:,:,info.ImageIndex)=I;
end
```

Write voxel values from Bq/mL to SUV

```
% Calculate SUV values for the volume using patient weight and injected dose
vol_SUV = vol .* (PET_info(i,2) ./ PET_info(i,4) ./ 1000);
```

Interpolation of the image volume by additional interpolation algorithm

```
vox = 3.3750;

vol_SUV_int = interpolation(vol_SUV, info.PixelSpacing, info.SliceThickness, vox);
```

Read VOI

```
path.voi = ['Z:\PLUTARCh\FLAB\FLAB_NRRDs\' subjects(i).name(1:2)];
cd(path.voi)

[voi, voimeta] = nrrdread('RTSTRUCT StructureSet-label.nrrd');
voi = voi >= 1;
```

Interpolation of the volume of interest by additional interpolation algorithm

Volumes of interest are boolean areas represented as integers. Interpolation function requires the volume in double precision. Therefore, the initial VOI is set to double precision and after interpolation the VOI is set to integers. This also corrects for grey level rounding (>=0.5)

```
voi_int = int16(interpolation(double(voi), info.PixelSpacing, info.SliceThickness, vox));
```

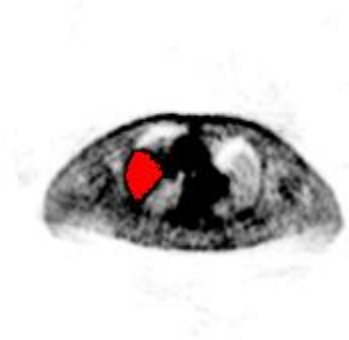
Write NRRD

```
pixelspacing = [3.3750 3.3750 3.3750];
origin = [0,0,0];

path.output = 'Z:\PLUTARCh\NRRD3Dvolvoi';
cd(path.output)

nrrdWriter(['img_' subjects(i).name(1:2) '.nrrd'], vol_SUV_int, pixelspacing, origin,
'raw');
nrrdWriter(['voi_' subjects(i).name(1:2) '.nrrd'], voi_int, pixelspacing, origin, 'raw');
```

Validate VOI

Visual evaluation of the image and the VOI

```
for p = 1:size(vol_SUV_int,3)
   figure, imshow(imoverlay(imcomplement(vol_SUV_int(:,:,p)),voi_int(:,:,p),'red'));
end
```



Write DICOM file (if necessary)

% Set volume to integers using almost the maximum value (32767) of integers % in the int16 class (factor 1.1). This will be corrected for in the % rescale slope vol_int16 = int16(vol_SUV_int .* 32767 ./ (max(vol_SUV_int(:)) .* 1.1));

% Add interpolated pixel spacing, slice thickness and number of slices to dicom info info.PixelSpacing = [vox vox]; info.SliceThickness = [vox]; info.NumberOfSlices = [size(vol_int,3)];

% Create output directory path.output = ['Z:\PLUTARCh\Interpolated\' subjects(i).name]; mkdir(path.output); cd(path.output);

% Clear unit Bq/mL from dicom info info.Units = ['SUV'];

```
% for slicenum=1:size(vol_int_SUV,3)
%     % Create filename
%     filenamedicom=['Static_PET' num2str(slicenum) '.dcm'];
%     % Add slice specific dicom info
%     info.ImageIndex = slicenum;
%     info.InstanceNumber = slicenum;
%     % Rescale slope corrects for maximum value int16 class
%     info.RescaleSlope = (max(vol_i_s(:)) .* 1.1) ./ 32767;
%     % Write the dicom file
%     dicomwrite(vol_int_SUV(:,:,slicenum), filenamedicom, info, 'CreateMode', 'copy') ;
% end

end
```

Published with MATLAB® R2017b

Appendix 3: MATLAB script pre-processing and interpolation of dynamic volumes

This script shows the pre-processing of the images from the Siemens Biograph Duo. A similar script was used for the images from the Siemens Biograph 40 mCT, taking into account a different matrix size and slice thickness.

The function ‘interpolation’ used in this script can be found in Appendix 4.

```
clear all
close all
clc

addpath('\\vf-d2-home\d2home$\wanoortman\MyDocs\MATLAB\Packages\nrrdWriter');
addpath('\\vf-d2-home\d2home$\wanoortman\MyDocs\MATLAB\Packages\nrrdread');
```

Load PET information: patient weight, injected activity

```
PET_info = xlsread('Z:\PLUTARCh\Databases\PET gegevens.xlsx');
```

Select patients

```
path.PET = 'Z:\PLUTARCh';
cd(path.PET);
subjects = dir('*_*');

for p = 1:length(subjects)
% Create a list with all subjects in the subject directory
Se{p} = subjects(p).name;
end

[pat_select, ok] = listdlg('PromptString', 'Select subjects', 'ListString', Se,
'SelectionMode', 'multiple'); % select one or more subjects
```

Loop through patients

```
for i = pat_select
```

Read volume of interest (VOI)

```
path.voi = ['Z:\PLUTARCh\FLAB\FLAB_NRRDs\' subjects(i).name(1:2)];
cd(path.voi)

[voi, voimeta] = nrrdread('RTSTRUCT StructureSet-label.nrrd');
voi = voi >= 1;

% Interpolate VOI

vox = 3.3750;
voi_int = int16(interpolation(double(voi), [2.65336 2.65336], 3.3750, vox)); % Biograph
```

```
Duo
%voi_int = int16(interpolation(double(voi), [1.59095 1.59095], 2.02700805664, vox)); %
Biograph mCT
```

Create 4D VOI

The FLAB VOI from the static image is used as VOI for the individual time frames

```
for z = 1:16
    voi4D(:,:,:,z) = voi_int;
end
```

Read DICOM file

```
% Read dicom files in directory
path.subj = [path.PET '\' subjects(i).name '\Recon_Plutarch_TvZ'];
cd(path.subj)
files = dir('*.ima');

% Find ImageIndex from file name. The image index determines the position
% of the slice in the stack. 1-47 are the z-slices for time frame 1, 48-94
% for frame 2, etc.
for j = 1:length(files)
    files(j).ImageIndex =
str2num(files(j).name(strfind(files(j).name,'.3.')+3:strfind(files(j).name,'.2015.0')-
1));
end

% Sort images by ImageIndex
[~,index] = sortrows([files.ImageIndex].');
files = files(index);
```

Create 4D volume

```
% Select frames (16 x 75 s frames (together), 8 x 150 s frames)
frame = [46:2:61 62:69];

for l = frame
    % Add 75 s frames per 2
    if l < 62
        for k = (l-1)*47+1 : (l)*47
            % Load first 75 seconds frame
            % Read dicom info for an individual slice
            info = dicominfo(files(k).name);
            % Read dicom file, correcting for rescale slope and set class to double
            I1 = info.RescaleSlope .* double(dicomread(files(k).name));
            % Create volume from slices, order based on image index
            vol1(:,:,48-(k-(l-1)*47)) = I1;

            % Load second 75 seconds frame
            info2 = dicominfo(files(k+47).name);
            I2 = info2.RescaleSlope .* double(dicomread(files(k+47).name));
            vol2(:,:,48-(k-(l-1)*47)) = I2;
```

```
        end

    % Take average of two volumes
    vol = (vol1 + vol2) ./ 2;

    else
        % Load 150 s frame
        for k = (l-1)*47+1 : (l)*47
            info = dicominfo(files(k).name);
            I = info.RescaleSlope .* double(dicomread(files(k).name));
            vol(:,:,48-(k-(l-1)*47)) = I;
        end
    end

    % Write voxel values from Bq/mL to SUV using patient weight and injected dose
    vol_SUV = vol .* (PET_info(i,2) ./ (PET_info(i,4)) ./ 1000);

    % Interpolation algorithm (centres alligned)
    vox = 3.3750;
    vol_int = interpolation(vol_SUV, info.PixelSpacing, info.SliceThickness, vox);

    % Determine interquartile range for individual time frames
    binw(i,1) = i;
    binw(i,2) = sum(voi_int(:));
    binw(i,l-43) = iqr(vol_int(logical(voi_int)));

