Methods for the integration of combined PET/MR into radiotherapy planning

Dissertation
der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Sara Leibfarth
aus Bad Urach

Tübingen
2015

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen

Tag der mündlichen Qualifikation:	01.06.2016
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Dr. Fritz Schick
2. Berichterstatter:	Prof. Dr. Daniela Thorwarth

Declaration of the framework of the collective work

Parts of this thesis have been published as collective work, or publication is pending. In the following, the contributions of each author are specified.

I. A strategy for multimodal deformable image registration to integrate PET/MR into radiotherapy treatment planning

published in Acta Oncologica 2013, **52**: 1353–1359.

Sara Leibfarth	first author, conception and execution of the study, data analysis, writing of the manuscript
David Mönnich	preparation of imaging data, data analysis
Stefan Welz	definition anatomical landmarks on MRI and CT
Christine Siegel	definition anatomical landmarks on MRI and CT, quality assessment of anatomical segmentations
Nina Schwenzer	PET/MR data acquisition
Holger Schmidt	PET/MR data acquisition
Daniel Zips	critical discussion of the results, proof-reading of the manuscript
Daniela Thorwarth	supervision of the project, study conception

II. Automatic delineation of tumor volumes by co-segmentation of combined PET/MR data

published in Physics in Medicine and Biology 2015, **60**: 5399–5412.

Sara Leibfarth	first author, conception and execution of the study, data analysis, writing of the manuscript
Franziska Eckert	generation of manual tumor delineations
Stefan Welz	generation of manual tumor delineations
Christine Siegel	generation of manual tumor delineations
Holger Schmidt	PET/MR data acquisition
Nina Schwenzer	PET/MR data acquisition
Daniel Zips	critical discussion of the results, proof-reading of the manuscript
Daniela Thorwarth	supervision of the project, study conception

III. Analysis of pairwise correlations in multi-parametric PET/MR data for biological tumor characterization and treatment individualization strategies

published in European Journal of Nuclear Medicine and Molecular Imaging 2016, **43**: 1199–1208.

Sara Leibfarth	first author, conception and execution of the study, data analysis, writing of the manuscript
Urban Simoncic	generation of DCE parameter maps
David Mönnich	PET/CT data acquisition and image post-processing
Stefan Welz	principal investigator of the clinical study in which PET/CT data was acquired
Holger Schmidt	PET/MR data acquisition
Nina Schwenzer	PET/MR data acquisition
Daniel Zips	critical discussion of the results, proof-reading of the manuscript
Daniela Thorwarth	supervision of the project, study conception

List of abbreviations

ADC	apparent diffusion coefficient (DW-MRI)
BEP	bending energy penalty
CNR	contrast to noise ratio
CRT	conformal radiotherapy
CT	computed tomography
DCE-MRI	dynamic contrast enhanced MRI
DP	dose painting
DPBN	DP by numbers
DR	deformable registration
DSI	Dice similarity index
DW-MRI	diffusion-weighted MRI
EPI	echo planar imaging (MR sequence)
EES	extracellular-extravascular space
FDG	[^{18}F]-fluorodeoxyglucose
FMISO	[^{18}F]-fluoromisonidazole
Gd-DTPA	gadolinium-diethylenetriaminepentacetate
GMI	global MI
GTV	gross tumor volume
HN	head and neck
HNC	HN cancer
ITK	Insight Segmentation and Registration Toolkit
IMRT	intensity modulated RT
LMI	localized MI
MI	mutual information
MLC	multi leaf collimator
MLD	mean line distance
MPD	mean point distance
MVD	mean volume distance

MRI	magnetic resonance imaging
NCC	normalized cross correlation
NOF	non-overlapping fraction
OAR	organ at risk
OSEM	ordered subset expectation maximization (PET reconstruction)
PMT	photomultiplier tube
p.i.	post injection
PET	positron emission tomography
PF	prescription function
PMF	probability mapping function
RF	radiofrequency
ROI	region of interest
RR	rigid registration
RT	radiation therapy
RTP	RT treatment planning
SNR	signal to noise ratio
SSD	sum of squared differences
STIR	short time inversion recovery (MR sequence)
SUV	standardized uptake value
TBR	tumor to background ratio
thLS	threshold level set
TPS	treatment planning system
T_1	longitudinal relaxation time (MRI)
T_2	transversal relaxation time (MRI)
TE	echo time (MRI)
TR	repetition time (MRI)
UTE	ultra-short echo time (MR sequence)
VTK	Visualization Toolkit

Contents

1 Introduction

Radiation therapy (RT) is an established method for the treatment of cancer. The goal of RT is the delivery of a high dose of ionizing radiation to the tumor, with the purpose of killing the cancer cells, while minimizing dose to the surrounding normal tissue. It is estimated that 50 % of the total amount of cancer patients would benefit from a RT treatment, either as single treatment modality or in combination with chemotherapy or surgery [1].

The foundation of RT was laid by the discovery of x-rays by Wilhelm Conrad Röntgen in 1895, and the following discovery of natural radioactivity by Henry Becquerel in 1896. The potential medical application of these types of ionizing radiation in terms of diagnosis and therapy was recognized quickly, with first x-ray treatments starting as soon as 1896 [2]. In 1899, the probably first patient, a woman with skin cancer, was cured from a malignant tumor with ionizing radiation [3].

The advantage of fractionated treatment was elaborated by Régaud and Coutard, beginning in 1919 [4]. They found that if radiation dose was applied not in a single treatment, but in various fractions, reduced side-effects for normal tissue could be achieved. This radiobiological concept, which is based on the reduced DNA repair capability of cancer cells compared to normal cells, is still exploited today.

In the beginning, external RT was performed with x-ray tubes, similar to those used in diagnostic radiology. With this technology, tube voltage was limited to about 250–300 kV. This resulted in a high skin dose with a following steep decrease in depth dose, as well as a high absorption in bone. Thus, treatment was limited to superficial tumors. With the emergence of betatrons in the 1940's, MeV x-ray energies became available. In the 1950's, cobalt-60, an emitter of gamma rays with discrete energies of 1.17 and 1.33 MeV, was introduced as an alternative radiation source. With higher radiation energies, energy disposition in deeper body regions became possible. In the 1970's, linear accelerators (linacs) were introduced into clinical routine, which are still by far the most common treatment modality used today. Linacs allow for the application of x-ray energies of about 4–40 MeV [5].

In the phase of early RT, the treatment fields had a very simple, usually rect-

angular, shape and were planned based on bony landmarks seen on x-ray images, i.e. two-dimensional projections of the patients anatomy. In the late 1980's, computed tomography (CT) scanners became increasingly available, allowing for the acquisition of three-dimensional (3D) representations of the patient. With the simultaneous development of computerized *treatment planning systems* (TPS) 3D conformal radiotherapy (3D–CRT) became possible [6]. With this method, radiation fields are typically applied from different angles around the patient, and can be adjusted for each angle in geometrical field shape and the contribution to total dose. Different field shapes are realized either by means of *multi leaf collimators* (MLCs) or by specifically shaped blocks. In addition, non-uniform fields can be achieved by positioning wedges in the radiation field. This allows for a high conformity of dose to the tumor while sparing nearby *organs at risk* (OARs). Organ specific dose constraints exist according to the dose-effect relationship. A prerequisite for 3D-CRT is reproducible patient positioning with adequate patient fixation systems. For reliable dose delivery, methods for imaging of the patient in the treatment room emerged, including cone-beam CT scanners attached to the linac.

In the end of the 1990's, intensity modulated RT (IMRT) has been introduced in the clinic, adding more flexibility to achievable dose distributions compared to conventional 3D-CRT [7]. With this technique, not only the fields from different angles are shaped by the MLC, but also the fluence patterns of the fields, including the possibility to deliver concave fluence cross profiles. This results in a higher degree of flexibility and complexity compared to conventional 3D-CRT.

In 3D-CRT field shapes and wedges are adjusted by the user until a satisfying dose distribution is obtained, a process called *forward planning*. The new degrees of freedom provided by the concept of IMRT raised the necessity of *inverse planning* algorithms. In this planning concept, the user defines objectives for the target volume and sets dose constraints for OARs. The treatment planning algorithm then determines MLC configurations for which the resulting dose distribution is optimal with respect to these constraints. For dose calculation, Monte Carlo algorithms [8] emerged, which provide most accurate results, but typically are only used in the final phase of the planning process due to their high computational demands.

In addition to photons, also protons and heavier ions, such as carbon ions, are being evaluated for treatment [9, 10]. Such *particle therapy* is already applied clinically, although to a much lesser extent. Proton and ion radiation is characterized by a clinically favorable depth dose profile, consisting of a slowly increasing energy disposition up to the very sharp *Bragg peak*. However, potential advantages of particle therapy over photon therapy depend on the treatment

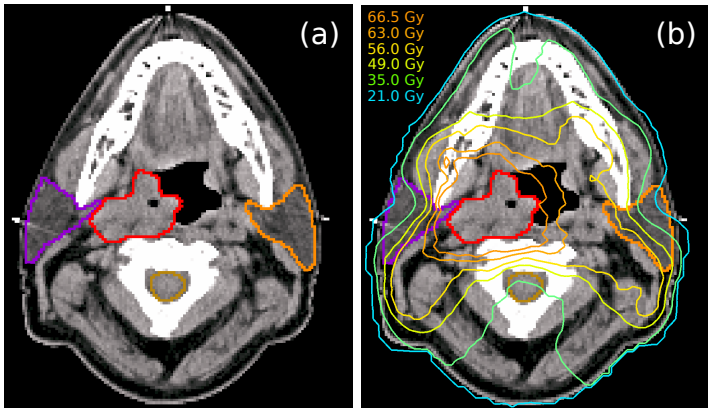

Figure 1.1: (a) Planning CT of a HNC patient with delineation of the gross tumor volume (GTV) (red) and OARs. OARs visible in this axial representation consist of left parotid gland (orange), right parotid gland (purple) and myelon (gold). Patient fixation in treatment position is obtained using thermoplastic masks. (b) Visualization of the planned total dose distribution for IMRT treatment of this patient by corresponding dose isocontours. A total dosis of 70 Gy was prescribed to the GTV. The dose sparing of the OARs can be appreciated.

site and have to be carefully evaluated in clinical studies.

In current clinical practice, standard of care photon treatment is delivered a fractionated setting with linacs using IMRT. Before the start of RT, a CT is acquired in treatment position, providing the electron density information needed for dose calculation. On this *planning CT*, the *gross tumor volume* (GTV) and OARs are manually delineated (cf. Figure 1.1 a). With the prescription of a high homogeneous dose to the target volume consisting of the GTV and additional safety margins, and specific dose constraints for the OARs, a plan is created with the TPS (cf. Figure 1.1 b). As an example, a typical RT treatment scheme for patients with head and neck cancer (HNC) consists of a dose prescription of 70 Gy to the tumor, applied in 35 daily fractions of 2 Gy each.

Even with the treatment concepts and techniques applied today, there are still tumor types for which a high fraction of recurrences is observed following RT. For HNC patients with locally advanced primary tumors, treated with a state of the art combination of radio- and chemotherapy, up to 50 % local recurrences are observed [11]. Given the high accuracy in dose planning and dose delivery available today, uncertainties in image-based tumor delineation and the missing consideration of biological heterogeneity of tumors in current treatment concepts seem to be major factors limiting the success of RT [12, 13]. In this regard,

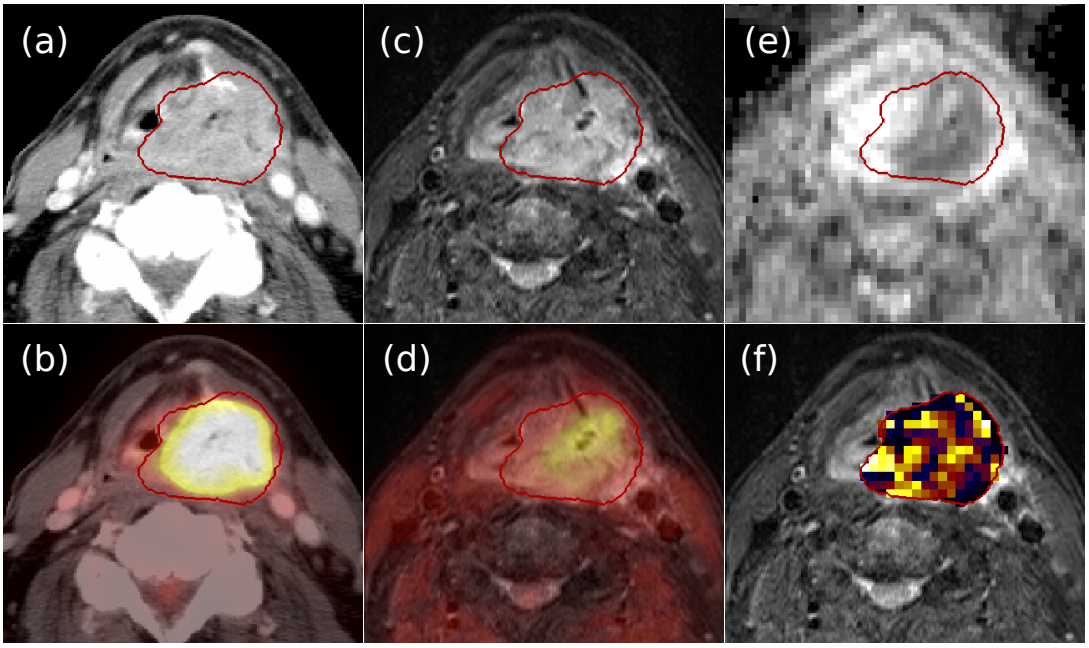

Figure 1.2: Exemplary CT-, MR-, and PET-based imaging data of a HNC patient with cancer in the hypopharynx, with GTV shown in red. Contrast-enhanced CT (a), fusion of FDG-PET and CT (b), anatomical MR (T2w STIR) (c), fusion of FMISO-PET 3 h after tracer injection and MR (d), ADC-map from DW-MR (e), K^{trans} parameter map derived from DCE-MR within the GTV (f). FDG-PET and CT were acquired in PET/CT scanner, all other datasets were acquired in integrated PET/MR scanner.

advanced multimodal imaging methods as well as accordant image processing algorithms may be of high potential for improving RT outcome.

Both magnetic resonance imaging (MRI) and positron emission tomography (PET) are valuable imaging modalities for complementing the CT acquired for treatment planning with additional anatomical and functional information. In MR imaging, anatomical, as well as functional imaging, such as diffusion weighted MR (DW-MR) and dynamic contrast enhanced MR (DCE-MR), are promising [14–16]. PET tracers with potential benefit for RT are the glucose analog [^{18}F]-fluorodeoxyglucose (FDG) and dedicated hypoxia tracers such as [^{18}F]-fluoromisonidazole (FMISO) [17–19] (cf. Section 3.1). To obtain a visual impression of the different imaging modalities and methods, Figure 1.2 provides exemplary slices of imaging data of a HNC patient.

PET and MR data may be beneficial in various steps of the RT workflow. On the one hand, it is expected that these datasets will improve the accuracy of

tumor volume delineation, which in actual clinical practice is often based on the planning CT only. For some tumor entities, the usage of FDG-PET or anatomical MR images for tumor volume delineation in addition to the CT is already standard of care. As an example, anatomical MRI is used to support the delineation of meningiomas [20], and FDG-PET the delineation of target volumes in non-small cell lung cancer [21]. Moreover, biological differences of tumors as revealed by functional imaging are currently investigated as explanation for heterogeneous treatment response. PET- and MR-based functional imaging data may be used to stratify patients and individualize treatment, hopefully resulting in increased treatment response rates and lower toxicities. In this context, one may switch from the paradigma of a homogeneous tumor dose to the adaption of radiation dose according to local biological characteristics of the tumor. Such biological treatment individualization based on functional imaging is currently evaluated in clinical studies [22–24].

Recently, combined PET/MRI scanners became available, allowing for the acquisition of PET and MR data in a single imaging session. This may facilitate the integration of PET and MR information into RT [25]. However, to fully exploit the potential of this modality, adequate image processing algorithms have to be developed. This thesis focuses on the development and evaluation of algorithms for the spatial registration of PET/MR data to the planning CT and automatic PET/MR-based co-segmentation to standardize tumor volume delineation. In addition, an analysis of functional PET and MR imaging datasets for complementarity is performed. It is intended to provide a basis for the integration of multimodal imaging data into RT and as a step towards biological treatment individualization.

2 Objectives and outline of the thesis

This thesis consists of three parts, each dealing with one aspect concerning the integration of combined PET/MR into RT treatment planning. The studies were performed on image datasets of HNC patients, but the methodology could in general also be extended to other tumor sites. Image processing algorithms were implemented in C++ with the usage of the libraries ITK (Insight Segmentation and Registration Toolkit, www.itk.org) and VTK (Visualization Toolkit, www.vtk.org). In the following, an overview of the performed studies is provided.

- **Part I (Chapter 4)** deals with the development of an algorithm for the deformable registration of MR and CT datasets in the HN region. The motivation of this study was to enable the integration of PET/MR-derived information into RT treatment planning (cf. Section 3.3.2). Previous literature addressing such multimodal deformable registration methods were rare, especially for the HN region no thoroughly evaluated algorithm was available (cf. Section 4.1). Thus, the objective of this project was to develop and evaluate an algorithm with adequate accuracy for RT purposes. Registration methods with a B-spline parametrized transform were considered, using a metric composed of a mutual information based similarity measure and a regularization term. Quantitative measures were introduced for a thorough evaluation of registration accuracy.

- **Part II (Chapter 5)** For a high-precision RT treatment, an accurate tumor volume delineation is essential. For this purpose, the combination of information from PET and MR datasets may be highly beneficial. Moreover, compared to a purely manual delineation of the tumor, an automatic algorithm may be valuable for standardization of this task. Thus, an algorithm for segmenting tumor regions by simultaneously considering information from FDG-PET and anatomical T2-weighted MRI was developed. The algorithm contains both connected threshold-based and model-based segmentation approaches (cf. Section 3.3.3). PET information is used to inform MR segmentation, which is intrinsically difficult due to the lack of a standard intensity scale in MRI. The algorithm was quantitatively eval-

uated against manual contours from three different radiation oncologists, by comparing the variability between algorithm and each observer to the variability between the observers.

- **Part III (Chapter 6)** Tumor characterization based on functional PET and MR data might be valuable for biologically individualized RT. For an effective integration of the various available functional datasets into RT, it is interesting to evaluate if these datasets are completely complementary or partially redundant (cf. Section 3.3.4). As a first step for addressing this question we performed a pairwise correlation analysis of functional multi-parametric PET and MR data, comprising FDG-PET, FMISO-PET, DW-MR, DCE-MR datasets. This work extends previous studies, which performed such analysis on a limited set of functional imaging information (cf. Section 6.1). The objective of the study was to quantify the level of correlation between the available functional datasets, as well as the variability between patients. It is intended as an hypothesis generating pilot study for further analysis in this field.

3 Scientific background

3.1 Physics of MR and PET imaging and RT-specific requirements

3.1.1 Magnetic resonance imaging

In the following, a short overview of the basics of MR imaging is provided. For further details, the reader is referred to MR standard literature such as [26, 27].

MR imaging is based on atomic nuclei with non-zero spin $\boldsymbol{S}$, which are associated with a non-vanishing magnetic moment $\boldsymbol{\mu} = \gamma \boldsymbol{S}$. The proportionality constant γ is the nucleus-specific *gyromagnetic ratio*. Clinical imaging is typically based on hydrogen nuclei (i.e. protons), which possess a spin quantum number of 1/2. Proton-based MRI is of advantage due to the natural abundance of hydrogen in the body and the high proton gyromagnetic ratio.

In the MR tomograph, a high static magnetic field $\boldsymbol{B}_0$ is created by superconducting coils, with a typical magnitude of 1.5 T or 3 T for clinical systems. As a consequence, nuclear spins in the tomograph precess around the field direction with the *Larmor frequency* $\omega_\mathrm{L} = \gamma B_0$. Each proton spin can either be oriented parallel or antiparallel to the magnetic field, with the associated potential energy $E = -\boldsymbol{\mu}\boldsymbol{B}_0$ yielding $E = -\frac{1}{2}\gamma\hbar B_0$ and $E = +\frac{1}{2}\gamma\hbar B_0$, respectively. In thermodynamic equilibrium, these energy levels are populated according to the Boltzmann distribution, with an occupation surplus of the more favorable parallel orientation of about 1/10 000 for 3 T systems at room temperature. This leads to a longitudinal macroscopic magnetization parallel to the magnetic field. By applying a radiofrequency (RF) pulse of resonant frequency $\omega_\mathrm{HF} = \omega_\mathrm{L}$ perpendicular to $\boldsymbol{B}_0$, the macroscopic magnetization is tilted towards the transverse plane. The resulting transversal magnetization leads to the induction of an alternating voltage to receiver coils, which constitutes the measured signal of the imaging system. Without further disturbance, a relaxation of the spins back into the thermodynamic equilibrium occurs.

There are two physical mechanisms of relaxation. The first, the *longitudinal relaxation*, causes the longitudinal magnetization to return to equilibrium. This relaxation mechanism is associated with an energy transfer to the surrounding.

The time period until the equilibrium magnetization is reached to a factor of 1/e is called *longitudinal relaxation time* T_1. The second relaxation mechanism is the *transversal relaxation* with the associated *transversal relaxation time* T_2. The underlying mechanism is caused by spin-spin interactions, resulting in a dephasing of the spins, and therefore the decay of transversal magnetization. Both T_1 and T_2 differ locally within heterogeneous objects.

Spatial encoding is realized by superimposing the static magnetic field $\boldsymbol{B}_0$ with linear magnetic field gradients, which are produced by dedicated *gradient coils* within the MR tomograph. These gradients lead to a spatial variation of Larmor frequencies within the measured object. For three-dimensional spatial encoding, image gradients are applied in the three orthogonal directions during the imaging process.

Contrast in anatomical MR images is determined by tissue-dependent differences in the physical properties influencing MR signals, i.e. the relaxation times T_1 and T_2, and the proton density ρ. In addition to anatomical imaging, also *functional imaging* techniques are available, revealing information about physical and physiological parameters rather than anatomical information.

In the following, types of MR images and potential RT-related fields of application are outlined.