    % Create 4D volume with the frames as the 4th dimension
    vol4D(:,:,:,find(frame == l)) = vol_int;

end
```

Save 4D volume as NRRD

```
addpath('\\vf-d2-home\d2home$\wanoortman\MyDocs\MATLAB\Packages\nrrd_read_write_rensonnet\nrrd_read_write_rensonnet')

nrrd.content = 'vol';
nrrd.data = vol4D;
nrrd.dimension = 4;
nrrd.space = 'left-posterior-superior';
nrrd.sizes = [size(nrrd.data,1) size(nrrd.data,2) size(nrrd.data,3) size(nrrd.data,4)];
nrrd.spacedirections = { '(3.3750,0,0)', '(0,3.3750,0)', '(0,0,3.3750)', 'none'};
nrrd.spacedirections_matrix = [vox 0 0; 0 vox 0; 0 0 vox];
nrrd.kinds = {'domain', 'domain', 'domain', 'list'};
nrrd.encoding = 'gzip';
nrrd.endian = 'little';
nrrd.spaceorigin = [0;0;0];

nhdr_nrrd_write('4Dvol.nrrd',nrrd,'true')
```

Save 4D VOI as NRRD

```
nrrd.content = '4Dvoi';
nrrd.data = voi4D;
nrrd.dimension = 4;
nrrd.space = 'left-posterior-superior';
nrrd.sizes = [size(nrrd.data,1) size(nrrd.data,2) size(nrrd.data,3) size(nrrd.data,4)];
nrrd.spacedirections = { '(3.375O,0,0)', '(0,3.375O,0)', '(0,0,3.375O)', 'none'};
nrrd.spacedirections_matrix = [vox 0 0; 0 vox 0; 0 0 vox];
nrrd.kinds = {'domain', 'domain', 'domain', 'list'};
nrrd.encoding = 'gzip';
nrrd.endian = 'little';
nrrd.spaceorigin = [0;0;0];

nhdr_nrrd_write('4Dvoi_16f.nrrd',nrrd,'true')
```

```
clear nrrd, clear vol4D, clear vol_SUV, clear vol_int, clear vol, clear vol1, clear vol2
```

```
end
```

Published with MATLAB® R2017b

Appendix 4: MATLAB interpolation function

```
function [vol_int] = interpolation(vol, pixelspacing, slicethickness, vox)
% The interpolation algorithm performs trilinear interpolation by centre
% creating isotrope voxels in given size.

% Input:
%   - vol: image (double)
%   - pixelspacing: voxel size in x and y direction
%   - slicethickness: voxel size in z direction
%   - vox: required voxel size in 1 dimension (isotropic voxels)

% Output:
%   - vol_int: interpolated volume

% Create mesh grid for the original image (rows, columns, slices = y, x, z)
[Xa,Ya,Za] = meshgrid(1:1:size(vol,2),1:1:size(vol,1),1:1:size(vol,3));

% Create interpolation mesh grid
[Xb,Yb,Zb] =
meshgrid(vox/pixelspacing(2):vox/pixelspacing(2):size(vol,2),vox/pixelspacing(1):vox/pixe
lspacing(1):size(vol,1),vox/slicethickness:vox/slicethickness:size(vol,3));

% Perform linear interpolation (default)
vol_int = interp3(Xa,Ya,Za,vol,Xb,Yb,Zb);

end
```

Published with MATLAB® R2017b

Appendix 5: Python code for extraction of dynamic GLCM features

```
from pathlib import Path
import os
import csv
import numpy
import SimpleITK as sitk
import pywt
import six
from six.moves import range

## Loop through patients
# Set root directory to PLUTARCh
rootdir = Path('Z:/PLUTARCh')
# Find patient directories in the root directory
patients = sorted(os.listdir(rootdir))

for patient in patients[0:37]:

## Load 4D image and volume of interest

    # Load image
    voldir = str(rootdir / patient / 'Recon_Plutarch_TvZ/4Dvol.nrrd')
    vol4D = sitk.GetArrayFromImage(sitk.ReadImage(voldir))

    # Load volume of interest
    voidir = str(rootdir / 'FLAB/FLAB_NRRDs' / patient[0:2] / '4Dvoi_16f.nrrd')

    voi4D = sitk.GetArrayFromImage(sitk.ReadImage(voidir))
    # Create boolean array where the voxels that are part of the VOI have value TRUE and other values are FALSE
    voibl4D = sitk.GetArrayFromImage(sitk.ReadImage(voidir)) == 1

    # Show image
    #sitk.Show(sitk.GetImageFromArray(vol4D))
    #sitk.Show(sitk.GetImageFromArray(voi4D))


    for frame in range(vol4D.shape[3]):
        # Obtain 3D for different frames from volume and VOI
        vol3D = vol4D[:, :, :, frame]
        voibl3D = voibl4D[:,:,:,frame]

        # Binning
        # binWidth per time frame
        binWidth = [0.26, 0.29, 0.31, 0.34, 0.36, 0.36, 0.39, 0.40, 0.42, 0.44, 0.48, 0.49, 0.50, 0.51, 0.53, 0.54]
        # Start binning form the first value lesser than or equal to the minimum value and evenly dividable by binwidth
        lowBound = min(vol3D[voibl3D]) - (min(vol3D[voibl3D]) % binWidth[frame])
        # Add + binwidth to ensure the maximum value is included in the range generated by numpy.arange
        highBound = max(vol3D[voibl3D]) + binWidth[frame]
        # Determine binEdges
        binEdges = numpy.arange(lowBound, highBound, binWidth[frame])
        # Binning
        vol3D[voibl3D] = numpy.digitize(vol3D[voibl3D], binEdges)

        # Combine frames in 4D volume
        vol4D[:, :, :, frame] = vol3D
```

```
54. 
55. ## GLCM
56.     # Exclude voxels outside segmentation, due to binning, no negative values w
    ill be encountered inside the mask
57.     vol4D[voi4D == 0] = -1
58. 
59.     # Angle for the calculation of the GLCM
60.     angles = numpy.array([(0,0,0,1)]) # GLCM is calculated in the time directio
    n (over the time frames)
61. 
62.     # Create empty GLCM, in the x- and y-direction the graylevels, in the z-
    direction the angles
63.     grayLevels = numpy.unique(vol4D[voibl4D])
64.     GLCM = numpy.zeros([len(grayLevels), len(grayLevels), len(angles)], dtype='
    float64')
65. 
66.     # Iterate over gray levels for center voxel
67.     for i_idx, i in enumerate(grayLevels):
68.         # Get the indices to all voxels which have the current gray level i
69.         i_indices = numpy.where(vol4D == i)
70. 
71.         # Iterate over gray levels for neighbouring voxel
72.         for j_idx, j in enumerate(grayLevels):
73.             # Get the indices to all voxels which have the current gray level j

74.             j_indices = set(zip(*numpy.where(vol4D == j)))
75. 
76.             for a_idx, a in enumerate(angles):
77.                 # Get the corresponding indices of the neighbours for angle a
78.                 neighbour_indices = set(zip(*(i_indices + a[:, None])))
79. 
80.                 # The following intersection yields the indices to voxels with
    gray level j
81.                 # that are also a neighbour of a voxel with gray level i for an
    gle a.
82.                 # The number of indices is then equal to the total number of pa
    irs with gray level i and j for angle a
83.                 count = len(neighbour_indices.intersection(j_indices))
84.                 GLCM[i_idx, j_idx, a_idx] = count
85. 
86. ## GLCM options
87.     # No weighting is applied
88.     # GLCM is not made symmetrical, because time is not rotationally invariant

89. 
90.     # Save GLCM for visual evaluation
91.     numpy.save('GLCM16f' + str(patient[0:2]) + '.npy',GLCM)
92. 
93.     # Normalize GLCM (divide by the sum)
94.     GLCM =  GLCM / numpy.sum(GLCM, (0, 1))
95. 
96. ## Calculate coefficients
97. 
98.     Ng = numpy.size(grayLevels) # the number of discrete intensity levels in th
    e image
99.     eps = numpy.spacing(1) # arbitrarily small positive number
100. 
101.           NgVector = grayLevels.astype('float')
102. 
103.           # shape = (Ng, Ng)
104.           i, j = numpy.meshgrid(NgVector, NgVector, indexing='ij', sparse=True
    )
```