- **Anatomical MR** images provide superior soft tissue contrast compared to CT acquisitions, and may thus facilitate tumor delineation, and possibly also the delineation of OARs, for RT treatment planning. Another advantage of MR compared to CT is the reduced extent of dental-inlay artifacts [28]. A wide variety of different image contrasts can be obtained by using different MR sequences. Depending on which of the physical contrast-influencing parameters prevails in the acquired image, the image is termed T_1-weighted, T_2-weighted, or proton density (ρ)-weighted.
- **Diffusion weighted MR** (DW-MR) is a functional imaging technique allowing for the quantification of the diffusion of water molecules in tissue. For diffusion weighting, an additional gradient followed by a reversing gradient of the same magnitude and duration is applied between the RF excitation and signal collection. For stationary spins this gradient pair has no effect, whereas for diffusing spins dephasing occurs in an extent related to the local diffusion coefficient. Since DW-MR images are highly sensitive to motion, rapid image acquisition is necessary. This is typically realized by using *echo-planar imaging* (EPI) sequences. From the acquisition of images with different b-values, i.e. different strengths of diffusion weighting, a physical parameter map containing apparent diffusion coefficients

(ADCs) can be derived (cf. Section 3.3.4). ADCs represent the effective water diffusion coefficients, being influenced by the biological environment such as membranes, and can be related to cellular density [29]. ADC maps derived from DW-MRI have been found to be of value for the staging of lymph nodes [30], as well as for prognosis of RT outcome [14].

- **Dynamic contrast-enhanced MR** (DCE-MR) is a functional imaging technique based on the injection of an MR contrast agent, such as the Gadolinium (Gd) chelate Gd-DTPA. Since Gd is paramagnetic, it causes a decrease of T_1 relaxation times in its surrounding, resulting in a concentration-dependent signal increase in T_1-weighted images. In DCE-MR imaging, a series of T_1-weighted MR images are acquired, from which the dynamics of the distribution of the contrast agent in blood pool and tissue can be determined. By fitting the resulting signal time curves with compartment models, parameter maps revealing a quantitative description of microvasculature properties can be derived (cf. Section 3.3.4). Characterization of the deficient vasculature of tumors may yield valuable information for RT. Compartmental analysis of DCE-MRI has proven to provide parameters associated with locoregional control after RT [15, 16].

Compared to MR images acquired for diagnostic purposes, MR images for RT may have special requirements. Diagnostic images are often acquired with two-dimensional sequences, in which imaging data is collected slice-by-slice. These images are usually characterized by a high resolution in-plane, and a low resolution in slice-encoding direction. Three-dimensional imaging sequences, which enable a high resolution in all three dimensions, may be more adequate for tumor volume delineation. A further difficulty may arise due to geometrical inaccuracies associated with MR imaging. Such inaccuracies can arise due to inhomogeneities in the static field $\boldsymbol{B}_0$, due to deviations in the linearity of gradient fields, or due to susceptibility or chemical shift artifacts. For anatomical sequences and field strengths in current clinical scanners these inaccuracies may be negligible [31, 32], whereas for EPI-based images, such as DW-MR images, geometrical distortions may hamper usage in RT without an adequate distortion correction scheme [33]. Another difficulty, especially for image processing algorithms, may be caused by MR signal intensity inhomogeneities, caused by locally differing receiver coil sensitivities or inhomogeneous RF excitations [34,35]. For facilitating the integration of MRI into the RT treatment planning process, the possibility of acquiring MR images in RT-specific position is being evaluated [28, 36].

3.1.2 Positron emission tomography

PET is a functional imaging technique, which allows to visualize metabolic processes in the body. For image acquisition, dedicated *tracers*, i.e. specific molecules labeled with positron emitting radioisotopes, are injected into the patient. Positrons originating from the radioactive decay annihilate after a short free path with an electron in the tissue. As a result, two 511 keV photons are emitted under an angle of approximately 180 degrees. For detecting these photons, the PET scanner is equipped with a ring of independent detectors. Photon detection allows for the reconstruction of the three-dimensional tracer distribution within the patient. For reconstruction, each coincidence detection, i.e. each detection of two photons within a pre-defined short time interval, is attributed to an annihilation event on the line connecting the two involved detectors. In modern PET systems, iterative reconstruction algorithms are used, such as the ordered subset expectation maximization (OSEM) algorithm.

Different PET tracers exist with which different medical questions can be addressed. The overall molecular structure of the PET tracer determines its distribution within the patient and potential involvement in biological pathways. An important physical property for the choice of the radioactive isotope is its half life. Moreover, the distribution of the kinetic energies of emitted positron is of relevance, since higher positron energies translate into longer mean free path lengths before annihilation, therefore restricting spatial resolution of PET images.

Two PET tracers which are of interest for RT are described in the following. The radioactive isotope for both tracers is ^{18}F, which has a half life of 110 min.

- **FDG** Compared to normal cells, tumor cells show an increased amount of glycolysis. This can be both in the case of deprivation of oxygen (Pasteur effect) [37], and under normal oxygenated conditions (Warburg effect) [38, 39]. Glucose metabolism can be monitored using the tracer and glucose analog [^{18}F]-fluorodeoxyglucose (FDG), which differs from glucose by the replacement of one hydroxyl group by the positron emitter ^{18}F. FDG-PET can be of value in RT in the field of tumor volume delineation and lymph node staging [40]. Furthermore, the metabolic tumor volume (MTV) derived from FDG-PET, as well as the total lesion glycolysis (TLG) and the SUV_{max} have been found to be related to treatment outcome [17, 41].

- **FMISO** The irregular growth of tumors is associated with a deficient microvasculature, due to which the supply of tumor cells with oxygen may

be hampered. The state of reduced oxygen levels in tissue is termed *hypoxia*. Hypoxic conditions can be caused by insufficient vascularization density (chronic hypoxia) or by dynamic changes in perfusion (acute hypoxia). Hypoxia has been related with radioresistance [42]. The hypoxic status of the tumor can be assessed using dedicated tracers such as FMISO, which selectively binds to macromolecules in cells with reduced oxygen level. Static PET protocols typically envisage image acquisition at about 4 hours post injection (p.i.), in the late retention phase [18, 19, 43]. However, also dynamic scan protocols exist, in which an additional dynamic scan is acquired during the first minutes p.i. Both static FMISO-PET and parameters derived from compartmental analysis of dynamic FMISO-PET (cf. Section 3.3.4) have been related to RT outcome [18, 19, 43, 44].

Today, stand-alone PET scanners are increasingly being replaced by PET/CT scanners [45]. The additional CT acquired with these combined systems provides anatomical information which is largely missing in PET images, and can further be used for CT-based PET attenuation correction (cf. Section 3.2.3).

Clinical PET images show low spatial resolution compared to MR and CT images (cf. Section 3.3.1), which is approximately 5 mm for clinical whole-body systems [45]. In comparison, the typical grid resolution for RT dose calculation is 3 mm. By the related partial volume effect, dimensions of small tracer-avid structures may be overestimated, with a simultaneous underestimation of the related activity. Moreover, PET images are usually characterized by a high level of noise due to the Poisson statistics of the radioactive decay, and limitations concerning the injected dose and the time interval of data collection. FMISO-PET images typically show low contrast to noise ratios due to the slow accumulation of FMISO in the cells combined with a slow clearance from blood and background [42]. Difficulties for the interpretation of FDG may arise due to physiological or inflammation-caused accumulation [46].

3.2 Combined PET/MR imaging

Combined PET/MR scanners have recently been developed [47–49], and are currently commercially available in two designs. A sequential design is provided by the Philips Ingenuity TF (Philips Healthcare, Hamburg, Germany), in which the PET and MRI scanners are separate, but the patient can be imaged on a common rotating bed which is transferred between these scanners. On the other hand, with the Siemens Biograph mMR (Siemens Healthcare, Erlangen, Germany) a fully integrated design exists in which the PET detector ring is

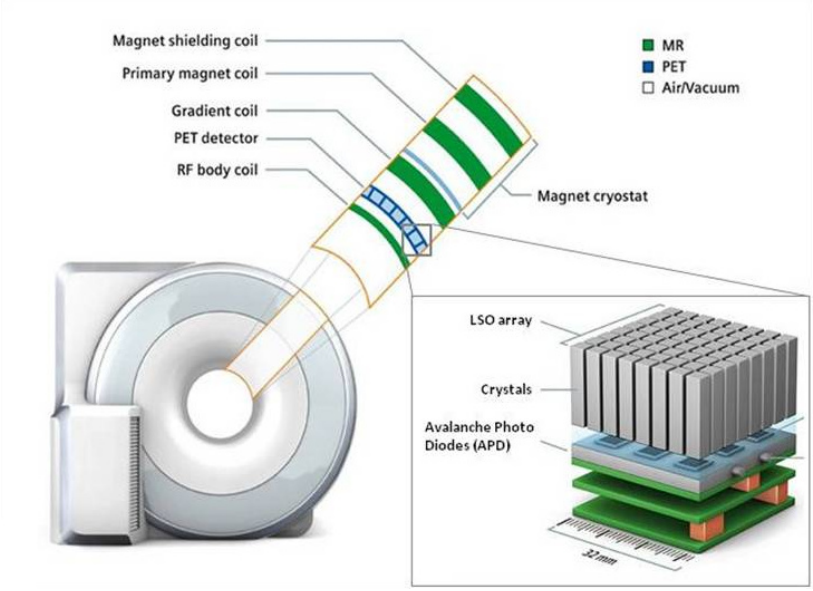

Figure 3.1: Schematic representation of integrated PET/MR scanner. Courtesy of Siemens Healthcare, Erlangen, Germany.

directly integrated into the MR tomograph. With both designs, intrinsically registered PET and MR images can be obtained, while the integrated design additionally allows for the simultaneous acquisition of PET and MR data. The configuration of a fully integrated PET/MR system, as realized by the Siemens Biograph mMR, is illustrated in Figure 3.1.

3.2.1 Perspectives for RT

The intrinsic registration provided by combined PET/MR facilitates the comparison and combined analysis of PET- and MR-derived data, and may thus be valuable in different steps of the RT workflow [25, 50]. Figure 3.2 illustrates a potential RT workflow under exploitation of combined PET/MR data. As displayed in the left column, PET/MR might be of benefit basically in all steps in which imaging data is required. The data processing part displayed in the central column will be referred to in Section 3.3.

PET/MRI might be already of value in the process of *diagnosis* and *staging*, on which the decision of the treatment modality as well as the indication for a palliative versus a curative treatment is based. Concerning the treatment modality, RT, eventually combined with chemotherapy, or other treatments like surgery may be applied.

Given an indication for RT treatment, anatomical MR images, potentially

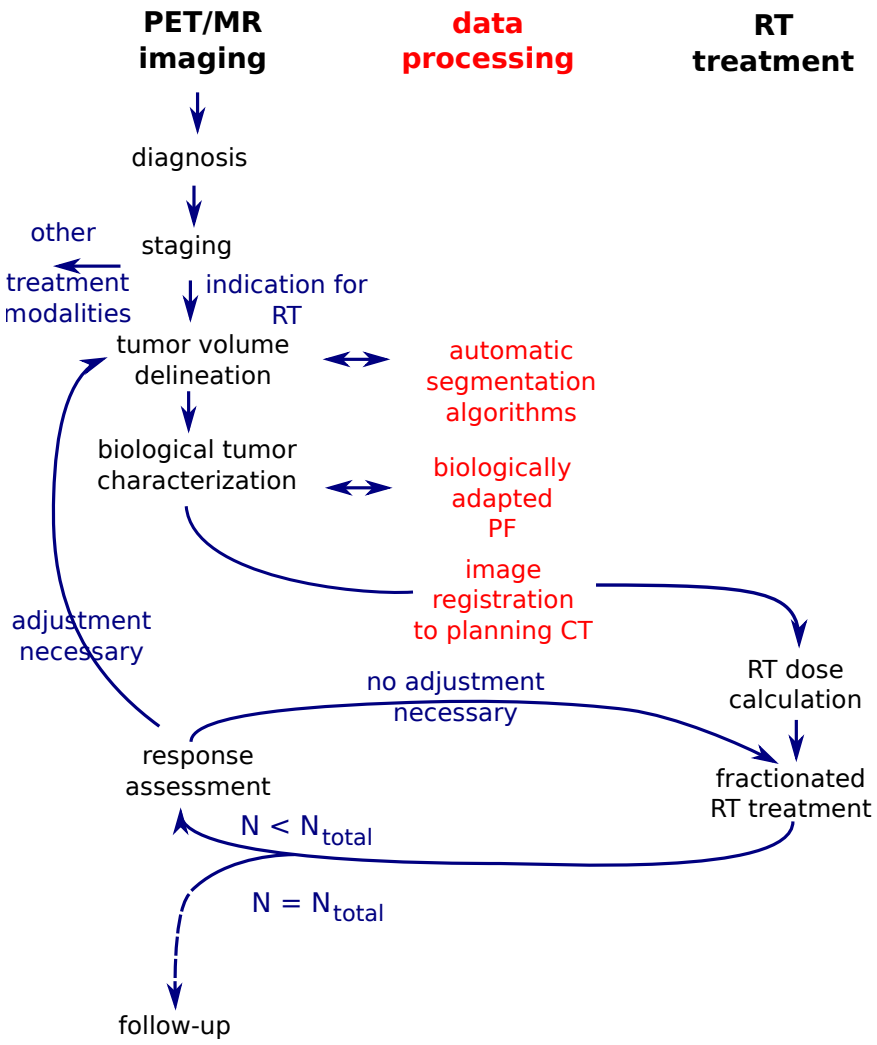

Figure 3.2: Potential RT workflow under exploitation of combined PET/MR imaging. The left column shows potential fields of application for which PET/MR imaging may be beneficial in the course of RT. The data processing steps in the central column are explained in detail in Section 3.3. After a given number N of RT fractions, the response to treatment may be assessed, and treatment plans potentially adapted. If total number of fractions N_{total} are applied, the RT treatment is completed. Treatment success is finally assessed in a later follow-up examination.

combined with functional PET and MR images could assist in an accurate *tumor volume delineation* (cf. Chapter 5). A standardization of this task may be supported by automatic segmentation algorithms.

Additionally, functional PET and MR images, such as FDG-PET, FMISO-PET, DW-MR, and DCE-MR may be used as basis for a *biological characterization* of the tumor, based on which the planned tumor dose might be locally adapted (cf. Chapter 6). For this purpose, a pre-defined biologically adapted dose prescription function (PF) has to be available. The objective of this PF is to achieve homogeneous control probabilities within the tumor, given the underlying heterogeneous biology.

To transfer this PET/MR-derived information to the planning CT, accurate image registration is a major prerequisite (cf. Chapter 4). RT treatment planning can then be performed under consideration of the PET/MR-based tumor delineation and dose adaptation strategies.

In the course of fractionated RT, additional PET/MR imaging sessions could help in the *assessment of treatment response*. For the early response assessment, functional imaging modalities may be superior to anatomical imaging [51–53]. On the other hand, anatomical imaging can be used to adapt treatment plans according to interfractional anatomical changes, as for example tumor shrinkage. Such *adaptive RT* concepts might increase complication-free tumor control, since the original plan may lead to underdosages in the tumor or overdosages in OARs due to the change in patient geometry [54, 55].

Finally, after the end of RT, PET/MR can assist in *follow-up* examinations [52, 56]. These examinations are performed several months after RT, in order to check if RT treatment was successful or if there is a recurrence of the tumor.

3.2.2 Technical aspects of the fully integrated design

The main challenge in the design of a fully integrated PET/MR scanner is the minimization of interactions between MRI and PET.

Conventional PET detectors consist of a scintillator such as a lutetium oxyorthosilicate (LSO), which convert the 511 keV photons into visible light, and photomultiplier tubes (PMTs), which subsequently translate the light into an electrical signal. While the scintillators are suited for operation in an MR tomograph, the PMTs are highly sensitive to magnetic fields, since electrons are deflected from their path between the dynodes by the Lorentz force. Thus, PET detectors of current PET/MR systems are build of scintillators combined with magnetically insensitive *avalanche photodiodes*. These semiconducter-based diodes have been shown to operate in high magnetic fields up to 9.4 T without

degradation of performance [57]. As a further point, copper screens are necessary for shielding PET electronics from gradients and RF pulses.

It was found that the integrated PET hardware of the Siemens Biograph mMR is competitive to the state-of-the-art photomultiplier tube-based hardware in PET/CT scanners [48]. In addition, the MR subsystem was found to perform essentially like a stand-alone system [48, 58]. However, the accuracy of PET quantification may be reduced in PET/MR compared to PET/CT systems, due to inaccuracies in MR-based PET attenuation correction [59] (cf. Section 3.2.3).

3.2.3 MR-based PET attenuation correction

Background

The photons created in the annihilation process have to transverse tissue, and potentially hardware components, before reaching the PET detectors. This leads to interaction processes via the photoelectric effect and Compton scattering, causing a decrease of the original photon intensity. The photon attenuation caused by the transit of a material with position-dependent linear attenuation coefficient $\mu(x)$ is in general described by the exponential law of Lambert-Beer

$$I = I_0 \exp\left(-\int \mathrm{d}x\, \mu(x)\right), \tag{3.1}$$

with I_0 being the original photon intensity, I remaining intensity, and the integration is over the photon path which would be followed without interaction processes. If PET image reconstruction is performed without considering this attenuation, the true PET activity concentrations will be underestimated. Thus, to yield correct quantitative PET information, an accordant *attenuation correction* has to be performed. For this purpose, the local attenuation coefficients μ for the photon energy of 511 keV have to be known. These coefficients are typically derived from additionally acquired images (see below) and provided in the form of an *attenuation map (μ-map)*, containing these coefficients in a voxel grid.

In stand-alone PET scanners, attenuation correction is based on a transmission scan from rotating Ge-68 sources, which, mediated by positron decay, emit 511 keV photons. In this case, attenuation information for the PET emission scan is readily provided at the appropriate photon energy. For the nowadays more common PET/CT scanners, an accordant μ-map is derived from the CT image. Since CT images are acquired with a continuous x-ray spectrum, a conversion to attenuation coefficients for photon energies of 511 keV has to be performed. The conversion can be readily performed using a bilinear function [60].

Object-related attenuation correction in PET/MR systems

Attenuation correction for PET/MR systems is more challenging than for PET/CT, since there is no simple relation of MR signal intensities to electron densities. However, different approaches have been developed performing MR-based PET attenuation correction, which either derive a *substitute CT* from a combination of MR images, or derive the μ-map directly.

On the one hand, there are segmentation approaches to derive μ-maps from MR images. The MR-based segments are intended to represent different tissue types, to which bulk attenuation coefficients are assigned accordingly. In a simple approach, fat and soft tissue (i.e. water) segmentations are identified using a Dixon sequence, and remaining voxels are classified as either lung tissue or air, or as an intermediate between fat and soft tissue* [61, 62]. To improve the accuracy of this approach, ideally a separate segmentation of bone tissue should be available. Since compact bone has very short transversal relaxation times of about $0.4 - 0.5\,\mathrm{ms}$ [63], they hardly show any signal in conventional MR sequences, which impedes the discrimination of bone from air. To obtain signal from bone, *ultra-short echo time (UTE) sequences* have to be applied [64].

In recent research, also the derivation of CT values on a voxel basis from the combination of several MR sequences is evaluated. To learn the relation between such multispectral MR data and CT values, a previous training dataset consisting of registered images has to be available. Training can be based either on classification [65, 66], or on regression [67, 68]. Once the training has been performed and validated, newly acquired sets of the same multispectral MR data can be readily assigned with CT values, allowing for the derivation of a μ-map for attenuation correction.

Another approach is to use deformable image registration to transfer image information from CT acquisitions of different patients to MR data acquired within the PET/MR imaging session. For this purpose, an *anatomical atlas*, consisting matched MR-CT acquisitions from different patients can be used as a basis [69,70]. The MR images of the atlas can be registered to the newly acquired MR data, which, due to the additionally available CT atlas data, allows for the assignment of substitute CT values to the newly acquired MR data.

Hardware-related attenuation correction in integrated designs

Hardware components being introduced into the PET detector's field of view (FOV), such as the MR table and RF receiver coils, have to be considered for

*The commercially available Siemens Biograph mMR (Siemens Healthcare, Erlangen, Germany), which was used for data acquisition in this work, uses this method for PET attenuation correction

attenuation correction in integrated PET/MR. For the MR table and rigid standard RF coils with pre-defined positions, such as standard head coils, manufacturer-provided μ-maps provided with the imaging software are usually available. For flexible RF coils with no reproducible positioning, attenuation correction is not straightforward due to the unknown coil location and geometry. Thus, these coils are usually neglected in attenuation correction. However, techniques have been presented for MR-based position determination of flexible coils, demonstrating that their attenuation correction is feasible [71].

3.3 Image processing for PET/MR data

For the integration of PET/MR data into RT treatment planning adequate image processing algorithms are mandatory. In the following, some of the associated challenges with focus on methods and algorithms which are of relevance for this thesis are described. We begin with a short mathematical description of medical images.

3.3.1 Characteristics of medical images

Description of medical images

A *continuous* image I_c is a function mapping from a spatial domain $\Omega_\mathrm{c} \subset \mathbb{R}^n, n \in \mathbb{N}^+$, to a value domain $W \subset \mathbb{R}$

$$\begin{aligned} I_\mathrm{c} : \Omega_\mathrm{c} &\mapsto W \\ \boldsymbol{x} &\mapsto I_\mathrm{c}(\boldsymbol{x}). \end{aligned} \tag{3.2}$$

By means of the image acquisition and reconstruction process, digital medical images are discretized in the spatial domain (*sampling*). The images can be spatially described by a regular grid $\Gamma_{\boldsymbol{\delta}}$ with lattice constant $\boldsymbol{\delta} = (\delta_1, \delta_2, \ldots, \delta_n)$. Image data is then provided in the discrete domain

$$\Omega = \Omega_\mathrm{c} \cap \Gamma_{\boldsymbol{\delta}}, \tag{3.3}$$

i.e.

$$\begin{aligned} I : \Omega &\mapsto W \\ \boldsymbol{x} &\mapsto I(\boldsymbol{x}), \end{aligned} \tag{3.4}$$

cf. Figure 3.3. The grid elements are termed *pixels* (picture elements), or, especially in the case $n = 3$, *voxels* (volume elements). The dimension of medical images from image modalities such as CT, MR and PET is usually $n = 3$.