```
105.
106.            # shape = (2*Ng-1)
107.            kValuesSum = numpy.arange(2, (Ng * 2) + 1)
108.            # shape = (Ng-1)
109.            kValuesDiff = numpy.arange(0, Ng)
110.
111.            # marginal row probabilities #shape = (Ng, 1, angles)
112.            px = GLCM.sum(1, keepdims=True)
113.            # marginal column probabilities #shape = (1, Ng, angles)
114.            py = GLCM.sum(0, keepdims=True)
115.
116.            # shape = (1, 1, angles)
117.            ux = numpy.sum(i[:, :, None] * GLCM, (0, 1), keepdims=True)
118.            uy = numpy.sum(j[:, :, None] * GLCM, (0, 1), keepdims=True)
119.
120.            # shape = (1, 1, angles)
121.            sigx = numpy.sum(GLCM * ((i[:, :, None] - ux) ** 2), (0, 1), keepdim
    s=True) ** 0.5
122.            # shape = (1, 1, angles)
123.            sigy = numpy.sum(GLCM * ((j[:, :, None] - uy) ** 2), (0, 1), keepdim
    s=True) ** 0.5
124.
125.            # shape = (2*Ng-1, angles)
126.            pxAddy = numpy.array([numpy.sum(GLCM[i + j == k], 0) for k in kValue
    sSum])
127.            # shape = (Ng, angles)
128.            pxSuby = numpy.array([numpy.sum(GLCM[numpy.abs(i - j) == k], 0) for
    k in kValuesDiff])
129.
130.            # entropy of px # shape = (angles)
131.            HX = (-1) * numpy.sum((px * numpy.log2(px + eps)), (0, 1))
132.            # entropy of py # shape = (angles)
133.            HY = (-1) * numpy.sum((py * numpy.log2(py + eps)), (0, 1))
134.            # shape = (angles)
135.            HXY = (-1) * numpy.sum((GLCM * numpy.log2(GLCM + eps)), (0, 1))
136.
137.            # shape = (angles)
138.            HXY1 = (-
    1) * numpy.sum((GLCM * numpy.log2(px * py + eps)), (0, 1))
139.            # shape = (angles)
140.            HXY2 = (-
    1) * numpy.sum(((px * py) * numpy.log2(px * py + eps)), (0, 1))
141.
142.        ## Calculate features and write to dictionary
143.            # Create empty dictionary
144.            GLCMfeatures = {}
145.
146.            # Autocorrelation (measure of the magnitude of the fineness and coar
    seness of texture)
147.            ac = numpy.sum(GLCM * (i * j)[:, :, None], (0, 1))
148.            GLCMfeatures['Autocorrelation'] = ac.mean()
149.
150.            # Joint average (mean gray level intensity of the i distribution)
151.            GLCMfeatures['JointAverage'] =ja = ux.mean()
152.
153.            # Cluster prominence (measure of the skewness and asymmetry of the G
    LCM. A higher values implies more asymmetry about the mean while a lower value
    indicates a peak near the mean value and less variation about the mean)
154.            cp = numpy.sum((GLCM * (((i + j)[:, :, None] - ux - uy) ** 4)), (0,
    1))
155.            GLCMfeatures['ClusterProminence'] = cp.mean()
156.
```

```
157.            # Cluster shade (measure of the skewness and uniformity of the GLCM.
     A higher cluster shade implies greater asymmetry about the mean.)
158.            cs = numpy.sum((GLCM * (((i + j)[:, :, None] - ux - uy) ** 3)), (0,
    1))
159.            GLCMfeatures['ClusterShade'] = cs.mean()
160.
161.            # Cluster tendency (measure of groupings of voxels with similar gray
    -level values)
162.            ct = numpy.sum((GLCM * (((i + j)[:, :, None] - ux - uy) ** 2)), (0,
    1))
163.            GLCMfeatures['ClusterTendency'] = ct.mean()
164.
165.            # Contrast (measure of the local intensity variation, favoring value
    s away from the diagonal :math:`(i = j)`. A larger value correlates with a grea
    ter disparity in intensity values among neighboring voxels.)
166.            cont = numpy.sum((GLCM * ((numpy.abs(i - j))[:, :, None] ** 2)), (0,
     1))
167.            GLCMfeatures['Contrast'] = cont.mean()
168.
169.            # Correlation (value between 0 (uncorrelated) and 1 (perfectly corre
    lated) showing the linear dependency of gray level values to their respective v
    oxels in the GLCM.)
170.            corm = numpy.sum(GLCM * (i[:, :, None] - ux) * (j[:, :, None] - uy),
     (0, 1), keepdims=True)
171.            corr = corm / (sigx * sigy + eps)
172.            corr[sigx * sigy == 0] = 1  # Set elements that would be divided by
    0 to 1.
173.            GLCMfeatures['Correlation'] = corr.mean()
174.
175.            # Difference average (measures the relationship between occurrences
    of pairs with similar intensity values and occurrences of pairs with differing
    intensity values)
176.            diffavg = numpy.sum((kValuesDiff[:, None] * pxSuby), 0)
177.            GLCMfeatures['DifferenceAverage'] = diffavg.mean()
178.
179.            # Difference entropy (measure of the randomness/variability in neigh
    borhood intensity value differences.)
180.            difent = (-1) * numpy.sum((pxSuby * numpy.log2(pxSuby + eps)), 0)
181.            GLCMfeatures['DifferenceEntropy'] = difent.mean()
182.
183.            # Difference variance (measure of heterogeneity that places higher w
    eights on differing intensity level pairs that deviate more from the mean.)
184.            diffavg = numpy.sum((kValuesDiff[:, None] * pxSuby), 0, keepdims=Tru
    e)
185.            diffvar = numpy.sum((pxSuby * ((kValuesDiff[:, None] - diffavg) ** 2
    )), 0)
186.            GLCMfeatures['DifferenceVariance'] = diffvar.mean()
187.
188.            # Dissimilarity (deprecated, since it is mathematically equal to dif
    ference average)
189.
190.            # Joint energy (Energy is a measure of homogeneous patterns in the i
    mage. A greater Energy implies that there are more instances of intensity value
     pairs in the image that neighbor each other at higher frequencies. Defined by
    IBSI as Angular Second Moment)
191.            ene = numpy.sum((GLCM ** 2), (0, 1))
192.            GLCMfeatures['JointEnergy'] = ene.mean()
193.
194.            # Joint entropy (measure of the randomness/variability in neighborho
    od intensity values)
195.            GLCMfeatures['JointEntropy'] = HXY.mean()
196.
```

```
197.            # Homogeneity 1 (deprecated, since it is mathematically equal to inv
    erse difference)
198.
199.            # Homogeneity 2 (deprecated, since it is mathematically equal to Inv
    erse Difference Moment)
200.
201.            # Informal Measure of Correlation (IMC) 1 (In the case where both HX
     and HY are 0 (as is the case in a flat region), an arbitrary value of 0 is ret
    urned to prevent a division by 0. This is done on a per-angle basis)
202.            div = numpy.max(([HX, HY]), 0)
203.            imc1 = (HXY - HXY1) / (div + eps)
204.            imc1[div == 0] = 0  # Set elements that would be divided by 0 to 0
205.            GLCMfeatures['IMC1'] =  imc1.mean()
206.
207.            # Informal Measure of Correlation (IMC) 2 (In the case where HXY = H
    XY2, an arbitrary value of 0 is returned to prevent returning complex numbers.
    This is done on a per-angle basis.)
208.            imc2 = numpy.sqrt(1 - numpy.exp(-2 * (HXY2-HXY)))
209.            GLCMfeatures['IMC2'] =  imc2.mean()
210.
211.            # Inverse Difference Moment (IDM) (measure of the local homogeneity
    of an image. IDM weights are the inverse of the Contrast weights (decreasing ex
    ponentially from the diagonal i=j in the GLCM).)
212.            idm = numpy.sum((GLCM / (1 + ((numpy.abs(i - j))[:, :, None] ** 2)))
    , (0, 1))
213.            GLCMfeatures['IDM'] = idm.mean()
214.
215.            # Inverse Difference Moment Normalized (IDMN)
216.            idmn = numpy.sum((GLCM / (1 + (((numpy.abs(i - j))[:, :, None] ** 2)
     / (Ng ** 2)))), (0, 1))
217.            GLCMfeatures['IDMN'] = idmn.mean()
218.
219.            ## Inverse Difference (ID) ((a.k.a. Homogeneity 1) is another measur
    e of the local homogeneity of an image. With more uniform gray levels, the deno
    minator will remain low, resulting in a higher overall value.)
220.            invDiff = numpy.sum((GLCM / (1 + (numpy.abs(i - j))[:, :, None])), (
    0, 1))
221.            GLCMfeatures['ID'] = invDiff.mean()
222.
223.            ## Inverse Difference Normalized (IDN) (IDN (inverse difference norm
    alized) is another measure of the local homogeneity of an image. Unlike Homogen
    eity1, IDN normalizes the difference between the neighboring intensity values b
    y dividing over the total number of discrete intensity values.)
224.            idn = numpy.sum((GLCM / (1 + ((numpy.abs(i - j))[:, :, None] / Ng)))
    , (0, 1))
225.            GLCMfeatures['IDN'] = idn.mean()
226.
227.            ## Inverse variance
228.            maskDiags = numpy.abs(i - j) > 0
229.            inv = numpy.sum((GLCM[maskDiags] / ((numpy.abs(i - j))[:, :, None] *
    * 2)[maskDiags]), 0)
230.            GLCMfeatures['InverseVariance'] = inv.mean()
231.
232.            ## Maximum Probability (occurrences of the most predominant pair of
    neighboring intensity values. Defined by IBSI as Joint maximum.)
233.            maxprob = GLCM.max((0, 1))
234.            GLCMfeatures['Maximum Probability'] = maxprob.mean()
235.
236.            ## Sum Average (relationship between occurrences of pairs with lower
     intensity values and occurrences of pairs with higher intensity values. When G
    LCM is symmetrical sumavg = 2*jointavg)
237.            sumavg = numpy.sum((kValuesSum[:, None] * pxAddy), 0)
```

```
238.            GLCMfeatures['SumAverage'] = sumavg.mean()
239.
240.            ## Sum variance (deprecated, since it is mathematically equal to Clu
    ster Tendency)
241.
242.            ## Sum entropy (sum of neighborhood intensity value differences)
243.            sumentr = (-1) * numpy.sum((pxAddy * numpy.log2(pxAddy + eps)), 0)
244.            GLCMfeatures['SumEntropy'] = sumentr.mean()
245.
246.            ## Sum of squares (measure in the distribution of neigboring intensi
    ty level pairs about the mean intensity level in the GLCM. only when GLCM is sy
    mmetrical. Defined by IBSI as Joint Variance)
247.            ss = numpy.sum((GLCM * ((i[:, :, None] - ux) ** 2)), (0, 1))
248.            GLCMfeatures['SumSquares'] = ss.mean()
249.
250.        ## Write features to CSV-file
251.            with open('GLCMfeatures16f' + str(patient[0:2]) + '.csv','w') as csv
    _file:
252.                writer = csv.writer(csv_file)
253.                for key, value in GLCMfeatures.items():
254.                    writer.writerow([key, value])
```

Appendix 6: Correlation clustering and principal component analysis

In this appendix correlation clustering and principal component analysis are briefly described. Correlation clustering was performed in MATLAB 2017b (Mathworks, Natick, Massachusetts). Figure 7 shows the variance matrix of all 232 in the feature set. Correlation clustering was performed: features with $R^2 > 0.56$ were grouped and from each group 1 feature was selected. Selected features from each group are shaded in table 6. Principal component analysis was performed to further reduce the number of features. Table 6 shows the features that were deprecated for an individual sampling adequacy measure below 0.5. The final PCA had a Kaiser-Meyer-Olkin Measure of Sampling Adequacy of 0.697, which is 'mediocre' (74). Three components had an eigenvalue above 1.0 and together they explained 81.4% of the total variance in the dataset. The features that correlated best with the components were the static volume, the static grey level cooccurrence matrix maximum probability and the static grey level coouccrrence matrix sum of squares.

Table 6: Complete dataset of features: static (indicated as SUV), parametric (indicated as MRg) and dynamic GLCM (indicated as dyn). The shaded features indicate the selected features from each group (numbers). Selection of the features was based on commonly used features in literature, a diverse selection for principal component analysis and experience of the authors. The features in bold have been deprecated from the PCA for an individual sampling adequacy measure below 0.5.

Feature group	Feature	Anti-image correlation	Feature group	Feature	Anti-image correlation
1	MRg_glszm_LargeAreaHighGrayLevelEmphasis	0.747	24	SUV_firstorder_10Percentile	
2	**SUV_firstorder_Energy**	**0.483**	24	SUV_firstorder_Mean	0.719
2	SUV_firstorder_TotalEnergy		24	SUV_firstorder_Median	
2	MRg_firstorder_Energy		24	SUV_firstorder_Minimum	
2	MRg_firstorder_TotalEnergy		24	SUV_firstorder_RootMeanSquared	
3	SUV_firstorder_Entropy	0.880	24	SUV_glcm_DifferenceAverage	
3	SUV_glcm_DifferenceEntropy		24	SUV_glcm_InverseVariance	
3	SUV_glcm_JointEntropy		24	SUV_glszm_ZonePercentage	
3	SUV_glcm_Idm		24	SUV_gldm_DependenceNonUniformityNormalized	
3	SUV_glcm_Id		24	SUV_gldm_SmallDependenceEmphasis	
3	SUV_glcm_SumEntropy		24	MRg_glcm_DifferenceAverage	
3	MRg_firstorder_Entropy		24	MRg_glcm_InverseVariance	
3	MRg_glcm_DifferenceEntropy		24	MRg_glszm_ZonePercentage	
3	MRg_glcm_JointEntropy		24	MRg_gldm_DependenceNonUniformityNormalized	
3	MRg_glcm_Idm		24	MRg_gldm_SmallDependenceEmphasis	
3	MRg_glcm_Id		24	Dyn_DifferenceAverage	
3	MRg_glcm_SumEntropy		24	Dyn_DifferenceEntropy	
3	Dyn_JointEntropy		24	Dyn_IDM	
3	Dyn_SumEntropy		24	Dyn_ID	
4	SUV_shape_LeastAxis		24	Dyn_InverseVariance	
4	SUV_shape_SurfaceArea	0.663	25	SUV_shape_MajorAxis	
4	SUV_glrlm_RunLengthNonUniformity		25	SUV_shape_Maximum2DDiameterColumn	
4	SUV_glszm_GrayLevelNonUniformity		25	SUV_shape_Maximum2DDiameterRow	
4	SUV_gldm_DependenceNonUniformity		25	SUV_shape_Maximum2DDiameterSlice	
4	MRg_shape_LeastAxis		25	SUV_shape_Maximum3DDiameter	

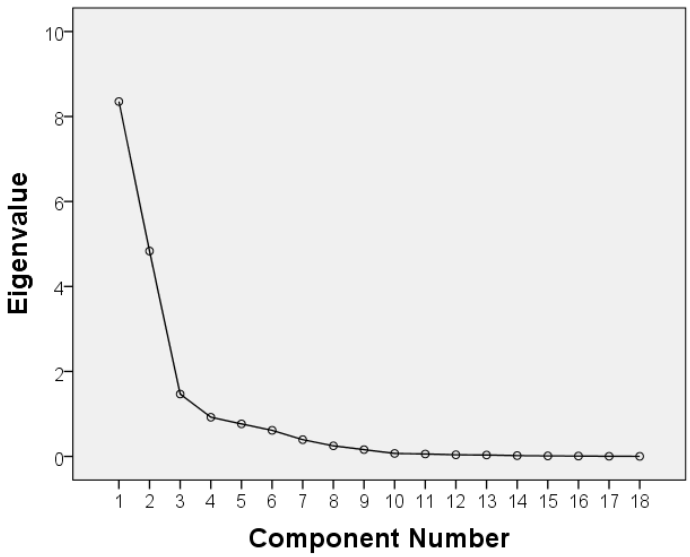
4	MRg_shape_SurfaceArea	
4	MRg_glrlm_RunLengthNonUniformity	
4	MRg_glszm_GrayLevelNonUniformity	
4	MRg_gldm_DependenceNonUniformity	
5	SUV_glszm_LargeAreaEmphasis	
5	SUV_glszm_LargeAreaHighGrayLevelEmphasis	
5	SUV_glszm_LargeAreaLowGrayLevelEmphasis	
5	**SUV_glszm_ZoneVariance**	**0.165**
5	MRg_glszm_LargeAreaEmphasis	
5	MRg_glszm_LargeAreaLowGrayLevelEmphasis	
5	MRg_glszm_ZoneVariance	
5	MRg_ngtdm_Busyness	
6	SUV_glcm_Imc1	0.759
6	MRg_glcm_Imc1	
7	SUV_firstorder_Variance	
7	SUV_glcm_SumSquares	0.654
7	SUV_glrlm_GrayLevelVariance	
7	SUV_glszm_GrayLevelVariance	
7	SUV_glszm_HighGrayLevelZoneEmphasis	
7	SUV_glszm_SmallAreaHighGrayLevelEmphasis	
7	SUV_gldm_GrayLevelVariance	
7	MRg_glcm_Autocorrelation	
7	MRg_glcm_Contrast	
7	MRg_glcm_DifferenceVariance	
7	MRg_glrlm_HighGrayLevelRunEmphasis	
7	MRg_glrlm_LongRunHighGrayLevelEmphasis	
7	MRg_glrlm_ShortRunHighGrayLevelEmphasis	
7	MRg_glszm_HighGrayLevelZoneEmphasis	
7	MRg_glszm_SmallAreaHighGrayLevelEmphasis	
7	MRg_gldm_HighGrayLevelEmphasis	
7	MRg_gldm_SmallDependenceHighGrayLevelEmphasis	
7	MRg_ngtdm_Complexity	
7	MRg_ngtdm_Strength	
7	Dyn_ClusterProminence	
7	Dyn_ClusterTendency	
8	SUV_glrlm_LongRunEmphasis	
8	SUV_glrlm_RunLengthNonUniformityNormalized	
8	SUV_glrlm_RunPercentage	
8	SUV_glrlm_RunVariance	0.715
8	SUV_glrlm_ShortRunEmphasis	
8	SUV_gldm_DependenceVariance	
8	SUV_gldm_LargeDependenceEmphasis	
8	SUV_ngtdm_Busyness	
8	MRg_glrlm_RunLengthNonUniformityNormalized	
8	MRg_glrlm_RunPercentage	
8	MRg_glrlm_ShortRunEmphasis	
25	SUV_shape_MinorAxis	
25	SUV_shape_Volume	0.643
25	MRg_shape_MajorAxis	
25	MRg_shape_Maximum2DDiameterColumn	
25	MRg_shape_Maximum2DDiameterRow	
25	MRg_shape_Maximum2DDiameterSlice	
25	MRg_shape_Maximum3DDiameter	
25	MRg_shape_MinorAxis	
25	MRg_shape_Volume	
26	MRg_firstorder_Uniformity	
26	MRg_glcm_JointEnergy	
26	MRg_glcm_MaximumProbability	
26	**MRg_glrlm_GrayLevelNonUniformityNormalized**	**0.345**
26	MRg_glrlm_LongRunEmphasis	
26	MRg_glrlm_LongRunLowGrayLevelEmphasis	
26	MRg_glrlm_RunVariance	
26	MRg_glszm_GrayLevelNonUniformityNormalized	
26	MRg_glszm_LowGrayLevelZoneEmphasis	
26	MRg_glszm_SmallAreaLowGrayLevelEmphasis	
26	MRg_gldm_LargeDependenceEmphasis	
26	MRg_gldm_LargeDependenceLowGrayLevelEmphasis	
27	**SUV_glcm_Correlation**	**0.199**
27	SUV_glcm_Imc2	
27	MRg_glcm_Correlation	
27	MRg_glcm_Imc2	
28	SUV_firstorder_90Percentile	
28	SUV_firstorder_InterquartileRange	
28	**SUV_firstorder_Maximum**	**0.498**
28	SUV_firstorder_MeanAbsoluteDeviation	
28	SUV_firstorder_Range	
28	SUV_firstorder_RobustMeanAbsoluteDeviation	
28	SUV_glcm_Autocorrelation	
28	SUV_glcm_JointAverage	
28	SUV_glrlm_HighGrayLevelRunEmphasis	
28	SUV_glrlm_LongRunHighGrayLevelEmphasis	
28	SUV_glrlm_RunEntropy	
28	SUV_glrlm_ShortRunHighGrayLevelEmphasis	
28	SUV_gldm_HighGrayLevelEmphasis	