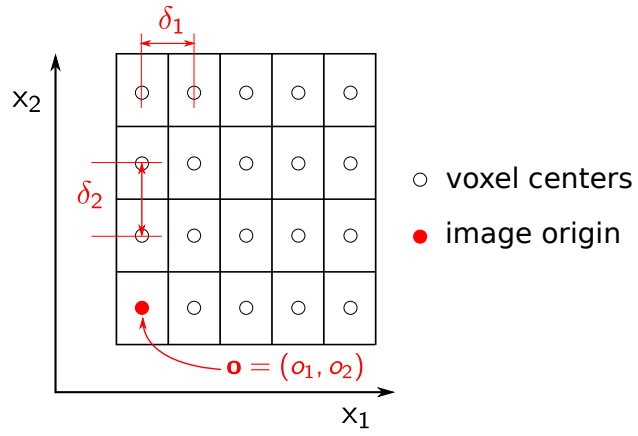

Figure 3.3: Schematic representation of a two-dimensional image. $\boldsymbol{\delta} = (\delta_1, \delta_2)$ and $\boldsymbol{o} = (o_1, o_2)$ denote voxel size and image origin, respectively.

The term *voxel size* describes the spatial extent of the voxels in each of the n spatial dimensions, which is identical with the lattice constant $\boldsymbol{\delta}$. Due to the specific image acquisition process, sampling densities, and resulting voxel sizes, are often anisotropic. Medical images are assigned to an absolute coordinate system. The image grid is fully described by the position of the image origin in the coordinate system, the number of voxels in each dimension, and the voxel size (cf. Figure 3.3).

Image data consists of the *image intensities* assigned to the center of each voxel. In digital images, intensities are also discretized (*quantization*) and defined within a value domain W. Data between voxel centers can be estimated by an *interpolation* method, such as *linear*, *B-spline*, or *Lanczos interpolation* [72].

Physical limitations and measures of image quality

The quality of medical images is determined by several factors. One such factor is the *spatial resolution*, denoting the minimum distance at which the imaging system is still able to depict two subject points as separate. Spatial resolution is related to the *point spread function* (PSF) of the imaging system, i.e. the response of the system to a single point stimulus. One factor influencing resolution in PET and CT imaging is the finite spatial extent of the detectors. For PET imaging, additional factors are caused by the underlying physics, such as the finite range of emitted positrons before annihilation, and deviations from the collinearity of the photon pairs created in the annihilation process. With respect to CT imaging, an additional factor influencing resolution is the finite focus size of the electrons reaching the anode in the X-ray source. In MRI,

spatial resolution in slice encoding direction is determined by the combination of gradient strength and width of the RF pulse spectrum. On the other hand, the finite extent of the k-space FOV is a limiting factor for the resolution in directions which are spatially resolved by frequency and phase encoding.

Image noise is an additional factor limiting image quality, and is usually present as random intensity fluctuations of zero mean. In case of a Gaussian noise distribution, the noise level can be quantified by the associated standard deviation σ. There are different origins of noise. In PET imaging, the number of positrons decaying in a fixed time interval is described by the discrete Poisson probability distribution, which has the general form

$$P_\lambda(k) = \frac{\lambda^k}{k!} \exp(-\lambda), \tag{3.5}$$

with k the dependent variable (i.e. the number of decays per time interval), and the parameter λ being the mean of the distribution. Similarly, in CT imaging, the number of photons being emitted from the X-ray source per time interval is Poisson distributed. For high values of λ, the Poisson distribution resembles a Gaussian distribution with $\sigma^2 = \lambda$. An additional source of noise in PET and CT imaging is the thermal noise of detectors and electronics. In MR imaging, noise originates from the thermal noise of the measured object, the receiver coils and electronics.

For the characterization of image quality, the noise level relative to the absolute measured signal S is of interest. For Gaussian noise, signal to noise ratios (SNRs) can be defined as

$$\mathrm{SNR} = \frac{S}{\sigma}. \tag{3.6}$$

If multiple acquisitions with identical acquisition and reconstruction parameters are available, SNRs can be determined on a voxel-by-voxel level, with S and σ given as the average voxel signal and standard deviation of the acquisitions, respectively. Since this approach requires many acquisitions for statistically reliable results, SNR is in practice usually determined from a single acquisition with evaluating signal statistics on two separate regions of interest (ROIs). The mean signal is evaluated in a ROI placed in the (homogeneous) tissue of interest, whereas the noise is quantified by the standard deviation within a ROI placed in the image background. A requirement for this approach is, however, a spatially homogeneous noise distribution.

Object detection depends on the object-to-background contrast as well as the noise level. Similar to the definition of the SNR, this can be accounted for by

the definition of a contrast to noise ratio (CNR)

$$\mathrm{CNR} = \frac{O - B}{\sigma}, \tag{3.7}$$

with O and B being the mean ROI-based object and background intensities, respectively.

3.3.2 Registration methods for integrating PET/MR into RT treatment planning

For calculation of dose distributions within the RT planning process, electron density information of the patient is required. To date, this information is provided only by a dedicated RT planning CT (cf. Section 1).

As pointed out in in Section 3.2.3 with respect to MR-based PET attenuation correction, there are attempts in recent research to derive electron density information directly from MR images. If electron densities could be derived with a sufficient accuracy using these approaches, RT planning based on PET/MR acquisitions only would be possible [68,70]. However, since this remains an area of research, a dedicated planning CT is still necessary. To integrate the PET/MR data, which is routinely acquired without a dedicated patient positioning system, into the planning process, the spatial transformation between PET/MR and CT dataset has to be found. Estimations of this spatial transformation can be derived by means of *image registration algorithms*.

Image registration algorithms

Let I_r and I_s be two images being defined in the spatial domains $\Omega_\mathrm{r} \subset \mathbb{R}^n$ and $\Omega_\mathrm{s} \subset \mathbb{R}^m$, with $n, m \in \mathbb{N}^+$, respectively,

$$\begin{aligned} I_\mathrm{r} : \Omega_\mathrm{r} &\mapsto W_\mathrm{r} & \qquad I_\mathrm{s} : \Omega_\mathrm{s} &\mapsto W_\mathrm{s} \\ \boldsymbol{x} &\mapsto I_\mathrm{r}(\boldsymbol{x}) & \qquad \tilde{\boldsymbol{x}} &\mapsto I_\mathrm{s}(\tilde{\boldsymbol{x}}) \end{aligned} \tag{3.8}$$

with some spatial correspondence. The objective of image registration is then to find the transformation

$$\begin{aligned} \boldsymbol{T} &: \Omega_\mathrm{r} \mapsto \mathbb{R}^m \\ \boldsymbol{T}(\boldsymbol{x}) &= \tilde{\boldsymbol{x}} \end{aligned} \tag{3.9}$$

mapping corresponding positions of the images (cf. Figure 3.4). With this transformation, $I_\mathrm{s}(\,\boldsymbol{T}(\boldsymbol{x})\,)$ yields an image with information provided from I_s which is spatially aligned to I_r. I_r and I_s are therefore also denoted as *reference* and *source image*, respectively.

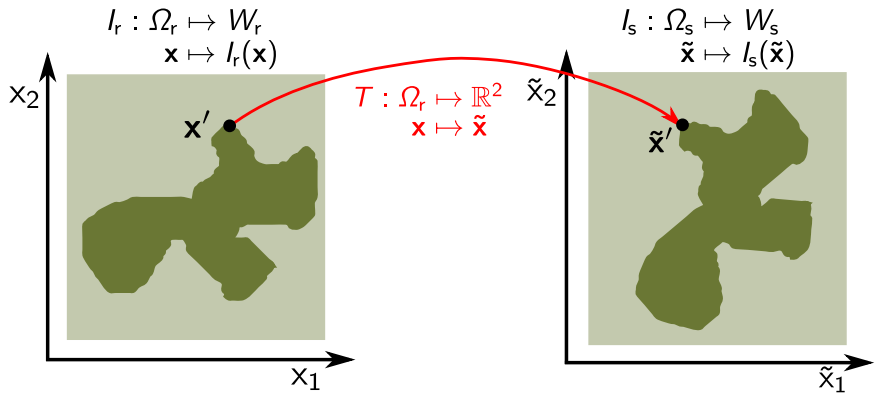

Figure 3.4: Schematic representation of the image registration problem in two dimensions. The input of an registration algorithm are two images with some spatial correspondence. I_s, the *source image* is defined in coordinates $\boldsymbol{x} = (x_1, x_2)$, whereas I_r, the *reference image*, is defined in coordinates $\tilde{\boldsymbol{x}} = (\tilde{x}_1, \tilde{x}_2)$. The objective of image registration is to find the best estimate for the transform $\boldsymbol{T}(\boldsymbol{x}) = \tilde{\boldsymbol{x}}$, which maps I_s to I_r.

In a *rigid registration*, the transformation is limited to translation and rotation. $\boldsymbol{T}$ can then be defined as

$$\boldsymbol{T}_{\text{rigid}}(\boldsymbol{x}) = \boldsymbol{R}_{\boldsymbol{\alpha}}\boldsymbol{x} + \boldsymbol{c}, \tag{3.10}$$

where $\boldsymbol{R}_{\boldsymbol{\alpha}}$ is the rotation matrix and $\boldsymbol{c}$ the translation vector. In three dimensions, this transformation consists of six degrees of freedom in total, defined by three rotation angles $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ and three translational coordinates. The associated parameter vector, containing the parameters to be optimized in the registration process, is thus given by $\boldsymbol{\mu} = (\boldsymbol{\alpha}, \boldsymbol{c})$.

If medical images from the same patient acquired in different scanners have to be registered, rigid registration can yield good results in areas where no, or only small, deformations are expected, such as the brain. However, for other anatomical sites a rigid registration may not yield satisfactory results. As an example, for acquisitions in the head and neck region, a *deformable registration* (DR) may be required to map differences in the bending of the cervical spine and deformations of soft tissue.

One approach for DR is the representation of the transform as a weighted sum of smooth basis functions. Such a parameterization-based approach will be described in more detail in the following. Examples for non-parametric approaches are optical flow based methods [73].

One possible parametrization of the transform can be realized using *cubic B-splines*, which have the advantage of being basis functions with local support.

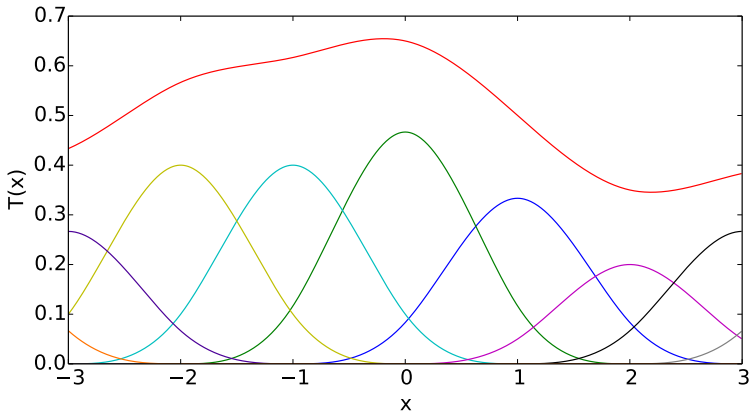

Figure 3.5: Illustration of the parametrization of a one-dimensional transform with cubic B-splines. B-spline control points are located at integer positions, which corresponds to a control point spacing of $d = 1$ (cf. Equation 3.12). The red curve is the final transformation, which is the sum of the B-splines centered at the control points (other colors).

With this parametrization, the transform can in three dimensions be defined as

$$\boldsymbol{T}_{\text{bspline}}(\boldsymbol{x}) = \boldsymbol{x} + \sum_{\boldsymbol{x}_k \in N_{\boldsymbol{x}}} \boldsymbol{p}_k \boldsymbol{\beta}^{(3)}(\boldsymbol{a}), \tag{3.11}$$

$$\text{with } a_i = \frac{(\boldsymbol{x} - \boldsymbol{x}_k)_i}{d_i} \quad \text{for } i \text{ in } \{1, 2, 3\}. \tag{3.12}$$

Here, $\boldsymbol{x}_k$ are control points, building up a regular grid. $\boldsymbol{d}$ is the control point spacing, which determines the flexibility of the transform. $\boldsymbol{\beta}^{(3)}$ the cubic multidimensional B-spline polynomial, given by the tensor product $\boldsymbol{\beta}^{(3)}(\boldsymbol{x}) = \beta^{(3)}(x_1) \otimes \beta^{(3)}(x_2) \otimes \beta^{(3)}(x_3)$ of the one-dimensional cubic B-spline

$$\beta^{(3)}(x) = \begin{cases} \frac{1}{6}(4 - 6x^2 + 3|x|^3), & 0 \leq |x| \leq 1 \\ \frac{1}{6}(2 - |x|^3), & 1 \leq |x| \leq 2 \\ 0, & 2 \leq |x|\,. \end{cases} \tag{3.13}$$

$N_{\boldsymbol{x}}$ is the set of all control points within the compact support of the B-spline at $\boldsymbol{x}$. $\boldsymbol{p}_k$ are coefficients which act as registration parameters. Thus, the parameter vector of this transform is given by $\boldsymbol{\mu} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_N)$, with N being the total number of control points. A schematic illustration of the B-spline parametrization is provided in Figure 3.5.

After the definition of a parametrized transform, the registration problem can be formulated as an optimization problem within which a pre-defined metric

$\mathcal{M}$, providing a measure of registration quality, is minimized with respect to the transform parameters $\boldsymbol{\mu}$,

$$\hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}} \mathcal{M}(\boldsymbol{\mu}; I_\mathrm{r}, I_\mathrm{s}). \tag{3.14}$$

$\mathcal{M}$ can generally be defined by a combination of a measure of image similarity $\mathcal{S}$, which depends on both the images and the transform parameters, and a regularization term $\mathcal{R}$, depending on the transformation parameters only,

$$\mathcal{M}(\boldsymbol{T}; I_\mathrm{r}, I_\mathrm{s}) = -\mathcal{S}(\boldsymbol{T}; I_\mathrm{r}, I_\mathrm{s}) + \gamma\mathcal{R}(\boldsymbol{T}). \tag{3.15}$$

For the determination of $\mathcal{S}$ interpolation of the source image data is necessary.

A simple image similarity measure for monomodal registration is the *sum of squared differences* (SSD),

$$\mathcal{S}_\mathrm{SSD}(\boldsymbol{T}; I_\mathrm{r}, I_\mathrm{s}) = \frac{1}{|\Omega_\mathrm{r}|} \sum_{\boldsymbol{x}\in\Omega_\mathrm{r}} \left(I_\mathrm{r}(\boldsymbol{x}) - I_\mathrm{s}(\boldsymbol{T}(\boldsymbol{x}))\right)^2, \tag{3.16}$$

where $|\Omega_\mathrm{r}|$ is the number of voxels contained in Ω and the summation is over all voxel centers. From the definition follows that the SSD becomes lower if image intensity values show a better spatial agreement.

For images from different modalities, as CT compared to MR, no simple relationship between image intensities exists. Thus, the similarity of the images given specific transform parameters has to be described in a more general, statistical way. Such a statistical measure is provided by the *mutual information* (MI), which has been introduced to image processing by Collignon et al. [74] and Viola and Wells [75]. This measure can be defined as

$$\mathcal{S}_\mathrm{MI}(\boldsymbol{T}; I_\mathrm{r}, I_\mathrm{s}) = \sum_{r,s} p(r,s) \log\left(\frac{p(r,s)}{p(r)p(s)}\right), \tag{3.17}$$

where r and s are the (binned) voxel values of the images I_r and I_s. $p(r)$ and $p(s)$ are marginal probabilities of r and s, respectively, and $p(r,s)$ their joint probability. From the definition in Equation 3.17 it can be inferred that MI measures the distance between the joint probability distribution of the two images, $p(r,s)$, and the joint distribution one would obtain if the two images were completely independent, i.e. $p(r)p(s)$. The definition of MI is related to the *Shannon entropy* from information theory [76]. Equation 3.17 can be rewritten

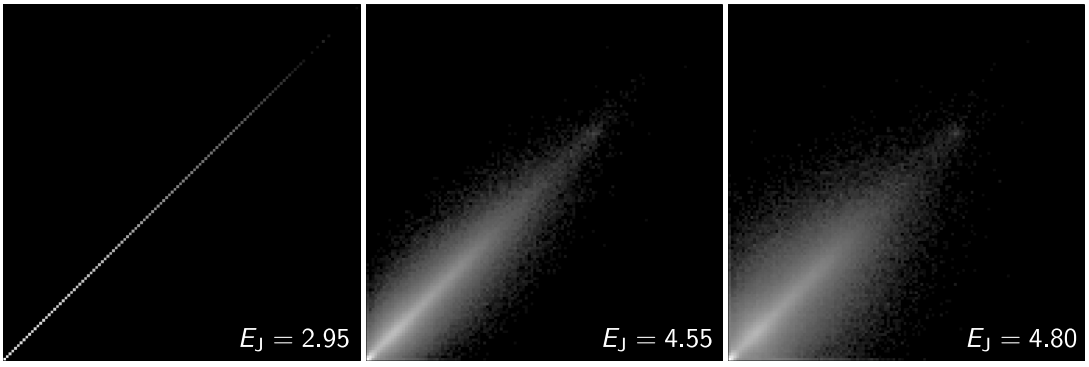

Figure 3.6: Joint histogram of an MR image of the head and neck with itself. From left to right: perfect overlap, translation by one voxel, translation by two voxels. Corresponding joint entropies E_J are provided in the plots. A joint histogram reveals the probabilities of co-occurrence of binned voxel values. With inreasing misregistration the joint histogram disperses and the joint entropy, which can be interpreted as a measure of this dispersion, increases.

as

$$\mathcal{S}_{\mathrm{MI}}(\boldsymbol{T}; I_\mathrm{r}, I_\mathrm{s}) = \underbrace{-\sum_s p(s)\log p(s)}_{E_\mathrm{s}} + \underbrace{(-\sum_r p(r)\log p(r))}_{E_\mathrm{r}} \tag{3.18}$$

$$- \underbrace{(-\sum_{r,s} p(r,s)\log p(r,s))}_{E_J}. \tag{3.19}$$

E_s and E_r are the Shannon entropy of the first and second image, respectively, whereas E_J is their *joint entropy*. An illustration of joint entropy is provided in Figure 3.6.

As regularization term $\mathcal{R}$, a common choice is the *bending energy penalty* (BEP), being in three dimensions defined as

$$\mathcal{R}_{\mathrm{BEP}}(\boldsymbol{T}) = \frac{1}{|\Omega_\mathrm{r}|} \sum_{\tilde{\boldsymbol{x}} \in \Omega_\mathrm{r}} \sum_{i=1}^{3} \sum_{j=1}^{3} \left(\frac{\partial^2 \boldsymbol{T}(\tilde{\boldsymbol{x}})}{\partial x_i \partial x_j} \right)^2, \tag{3.20}$$

with the summation being performed over the voxel centers. This term is motivated by the two-dimensional bending energy of a thin plate of metal [77], and favors smooth transformations.

The registration parameters $\boldsymbol{\mu}$ are then optimized with respect to $\mathcal{M}$. Generally, $\mathcal{M}$ is a complex function with multiple local minima. For optimization,

various algorithms are available in numerical literature. A comparison of different optimization strategies applied for image registration is provided in [78].

3.3.3 Dedicated tumor segmentation strategies

Image segmentation is the task of partitioning images into multiple segments according to the underlying content. For RT, automatic segmentation strategies for the distinction of tumor and non-tumor regions are of particular interest. As outlined in Section 3.1, combined PET/MR data may be beneficial for accurate tumor volume delineation.

A variety of segmentation algorithms already exist, which can be used as a basis for segmenting PET/MR data. Some algorithms use image intensities directly as input for segmentation, whereas others operate on image gradients or edges. On the other hand, segmentation algorithms can be discriminated by how much they are influenced by the image data and how much they impose a specific model on the segmentation. Model-driven algorithms can enforce closed and smooth boundaries irrespective of the image content.

Concerning PET segmentation, different methods have been published in literature. Comprehensive reviews of these methods are provided in [79, 80]. However, literature addressing the segmentation of anatomical MR images has mainly been focused on brain tumors [81, 82]. In the following, segmentation algorithms applied within this thesis are discussed.

Threshold-based segmentation

The basic threshold segmentation algorithm classifies each voxel according to its data value in a purely data-driven approach. In this concept, the segmentation of an object from background is in general based on the definition of a lower threshold L and an upper threshold U, setting the classification criterion of voxel i with intensity I_i to

$$m_i = \begin{cases} 1 & \text{if } L < I_i < U \\ 0 & \text{else} \end{cases} \tag{3.21}$$

with 1 and 0 denoting the object and background segments, respectively. A common method for tumor segmentation based on FDG-PET is a relative threshold segmentation with the lower threshold set to 40 % of the maximum SUV of the tumor [21].

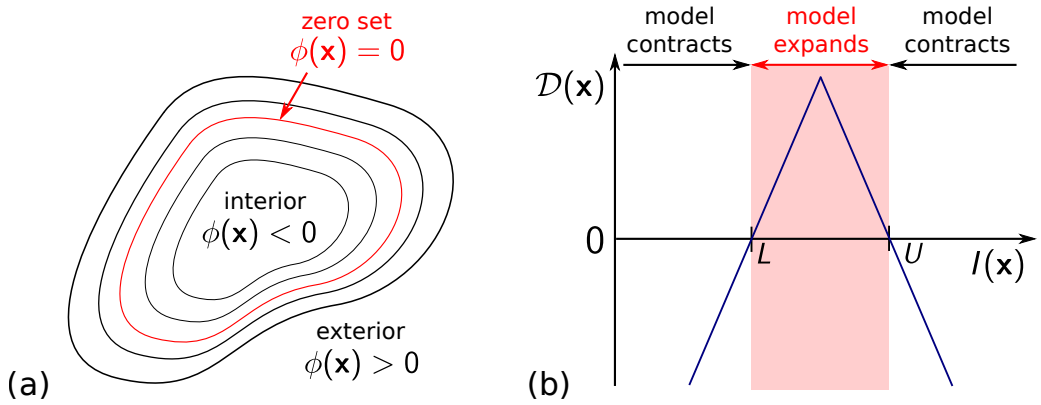

Figure 3.7: Illustrations of the threshold level set algorithm. (a) Level set segmentation is based on the evolution of a level set function $\phi(\boldsymbol{x})$. The segmentation boundary Γ is defined at each time point by $\Gamma = \{\boldsymbol{x}|\phi(\boldsymbol{x}) = 0\}$. (b) The propagation function $\mathcal{D}(\boldsymbol{x})$ used for the threshold level set algorithm is determined by a lower threshold L and an upper threshold U, see Equation 3.29. Reproduced from the ITK manual [83].