28	MRg_firstorder_10Percentile	
28	MRg_firstorder_90Percentile	
28	MRg_firstorder_InterquartileRange	
28	MRg_firstorder_Maximum	
28	MRg_firstorder_MeanAbsoluteDeviation	
28	MRg_firstorder_Mean	
28	MRg_firstorder_Median	
28	MRg_firstorder_Minimum	

8	MRg_gldm_DependenceVariance	
9	**MRg_glrlm_LowGrayLevelRunEmphasis**	**0.195**
9	MRg_glrlm_ShortRunLowGrayLevelEmphasis	
9	MRg_gldm_LowGrayLevelEmphasis	
10	**SUV_ngtdm_Strength**	**0.43**
11	SUV_glcm_Idmn	
11	SUV_glcm_Idn	0.776
11	MRg_glcm_Idmn	
11	MRg_glcm_Idn	
12	**Dyn_IDMN**	**0.16**
12	Dyn_IDN	
13	SUV_glcm_ClusterTendency	0.642
13	MRg_firstorder_Variance	
13	MRg_glcm_ClusterTendency	
13	MRg_glcm_SumSquares	
13	MRg_glrlm_GrayLevelVariance	
13	MRg_glszm_GrayLevelVariance	
13	MRg_gldm_GrayLevelVariance	
14	SUV_glrlm_LowGrayLevelRunEmphasis	
14	SUV_glrlm_ShortRunLowGrayLevelEmphasis	
14	SUV_gldm_LowGrayLevelEmphasis	0.844
15	SUV_gldm_DependenceEntropy	0.572
15	MRg_gldm_DependenceEntropy	
16	**SUV_shape_Sphericity**	**0.155**
16	MRg_shape_Sphericity	
17	SUV_glszm_SizeZoneNonUniformityNormalized	
17	SUV_glszm_SmallAreaEmphasis	0.534
17	MRg_glszm_SizeZoneNonUniformityNormalized	
17	MRg_glszm_SmallAreaEmphasis	
18	SUV_shape_SurfaceVolumeRatio	0.532
18	MRg_shape_SurfaceVolumeRatio	
19	**MRg_gldm_SmallDependenceLowGrayLevelEmphasis**	**0.172**
20	**SUV_gldm_SmallDependenceLowGrayLeveIEmphasis**	**0.121**
21	SUV_shape_Elongation	
21	**SUV_shape_Flatness**	**0.383**
21	MRg_shape_Elongation	
21	MRg_shape_Flatness	
22	SUV_glrlm_GrayLevelNonUniformity	0.682
22	SUV_gldm_GrayLevelNonUniformity	
22	MRg_glrlm_GrayLevelNonUniformity	
22	MRg_gldm_GrayLevelNonUniformity	
23	SUV_glszm_SizeZoneNonUniformity	
23	**SUV_TLG**	**0.484**
23	MRg_glszm_SizeZoneNonUniformity	
23	MRg_TLG	
28	MRg_firstorder_Range	
28	MRg_firstorder_RobustMeanAbsoluteDeviation	
28	MRg_firstorder_RootMeanSquared	
28	MRg_glcm_JointAverage	
28	MRg_glrlm_RunEntropy	
28	Dyn_Autocorrelation	
28	Dyn_JointAverage	
28	Dyn_SumAverage	
29	SUV_glszm_ZoneEntropy	0.683
29	MRg_glszm_ZoneEntropy	
30	SUV_glcm_Contrast	
30	SUV_glcm_DifferenceVariance	
30	SUV_gldm_SmallDependenceHighGrayLevelEmphasis	
30	**SUV_ngtdm_Complexity**	**0.423**
30	SUV_ngtdm_Contrast	
30	MRg_ngtdm_Contrast	
30	Dyn_Contrast	
30	Dyn_DifferenceVariance	
31	**Dyn_Correlation**	**0.195**
31	Dyn_IMC1	
31	Dyn_IMC2	
32	**SUV_glcm_ClusterProminence**	**0.432**
32	SUV_glcm_ClusterShade	
32	MRg_glcm_ClusterProminence	
32	MRg_glcm_ClusterShade	
32	Dyn_ClusterShade	
33	SUV_firstorder_Uniformity	
33	SUV_glcm_JointEnergy	
33	SUV_glcm_MaximumProbability	0.721
33	SUV_glrlm_GrayLevelNonUniformityNormalized	
33	SUV_glrlm_LongRunLowGrayLevelEmphasis	
33	SUV_glszm_GrayLevelNonUniformityNormalized	
33	SUV_glszm_LowGrayLevelZoneEmphasis	
33	SUV_glszm_SmallAreaLowGrayLevelEmphasis	
33	SUV_gldm_LargeDependenceLowGrayLevelEmphasis	
33	Dyn_JointEnergy	
33	Dyn_MaximumProbability	
34	SUV_ngtdm_Coarseness	0.767
34	MRg_ngtdm_Coarseness	
35	SUV_firstorder_Kurtosis	
35	**SUV_firstorder_Skewness**	**0.112**
35	MRg_firstorder_Kurtosis	
35	MRg_firstorder_Skewness	
36	**SUV_gldm_LargeDependenceHighGrayLeveIEmphasis**	**0.195**
36	MRg_gldm_LargeDependenceHighGrayLevelEmphasis	

Total Variance Explained

Component	Initial Eigenvalues			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	8.352	46.402	46.402	5.670	31.500	31.500
2	4.832	26.846	73.248	4.621	25.672	57.172
3	1.468	8.155	81.404	4.362	24.232	81.404
4	0.923	5.125	86.529			
5	0.766	4.253	90.782			
6	0.614	3.409	94.190			
7	0.394	2.189	96.379			
8	0.249	1.381	97.760			
9	0.160	0.888	98.648			
10	0.070	0.387	99.036			
11	0.057	0.318	99.354			
12	0.038	0.213	99.567			
13	0.033	0.181	99.748			
14	0.019	0.105	99.852			
15	0.012	0.068	99.920			
16	0.009	0.048	99.969			
17	0.004	0.020	99.988			
18	0.002	0.012	100.000			

Extraction Method: Principal Component Analysis.



Rotated Component Matrix[a]

	Component		
	1	2	3
SUV_shape_Volume	,940	,082	-,043
SUV_shape_SurfaceArea	,856	-,006	,078
SUV_glrlm_GrayLevelNonUniformity	,820	,435	-,184
SUV_ngtdm_Coarseness	-,820	-,040	-,131
SUV_glcm_Idn	,790	,371	-,042
SUV_gldm_DependenceEntropy	,783	-,296	,423
SUV_shape_SurfaceVolumeRatio	-,738	-,141	,288
MRg_glszm_LargeAreaHighGrayLevelEmphasis	,603	,372	-,279
SUV_glcm_MaximumProbability	,166	,925	-,123
SUV_glrlm_RunVariance	,428	,854	-,201
SUV_glszm_ZoneEntropy	,107	-,837	,447
SUV_gldm_LowGrayLevelEmphasis	-,012	,705	-,454
SUV_glcm_Imc1	-,543	-,688	,179
SUV_glcm_SumSquares	-,058	-,244	,924
SUV_glcm_ClusterTendency	,024	-,227	,881
SUV_glszm_SmallAreaEmphasis	,028	-,070	,826
SUV_firstorder_Mean	-,155	-,459	,805
SUV_firstorder_Entropy	-,134	-,676	,713

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 5 iterations.

Figure 10: Eigenvalues and percentage variance explained for all components in principal component analysis (left), scree plot of the components (right) and rotated component matrix (below). The first three components were selected for a eigenvalue larger than 1. Also, the 'elbow' of the scree plot can be seen between the third and fourth component. Selection of the components left from the elbow can also be used as a selection criterion for components explaining a sufficient amount of variance in the dataset. Together, the three selected components explain 81.4% of the total variance in the dataset. The rotated component matrix shows that the static volume, the static grey level cooccurrence matrix maximum probability and the static grey level coouccrrence matrix sum of squares show the best correlation with the components (oranges frames).

VERANTWOORDING

Na een succesvolle M2-stage op de nucleaire geneeskunde in het LUMC, startte ik afgelopen januari ook met mijn afstuderen op deze afdeling. Inmiddels zijn we bijna een jaar verder en blik ik terug op dit jaar. Het eerste deel van dit verslag zal gaan over mijn klinische ontwikkeling. Daarna zal ik uitweiden over mijn ontwikkeling als onderzoeker, onderwijs dat ik gevolgd heb en andere dingen waar ik mij mee bezig heb gehouden.

Kliniek

Gezien ik kwantitatieve analyses van scans maak in het kader van mijn onderzoek, wilde ik weten hoe de kwalitatieve (en voor een klein deel kwantitatieve) beoordeling van deze scans in zijn werk gaat. Daarnaast heb ik PET-scans altijd een ontzettend 'sophisticated' manier van medische beeldvorming gevonden, omdat het de fysiologie in beeld brengt en niet zozeer de anatomie. Daarom heb ik het grootste deel van mijn klinische werkzaamheden besteed aan het verslaan van scans en dit laten superviseren door nucleair geneeskundigen. Ik heb verschillende scans beoordeeld, waaronder skeletscintigrafieën, renogrammen en PET-scans voor verschillende indicaties. Ik heb veel geleerd over het systematisch kijken en de manier van verslaan. Ik merk dat ik, zeker bij FDG-PET scans, een redelijk beeld heb wat fysiologische – en wat pathologische opname is. Dit gaat veel beter dan in het begin, want toen had ik alleen al veel moeite met het ontdekken van opvallende opname. Dit valt mij eigenlijk mee, gezien ik van tevoren dacht dat ik het echt nooit zou kunnen. Wel is het een kwestie van veel scans beoordelen, zodat je een goede referentie hebt. Wat ik op dit moment het lastigst vind is het beschrijven van de klinische implicaties voor de patiënt. Ook hier speelt ervaring een grote rol en gezien ik veel scans voor verschillende indicaties beoordeeld heb, zie ik de parallellen minder. Globaal heb ik nu een goed beeld van wat er komt kijken bij de beoordeling van een PET-scan. De kennis die ik heb is zeker niet fundamenteel, maar gezien ik geen nucleair geneeskundige zal worden, is dit prima.