Level set segmentation

The *level set method* belongs to the implicit active contours/surfaces algorithms within the model-driven segmentation methods. Within these methods, the segmentation boundary is implicitly represented by a function in the space in which the boundary is embedded. A boundary Γ of $n-1$ dimensions can be defined by a auxiliary level set function $\phi(\boldsymbol{x})$ of n variables based on the relation

$$\Gamma = \{\boldsymbol{x}|\phi(\boldsymbol{x}) = 0\} \tag{3.22}$$

(cf. Figure 3.7 a). Γ is termed the *zero level set*. An adequate function for $\phi(\boldsymbol{x})$ is then, for example, the signed distance to the represented boundary. For volumetric medical images, segmentation boundaries are usually two-dimensional surfaces, as n equals 3.

During the segmentation process, the level set function, and thus also Γ, is evolved in time. The time evolution is governed by a differential equation, which couples the speed of deformation to the image data. Many variants of this basic concept have been proposed [84]. The generic form of the level set equation can be derived by building the total differential of the time-dependent defining equation

$$\phi(\boldsymbol{x}(t), t) = 0, \tag{3.23}$$

which yields

$$\frac{\partial \phi}{\partial t} + \nabla \phi \cdot \frac{\partial \mathbf{x}}{\partial t} = 0. \tag{3.24}$$

To impose a curve evolution in the normal direction, it can be forced that

$$\frac{\partial \boldsymbol{x}}{\partial t} = F\boldsymbol{N}, \tag{3.25}$$

where

$$\boldsymbol{N} = \frac{\nabla \phi}{|\nabla \phi|} \tag{3.26}$$

is the normal to the level set curve. F determines how fast the curve evolves locally, and is therefore termed *speed function*. With this condition, the boundary Γ thus evolves in the normal direction with the local speed determined by F.

With Equations 3.24 to 3.26, the generic partial differential equation (PDE) [85] of the level set segmentation is obtained

$$\frac{\partial \phi}{\partial t} = -F\,|\nabla \phi|\,. \tag{3.27}$$

The speed function F can in general be splitted into a data-driven term $\mathcal{D}$, and a model-driven term $\mathcal{M}$

$$F = \alpha\mathcal{D} + \beta\mathcal{M}, \tag{3.28}$$

with respective weights α and β, respectively. $\mathcal{D}$ can either be based on original image intensities, or on gradient or edge information.

For *threshold level set segmentation*, the data driven term depends on the position vector only, and can be defined as

$$\mathcal{D}(\boldsymbol{x}) = \begin{cases} I(\boldsymbol{x}) - L & \text{if } I(\boldsymbol{x}) < (U-L)/2 + L \\ U - I(\boldsymbol{x}) & \text{otherwise} \end{cases} \tag{3.29}$$

with an upper threshold U and a lower threshold L (cf. Figure 3.7 b). The *feature image* $I(\boldsymbol{x})$ can be either the original image to be segmented or some kind of preprocessed version thereof.

The model-driven term $\mathcal{M}$ is independent on the image data, and can be used to provide a regularization to the segmentation. A common choice is to set this term equal to the curvature $\kappa = \nabla \cdot \frac{\nabla \phi}{|\nabla \phi|}$, imposing a smooth segmentation boundary.

Equation 3.27 in combination with an initial level set function $\phi(\boldsymbol{x}, t = 0)$ can be considered as *initial value problem*. $\phi(\boldsymbol{x}, t = 0)$ can be provided by the signed distance of a contour obtained with some preliminary rough segmentation

approach. The level set function is subsequently evolved in time by numerical integration of a discretized form of Equation 3.27 until a pre-defined convergence criterion is achieved. The segmentation result is given by the final zero level set.

3.3.4 Analysis of physical properties in multi-parametric PET/MR data

PET/MR scanners allow for the acquisition of various types of functional images (cf. Section 3.1), which may allow for biologically adapted treatments. Potential applications are RT treatment individualization, the assessment of chemotherapy efficacy, and the application of targeted agents, such as hypoxia-sensitizing or antiangiogenic drugs [86, 87].

For RT treatment adaption, a dose PF has to be found, which ideally modulates the locally prescribed tumor dose in a way that equal control probabilities within the heterogeneous tumor tissue are obtained, thus compensating for different biological microenvironments. However, the definition of an adequate PF is not straightforward. Statistical analysis of functional imaging datasets and their relation to treatment outcome could provide evidence for the design of individually adapted treatment approaches [13].

Derivation of (semi-)quantitative parameter maps from functional images

Functional imaging data may be preprocessed to extract quantitative parameters, ideally with known physical and physiological meaning. Fit parameters from voxel-by-voxel analysis of the imaging data can be stored in *parameter maps*. Advantages of deriving quantitative parameters are increased inter-patient comparability as well as the dimensionality reduction of acquisitions with various time frames such as dynamic FMISO-PET and DCE-MRI. In the following, PET- and MR-based parameter maps are described.

- The comparability of **FDG-PET** activity concentrations between different patients can be enhanced by a normalization with respect to patient weight and decay-corrected injected activity. This normalization leads to *standardized uptake values* (SUVs) of the form

$$\mathrm{SUV} = \frac{c_\mathrm{A}}{A_0/m_\mathrm{pat}}, \tag{3.30}$$

  with c_A being the local tracer activity concentration as revealed by the imaging data, A_0 the decay-corrected injected activity, and m_pat the mass of the patient. However, SUVs are not fully quantitative, since they are influenced by factors such as body composition, and the plasma glucose level [88].

- In **static FMISO-PET** images, tumor voxel activities are usually normalized by the background signal in a deep neck muscle [19]. After defining a muscle ROI Ω_{M}, *tumor to background ratios* (TBRs) are given as

$$\mathrm{TBR} = \frac{c_{\mathrm{A}}}{\overline{c}_{\mathrm{A},\Omega_{\mathrm{M}}}}. \tag{3.31}$$

  where $\overline{c}_{\mathrm{A},\Omega_{\mathrm{M}}}$ is the mean activity concentration in Ω_{M}.

- From **dynamic FMISO-PET**, quantitative parameter maps can be derived by compartmental modeling (cf. Figure 3.8 a). It has been proposed to fit measured time-activity data on a voxel level with a two-compartment model including tracer retention [89]. Compartmental parameters derived from dynamic FMISO-PET may be superior for the characterization of hypoxia than imaging information from a single PET time frame, since they might also reveal severe hypoxia (existing along with necrotic areas) which is potentially missed in a static-only acquisition [44].

- Also **DCE-MR** data can be fitted by compartment models to yield quantitative parameter maps of physiological meaning (cf. Figure 3.8 b). Different models are available in literature, such as the Brix [91], extended Tofts [90], and the two-compartment exchange model (2CXM) [92]. In this work, the extended Tofts model is applied, yielding parameter maps of the volume fraction of the extracellular-extravascular space (EES), v_{e}, the volume fraction of the blood, v_{p}, and the volume transfer rate from blood plasma to the EES, K^{trans}.

- In **DW-MR**, ADC maps can be derived from the images acquired with different b-values, i.e. different degrees of diffusion weighting. The relation between signals S and b-values is usually modeled by a mono-exponential function [29]

$$S = S_0 \exp\left(-b \cdot \mathrm{ADC}\right), \tag{3.32}$$

  where S_0 is the signal without diffusion weighting. With Equation 3.32 ADC values can be obtained by fitting $\ln S$ over b with a linear model. Especially for low b-values, microcapillary perfusion may influence the measured signals. If the measured signal values S are influenced by both perfusion and diffusion effects, a bi-exponential fit is more adequate [93]. Calculation of ADC maps is typically integrated in manufacturer-provided software of the MR scanner.

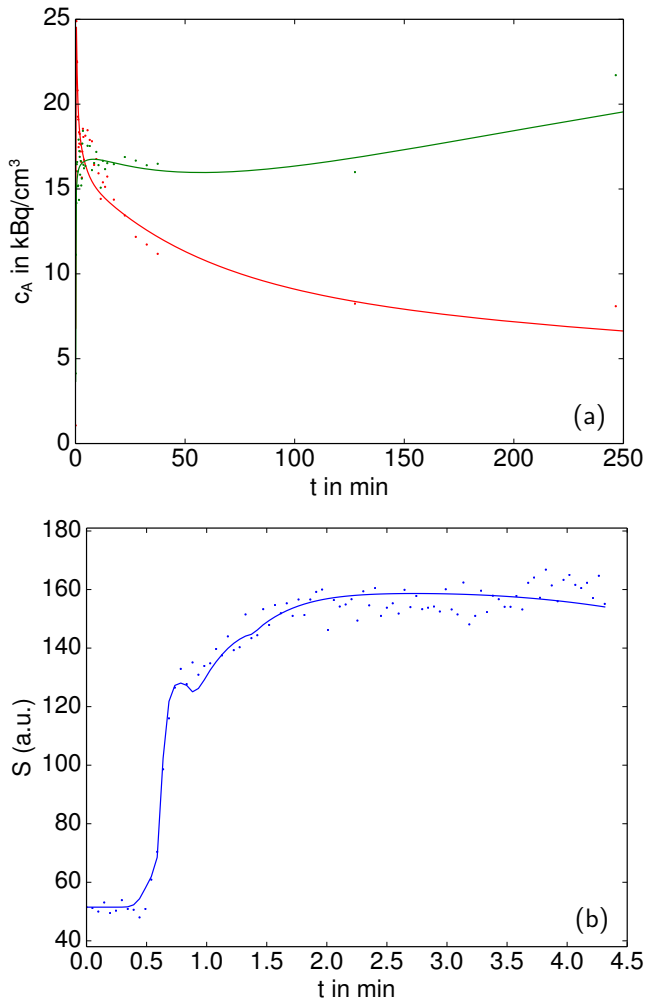

Figure 3.8: Dynamic functional data (*points*) derived from individual voxels within the GTV of a HNC patient, shown along with fits by compartment models (*lines*). (a) FMISO-PET activity concentration c_A over time, and fit by two-compartment model [89], for two different voxels (red, green). The data acquisition protocol consisted of dynamical data collection 0–35 min p.i., and subsequent static scans at 2 h p.i. and 4 h p.i.. The red colored data is characterized by a high initial peak due to a good vascularization, as well as missing tracer retention, indicating a *non-hypoxic* voxel. The green colored data shows increasing activity in late times, thus indicating tracer retention due to *hypoxic* conditions. (b) Signal S from DCE-MR over time for an exemplary voxel in the GTV of a HNC patient. The fit was performed with the extended Tofts model [90]. Figures kindly provided by Urban Simončič.

Statistical analysis and statistical learning methods

Inter-dependence of two variables, such as pairs of functional image data, can be examined by means of *statistical analysis* methods.

With the *Pearson product-moment correlation coefficient* $\rho_{X,Y}$ it can be quantified to which extent two variables X and Y are linearly correlated. $\rho_{X,Y}$ is defined by

$$\rho_{X,Y} = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)\,\mathrm{Var}(Y)} \tag{3.33}$$

$$= \frac{E[(X - E[X])(Y - E[X])]}{\sqrt{E[(X - E[X])^2]}\sqrt{E[(Y - E[Y])^2]}}, \tag{3.34}$$

where $\mathrm{Cov}(X,Y)$ is the *covariance* between X and Y, $\mathrm{Var}(\cdot)$ is the *variance*, and $E[\cdot]$ the *expected value* of the respective variable.

Values of $\rho_{X,Y}$ range between -1 and 1, where -1 and 1 are obtained for perfect positive and negative linear correlations, respectively. A value of 0 is obtained for totally uncorrelated variables.

If samples of X and Y are available, sample-based estimates of covariances and variances yield the expression

$$r = r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}, \tag{3.35}$$

where $\{x_i \,|\, i = 1, \ldots, n\}$ and $\{y_i \,|\, i = 1, \ldots, n\}$ are the samples taken from X and Y, respectively, and $\bar{x}$ and $\bar{y}$ are the sample averages.

The *Spearman rank correlation coefficient* quantifies the degree of monotonicity between to variables, irrespective of the exact form of the relation. This is achieved by modifying Equation 3.35 by replacing the sample values with their respective *rang* $\mathrm{rg}(\cdot)$ within the sample

$$x_i \to \mathrm{rg}(x_i), y_i \to \mathrm{rg}(y_i). \tag{3.36}$$

Since the Spearman coefficient describes the general degree of monotonicity, it has broader scope than the Pearson coefficient. By definition, the Spearman coefficient is also much less sensitive to outliers.

In addition to the mathematically simple correlation coefficients for pairs of variables, also more advanced methods for statistical analysis exist, with which statistical relations between three or more variables can be inferred. Such *machine learning algorithms* can be applied to higher-dimensional data.

For exploratory data analysis without the availability of response variables, *clustering* approaches can be applied. These allow for the detection of patterns in a multidimensional dataset [94].

If categorical response variables (i.e. class labels) are available, *classification* can be performed. In general, the purpose of statistical classifiers is to derive a relationship between the feature vectors and class labels using a training dataset. The goal of the classification is to predict the class label of feature vectors of an independent test dataset with a high accuracy [94, 95]. Examples for non-linear classification methods include the *support vector machine*, the *random forest* method and *neural networks*.

If, on the other hand, the response variable is not categorical but quantitative, *regression* instead of classification can be applied to learn the relation between feature vectors and response variables. Non-linear regression can be performed with the *random forest* method, which is also suited for classification purposes.

During the training process of classification as well as regression methods, the relationship between feature vectors and response variable is learned using the dedicated *training dataset*. Since there is a risk of overfitting noisy or unrepresentative data, a *test dataset* has to be available to validate the trained relationship. As an alternative, *cross-validation* schemes can be applied.

4 Part I: A strategy for multimodal deformable image registration to integrate PET/MR into radiotherapy treatment planning

Sara Leibfarth[1*], David Mönnich[1], Stefan Welz[2], Christine Siegel[2], Nina Schwenzer[3], Holger Schmidt[3,4], Daniel Zips[2], Daniela Thorwarth[1]

[1] Section for Biomedical Physics, Department of Radiation Oncology, University Hospital Tübingen
[2] Department of Radiation Oncology, University Hospital Tübingen
[3] Department of Diagnostic and Interventional Radiology, University Hospital Tübingen
[4] Laboratory for Preclinical Imaging and Imaging Technology of the Werner Siemens Foundation, Department of Preclinical Imaging and Radiopharmacy, Tübingen

published in
Acta Oncologica 2013, **52**: 1353–1359.

Abstract

Background: Combined PET/MRI is highly promising for biologically individualized radiotherapy (RT). Hence, the purpose of this work was to develop an accurate and robust registration strategy to integrate combined PET/MR data into RT treatment planning.
Material and Methods: Eight patient datasets consisting of an FDG PET/CT and a subsequently acquired PET/MR of the head and neck (HN) region were available. Registration strategies were developed based on CT and MR data only, whereas the PET components were fused with the resulting deformation field. Following a rigid registration, deformable registration was performed with a transform parametrized by B-splines. Three different optimization metrics were investigated: global mutual information (GMI), GMI combined with a bending energy penalty (BEP) for regularization (GMI+BEP) and localized mutual information with BEP (LMI+BEP). Different quantitative registration quality measures were developed, including volumetric overlap and mean distance measures for structures segmented on CT and MR as well as anatomical landmark distances. Moreover, the local registration quality in the tumor region was assessed by the Normalized Cross Correlation (NCC) of the two PET datasets.
Results: LMI+BEP yielded the most robust and accurate registration results. For GMI, GMI+BEP and LMI+BEP, mean landmark distances (standard deviations) were 23.9 mm (15.5 mm), 4.8 mm (4.0 mm) and 3.0 mm (1.0 mm), and mean NCC values (standard deviations) were 0.29 (0.29), 0.84 (0.14) and 0.88 (0.06), respectively.
Conclusion: Accurate and robust multimodal deformable image registration of CT and MR in the HN region can be performed using a B-spline parametrized transform and LMI+BEP as optimization metric. With this strategy, biologically individualized RT based on combined PET/MRI in terms of dose painting is possible.

4.1 Introduction

Recently, combined imaging of Positron Emission Tomography (PET) and Magnetic Resonance Imaging (MRI) has become available [96]. This new imaging technique allows for the simultaneous acquisition of functional MR and molecular PET information, resulting in accurately co-registered datasets. As a consequence, integrated PET/MR may be highly beneficial for radiotherapy (RT) individualization [25]. On the one hand, the molecular, functional and anatomical information from PET/MR data be can be used in RT treatment planning (RTP) for improving the precision of tumor volume delineation [79, 97]. Moreover, the molecular information on tumor physiology, biology, and radioresistance may serve as a basis for biologically individualized RT in terms of dose painting (DP) [98, 99]. Furthermore, combined PET/MR imaging might be also valuable for the assessment of treatment response and follow-up after therapy [100].

However, for the integration into RTP, PET/MR data has to be fused with the Computed Tomography (CT) image generally used for treatment planning. For an accurate fusion of PET/MR and CT data, the anatomical information of CT and MR can be used for registration, and the PET component of the PET/MR can be co-transformed with the resulting deformation field. Due to the different physical principles of image acquisition, CT and MR do not show a simple relationship between image intensities. Moreover, to date it is not possible to acquire combined PET/MR in the head and neck (HN) region with dedicated RT positioning aids. Therefore a suitable multimodal deformable registration (DR) algorithm is required. Another potential application of DR of MR and CT is the attenuation correction of PET/MR data using deformed CT images [70].

Previous strategies for the DR of CT and MR images have been published for different anatomical sites such as liver [101], breast [102] and prostate [103]. For the HN region, the development of DR strategies has so far mainly been focused on monomodal DR of CT images [104]. Nevertheless, some studies also investigated DR algorithms for the fusion of multimodal imaging data. In the study published by Söhn et al, a multimodal DR algorithm based on rigidly matching local image subvolumes was applied to fuse CT and MR data of the HN region [105]. Nevertheless, this study was intended as a proof of concept without quantitative evaluation. Another study applied a different algorithm matching the bony structures in combination with a linear elastic biomechanical finite element model for image registration of CT and MR in the HN region [106], and a landmark-based evaluation for four datasets was provided. A common approach for multimodal DR is to use an algorithm consisting of a B-spline parametrized transform and mutual information (MI) as similarity measure [77,107]. However,

also localized versions of MI have been considered recently, where MI is evaluated in subregions of the images only [108, 109].

In this study, B-spline based algorithms with both global and localized forms of MI are applied for DR of CT and MR in the HN region and the registration accuracy is assessed by means of quantitative measures. The aim is to develop an accurate and robust registration method for potential future integration of PET/MR image data into RTP.

4.2 Material and Methods

4.2.1 Patient data

Eight patient datasets acquired within a clinical imaging study were available. Each dataset consists of a PET/CT and subsequently acquired PET/MR, with the PET-tracer [^{18}F]-Fluorodeoxyglucose (FDG) (cf. Figure 4.5). The mean injected activity of FDG was 351 MBq (range: 320 MBq – 388 MBq). PET/CT data were acquired after a mean time of 81 min p.i. (range: 76 min – 94 min) and PET/MR data after 136 min p.i. (range: 120 min – 166 min). Detailed information about the patient characteristics is given in Table 4.2.

The PET/CT was acquired without fixation with a Siemens Biograph mCT (Siemens Healthcare, Knoxville, TN). The CT consisted of a low-dose and contrast enhanced acquisition. For PET reconstruction, 3D OSEM with 2 iterations, 21 subsets and a 3D Gaussian filter of 2 mm was used. Approximate voxel sizes of the CT and the PET images were $0.8 \times 0.8 \times 3.0\,\mathrm{mm}^3$ and $1.6 \times 1.6 \times 3.0\,\mathrm{mm}^3$, respectively.

The PET/MR was acquired without fixation with the Siemens Biograph mMR (Siemens Healthcare, Erlangen, Germany). The MRI datasets taken into account for this study were T2-weighted acquisitions using a Short Time Inversion Recovery (STIR) sequence taken in the transversal direction. For PET reconstruction, 3D OSEM with 3 iterations and 21 subsets and a 3D Gaussian filter of 3 mm was used. MR attenuation correction was performed based on a segmentation approach on basis of spoiled gradient-echo sequences with DIXON-based fat-water separation. Approximate voxel sizes of the STIR and the PET images were $0.7 \times 0.7 \times 4.8\,\mathrm{mm}^3$ and $2.8 \times 2.8 \times 2.0\,\mathrm{mm}^3$, respectively.

4.2.2 Image Registration

Image registration was performed using information from CT and MR only, with the CT serving as fixed image and the MR serving as moving image. After reg-

istration, both MR and PET from the PET/MR examination were transformed to the PET/CT data. Thus, after registration, the CT and the MR as well as the two PET images were defined in the same coordinate system.

First, a rigid registration (RR) was performed. The resulting transform was then used for the initialization of the DR methods, where three different strategies were investigated. Since the available datasets had different Fields of View (FOV), they were cropped after RR in order to cover the same anatomical region.

Both RR and DR were performed with the freely available, open source registration package *elastix* [110] which is based on the ITK (Insight Segmentation and Registration Toolkit, www.itk.org). DR was performed with a three level multi resolution approach, using only Gaussian smoothing without downsampling. Smoothing scales were chosen as $\sigma = 8.0, 4.0, 1.0$ in x and y direction and $\sigma = 2.0, 1.0, 0.5$ in z direction, to account for voxel anisotropy.

For each DR strategy a transform parametrized by B-splines was used [77]. For the optimization of the metric, a stochastic gradient descent optimizer [111] was applied. In this optimization strategy, a sampler randomly choses a given number of image positions every iteration for metric estimation.

As similarity measure, firstly MI calculated from samples over the whole image domain was used, which is referred to as global MI (GMI). Secondly, a localized version of MI was considered, where the localization is achieved as described in [108]. Briefly, sampling is constrained to a cubic subregion of the image of length L_{sub}, which is chosen randomly in every iteration step from the fixed image domain. This method is referred to as localized MI (LMI). For regularization of the transform a bending energy penalty (BEP) term was applied [77], defined as

$$\text{BEP} = \frac{1}{P} \sum_{\tilde{\boldsymbol{x}}_p} \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=1}^{3} \left(\frac{\partial^2 T_i(\tilde{\boldsymbol{x}}_p)}{\partial x_j \partial x_k} \right)^2 , \tag{4.1}$$

where $\tilde{\mathbf{x}}_p$ are the voxel positions, and P is the number of voxels. T_i denotes the ith component of the transform, and x_1, x_2 and x_3 are the coordinates in x, y and z direction, respectively. The BEP favors a smooth deformation field.