Daarnaast ben ik regelmatig naar het MDO voor longtumoren geweest om een beeld te krijgen van de diagnostiek, behandeling en betrokken specialismen. In dit MDO schuiven longartsen, oncologen, chirurgen, radiotherapeuten, pathologen en nucleair geneeskundigen/radiologen aan en worden de patiënten in een rap tempo besproken. In het begin had ik daarom moeite met het volgen van de redenatie en de rol van de verschillende specialismen hierin. Naarmate ik vaker aanwezig was en de betrokken partijen leerde kennen, snapte ik de manier van redeneren beter en was ik beter in staat mee te denken. Het hielp ook om het MDO voor te bespreken met de nucleair geneeskundige gespecialiseerd in longaandoeningen. Hij kon van tevoren precies uitleggen waar de discussie over zou gaan. Ook heb ik hem veel vragen kunnen stellen over longaandoeningen op PET/CT. Ik bewonder dit MDO voor zijn dynamiek en de soepele samenwerking tussen de verschillende specialismen.

Om mij ook te bekwamen in het patiëntcontact heb ik patiënten gezien op de schildklierpoli. Tijdens mijn M2 had ik al een aantal keer meegekeken met deze poli en ik vond de casuïstiek erg interessant. Ik was vooral gefascineerd door de beleidsfase, omdat daar veel bij komt kijken. Patiënten moeten namelijk helder geïnstrueerd worden over de voorbereiding, eventuele opname en leefregels rond de behandeling. Daarbij komt nog dat straling bij veel patiënten tot de verbeelding spreekt, waardoor hier extra aandacht voor nodig is. Ik heb de beleidsfase altijd het leukste deel van een consult gevonden, omdat ik het interessant vind om bezig te zijn met deze informatieoverdracht en hoe je die aanpast aan het niveau van de patiënt. In het begin vond ik dit wel pittig, omdat ik de verschillende stappen van het beleid nog niet overzag, maar dit ging steeds beter. Ook vond ik het lastig om het consult te structureren, voornamelijk door de grote hoeveelheid informatie. Mijn begeleiders doen de verschillende fases van het consult door elkaar, maar ik houd toch liever een vaste volgorde aan om zo de rode draad niet kwijt te raken. Wat ik uiteindelijk het

lastigst vond was hoe ik moest doorvragen als een bepaald symptoom aanwezig bleek tijdens de anamnese. Ik werd onzeker als ik voor mijn gevoel niet snel genoeg kon beredeneren of en hoe dit symptoom kon passen bij de aandoening en hoe ik hier vervolgens op door moest vragen. Uiteindelijk speelt ervaring hier een belangrijke rol in. Al met al vond ik de consulten leuk om te doen en merkte ik dat het patiëntcontact goed verliep.

Daarnaast heb ik het prikken van infusen herhaald. Tijdens verschillende M2 stages had ik wel eens venapuncties gedaan en infusen geprikt, maar ik merkte dat ik mij niet bekwaam voelde. Hierin speelde mee dat ik dan meestal maar één of twee patiënten prikte, waardoor het lastig was om er gevoel voor te krijgen. Daarom heb ik geregeld dat ik een dag kon meelopen bij de CT-scanners om daar meerdere infusen achter elkaar te kunnen prikken. Zo kon ik actief iets doen met de feedback en kon ik erachter komen wat voor mij een fijne manier was om het infuussysteem vast te houden. Eerder had ik gemerkt dat mijn grootste drempel wat betreft voorbehouden handelingen is dat een supervisor mij teveel op mijn vingers kijkt, vooral als het gaat om een handeling die ik in feite wel kan. Ik merk dat ik dan erg nerveus word, omdat ik het graag goed wil doen, en de handeling daar uiteindelijk niet beter van gaat. Natuurlijk begrijp ik ook dat het voor een supervisor wel fijn is om eerst even mee te kijken om te kijken of iemand geen gekke dingen doet. Ik heb dit gecommuniceerd naar laboranten en doktersassistenten en dit werkte wel heel prettig. Eerst heb ik wel een paar keer onder directe supervisie geprikt, maar de rest van de dag waren supervisors in de ruimte aanwezig, zodat ik ze wel om hulp kon vragen als dat nodig was.

Gezien het toenemende aantal TG'ers en krappe bezetting medisch specialisten leek het een goed idee om fietsproeven over te laten nemen door technisch geneeskundigen. Hiervoor heb ik een plan opgesteld in overleg met afdelingshoofd en laborant en heb ik een aantal keer meegelopen met de fietsproeven om hier een beeld van te krijgen. Uiteindelijk bleek dat de myocardperfusiescans in de nabije toekomst worden verplaatst naar het Alrijne Ziekenhuis in Leiderdorp, waardoor het niet meer de moeite waard was om TG'ers in te werken voor deze procedure.

Tenslotte is er elke dag heilig uur, onderwijs voor arts-assistenten radiologie. Hier wordt klinische casuïstiek vanuit de verschillende secties van de radiologie besproken en soms worden er ook wetenschappelijke artikelen besproken. Ik probeer hier ongeveer eenmaal per week naartoe te gaan om mijn klinisch redeneren te verbeteren. Het is grappig om te zien dat artsen een andere manier van redeneren hebben dan technisch geneeskundigen. Artsen redeneren meer vanuit de differentiaal diagnose, terwijl technisch geneeskundigen meer redeneren vanuit de afwijking op de scan en van daaruit redeneren wat voor gevolgen dit heeft. Ik merkte ook dat ik gedurende dit jaar beter ben geworden in het aanwijzen van de afwijkingen op de scans, in ieder geval op PET en CT, op MRI blijf ik dit toch lastig vinden.

Al met al denk ik dat ik meer uit mijn klinische ontwikkeling had kunnen halen. Dit heb ik niet gedaan, omdat ik mijn toegevoegde waarde niet zozeer zag in de kliniek, waardoor er niet echt een uitdaging in zat voor mij. Daarom heb ik de uitdaging elders gezocht: in mijn onderzoek en het geven van onderwijs. Ik merk namelijk dat ik mijn toegevoegde waarde veel meer vind in het doen van onderzoek dan in de kliniek, tenminste, de kliniek die ik in de afgelopen twee jaar tegengekomen ben. Ik vind het zeker leuk om zo af en toe een scan te verslaan of een patiënt te zien op de poli, maar ik zie dit meer als afwisseling van mijn onderzoek. Ik zie mijzelf dan ook meer als wetenschapper. Ik zou in de toekomst ook zeker kliniek willen doen, maar ik moet wel het gevoel hebben dat ik dat doe met een bepaalde reden en dat niet elke arts dat ook zou kunnen doen. Als er iets op mijn pad komt wat bij mij past, duik ik daar overigens graag in.

Onderzoek

Na vier korte stages met korte onderzoeken keek ik ernaar uit om tijdens mijn M3 meer tijd te hebben om een groter project op te pakken en uitgebreider onderzoek te kunnen doen. Ik vond