The following combinations were investigated for the registration metric $\mathcal{M}$:

$$\mathcal{M} = \begin{cases} -\text{GMI} & \textbf{(GMI)} \\ -\text{GMI} + \lambda\,\text{BEP} & \textbf{(GMI+BEP)} \\ -\text{LMI} + \lambda\,\text{BEP} & \textbf{(LMI+BEP)} \end{cases} , \tag{4.2}$$

where λ determines the weight of the regularization with respect to the similarity measure.

For some of the free parameters of the DR methods, such as the B-spline grid spacing, the number of iterations and the number of samples used for metric evaluation during optimization, predetermined parameters were used. The remaining free parameters, i.e. the number of histogram bins used for MI calculation, λ, and L_{sub}, were optimized independently for each DR method according to the registration quality measures derived from segmentations described below. The final parameter sets are provided in Table 4.3.

4.2.3 Visual assessment of registration quality

A visual assessment of registration quality was performed based on fusion images of CT and MR and as well as of two PET images. Moreover, the deformation fields obtained from the different DR methods were assessed qualitatively by the corresponding map of Jacobian determinants I_{JD} [112,113]. Jacobian determinant values smaller than 1 are associated with a local volume compression, values equal to 1 with local volume preservation, and values greater than one with local volume expansion. Negative values of the Jacobian determinant are obtained in the presence of foldings, which are unrealistic for medical images.

4.2.4 Quantitative assessment of registration accuracy

Several measures for the determination of registration accuracy were implemented using ITK and VTK (Visualization Toolkit, www.vtk.org). For the structures skin, carotids, and respiratory tract segmented on CT and MR, both the Dice Similarity Index (DSI) and mean distances after registration were evaluated. As mean distance measures, the Mean Volume Distance (MVD) was determined for the skin as well as for the respiratory tract structure, whereas for the carotids the Mean Line Distance (MLD) was evaluated. In addition, the Non Overlapping Fraction (NOF) of the bony structures segmented from CT and the spinal canal segmented from the MR image was calculated. Moreover, the mean residual distance (Mean Point Distance, MPD) of anatomical landmarks defined by two experienced radiation oncologists was evaluated after registration.

A validation of the local registration accuracy in the tumor region based on the image information from CT and MR is difficult, since the intra-tumor region can be of low contrast in these imaging modalities. As the PET images provide local image information in this region, the registration accuracy in the tumor region was assessed by Normalised Cross Correlation (NCC) of the fused PET images. Detailed information about the quantitative registration accuracy measures is provided in the Section 4.6.

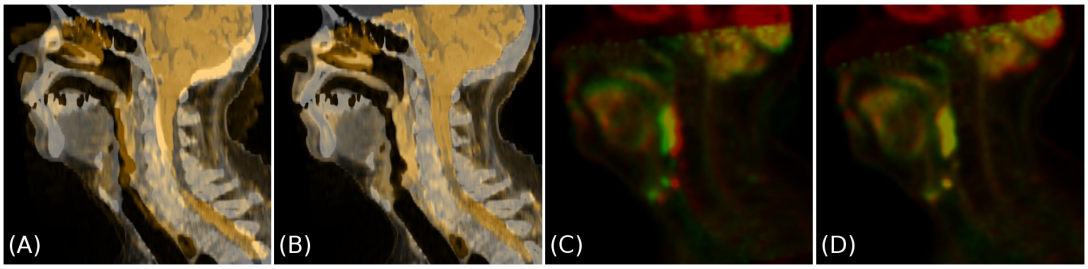

Figure 4.1: Fusion images after RR and DR for Patient 4. Fusion of the original CT (grey) and the deformed MR (orange) after RR (A) and after DR with LMI+BEP (B). Fusion of the PET of the PET/CT (red) and the deformed PET of the PET/MR (green) after RR (C) and after DR with LMI+BEP (D).

4.3 Results

4.3.1 Qualitative Evaluation

Fusion images of CT and MR as well as of the PET images are displayed in Figure 4.1 exemplarily for Patient 4 after both RR and DR with LMI+BEP. The fusion images after RR showed large remaining misalignments due to different patient positioning. After DR, a visually good alignment of both the anatomical and the functional images could be obtained.

Figure 4.2 shows I_{JD} shows and the corresponding deformed MR images obtained by the three DR methods obtained for Patient 1. GMI led to Jacobian determinant values with high absolute differences to 1, and also foldings were present. Compared to RR, the corresponding deformed MR was better aligned to the shown anatomical contours, but unrealistic deformations occurred. The deformed MR images obtained by the registration strategies that apply a regularization term (GMI+BEP and LMI+BEP) did not suffer from unrealistic deformations and corresponding I_{JD} showed only moderate and smooth volume expansions and compressions. Moreover, also a good alignment to the anatomical contours was achieved.

Figure 4.3 shows axial slices of the original CT and the transformed MR after RR as well as after DR with LMI+BEP for Patient 2, in addition to contours defined on basis of the CT image. For the rigidly transformed MR, large misalignments to the CT contours remained, whereas the DR with LMI+BEP led to a significant improvement. However, for the skin some misalignments remained in regions where large local deformations are present.

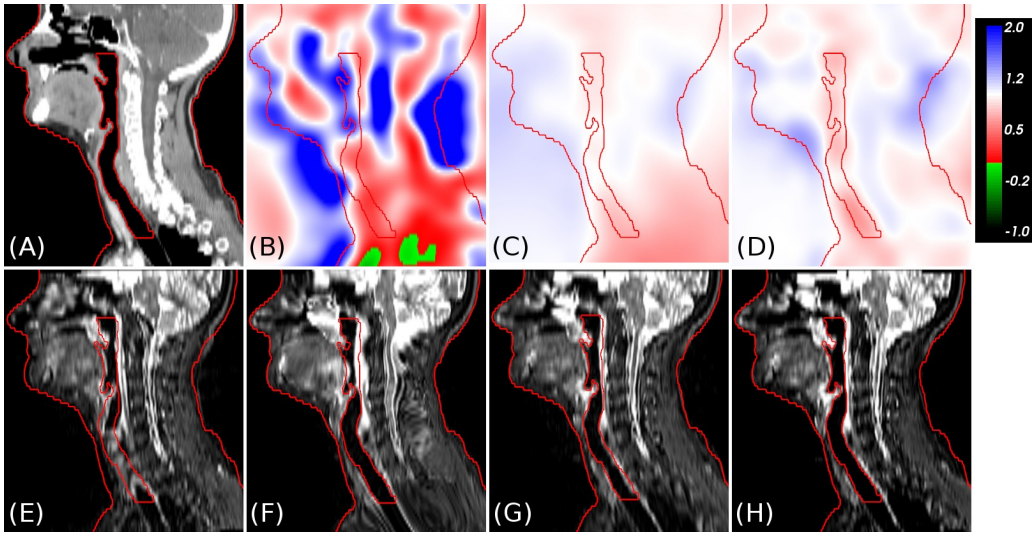

Figure 4.2: Map of Jacobian determinants (I_{JD}) and corresponding deformed MR from different registration methods for Patient 1. Original CT (A), I_{JD} from DR with GMI (B), GMI+BEP (C) and LMI+BEP (D). Transformed MR from RR (E), deformed MR from DR with GMI (F), GMI+BEP (G) and LMI+BEP (H). The structures skin and respiratory tract segmented on the original CT are shown as red contours.

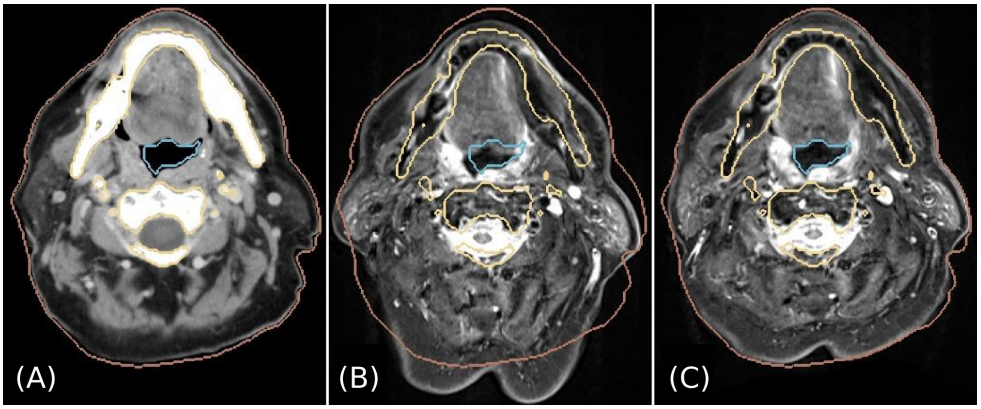

Figure 4.3: Axial slices of the original CT (A), transformed MR from RR (B) and from DR with LMI+BEP (C) for Patient 2. Contours of skin (brown), bones (yellow) and respiratory tract (blue) derived from the original CT are also shown.

Table 4.1: Quantitative results of the registration methods as mean (standard deviation) over all patients.

	RR	GMI	GMI+BEP	LMI+BEP
DSI skin	0.94 (0.02)	0.97 (0.01)	0.97 (0.00)	0.98 (0.00)
DSI carotids	0.14 (0.16)	0.07 (0.09)	0.37 (0.26)	0.59 (0.18)
DSI respiratory tract	0.39 (0.22)	0.49 (0.31)	0.59 (0.30)	0.76 (0.07)
NOF bone/spinal canal	0.50 (0.24)	0.55 (0.12)	0.72 (0.27)	0.92 (0.06)
NCC PET	0.67 (0.31)	0.29 (0.29)	0.84 (0.14)	0.88 (0.06)
MVD skin (mm)	4.07 (1.58)	1.70 (0.71)	1.75 (0.49)	1.47 (0.44)
MLD carotids (mm)	7.27 (3.17)	15.82 (6.79)	5.41 (5.22)	1.92 (0.61)
MVD respiratory tract (mm)	4.45 (2.47)	5.92 (7.87)	3.92 (5.43)	1.33 (0.46)
MPD landmarks (mm)	6.97 (4.95)	23.86 (15.51)	4.78 (4.02)	2.96 (1.02)

4.3.2 Quantitative Evaluation

Quantitative results for all patients are summarized in Table 4.1 as well as in Figure 4.4. For GMI, the measures for the skin improved compared to RR, but unrealistic deformations as shown in Figure 4.2 led to worse quantitative registration results for other measures. For GMI+BEP, all registration accuracy measures improved compared to RR. However, considerable variations remained, showing that the method is not robust with respect to inter-patient variation between different datasets.

In contrast, LMI+BEP showed a robust behavior and a good registration accuracy. Residual distances were in the order of half the voxel size of fixed and moving image in z-direction (voxel size CT: 3.0 mm, MR: 4.8 mm) (cf. Table 4.1). Moreover, the NCC of the PET images in the tumor region improved in line with the registration accuracy measures defined based on CT and MR.

4.4 Discussion

In this work, three different DR strategies for the fusion of CT and MR in the HN region were evaluated qualitatively and quantitatively, with the purpose of integrating PET/MR data into RTP. All algorithms used a multi-resolution approach and a B-spline parametrized transform, whereas the different metrics GMI, GMI+BEP and LMI+BEP were applied. Compared to RR, using the metric GMI in most cases resulted in worse registration results in addition to

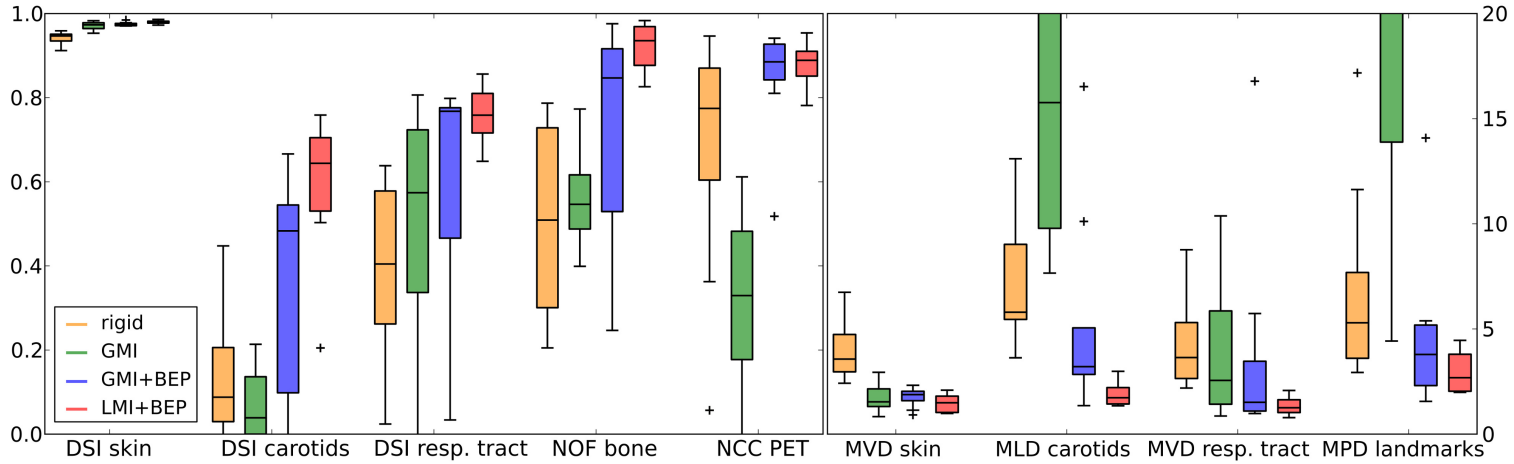

Figure 4.4: Quantitative results of the registration methods. Left: Boxplots of quality measures ranging between 0 and 1, with 1 being the best value. Right: Boxplots of distance quality measures, with values given in mm. Low values indicate a good registration accuracy.

unrealistic deformations. In principle, the degree of unrealistic deformations could be lowered by a smaller B-spline grid spacing, but this would simultaneously decrease the degree of freedom of the transform. For DR with GMI+BEP, registration accuracy improved compared to RR, but a lack of robustness was observed. Using LMI+BEP as optimization metric, accurate and robust results were obtained, even for patients with large positioning differences in CT and MR. Distance quality measures showed that residual distances were in the order of half the voxel size of the CT and MR in z-direction, indicating a high geometric accuracy.

Due to the small diameter of the carotids, the DSI for this structure can decrease severly even for small misregistrations. Therefore, the MLD should provide a better measure of registration accuracy for this structure, showing that the registration accuracy in the carotid region is similar to the other structures. In contrast to the other registration methods, for LMI+BEP no outliers were present apart from the outlier observed for the DSI measure of the carotids (cf. Figure 4.4). However, since this registration method has been evaluated on a limited dataset, at least a visual examination of the registration results obtained with LMI+BEP should be performed for additional datasets.

However, there are also inherent limitations of the algorithm as shown in Figure 4.3. Since the B-spline parametrization provides only a limited degree of freedom of the transform, the algorithm is not able to map large local deformations. This ability is further reduced by the BEP. However, both the parametrization and the BEP favor a smooth transform which generally is a reasonable assumption for medical images, particularly in low-contrast regions where little anatomical information is available [114].

Especially in the case of MRI data, spatial intensity distortions may be present. Hence, using the localized instead of the global form of MI is preferential in this case as it evaluates the MI in subregions of the images only [109]. LMI is advantageous also for multimodal registration if one intensity class corresponds to a specific tissue type in one imaging modality and to different tissue types in the other imaging modality [109]. Finally, if only a limited number of samples is chosen during optimization to evaluate MI as performed in this study, choosing the samples from a localized region may improve the statistical power of the method.

For clinical application, a high registration accuracy in the tumor region is of major importance. Since there was a temporal delay between the acquisitions of the two PET images in our study, it is not expected that voxels of the same anatomical position have the same intensity values, but the assessment of the correlation of the PET images after registration still provides a meaningful mea-

sure of correspondence. The high NCC between the PET images indicate that surrounding anatomical structures in CT and MR give sufficient information for the registration in the low contrast tumor region. In principle, the proposed registration method could also be applied to other anatomical sites. However, a re-optimization of the free parameters of the registration method may be favorable in this case to meet the different conditions of the specific body region.

PET/MR data can be integrated into biologically adapted RT in different ways. Using contours derived from PET and MR [115] requires a high registration accuracy around the contour boundaries only. The situation is different if data is to be integrated on a voxel basis as for Dose Paining By Numbers (DPBN) [99]. In that case, intra-tumor registration accuracy is of highest importance.

To date, it is not possible to acquire simultaneous PET/MR data in the HN region with RT positioning aids. Therefore, it is beneficial that LMI+BEP yields robust results also in the case of larger positioning differences between CT and MR. Nevertheless, dedicated RT positioning aids compatible with PET/MR would be favorable, since the final geometric accuracy achieved with DR will be improved by a better initial alignment. By using appropriate positioning aids strong local deformations could be avoided, which can not be correctly mapped by the proposed algorithm due to its design. However, even with positioning aids it is likely that slight deformations remain, and a DR in addition to a RR may still be favorable. Particularly for the evaluation of treatment response, DR allows to account for a shrinkage of the tumor. For final integration of PET/MR data into RTP, the definition of PET as well as MR acquisition parameters should be optimized to meet special RT requirements, such as a high, isotropic resolution and correspondingly adjusted voxel sizes.

In conclusion, this study showed that DR with a B-spline parametrized transform combined with LMI+BEP as optimization metric yields accurate and robust results for registration of CT and MR in the HN region. As a consequence, this strategy for deformable multimodal image registration provides a basis for the integration of individual molecular, functional and anatomical PET/MR data into RTP.

4.5 Acknowledgements

This work was supported by the German Research Foundation, grant no. AL 877/1-3 and by intramural funding of the University Hospital Tübingen, fortüne grant no. 1945-0-0. DT was financially supported by the European Social Fund

and the Ministry of Science, Education and the Arts Baden-Württemberg.

4.6 Supplementary Material

Patient characteristics

Table 4.2: Patient characteristics

Patient #	Age	Tumor localization	TNM	Volume [cm^3]
Patient 1	57	Oropharynx	cT3 cN2c M0	110.3
Patient 2	70	Hypopharynx	cT4 cN2b cM0	52.4
Patient 3	62	Hypopharynx	cT4 cN2c M0	89.0
Patient 4	64	Oropharynx	cT4 cN2c M0	89.0
Patient 5	53	Oropharynx	cT3 cN2c M0	104.0
Patient 6	44	Cervical lymph node	rcN2b M0	361.1
Patient 7	57	Oropharynx	cT2 cN2b M0	30.7
Patient 8	63	Mandibula	cT2 cN0 cM0	23.2

Quantitative assessment of registration accuracy

Several measures for the determination of registration accuracy were implemented using ITK and VTK (Visualization Toolkit, www.vtk.org). Measures based on CT and MR as well as a measure derived from the PET images were defined.

Several anatomical structures that were both visible on CT and MR were segmented automatically if possible or manually otherwise, yielding volumes C and M, respectively. These structures were skin, left and right carotid, and parts of the respiratory tract. For all volumes C and M, polygonal surface meshes $\mathcal{C}$ and $\mathcal{M}$ were extracted using the marching cubes algorithm [116]. Resulting $\mathcal{C}$ were transformed to the coordinate system defined by the MR with the transform obtained by registration, yielding $\tilde{\mathcal{C}}$, from which the corresponding volume $\tilde{C}$ was derived. For each pair of corresponding structures, the Dice Similarity Index (DSI) was calculated as

$$\mathrm{DSI}(\tilde{C}, M) = \frac{2|\tilde{C} \cap M|}{|\tilde{C}| + |M|}. \tag{4.3}$$

Moreover, also mean distance measures were defined. For both skin and res-

piratory tract, mean volume distances (MVD) were calculated as

$$\mathrm{MVD}(\tilde{\mathcal{C}}, \mathcal{M}) = \frac{1}{2}\left(\frac{1}{N_\mathrm{C}}\sum_{i=1}^{N_\mathrm{C}} d(\tilde{\mathbf{c}}_i, \mathcal{M}) + \frac{1}{N_\mathrm{M}}\sum_{i=1}^{N_\mathrm{M}} d(\boldsymbol{m}_i, \tilde{\mathcal{C}})\right), \tag{4.4}$$

where $\tilde{\mathbf{c}}_i$ ($\boldsymbol{m}_i$) and N_C (N_M) are the spatial center of mesh element i and the total number of mesh elements contained in $\tilde{\mathcal{C}}$ ($\mathcal{M}$), respectively, and $d(\mathbf{p}_i, \mathcal{S})$ is the shortest distance of point $\mathbf{p}_i$ to the polygonal surface mesh $\mathcal{S}$.

For the carotid volumes C and M, lines $\mathcal{L}^\mathrm{C}$ and $\mathcal{L}^\mathrm{M}$ defining the axial centre of the structures were extracted and $\mathcal{L}^\mathrm{C}$ was transformed according to the registration result, yielding $\tilde{\mathcal{L}^\mathrm{C}}$. For each slice i in the original MR containing both a point $\tilde{\boldsymbol{c}}_i$ from $\tilde{\mathcal{L}^\mathrm{C}}$ and $\boldsymbol{m}_i$ from $\mathcal{L}^\mathrm{M}$, the Euclidean distance $d(\tilde{\boldsymbol{c}}_i, \boldsymbol{m}_i)$ was calculated and the Mean Line Distance (MLD) was obtained as

$$\mathrm{MLD}(\tilde{\mathcal{L}^\mathrm{C}}, \tilde{\mathcal{L}^\mathrm{M}}) = \frac{1}{N_\mathrm{LP}}\sum_{i=1}^{N_\mathrm{LP}} d(\tilde{\boldsymbol{c}}_i, \boldsymbol{m}_i), \tag{4.5}$$

where N_LP is the total number of valid slices.

In addition, the bony structures were segmented from the CT as well as the spinal canal from the MR image. For these two structures, the Non Overlapping Fraction (NOF) was determined

$$\mathrm{NOF}(\tilde{C}, M) = 1 - \frac{|\tilde{C} \cap M|}{|M|}, \tag{4.6}$$

where $\tilde{C}$ is the transformed volume of the bones segmented from the CT and M is the volume of the spinal canal segmented from the MR.