het leuk om nu de mogelijkheid te hebben om fundamenteler in de stof te kunnen duiken en dingen uit te proberen. Al is dit afstudeerjaar zeker ook voorbij gevlogen.

In het begin vond ik het lastig om het project te overzien, vooral gezien de grootte. Daarnaast had ik ook het gevoel dat ik ook al een plan moest maken voor mijn gehele promotietraject, waardoor ik het allemaal nog minder overzag. Ik heb dan ook behoorlijk gezwommen, omdat ik niet zo goed wist waar ik naartoe moest. Daarnaast had ik voor dit afstuderen plannen gemaakt om te gaan programmeren in Python, een taal die ik nog niet kende. Hierdoor vond ik het moeilijk een inschatting te maken hoe lang het allemaal zou duren. Ik ben er wel achter gekomen dat veel dingen toch echt langer duren dan ik denk. Gezien ik alles altijd meer, beter en sneller wil, viel dit me soms ook behoorlijk tegen. Soms had ik dan ook het idee dat dit afstudeerjaar meer een persoonlijk project was dan een wetenschappelijk inhoudelijk project. Ik weet ook dat ik vrij streng op mezelf ben. Wat dat betreft was dit jaar ook wel een cursus verwachtingsmanagement en ben ik er absoluut achter gekomen dat je van tevoren niet kunt voorspellen waar je allemaal tegenaan zult lopen.

Ik merkte dat ik het soms lastig vond om naar mijn begeleiders uit te spreken dat bepaalde onderdelen minder goed liepen dan verwacht. Ik heb dan in zekere zin het gevoel dat ik gefaald heb en dat ik het niet goed genoeg doe. Nu typeert het mij ook wel dat ik zelfstandig ben en dingen graag uitzoek, maar ik denk dat ik hierdoor soms wel langer in bepaalde dingen ben blijven hangen dan noodzakelijk was. Dit is ook een punt dat aan bod is gekomen bij de 360 graden feedback rondes. Naar aanleiding van de eerste 360 graden feedback gesprekken hebben we dan ook een extra inhoudelijk overleg te plannen in de week dat we geen promotieoverleg hadden. Dit heeft mij erg geholpen, omdat Dennis en Floris op deze manier meer inzicht hadden in waar ik mee bezig was en zo beter konden bijsturen. De toegevoegde waarde van dit gesprek hebben we ook geëvalueerd tijdens de tweede feedbackronde. Hieruit volgde een gesprek over hoe belangrijk het is om inhoudelijke maar ook persoonlijke problemen te blijven uiten naar je begeleiders toe, zodat ze hierbij kunnen helpen als dat mogelijk is en anders in ieder geval op de hoogte zijn. Hoewel dit, zoals eerder beschreven, niet in mijn natuur zit, merk ik wel dat de drempel hiervoor lager begint te worden. Ik merk namelijk dat er vertrouwen in mij is en ben daardoor minder bang om aan te geven dat iets niet gelukt is. Daarnaast besef ik ook dat tijd besteed aan ‘mislukte’ dingen niet per se weggegooid is, omdat het nog steeds leerzaam kan zijn en ik er eigenlijk nu juist tijd voor heb. Kortom: ik mag best wat meer dingen uitproberen en fouten durven maken.

Naast dit persoonlijke project heb ik ook inhoudelijk veel geleerd. Qua vaardigheden ben ik veel bezig geweest met image processing, heb ik mijn MATLAB skills verbeterd en Python en R geleerd. Inhoudelijk heb ik vooral veel geleerd over (dynamische) PET-scans, de toepassing van radiomics en hoger dimensionele statistiek. Vooral dit laatste is iets waar ik mijn hoofd behoorlijk over gebroken heb, doordat het zo abstract is. Soms vind ik het lastig in te schatten in hoe verre ik tijd moet steken in het fundamenteel snappen van dit soort dingen en heb ik het gevoel dat ik opgeslokt kan raken door details. Het zou dan efficiënter zijn om een statistische methode gewoon te gebruiken en zonder de onderliggende wiskunde uit te diepen. Wat dat betreft weet ik niet of ik altijd de juiste prioriteiten stel. Aan de andere kant merk ik wel dat dit onderliggende begrip helpt om een bepaalde methode beter te begrijpen. Daarom denk ik ook niet dat het erg is om mij, nu ik de tijd ervoor heb, hierin te verdiepen.

Iets wat ik op dit moment lastig vind is het leggen van nieuwe onderzoekscontacten. Dit is beide keren teruggekomen bij 360 graden feedback. Ik merk dat ik het lastig vind om actief contact te leggen, omdat ik onzeker ben over mijn kennis als professional. In eerste instantie hebben we daarom besloten dat ik een presentatie zou geven bij het heilig uur wetenschap op de radiologie. Dit was succesvol, omdat ik na het praatje werd aangesproken door een associate professor van het Laboratorium voor Klinische en Experimentele Beeldverwerking (LKEB) of ik een keer langs wilde komen voor een mogelijke samenwerking. Gezien ik toen met andere dingen bezig was, heb ik deze afspraak nog een tijdje uitgesteld, maar wat denk ik ook meespeelde is dat ik mij eerst nog

wel wat beter wilde verdiepen in datgeen wat we gingen bespreken. Ergens schuif ik dit soort dingen dus ook wel een beetje voor me uit. Toen dit gesprek er een aantal weken later wel kwam, was een grappig detail dat ik mij tijdens dit gesprek getoetst voelde als student door alle vragen die de associate professor stelde. Hij bemerkte dit en gaf aan dat hij die vragen stelde uit interesse en dat hij de antwoorden ook niet wist. Dit geeft wel aan dat ik wel wat meer vertrouwen mag hebben in mijn eigen kennis en kunnen. Ik ben de expert op mijn onderzoeksgebied of ik begin het in ieder geval te worden. Bovendien is het niet erg om een keer iets niet te weten. Dit is nog een gebied waar ontwikkelingsmogelijkheden liggen. Om in ieder geval op de hoogte te zijn van soortgelijk onderzoek binnen het LUMC ga ik zo af en toe naar de research meetings van de LKEB en de MRI groep. Zo heb ik een globaal beeld waar men mee bezig is en zou ik eens contact kunnen zoeken. Dit heeft tot dusver nog niet geleid tot concrete samenwerkingen.

Voorafgaand aan dit jaar leek de wisselwerking tussen mijn afstudeerjaar en promotietraject mij een uitdaging. Ik was bang dat mijn afstudeerproject onder zou sneeuwen door de nevenprojecten die ik alvast op zou zetten voor mijn promotie. In mijn persoonlijk ontwikkelingsplan gaf ik aan dit te willen afbakenen door een heldere planning te maken. Uiteindelijk bleek deze zorg overbodig. Ik heb wel een aantal projecten gedaan die niet direct toegevoegde waarde leverden aan mijn afstudeeronderzoek, maar ze hebben allemaal wel bijgedragen aan bijvoorbeeld mijn persoonlijke ontwikkeling of wetenschappelijke kennis. Uiteindelijk denk ik dat nevenprojecten mij juist wel scherp houden, omdat ik op die manier wat afwisseling kan zoeken qua werkzaamheden. Concreet heb ik een aantal weken besteed aan de verdere uitwerking van mijn M2-onderzoek, heb ik twee onderzoeksprotocollen geschreven voor datatransfers voor retrospectieve analyse en heb ik de Basiscursus Regelgeving en Organisatie voor Klinisch onderzoekers (BROK) cursus gevolgd.

Onderwijs en andere activiteiten

Bij persoonlijke ontwikkeling hoort voor ook het leren van nieuwe dingen, niet eens zo zeer verdiepend, maar meer verbredend. Daarom heb ik afgelopen jaar een aantal cursussen gedaan en aantal symposia en ander onderwijs bijgewoond.

Zo heb ik begin dit jaar het vak multivariate analysis and multidimensional data analysis van de masters Computer Science en Statistical Science for the Life & Behavioural Sciences aan de Universiteit Leiden gevolgd. Ik merkte namelijk dat mijn kennis op het gebied van statistiek niet toereikend was. De statistiekvakken uit de bachelor waren namelijk vrij algemeen en hier was ook niet heel veel van blijven hangen. Daarnaast kwam ik binnen mijn onderzoek in aanraking met hoger dimensionele statistiek en hier wist ik helemaal niks van. Gezien het ging om een vak van 6 EC met veel tussentijdse opdrachten, heb ik ervoor gekozen om geen tentamen te maken en alleen naar de hoorcolleges en een deel van de werkgroepen te gaan. Uiteindelijk was ik blij met deze keuze, gezien ik de hoeveelheid fundamentele wiskunde die naar voren kwam in dit vak enigszins heb onderschat. Uiteindelijk heb ik mij voornamelijk gefocust op de toepassing van de methodes. Tijdens dit vak heb ik ook veel geleerd over het gebruik van R statistics. Er werd namelijk van uitgegaan dat studenten hier al veel van wisten, terwijl SPSS juist werd uitgelegd. Zelf had ik meer ervaring met SPSS, waardoor ik voor R nog wel flink moeite heb moeten doen. Uiteindelijk heeft dit vak me een beeld gegeven over statistische methodes in hoger dimensionele statistiek en de basale toepassing hiervan in R en SPSS.

Ook heb ik afgelopen jaar, alvast ter voorbereiding op mijn PhD, de Basiscursus Regelgeving en Organisatie voor Klinisch onderzoekers (BROK) gevolgd en afgerond. Dit bleek zeker nuttig om inzicht te krijgen in de wet- en regelgeving waar je als klinisch onderzoeker mee te maken krijgt en wat voor proces er vooraf is gegaan aan de data waar ik nu mijn analyses op gedaan heb. Dit heeft me geholpen bij het schrijven van een niet WMO-plichtig onderzoeksprotocol voor retrospectieve data-analyse voor een dataset tijdens mijn PhD.

Daarnaast ben ik naar een aantal symposia geweest. In februari ben ik naar het afscheidssymposium van professor Adriaan Lammertsma, hoogleraar medische fysica, in het

VUmc geweest. Dit symposium was gericht op kwantitatieve en dynamische PET en de toekomst daarvan. Hier heb ik vooral veel geleerd over de historie van beeldvorming met PET. Daarnaast ben ik naar het symposium taakherschikking in de zorg georganiseerd door het MUMC+ geweest. Dit symposium werd georganiseerd naar aanleiding van het evalutatieonderzoek voor technisch geneeskundigen in de wet BIG. Daarnaast ging het over de rol van de physician assistant en verpleegkundig specialist, het wettelijk kader en voorbehouden handelingen. Ook was er afgelopen maand een artificial intelligence symposium vanuit Philips op de afdeling radiologie. Hier kwamen zeven Europese startups hun ideeën pitchen en gingen zij vervolgens in discussie met de medisch specialisten om het hebben over de klinische implementatie. Het was interessant om te zien is hoever sommige bedrijven van de klinische relevantie af stonden. Dit deed me wel weer beseffen dat technisch geneeskundigen zeker nodig zijn om een klinische vraag te vertalen naar een technische oplossing.

Ook ben ik in het kader van mijn onderzoek twee dagen naar Siemens in Zoetermeer geweest. Ik had contact gelegd met een product manager op het gebied van PET/CT van Siemens om additionele reconstructies van de scans te maken. Helaas is dit nier gelukt, maar ik heb wel een kijkje in de keuken gekregen over hoe product ondersteuning door de fabrikant in zijn werk gaat en geleerd over de reconstructie van PET-scans.

Tijdens mijn M2-stage heb ik gebruik gemaakt van een dataset van de EORTC voor radiomics onderzoek. Omdat ik eigenlijk weinig inzicht had in deze organisatie, ben ik naar een symposium gegaan waarin werd uitgelegd wat de EORTC is en waar ze bij kunnen helpen. Dit symposium heeft mij inzicht gegeven in wat er komt kijken bij het doen van wetenschappelijk onderzoek op gebied van onderzoeksopzet, statistisch design, datavergaring, regelgeving, voorwaarpen voor het doen van onderzoek en nog veel meer.

Daarnaast hebben Pim Hendriks en ik halverwege dit jaar een 6-wekelijkse TG-lunch opgezet. We zijn inmiddels met bijna tien M3'ers en afgestudeerden in het LUMC en daarom leek het ons zinvol met elkaar in gesprek te gaan. Tijdens deze lunches bespreken we ontwikkelingen binnen het vakgebied, bureaucratie binnen het LUMC en is er de mogelijkheid te sparren over ons onderzoek. Ook M2-studenten uit Delft zijn betrokken bij deze meetings.

Ook zijn we sinds kort met de studenten en onderzoekers op de afdeling gestart met een researchmeeting. Elke twee weken geeft iemand een update over zijn of haar onderzoek en iemand presenteert een artikel. Op deze manier proberen we interactie tussen de verschillende onderzoeken te vergroten en op de hoogte te blijven van de ontwikkelingen binnen de radiologie en nucleaire geneeskunde.

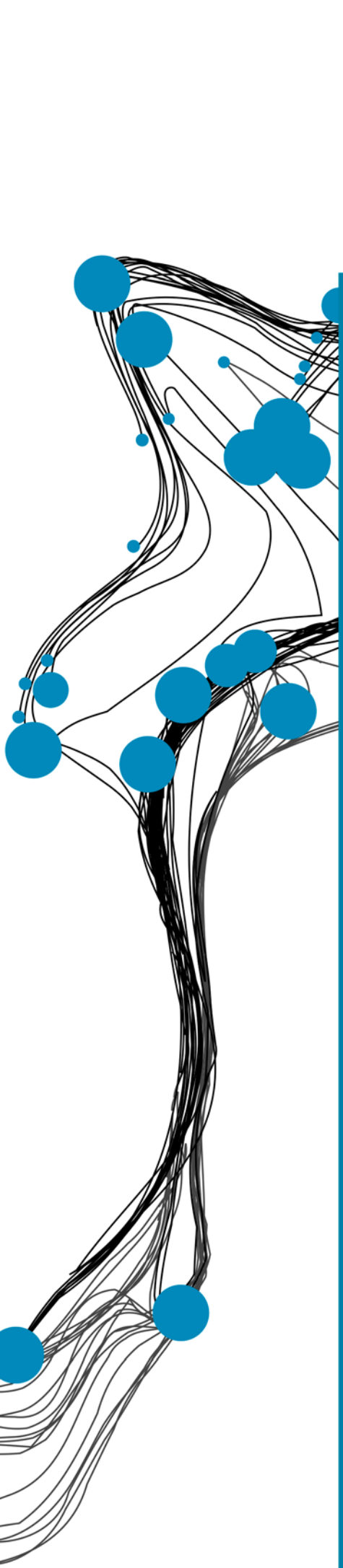
Onderwijs geven

Naast het volgen van onderwijs, vind ik het ook erg leuk om onderwijs te geven. Ik vind het leuk om na te denken over manieren van kennisoverdracht en hoe figuren en animaties hierbij kunnen helpen om een concept uit te leggen. Om wat meer interactie te krijgen met de AIOS'en en andere onderzoekers heb ik een presentatie over radiomics in de bredere zin gegeven bij het heilig uur. Dit was een succesvolle presentatie, want van hieruit kwam ik in contact met de LKEB om deep learning te gaan verkennen en werd ik gevraagd om een uur college te verzorgen binnen de geneeskundeminor Biomedical Imaging. Ook dit college ging over radiomics in het algemeen. Ik vond het leuk om te zien hoe sommige studenten ontzettend geïnteresseerd waren en anderen zich echt afvroegen waar ze beland waren. Ik wil mij zeker verder ontwikkelen op het gebied van onderwijs. Lioe-Fee en ik hebben dan ook de mogelijkheden hiervoor besproken tijdens de 360 graden feedback. Er zijn zeker mogelijkheden om onderwijs te verzorgen aan de geneeskundestudenten in Leiden en misschien ook wel bij klinische technologie in Delft. Daarnaast zijn er cursussen die ik kan volgen, zoals voor het geven van hoorcolleges of de basiskwalificatie onderwijs (BKO) cursus.

Afsluitend

Het afgelopen jaar is voorbij gevlogen. Dat is een goed teken, want ik heb een ontzettend leuke tijd gehad. Ik ben op een fijne plek terecht gekomen met collega's waar ik heel hard mee kan lachen, maar die me ook een spiegel kunnen voorhouden, en begeleiders met wie ik prettig samenwerk en ook door wie ik ook word uitgedaagd op wetenschappelijk - en persoonlijk vlak. Ruim zes jaar geleden startte ik met mijn bachelor technische geneeskunde. Destijds had ik niet durven voorspellen dat ik hier nu zou staan. De afgelopen jaren heb ik veel geleerd en ben ik gegroeid als persoon, als technisch geneeskundige. Ik ben trots op de ontwikkeling die heb doorgemaakt. Gelukkig gaat groei en ontwikkeling altijd door en ik kijk er dan ook naar uit om in januari te starten met mijn promotietraject. Ik ben benieuwd wat de toekomst zal brengen.

Wyanne Noortman,
november 2018



07-12-2018

UNIVERSITY OF TWENTE.