Anatomical landmarks were defined on basis of the CT and the MR by two experienced radiation oncologists, yielding landmarks $\{\boldsymbol{c}_i\}_{i=1}^{N_\mathrm{AL}}$ and $\{\boldsymbol{m}_i\}_{i=1}^{N_\mathrm{AL}}$. The mean number of available landmarks per patient was $N_\mathrm{AL} = 7.5$ (range: 4 – 10). For corresponding pairs of landmarks, the Mean Point Distance (MPD) after registration was evaluated

$$\mathrm{MPD}\left(\{\tilde{\boldsymbol{c}}_i\}_{i=1}^{N_\mathrm{AL}}, \{\boldsymbol{m}_i\}_{i=1}^{N_\mathrm{AL}}\right) = \frac{1}{N_\mathrm{AL}}\sum_{i=1}^{N_\mathrm{AL}} d(\tilde{\boldsymbol{c}}_i, \boldsymbol{m}_i). \tag{4.7}$$

An additional measure was introduced to quantify the registration accuracy in the tumor region. For this region, a validation of the local registration accuracy is difficult based on the image information from CT and MR, since the intratumor region can be of low contrast in these imaging modalities. However, the additional PET images available in this study provide local information

in this region. Therefore, the main tumor volume was derived from the PET intensities in the PET/CT (I_{PC}) using a threshold based segmentation method. The resulting volume was expanded by a margin of 7 mm, yielding the final volume Ω_{T}. In this volume, the Normalized Cross Correlation (NCC) of I_{PC} and the transformed PET of the PET/MR ($\tilde{I}_{\mathrm{PM}}$) is calculated:

$$\mathrm{NCC}(I_{\mathrm{PC}}, \tilde{I}_{\mathrm{PM}}) = \sum_{\boldsymbol{x}_i \in \Omega_{\mathrm{T}}} \frac{(I_{\mathrm{PC}}(\boldsymbol{x}_i) - \overline{I_{\mathrm{PC}}})(\tilde{I}_{\mathrm{PM}}(\boldsymbol{x}_i) - \overline{\tilde{I}_{\mathrm{PM}}})}{\sqrt{\sum_{\boldsymbol{x}_i \in \Omega_{\mathrm{T}}} (I_{\mathrm{PC}}(\boldsymbol{x}_i) - \overline{I_{\mathrm{PC}}})^2 \sum_{\boldsymbol{x}_i \in \Omega_{\mathrm{T}}} (\tilde{I}_{\mathrm{PM}}(\boldsymbol{x}_i) - \overline{\tilde{I}_{\mathrm{PM}}})^2}}. \tag{4.8}$$

The NCC ranges from 0 to 1, where 0 means that the images are totally uncorrelated, and 1 that there is a perfect positive linear correlation. In the case of a high local registration accuracy of CT and MR, and thus also of the PET images, the NCC is expected to have a value close to 1.

Exemplary dataset

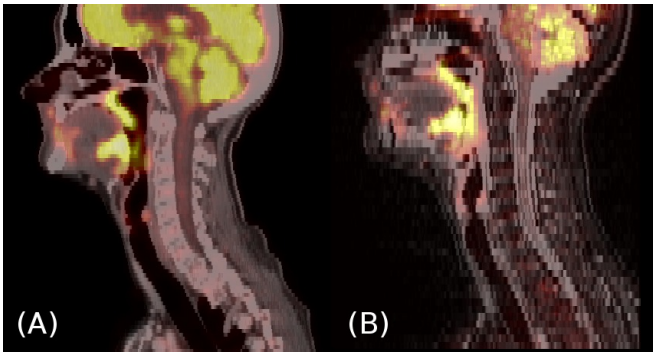

Figure 4.5: Dataset of Patient 1. (A) PET/CT, (B) PET/MR.

Registration parameters

Table 4.3: Parameters used in the multi-resolution approach of DR.

	GMI	GMI+BEP	LMI+BEP
B-spline grid spacing (mm)	60/30/15	60/30/15	60/30/15
# iterations	5000	5000	5000
# samples	10000	10000	10000
# histogram bins	60	60	60
λ	–	250	50
L_{sub} (mm)	–	–	40

The parameters below the dashed line were optimized with regard to the registration quality measures defined above. If only one value is given, this value was used in all resolution levels. For GMI+BEP and LMI+BEP different optimal values for λ were determined. For GMI+BEP, $\lambda = 250$ yielded better quantitative results than $\lambda = 50$, for which unrealistic deformations were observed for some of the patients.

5 Part II: Automatic delineation of tumor volumes by co-segmentation of combined PET/MR data

Sara Leibfarth[1*], Franziska Eckert[2], Stefan Welz[2], Christine Siegel[2], Holger Schmidt[3,4], Nina Schwenzer[3], Daniel Zips[2], Daniela Thorwarth[1]

[1] Section for Biomedical Physics, Department of Radiation Oncology, University Hospital Tübingen
[2] Department of Radiation Oncology, University Hospital Tübingen
[3] Department of Diagnostic and Interventional Radiology, University Hospital Tübingen
[4] Laboratory for Preclinical Imaging and Imaging Technology of the Werner Siemens Foundation, Department of Preclinical Imaging and Radiopharmacy, Tübingen

published in
Physics in Medicine and Biology 2015, **60**: 5399–5412.

Abstract

Combined PET/MRI may be highly beneficial for radiotherapy treatment planning in terms of tumor delineation and characterization. To standardize tumor volume delineation, an automatic algorithm for the co-segmentation of head and neck (HN) tumors based on PET/MR data was developed. Ten HN patient datasets acquired in a combined PET/MR system were available for this study. The proposed algorithm uses both the anatomical T2-weighted MR and FDG-PET data. For both imaging modalities tumor probability maps were derived, assigning each voxel a probability of being cancerous based on its signal intensity. A combination of these maps was subsequently segmented using a threshold level set algorithm. To validate the method, tumor delineations from three radiation oncologists were available. Inter-observer variabilities and variabilities between the algorithm and each observer were quantified by means of the Dice similarity index and a distance measure. Inter-observer variabilities and variabilities between observers and algorithm were found to be comparable, suggesting that the proposed algorithm is adequate for PET/MR co-segmentation. Moreover, taking into account combined PET/MR data resulted in more consistent tumor delineations compared to MR information only.

5.1 Introduction

Modern radiation therapy (RT) treatment techniques like intensity modulated RT (IMRT) allow for the application of sophisticated dose prescriptions [117]. However, to fully exploit the potential of these developments an accurate tumor volume delineation is mandatory. Furthermore, with characterization of the biological heterogeneity, non-uniform tumor dose prescriptions in terms of local dose escalation are feasible [118, 119].

For accurate tumor delineation and characterization the combination of information from different, complementary imaging modalities seems to be highly beneficial. Anatomical magnetic resonance imaging (MRI) provides superior soft tissue contrast compared to computed tomography (CT), and may thus facilitate tumor delineation for RT treatment planning. Another advantage of MRI compared to CT is the reduced extent of dental-inlay artifacts [28]. Moreover, functional MR imaging sequences like diffusion weighted MRI (DW-MRI) and dynamic contrast enhanced MRI (DCE-MRI) may contribute to a biological characterization of the tumor [98]. On the other hand, positron emission tomography (PET) imaging with tracers like [^{18}F]-fluorodeoxyglucose (FDG) or [^{18}F]-fluoromisonidazole (FMISO) may provide information about glucose metabolism or hypoxia status of the tumor [99].

In general, PET and MRI provide complementary information due to the different underlying biophysical processes. Recently, combined PET/MR systems have become available [47, 48], allowing for the simultaneous acquisition of PET and MR data. Combined PET/MRI provides intrinsically coregistered and complementary anatomical and functional information of the tumor which may be highly valuable for RT treatment planning [25, 50].

With respect to advanced head and neck (HN) carcinomas, different studies have shown that manual delineation of tumor volumes on anatomical MR images yields smaller volumes with less inter-observer variation compared to CT [120, 121]. Other studies showed that anatomical MR data may provide more accurate target volume delineation compared to CT [122, 123]. On the other hand, in the study of Daisne et al. [124] comparing information of CT, MRI and FDG-PET to histology for oropharyngeal, laryngeal and hypopharyngeal cancer sites, tumor delineations derived from FDG-PET most accurately represented the histological ground truth.

Manual delineation of tumor volumes often results in high inter-observer variabilities [125]. For improved, standardized target volume definition automatic segmentation algorithms may be beneficial. Concerning the automatic segmentation of tumor volumes from FDG-PET images, different methods have been

published previously [79,80]. However, literature addressing the segmentation of anatomical MR images has mainly been focused on brain tumors [81,82], rather than on tumors in the HN region [126]. Only few investigations have focused on multi-modality co-segmentation, where most methods have been developed with the aim of co-segmenting PET/CT data [127–129]. However, the random walk based image co-segmentation algorithm of Bagci et al. was recently applied to PET/MR data [130].

There are different challenges for the automatic segmentation of PET/MR images. For HN cancer, the tumor is often surrounded by various tissues, which may show various signal intensities in the MR image, and there may also be pronounced intra-tumor variations. Moreover, standard MR images are not quantitative, which complicates the assignment of signal intensities to underlying tissue [131]. Additionally, there may not always be sharp edges between tumor and normal tissue, but also diffuse transitions. Finally, MR images often have anisotropic voxel sizes with a low resolution in one image dimension, especially for diagnostic acquisitions. PET images are often characterized by a high level of noise and low spatial resolution, which may impede tumor volume definition especially in small uptake regions [132]. With respect to FDG-PET, there may be physiological uptake and tumor subvolumes with low tracer uptake. Finally, an appropriate way of deriving a tumor delineation respecting the information of both modalities has to be identified, which is especially demanding when there is conflicting information of PET and MR data.

In this work we present a co-segmentation algorithm for the automatic derivation of tumor delineations from combined PET/MR data comprising anatomical T2-weighted MR and FDG-PET information. The goal is to improve standardization and accuracy of target volume definition in the HN region by automating the delineation process and by integrating information from complementary modalities. The method is validated by comparing the automatically derived tumor delineations with manual contours of three experienced radiation oncologists.

5.2 Material and Methods

5.2.1 Patient data

Ten datasets from patients with HN cancer acquired in a combined PET/MR scanner were available. Information about the patient characteristics is given in Table 5.1. All patients gave their written informed consent to the PET/MR examination and the use of their data for scientific purposes. The study was

approved by the local institutional review board. PET/MR data were acquired with the Siemens Biograph mMR (Siemens Healthcare, Erlangen, Germany). A mean FDG activity of 354 MBq (range: 320 – 388 MBq) was injected. PET/MR data were acquired after a mean time of 135 min post injection (p.i.) (range: 120 – 166 min p.i.). PET images were reconstructed using OSEM 3D with 3 iterations and 21 subsets and a 3D Gaussian filter of 3 mm. MR-based PET attenuation correction was performed by a segmentation approach based on spoiled gradient-echo sequences with DIXON-based fat-water separation [61]. The MR datasets taken into account for this study were transversal T2-weighted acquisitions using a short time inversion recovery (STIR) sequence (repetition time/echo time/inversion time 4830/37/220 ms; flip angle 160 °). Voxel sizes of the STIR and the PET images were $0.7 \times 0.7 \times 4.8\,\mathrm{mm}^3$ and $2.8 \times 2.8 \times 2.0\,\mathrm{mm}^3$, respectively.

Table 5.1: Patient characteristics.

Patient	Gender	Age	Tumor localization	Histology	TNM
1	m	70	Hypopharynx	SCC[a]	cT4 cN2b cM0
2	m	41	Sinonasal	STCA[b]	rcT4a cN0 cM0[d]
3	f	57	Oropharynx	SCC	cT3 cN2c M0
4	m	62	Larynx	SCC	cT4 cN2c M0
5	m	64	Oropharynx	SCC	cT4 cN2c M0
6	m	53	Oropharynx	SCC	cT3 cN2c M0
7	m	44	Cervical lymph node	SCC	cT0 rcN2b M0
8	f	57	Oropharynx	SCC	cT2 cN2b M0
9	f	77	Nasopharynx	SCC	rcT4 cN0 M0
10	f	63	Mandibula	ES[c]	cT2 cN0 cM0

[a] squamous cell carcinoma
[b] sinonasal teratocarcinosarcoma
[c] Ewing sarcoma
[d] in analogy to sinonasal carcinomas

5.2.2 Manual segmentations

For each patient, manual delineations of the gross tumor volume were provided by three experienced radiation oncologists (FE, SW, CS), serving as a reference for assessing the accuracy of the automatic segmentation algorithm. The radiation oncologists were advised to delineate for each patient a volume considering

MR data only, and a volume considering both MR and PET data simultaneously.

5.2.3 Co-segmentation algorithm

A schematic overview of the design of the proposed algorithm for automatic PET/MR co-segmentation is shown in Figure 5.1. The PET image was resampled to the MR image in a preprocessing step using B-spline interpolation of third order. In addition, for each patient a region including the tumor was roughly pre-defined. This was required to determine the maximal PET activity concentration within the tumor (see below).

In a first step, tumor probability maps were derived for both the PET and the MR image, assigning each voxel a probability of being cancerous based on the respective signal intensity. For this purpose, a heuristic probability mapping function (PMF) was defined specifically for each modality, relating intensity and probability values on a voxel basis.

The PMF for the PET data ($\mathrm{PMF}_{\mathrm{pet}}$) was assumed to be a logistic function of the PET activity concentration a, providing the sigmoidal relationship

$$\mathrm{PMF}_{\mathrm{pet}}(a) = \frac{1}{1 + s \cdot \exp(a - t_1)}. \tag{5.1}$$

For setting parameters s and t_1 the two conditions

$$\mathrm{PMF}_{\mathrm{pet}}(a_{\max}) = 0.95 \tag{5.2}$$

$$\mathrm{PMF}_{\mathrm{pet}}(0.4 \cdot a_{\max}) = 0.50 \tag{5.3}$$

were enforced, with $a_{\max}$ being the maximum activity concentration occurring in the tumor containing region (Figure 5.1 b0). The condition given in Eq. 5.3 was motivated by current threshold-based PET tumor segmentation methods [79,80]. The application of $\mathrm{PMF}_{\mathrm{pet}}$ to the PET image data yields the corresponding tumor probability map (Figure 5.1 c0). From this map, a PET tumor volume V_{petGTV} was automatically derived using a probability of 0.5 as threshold, which is equivalent to performing a threshold segmentation on the original PET data at 40 % $\mathrm{SUV}_{\max}$. Moreover, an isotropic margin of 10 mm (V_{petMarg}) was defined expanding V_{petGTV} followed by the subtraction of the original volume from the expanded volume (Figure 5.1 d0).

The obtained volumes V_{petGTV} and V_{petMarg} were subsequently used to support an adequate definition of the PMF for the MR data ($\mathrm{PMF}_{\mathrm{mr}}$). For this purpose, histograms of the MR signal intensities in both volumes were generated, and corresponding probability density functions $\mathrm{PDF}_{\mathrm{mrGTV}}$ and $\mathrm{PDF}_{\mathrm{mrMarg}}$ were derived by kernel density estimation (Figure 5.1 a1). The MR value of highest

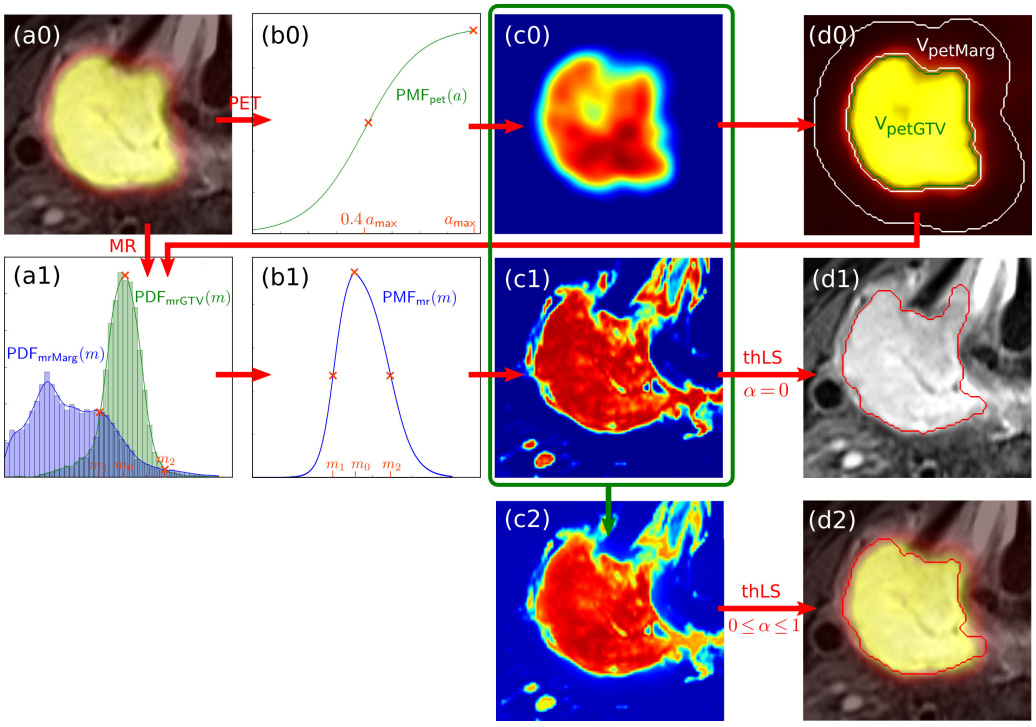

Figure 5.1: Schematic illustration of the proposed co-segmentation method. Based on the PET/MR data (a0) a PMF is defined relating PET activity concentrations a to the probability of being caused by cancerous tissue (b0). From the corresponding tumor probability map (c0), a PET tumor volume V_{petGTV} and a margin V_{petMarg} are derived (d0). With the histograms of MR signal intensities m in these volumes (a1) a PMF for the MR data is defined (b1). A combined probability map (c2) is finally obtained by the weighted sum of MR (c1) and PET probability maps (c0) according to equation 5.8. Segmenting this map with a thLS algorithm yields the final tumor volume (d2). The special case $\alpha = 0$ is equivalent to segmenting the MR probability map only (d1).

probability according to $\mathrm{PDF}_{\mathrm{mrGTV}}$ was defined as m_0, whereas the lower and the higher intersection point of $\mathrm{PDF}_{\mathrm{mrGTV}}$ and $\mathrm{PDF}_{\mathrm{mrMarg}}$ were defined as m_1 and m_2, respectively. $\mathrm{PMF}_{\mathrm{mr}}$ (Figure 5.1 b1) was finally defined as a composite of two logistic functions of MR signal intensity m

$$\mathrm{PMF}_{\mathrm{mr}}(m) = \begin{cases} \frac{1}{1+s_1 \cdot \exp(m-t_1)}, & m \leq m_0 \\ 1 - \frac{1}{1+s_2 \cdot \exp(m-t_2)}, & m \geq m_0 \end{cases} \tag{5.4}$$

Parameters s_1, s_2, t_1 and t_2 were set to meet the following conditions

$$\mathrm{PMF}_{\mathrm{mr}}(m_0) = 0.95 \tag{5.5}$$

$$\mathrm{PMF}_{\mathrm{mr}}(m_1) = 0.50 \tag{5.6}$$

$$\mathrm{PMF}_{\mathrm{mr}}(m_2) = 0.50. \tag{5.7}$$

Thus, tumor probabilities of 0.50 were assigned to MR signal intensities occurring with the same frequency in the PET-derived tumor volume and surrounding. The corresponding tumor probability map was then derived by the application of $\mathrm{PMF}_{\mathrm{mr}}$ to the MR image (Figure 5.1 c1).

Finally, a combined probability map was defined (Figure 5.1 c2) by a weighted sum of the single maps. The probability $p(m, a)$ assigned to a voxel of MR signal intensity m and PET activity concentration a in the combined map was set according to

$$p(m, a) = (1 - \alpha)\, \mathrm{PMF}_{\mathrm{mr}}(m) + \alpha \mathrm{PMF}_{\mathrm{pet}}(a), \quad 0 \leq \alpha \leq 1. \tag{5.8}$$

The parameter α determines the weight of the PET information relative to the MR information and can be adapted to meet specific requirements.

The tumor delineation was subsequently derived with the threshold level set (thLS) segmentation algorithm provided by the Insight Segmentation and Registration Toolkit (ITK)*, version 4.5.2 (Figure 5.1 d2). In this method, the tumor surface is represented by an isocontour of an implicit function which is iteratively evolved according to voxel-based information of a feature image and a regularization term [133]. In this particular case, the combined tumor probability map was used as feature image and the algorithm was initialized with the signed distance map of the segmented PET volume V_{petGTV}.

For all segmentations, the curvature scaling parameter determining the amount of regularization was set to 0.2, whereas the maximum number of iterations was set to 1000. To derive an optimal value of α, segmentations were performed on different probability maps, comprising α values ranging from 0 to 1 with a step size of 0.1. Moreover, the threshold value of the thLS algorithm ($\mathrm{th}_{\mathrm{LS}}$) was optimized within the range from 0.30 to 0.55, with a step size of 0.05.

*www.itk.org

5.2.4 Evaluation of segmentation accuracy

Due to the limited number of patient datasets, no independent validation dataset was available. To ensure that the optimized parameters are reasonable in terms of not being overfitted, a leave-one-out cross-validation was performed. Parameters were optimized separately for each subset of nine patients and tested on the remaining patient.

For evaluating the accuracy of the proposed co-segmentation algorithm, the automatically obtained delineations were compared to those manually defined by the three observers. In addition to the qualitative visualization of the manually and automatically derived delineations, also a quantitative evaluation of both inter-observer variability and variability between manually and automatically defined segmentations was performed. Mutual scores of Dice similarity index (DSI) as well as mutual distances between the surfaces of the tumor volumes were evaluated. For distance quantification, a surface mesh was derived from each volume by means of the marching cubes algorithm [116] using the Visualization Toolkit (VTK)[†]. The comparison of two volumes V_1 and V_2 was subsequently performed by calculating the smallest distances of the centers of each mesh element from the surface mesh of V_1 to the surface mesh of V_2. To treat the volumes symmetrically, calculations were also performed vice versa. The fraction of distances smaller than 3 mm was used as measure of the agreement between the tumor volumes. Thus, distances smaller than 3 mm are regarded as negligible, which is motivated by the typical size of the grid used for RT dose calculation.

On the one hand, DSI and distance measures were evaluated for tumor delineations derived from combined PET/MR information, where for the automatically derived segmentations parameters α and $\mathrm{th}_{\mathrm{LS}}$ were optimized towards optimal agreement with the manual delineations. On the other hand, manually defined delineations from MR-only information were compared to segmentations derived automatically with $\alpha = 0$ and the $\mathrm{th}_{\mathrm{LS}}$ obtained by PET/MR-based optimization.

Tumor delineations derived from combined PET/MR information were tested for statistically significant differences to MR-only derived delineations with a two-sided paired Wilcoxon test. p-values lower than 0.05 were considered statistically significant.

[†]www.vtk.org

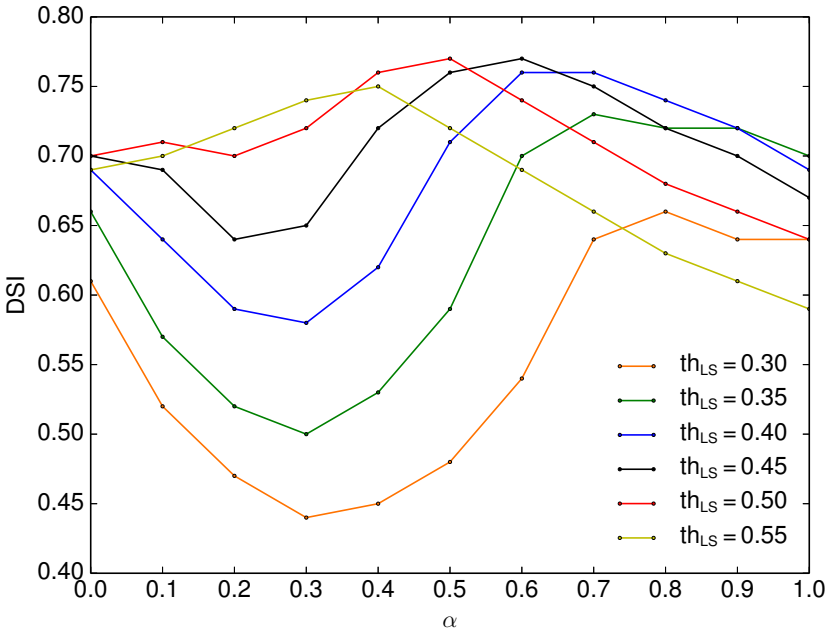

Figure 5.2: DSI scores between tumor delineations defined manually using combined PET/MR information and delineations derived automatically by co-segmentation with different combinations of α and $\mathrm{th}_{\mathrm{LS}}$. Each point represents the median DSI value between the automatic segmentation and the manually derived volumes averaged over all patients.

5.3 Results

In Figure 5.2, the tumor delineations yielded by the segmentation algorithm are compared to those of the observers for different values of the parameters α and $\mathrm{th}_{\mathrm{LS}}$. Each point represents the median DSI value between the automatic segmentation and the manually derived delineations averaged over all patients. Using two significant digits, best agreement was obtained for the parameter combinations $(\alpha, \mathrm{th}_{\mathrm{LS}}) = (0.5, 0.5)$ and $(\alpha, \mathrm{th}_{\mathrm{LS}}) = (0.45, 0.6)$. Results of the leave-one-out cross-validation are shown in Table 5.2. Since the determined optimal parameters are very similar across the training datasets, the cross-validation indicates that the parameter optimization is robust and may also perform well for an independent validation dataset. $(\alpha, \mathrm{th}_{\mathrm{LS}}) = (0.5, 0.5)$ was subsequently used for a comprehensive comparison of the algorithm and the delineations defined manually in the PET/MR images, whereas $(\alpha, \mathrm{th}_{\mathrm{LS}}) = (0, 0.5)$ was evaluated against MR-only derived delineations.

Figure 5.3 shows manually and automatically derived tumor contours for pa-

Table 5.2: Results of the leave-one-out cross-validation.

Patient	$(\alpha, \mathrm{th_{LS}})^{a}$	DSI training	DSI validation
1	1, 2	0.76	0.86, 0.87
2	1, 2	0.76	0.87, 0.89
3	2	0.78	0.76
4	1, 2	0.77	0.78, 0.81
5	2, 3	0.77	0.82, 0.71
6	1, 2	0.79	0.54, 0.59
7	1, 2	0.79	0.60, 0.62
8	2	0.79	0.67
9	2	0.77	0.83
10	2	0.76	0.88

[a] $(\alpha, \mathrm{th_{LS}}) = (0.5, 0.5)$, 2: $(\alpha, \mathrm{th_{LS}}) = (0.45, 0.6)$, 3: $(\alpha, \mathrm{th_{LS}}) = (0.45, 0.5)$.
Note: For each patient, the optimal parameter combination $(\alpha, \mathrm{th_{LS}})$ for the complementary training dataset, consisting of the other nine patients, is shown together with the respective patient-averaged DSI value (DSI training, two significant digits). Moreover, the DSI obtained for the respective patient with the obtained parameter combination(s) is given (DSI validation). The DSI score taken into account is the median value between the automatic and the manually derived delineations.

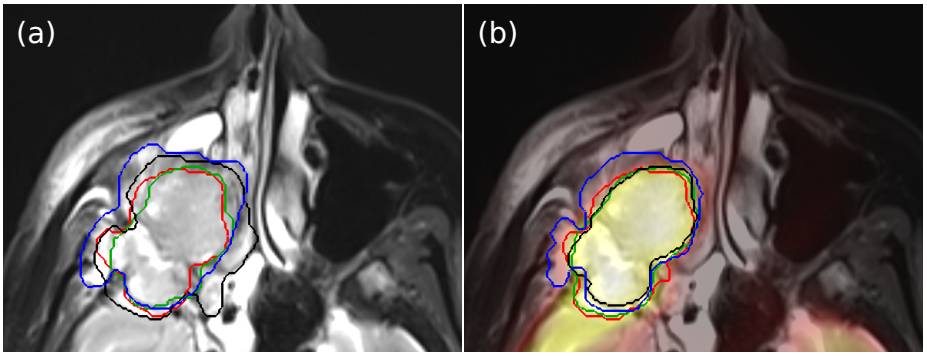

Figure 5.3: Exemplary tumor contours for patient 9. Contours of MR-based tumor delineations from the three observers (green, black, blue) and co-segmentation algorithm (red) with $(\alpha, \mathrm{th}_{\mathrm{LS}}) = (0,\ 0.5)$ are shown (a), as well as PET/MR-derived contours from the observers and the co-segmentation algorithm with $(\alpha, \mathrm{th}_{\mathrm{LS}}) = (0.5,\ 0.5)$ (b).

tient 9. For both MR-only and for PET/MR contours inter-observer variations as well as variations between the observers and the algorithm were observed. For a more detailed illustration of accordances and discrepancies of PET/MR-derived delineations, Figure 5.4 shows several slices with contours of both the observers and the automatic algorithm. From left to right variabilities between the contours increase, with high discrepancies in the last slices.

The volumes of the manually and automatically derived tumor segmentations using MR-only as well as combined PET/MR information are visualized in Figure 5.5. High differences in the volumes occurred among the observers when tumor boundaries could not be clearly identified on the image data (cf. patient 7). Other cases showed only minor variations for both MR-only and combined PET/MR derived tumor volumes (cf. patient 10). In general, additional PET information can both lead to an increase and a reduction of the defined tumor volume. However, volumes tended to decrease as shown by the volumes averaged over all patients.

Quantitative results of the variability between tumor delineations in terms of DSI and distance evaluations are shown in Figure 5.6 for both MR-based delineations and PET/MR-based delineations. A DSI value of 0.67 ± 0.17 (mean $\pm$ stddev) was obtained for MR-based delineations when combining the result of the three pairs of observers, whereas 0.69 ± 0.14 was obtained for the comparison of manually and automatically defined delineations. For combined PET/MR information values were 0.72 ± 0.13 and 0.75 ± 0.12, respectively. For the fraction of distances smaller than 3 mm, MR-only derived delineations yielded 0.78 ± 0.17

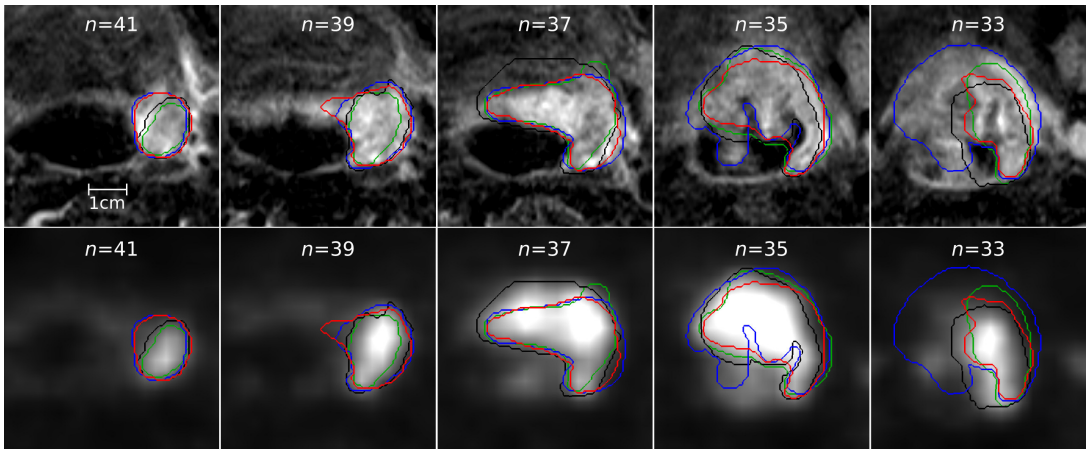

Figure 5.4: Tumor contours obtained for patient 4. PET/MR-derived contours from the observers (green, black, blue) and the co-segmentation algorithm (red) with $(\alpha, \mathrm{th}_{\mathrm{LS}}) = (0.5,\ 0.5)$ are shown overlayed on the MR (first row) and PET images (second row), with n indicating the corresponding slice number.

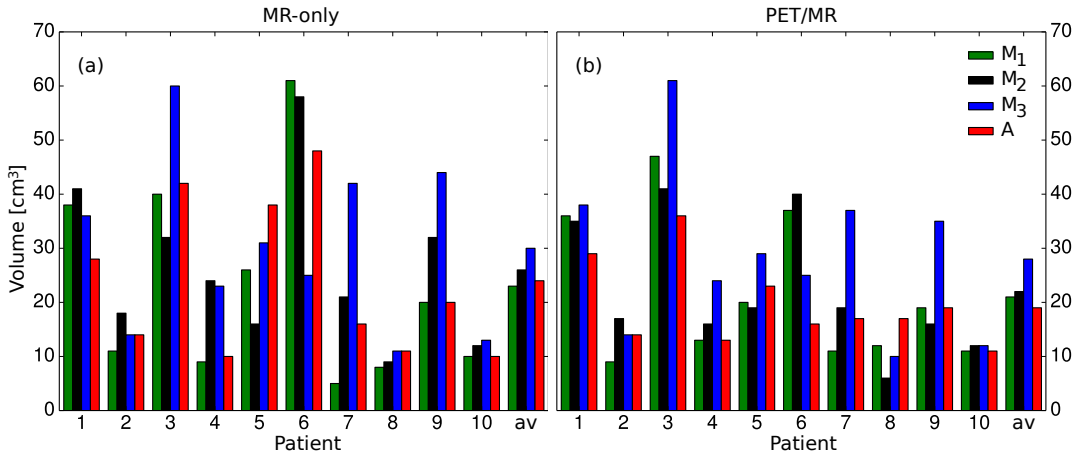

Figure 5.5: Volumes of tumor delineations derived by the observers ($\mathrm{M}_1/\mathrm{M}_2/\mathrm{M}_3$) and the co-segmentation algorithm (A) for each patient. 'av' denotes the volume averaged over all ten patients. Both MR-derived volumes with $(\alpha, \mathrm{th}_{\mathrm{LS}}) = (0,\ 0.5)$ for automatic segmentation (a) and PET/MR-derived volumes with $(\alpha, \mathrm{th}_{\mathrm{LS}}) = (0.5,\ 0.5)$ (b) are shown.

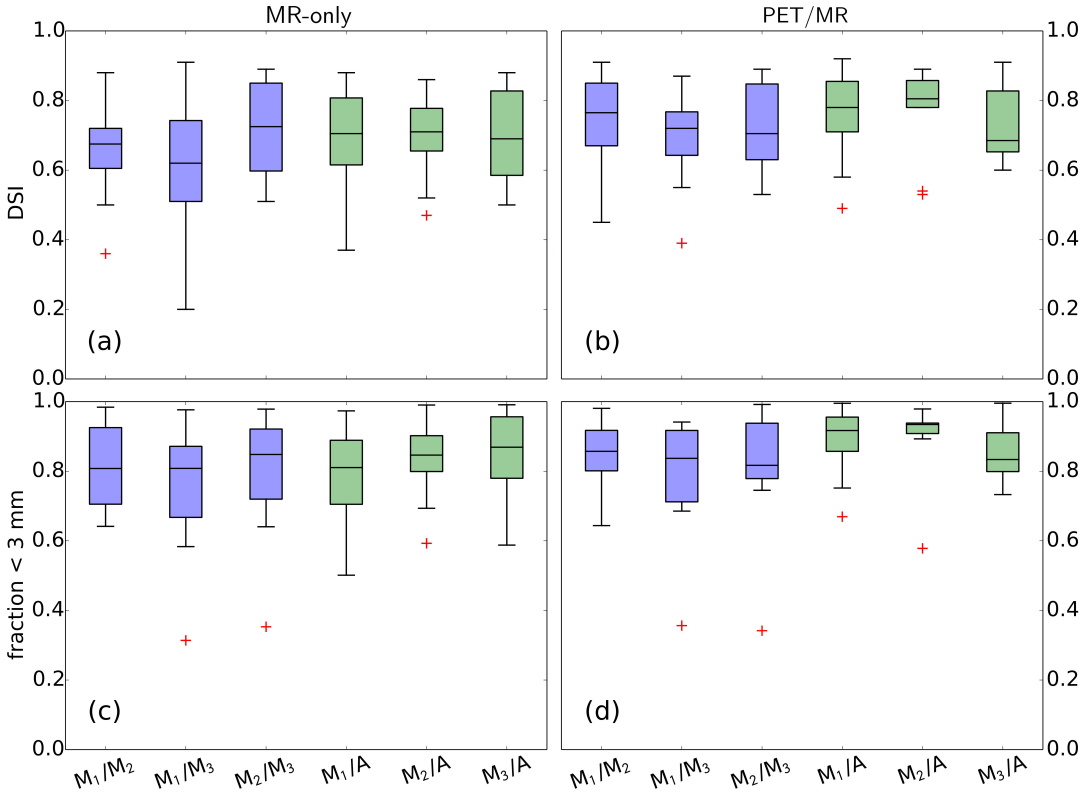

Figure 5.6: Comparison of tumor volumes from the three different observers ($M_1/M_2/M_3$) and the automatic method (A). Boxplots show mutual DSI values (a,b) and fractions of inter-surface distances within 3 mm (c,d) obtained for the patient cohort ($N = 10$). Both MR-only volumes with $(\alpha, \mathrm{th}_{\mathrm{LS}}) = (0,\ 0.5)$ for automatic segmentation and PET/MR-derived volumes with $(\alpha, \mathrm{th}_{\mathrm{LS}}) = (0.5,\ 0.5)$ are shown.

for inter-observer comparison and 0.82 ± 0.14 for comparison between automatically and manually derived delineations. For combined PET/MR information values changed to 0.81 ± 0.16 and 0.88 ± 0.10, respectively. Both the evaluation of DSI and distances show that substantial differences in tumor delineation were obtained for both MR-only and PET/MR derived volumes. However, both the concordance between manual volumes and between manually and automatically derived delineations increased significantly when complementing MR information by PET information according to DSI as well as distance evaluation ($p < 0.05$). Most importantly, variabilities between automatic and manual contours are comparable to the inter-observer variabilities for both MR-only and PET/MR derived delineations.

5.4 Discussion

In this work, an algorithm for co-segmenting tumor volumes based on PET/MR information was presented and evaluated against manual delineations of three observers. Since there are free parameters, the proposed method has potential for calibration. We optimized the parameter defining the relative weight of the probability maps, α, and the threshold used in the thLS algorithm, $\mathrm{th_{LS}}$, with respect to maximum agreement of the automatically defined volumes with those of the three observers. This yielded the parameter combinations $(\alpha, \mathrm{th_{LS}}) = (0.5,\ 0.5)$ and $(\alpha, \mathrm{th_{LS}}) = (0.6,\ 0.45)$, respectively, using two significant digits for comparing DSI scores. With $\alpha = 0.5$ the PET and MR data contribute equally to the segmentation. Since α was optimized with respect to best agreement with the three observers, this indicates that the observers tended to incorporate MR- and PET-based information to a similar extend for deriving contours from PET/MR data. However, also other values for α would have been plausible, because an increased overall reliance on one or the other modality could in general be reasonable. $\mathrm{th_{LS}} = 0.5$ means that the segmentation border is determined by voxels having equal probabilities of being cancerous or not cancerous. An optimal value of about 0.5 thus suggests that the PMFs are defined in a reasonable way.

The leave-one-out cross-validation indicates robust parameter optimization without overfitting to the available patient cohort. Moreover, MR-based segmentation ($\alpha = 0$) yields results which are within inter-observer variability using only one optimized parameter ($\mathrm{th_{LS}}$). However, an independent validation of the algorithm should be performed when more patient datasets are available. Also the performance of the algorithm on other tumor sites could be evaluated in a further work.

Tumor volumes obtained by the co-segmentation algorithm tended to be smaller than volumes defined by the observers (cf. Figure 5.5), especially for delineations derived from combined PET/MR information. In general, larger volumes could be obtained with a $\mathrm{th_{LS}}$ value lower than 0.5. However, $\mathrm{th_{LS}}$ was optimized with respect to DSI, which not only compares sizes of derived tumor volumes, but also the spatial concordance. For further assessment of this tendency, the evaluation on more patient datasets would be preferable.

Comparison of mutual DSI scores and mutual distances of tumor volumes showed that variabilities between different observers are similar to variabilities between manually and automatically derived volumes. This suggests that the automatic co-segmentation algorithm yields results with an accuracy comparable to clinical practice. Compared to MR-only information, using combined

PET/MR information resulted in a statistically significant improvement of the agreement between observers, as well as between observers and the automatic algorithm, supporting the assumption that the usage of multi-modality data increases the consistency of tumor definition [134]. However, both inter-observer variabilities and variabilities between observers and the automatic segmentation method are still considerable. These variabilities are likely to be reduced when more complementary imaging information is available, e.g. from additional MR sequences. Since the proposed co-segmentation method is modular, additional imaging data could easily be integrated by defining appropriate probability maps. As an example, quantitative functional data derived from DW- or DCE-MRI could help to obtain more accurate delineations. Compared to anatomical MR data, the definition of PMFs based on quantitative image data is simplified, since their voxel values have a well-defined meaning.

Segmentation of anatomical MR images is especially challenging since there is no standard intensity scale in MRI, which impedes a straightforward assignment of signal intensities to underlying tissue [131]. Moreover, high signal intensities in T2-weighted images may be caused by increased water content due to inflammation or edema [135]. Thus, contrarily to PET, a PMF monotonically increasing with signal intensity may not be adequate for MR images. However, the results obtained by the proposed algorithm for MR-only segmentation show that information from PET data can guide the derivation of an appropriate PMF for MR data. Though providing complementary information, FDG-PET seems to be sufficiently correlated with the MR information for this purpose. The method proposed in this study to include MR information in a margin V_{petMarg} encompassing the PET-derived tumor segmentation (cf. Figure 5.1 d0) was found to be crucial for defining an adequate PMF for the MR data. Thus, information about the contrast between tumor and surrounding tissue seems to be mandatory for this purpose.

An additional issue concerning MR images is the high level of anisotropy with a low resolution in one direction, which is often present especially in diagnostic images. The MR images available in this study had an out-of-plane voxel spacing of 4.8 mm. Since this limits segmentation accuracy, high resolution isotropic acquisitions would be preferable [136].

Concerning the PET images, the relative threshold segmentation and consequently also the tumor probability maps may be influenced by the method used for PET image reconstruction [137]. When using different image reconstruction methods, the influence on the algorithm should be evaluated and thresholds should be adapted, if necessary. Moreover, PET images from combined PET/MRI are influenced by MR-based attenuation correction. This constitutes

a potential source of uncertainty for PET quantification. In particular, the PET data is not completely independent of MR-based information [138].

In the proposed method, the thLS algorithm is used for segmenting the tumor probability maps. In contrast to a simple threshold segmentation, thLS is not only based on bare voxel intensities but also includes a regularization of the segmented surface. Moreover, the risk of "leakage" of the segmentation is reduced compared to the conventional level set algorithm [139], for which difficulties occur at weak tumor boundaries since it depends on edge information. Level set algorithms allow for topological changes during optimization, which renders the algorithm adaptive to diverse tumor geometries such as a tumor enclosing the trachea. Moreover, the regularization is performed in three dimensions, therefore the segmentations between adjacent slices are interdependent.

In general, standardization of tumor volume delineation is highly beneficial. This supports the usage of automatic tumor segmentation algorithms. However, a visual inspection is still mandatory and may imply manual corrections. Nonetheless, if a common starting segmentation is provided, inter-observer variations are supposed to decrease.

For a further validation of the algorithm, a comparison to histology would be preferable [124], as histology most closely reveals the ground truth. Moreover, histological details would also provide more evidence for the definition of PMFs. Such an evidence-based PMF definition could be guided by the sensitivities and specificities of the respective imaging modalities.

To finally integrate results into RT treatment planning, the tumor segmentation derived from PET/MR has to be transferred to the CT acquired for treatment planning. In general, an accurate deformable registration strategy is necessary for spatially matching the PET/MR data to the planning CT [140]. However, these algorithms are hard to validate and spatial inaccuracies remain. Since registration accuracy increases with identical patient positioning, a solution allowing for RT specific positioning of HN patients in the PET/MR system [141] would be highly preferable. Another perspective is PET/MR-only treatment planning. For this approach, CT equivalent information has to be derived from MRI for RT dose planning as well as for PET attenuation correction [67]. In combination with an accurate PET/MR-based tumor segmentation strategy and methods for adapting the dose prescription within the tumor according to functional datasets and associated locoregional control probabilities, a biologically individualized PET/MR-only based treatment planning could be possible in the future.

5.5 Conclusion

The proposed PET/MR co-segmentation algorithm yielded tumor volumes which were comparable to manually defined volumes when taking into account mutual variabilities. Following further evaluation of the algorithm it may be used as a tool to define a standardized starting point for RT tumor volume definition, which consequently may reduce inter-observer variations. Additional imaging information can readily be incorporated in the method. Therefore, it provides a basis for integrating multimodal imaging information into RT planning.

5.6 Acknowledgements

This project was financially supported by the Ludwig-Hiermaier-Stiftung of the Comprehensive Cancer Center Tübingen.

6 Part III: Analysis of pairwise correlations in multi-parametric PET/MR data for biological tumor characterization and treatment individualization strategies

Sara Leibfarth[1*], Urban Simoncic[1,4,5], David Mönnich[1], Stefan Welz[2] , Holger Schmidt[3], Nina Schwenzer[3], Daniel Zips[2], Daniela Thorwarth[1]

[1] Section for Biomedical Physics, Department of Radiation Oncology, University Hospital Tübingen
[2] Department of Radiation Oncology, University Hospital Tübingen
[3] Department of Diagnostic and Interventional Radiology, University Hospital Tübingen
[4] Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana, Slovenia
[5] Jozef Stefan Institute, Ljubljana, Slovenia

published in
European Journal of Nuclear Medicine and Molecular Imaging 2016, **43**: 1199–1208.

Abstract

Purpose The aim of this pilot study was to explore simultaneous functional PET/MR for biological characterization of tumors and potential future treatment adaptations. To investigate the extent of complementarity between different PET/MR-based functional datasets, a pairwise correlation analysis was performed.
Methods Functional datasets of N=15 head and neck (HN) cancer patients were evaluated. For patients of group A (N=7), combined PET/MR datasets including FDG-PET and ADC maps were available. Patients of group B (N=8) had FMISO-PET, DCE-MRI and ADC maps from combined PET/MRI, an additional dynamic FMISO-PET/CT acquired directly after FMISO tracer injection as well as an FDG-PET/CT acquired a few days earlier. From DCE-MR, parameter maps K^{trans}, v_{e} and v_{p} were obtained with the extended Tofts model. Moreover, parameter maps of mean DCE enhancement, $\overline{\Delta S}_{\mathrm{DCE}}$, and mean FMISO signal 0-4 min p.i., $\overline{A}_{\mathrm{FMISO}}$, were derived. Pairwise correlations were quantified using the Spearman correlation coefficient (r) on both a voxel and a regional level within the gross tumor volume.
Results Between some pairs of functional imaging modalities moderate correlations were observed with respect to the median over all patient datasets, whereas distinct correlations were only present on an individual basis. Highest inter-modality median correlations on the voxel level were obtained for FDG/FMISO ($r = 0.56$), FDG/$\overline{A}_{\mathrm{FMISO}}$ ($r = 0.55$), $\overline{A}_{\mathrm{FMISO}}$/$\overline{\Delta S}_{\mathrm{DCE}}$ ($r = 0.46$), and FDG/ADC ($r = -0.39$). Correlations on the regional level showed comparable results.
Conclusion The results of this study suggest that the examined functional datasets provide complementary information. However, only pairwise correlations were examined, and correlations could still exist between combinations of three or more datasets. These results might contribute to the future design of individually adapted treatment approaches based on multiparametric functional imaging.

6.1 Introduction

Biological tumor characterization based on functional and molecular imaging might be highly valuable for radiotherapy (RT). On the one hand, it could allow for an improved target volume definition and an individualized dose prescription within the tumor according to local biological characteristics. Such dose painting strategies can be readily applied with the technical availability of intensity modulated RT (IMRT). Moreover, functional imaging might be of high value for early response assessment and potential treatment adaptation in the course of fractionated RT [52, 53]. Other fields of application are the assessment of chemotherapy and the application of targeted agents, such as hypoxia-sensitizing or antiangiogenic drugs [86, 87].

Both positron emission tomography (PET) and magnetic resonance imaging (MRI) may provide functional information beneficial for personalized treatment strategies. PET imaging using [^{18}F]-fluorodeoxyglucose (FDG) can be used to monitor glucose metabolism, whereas the hypoxic status of the tumor can be assessed using dedicated tracers such as [^{18}F]-fluoromisonidazole (FMISO). Diffusion weighted MRI (DW-MRI) provides the possibility to quantify the diffusion of water molecules, which is related to cellular density [29]. Dynamic contrast-enhanced MRI (DCE-MRI) yields a temporally varying signal due to the distribution of contrast agent in blood pool and tissue. By compartmental modeling estimates of quantitative physiological parameters can be derived [92].

With the advent of combined PET/MR imaging [47, 48] the acquisition of simultaneous, intrinsically registered PET and MR data has become possible. This facilitates the comparison and combined analysis of PET- and MR-derived functional imaging data. Simultaneous PET/MR may thus be of high potential for treatment individualization [25, 50].

Recent studies have associated different functional imaging information with RT outcome for head and neck (HN) cancer. This applies to FDG-PET [17, 41], static as well as dynamic FMISO-PET [18, 19, 43, 44], apparent diffusion coefficients (ADCs) inferred by DW-MRI [14], as well as DCE-MRI [15, 16]. These studies provide a rationale to adapt RT treatment plans according to functional imaging information.

It is not clear yet if datasets from different functional imaging modalities are completely complementary, or if information is to some extent redundant. Initial analyses of correlations between different functional datasets have already been performed in recent studies. The studies of Rajendran et al. [142] and Thorwarth et al. [143] revealed good voxel-by-voxel correlation of FDG and FMISO in some HN tumors, whereas others showed no clear correlation. The biological basis of

the observed correlations may be the hypoxia-inducible factor 1α (HIF 1α) [142]. Similar results were obtained by Zegers et al. [144] comparing uptake of FDG and the hypoxia PET tracer [^{18}F]-HX4 in patients with non–small cell lung cancer. Houweling et al. [145] quantified correlations between FDG and ADC maps of HN tumors on a voxel level, and found a negative correlation in most patients. Both Newbold et al. [146] and Donaldson et al. [147] found correlations between hypoxia derived from pimonidazole staining and DCE-derived parameter maps on a region-of-interest (ROI) level. A study by Jansen et al. [148] found that neck nodal metastases with positive FMISO uptake differed significantly in median K^{trans} values from those with no FMISO uptake.

Earlier studies have shown that a dynamic imaging protocol may be superior compared to a single time frame for hypoxia quantification using FMISO-PET [44]. However, in addition to a late static scan several hours post injection (p.i.), such a dynamic protocol requires a PET acquisition during tracer wash-in in the first minutes p.i. [149], which may hamper its usage in clinical routine. A positive correlation result between early FMISO and DCE information would potentially provide the possibility to infer early FMISO information from DCE, which would facilitate its clinical usage.

To address the question if available functional information of PET/MR is complementary or to some extend redundant, this study extends beyond existing studies by considering a comprehensive set of functional data. Correlations of various functional datasets are quantified on a voxel as well as on a regional level within HN tumors by means of the Spearman correlation coefficient. For the analysis, FDG-PET, FMISO-PET acquired in the wash-in, as well as in the retention phase, ADC maps extracted from DW-MRI, and DCE-MRI derived maps are taken into account. The study is a first explorative, hypothesis generating approach to investigate the utilization of integrated PET/MR for personalized treatment strategies.

6.2 Material and Methods

6.2.1 Patient data

Datasets from 15 HN cancer patients from two different studies were available in total, examined with combined PET/MR (Biograph mMR, Siemens Healthcare, Erlangen, Germany) and PET/CT (Biograph mCT, Siemens Healthcare, Erlangen, Germany) before the start of RT. The studies were approved by the local ethics committee. All patients gave written informed consent for participating in the imaging studies.

For 7 patients (group A) the PET/MR imaging session was performed about 2 h (120 – 166 min, median: 129 min) after injection of FDG (320 – 388 MBq, median: 357 MBq). The other 8 patients (group B) were imaged 0–40 min after injection of FMISO (165 – 377 MBq, median: 339 MBq) in PET/CT using a dynamic acquisition mode, with a subsequent PET/MR imaging session about 3 h p.i. (164 – 206 min, median: 174 min). For these patients, an additional FDG-PET/CT (307 – 354 MBq, median: 330 MBq) acquired 1 – 30 days earlier (median: 8 days) at about 1 h p.i. (55 – 81 min, median: 71 min) was also available. An overview of the patient cohort including the imaging data available for each patient is shown in Table 6.1.

PET images obtained from PET/MR were reconstructed to a voxel size of $2.8 \times 2.8 \times 2.0\,\mathrm{mm}^3$ using an OSEM 3D algorithm with 2 iterations and 21 subsets (2i21s) and a 3D Gaussian filter of 4 mm. MR-based PET attenuation correction was performed by a vendor-provided segmentation approach based on spoiled gradient-echo sequences with DIXON-based fat-water separation [61]. FMISO-PET images from the PET/CT were reconstructed to a voxel size of $4.1 \times 4.1 \times 5.0\,\mathrm{mm}^3$ using OSEM 3D with 4i8s and a 3D Gaussian filter of 5 mm. FDG-PET images from the PET/CT were reconstructed to a voxel size of $2.0 \times 2.0 \times 3.0\,\mathrm{mm}^3$ using OSEM 3D with 3i24s and a 3D Gaussian filter of 3 mm.

MRI acquisitions at the Biograph mMR were performed with the standard 16 channel head neck coil. An anatomical, transversal T2-weighted acquisition using a short time inversion recovery (STIR) sequence was acquired for each patient (repetition time (TR)/echo time (TE)/inversion time (TI) = 4830 ms/37 ms/220 ms, flip angle 160°, voxel size $0.7 \times 0.7 \times 4.8\,\mathrm{mm}^3$, bandwidth 220 Hz/px, 2 averages, acquisition time 3m58s).

In addition, DW-MR images were obtained using a single-shot spin-echo echo-planar imaging (TR/TE = 7400 ms/49 ms, b-values $50\,\mathrm{s/mm}^2$ and $800\,\mathrm{s/mm}^2$, bandwidth 2083 Hz/px, voxel size $2.1 \times 2.1 \times 5.0\,\mathrm{mm}^3$, 3 averages, spectral attenuated inversion recovery fat suppression, acquisition time 2m26s). ADC maps were obtained from the scanner software (Syngo MR B18P).

For N=5 patients also DCE-MR datasets were obtained. An axial T1-weighted fast spoiled gradient echo sequence (TWIST, TR/TE = 2.86 ms/1.01 ms, flip angle 12°, voxel size $1.1 \times 1.1 \times 4.0\,\mathrm{mm}^3$, temporal resolution 2.9 s, bandwidth 530 Hz/px, acquisition time 4m18s) was performed after an automatic fast bolus injection of 0.1 mmol Gd-DTPA per kg patient weight, followed by a saline flush. The field of view included the entire tumor and the common carotid arteries.

For the derivation of the native longitudinal relaxation times needed for DCE-quantification, additional VIBE sequences were acquired with two different flip

Table 6.1: Patient characteristics and acquired datasets.

Patient	Gender	Age	Tumor localization	volume of GTV [cm^3]	imaging modalities[a]	DCE-MRI[b]	DW-MRI[c]	# samples voxel level	# samples regional level
1	m	70	Hypopharynx	35	A	-	x	453	7^+
2	f	57	Oropharynx	41	A	-	x	2519	24
3	m	62	Larynx	16	A	-	x	1050	11
4	m	64	Oropharynx	19	A	-	x	1172	13
5	m	44	Cervical lymph node	19	A	-	x	1176	12
6	f	77	Nasopharynx	16	A	-	x	1005	12
7	m	52	Hypopharynx	16	A	-	x	743	6^+
8	f	62	Base of tongue	16	B	x	(x)*	1008	13
9	f	57	Oropharynx	23	B	x	x	423	21
10	m	48	Oropharynx	33	B	x	(x)*	2112	30
11	m	56	Hypopharynx	50	B	x	x	3148	45
12	m	64	Oropharynx	167	B	-	x	10462	166
13	m	57	Oropharynx	10	B	-	x	638	6^+
14	f	55	Oropharynx	56	B	x	x	3556	52
15	f	53	Oropharynx	32	B	-	x	2005	27

[a] A: FDG-PET/MR 2 h p.i., B: FDG-PET/CT 1 h p.i. + dynamic FMISO-PET/CT 0–40 min p.i. + FMISO-PET/MR 3 h p.i., [b] x: DCE-MRI acquired, -: no DCE-MRI available, [c] x: ADC-map acquired and evaluated, (x)* ADC-map omitted due to spatial distortions (according to visual assessment), $^+$ not used for evaluation due to small sample size

angles ($\alpha_1 = 2\,°, \alpha_2 = 12\,°$) before contrast agent injection (TR/TE = 4.04 ms/1.52 ms). The image grid was identical to the one of the DCE-MR acquisitions.

6.2.2 Calculation of parameter maps

The activity of the dynamic FMISO-PET datasets acquired during tracer wash-in was integrated for each voxel between 0 to 4 min p.i. using the rectangle method. By normalizing with respect to the acquisition time range of 4 min, a map of mean activity, $\overline{A}_{\mathrm{FMISO}}$, was obtained. For DCE images, the time-dependent signal enhancement of each voxel was calculated by subtraction of the mean signal before contrast agent injection. Subsequently, maps of mean signal enhancement $\overline{\Delta S}_{\mathrm{DCE}}$ from 0 to 4 min p.i. were calculated analogous to the $\overline{A}_{\mathrm{FMISO}}$ maps.

Before compartmental analysis, DCE images were resampled to the FMISO image grid from PET/MR. Signal-time-curves of DCE were fitted with an in-house implemented software (Matlab R2014b) using the extended Tofts model [90] and the Levenberg-Marquardt least squares algorithm. The arterial input function (AIF) was derived for each patient independently from a fit of the signal-time curve in the common carotid artery. Parameter maps of the volume fraction of the extracellular-extravascular space (EES), v_{e}, the volume fraction of the blood plasma, v_{p}, and the volume transfer rate from plasma to the EES, K^{trans}, were obtained. For regional analysis (see below) compartmental analysis was performed separately on the regional level of $3 \times 3 \times 4$ voxels of the resampled DCE images.

6.2.3 Image registration and tumor volume delineation

For performing the correlation analysis, the GTV of each patient from group A was delineated by an experienced radiation oncologist based on combined information of the FDG-PET and the T2-weighted STIR image [150]. ADC maps were resampled to match the FDG image grid with b-spline interpolation of 3rd order using the Insight Segmentation and Registration Toolkit (ITK version 4.5.2, www.itk.org).

For delineation of the GTVs of group B, manual contours defined by an experienced radiation oncologist on corresponding planning CTs were transferred to the PET/MR datasets by deformable registration of the CT and STIR images. The registrations were performed with elastix [110] using a b-spline parametrized transform and mutual information as similarity measure. Details of the used deformable registration algorithm and the respective parameter set are given

in [140]. ADC images from PET/MR were resampled to match the FMISO image grid. Additionally, the FDG image from PET/CT was transformed to the FMISO image grid from PET/MR by deformable registration of the corresponding CT and STIR images with the method described above. The $\overline{A}_{\mathrm{FMISO}}$ map was registered to the PET/MR dataset in the same way.

6.2.4 Statistical analysis

Correlation analysis was performed for all available pairwise combinations of functional and parametric maps with Python 2.7.6 using the SciPy library (www.scipy.org). Correlations were quantified by evaluating for each patient seperately the Spearman correlation coefficient on a voxel as well as on a regional level within the GTV. For voxel-based analysis, samples were defined by the PET image grid from PET/MR. For regional analysis, samples were defined as averages from non-overlapping sub-regions of the GTV. Each sub-region was defined over $3 \times 3 \times 4$ voxels of the PET image grid, corresponding to a size of $8.4 \times 8.4 \times 8\,\mathrm{mm}^3$. Patients with less than ten subregions were excluded from regional analysis (cf. Table 6.1).

6.3 Results

All functional images and parameter maps used for pairwise correlation analysis, together with the anatomical STIR acquisition and the delineated GTV, are exemplarily visualized for Patient 11 in Figure 6.1.

Exemplary scatter plots of the voxel-based pairwise correlation analysis are shown in Figure 6.2, visualizing results of two exemplary patients. Scatter plots and corresponding correlation coefficients show that there were patients for which pairs of functional data which showed rather strong correlations (e.g. FDG/ADC, $\overline{\Delta S}_{\mathrm{DCE}}/\overline{A}_{\mathrm{FMISO}}$ for Patient 11), while for other patients the correlations of the same pairs were much lower (cf. Patient 14).

Figure 6.3 shows a correlation matrix representing the median Spearman correlation coefficients obtained over the available patient datasets for all pairwise combinations of functional data, both for voxel-based and regional analysis. Highest inter-modality median coefficients of the voxel-based analysis were obtained for the combinations FDG/FMISO ($r = 0.56$, range: 0.08 – 0.80, $N = 8$), FDG/$\overline{A}_{\mathrm{FMISO}}$ ($r = 0.55$, range: 0.19 – 0.76, $N = 8$), $\overline{A}_{\mathrm{FMISO}}/\overline{\Delta S}_{\mathrm{DCE}}$ ($r = 0.46$, range: 0.30 – 0.57, $N = 5$), and ADC/FDG ($r = -0.39$, range: -0.82 – 0.30, $N = 13$). For regional analysis, values changed to FDG/FMISO ($r = 0.51$, range: 0.06 – 0.86, $N = 7$), FDG/$\overline{A}_{\mathrm{FMISO}}$ ($r = 0.32$, range: -0.02 –

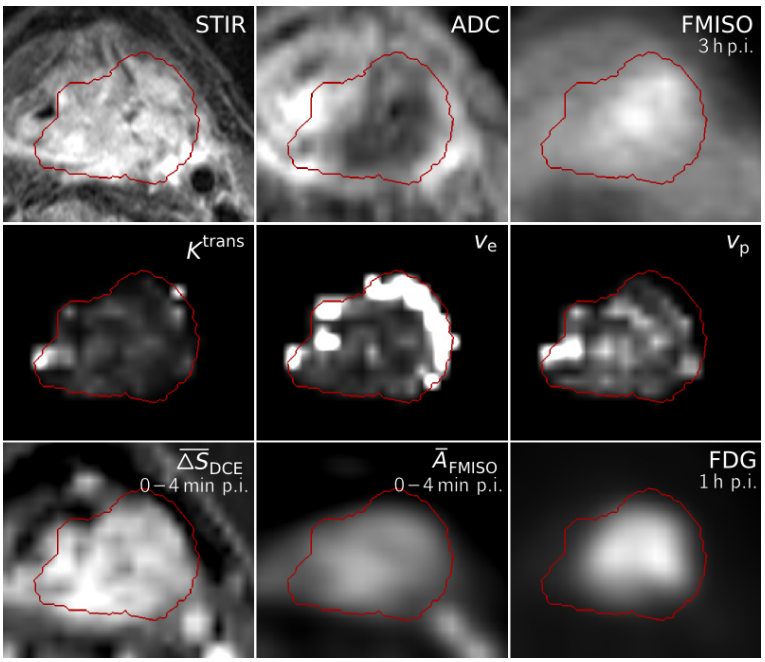

Figure 6.1: Dataset of Patient 11, showing transversal slices of the anatomical T2-weighted image (STIR), ADC map, FMISO image at 3 h p.i., the DCE-derived maps K^{trans}, v_e and v_p, the maps of mean signal enhancement $\overline{\Delta S}_{DCE}$ and mean FMISO activity $\overline{A}_{FMISO}$ in the time range of 0–4 min p.i, and the FDG image at 1 h p.i. All images and parameter maps were acquired in a single PET/MRI session, except for the FDG image and the $\overline{A}_{FMISO}$ map which were transferred to the PET/MR dataset by deformable registration. The delineation of the GTV is shown in red.

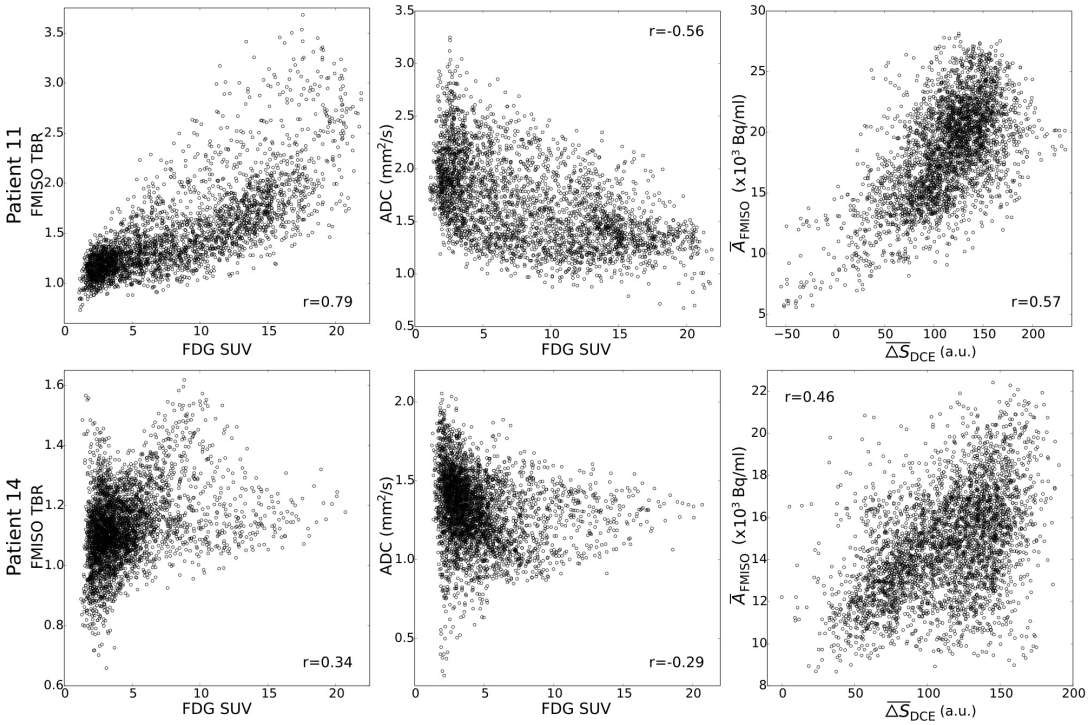

Figure 6.2: Exemplary scatter plots for Patients 11 (top) and 14 (bottom), with samples obtained on the voxel level. For increased comparability across patients, FDG activity concentrations were converted to standardized uptake values (SUVs). FMISO data was normalized by devision by the background signal in a deep neck muscle, resulting in the tumor to background ratio (TBR). The Spearman correlation coefficients r associated with the scatterplots are shown within each plot.

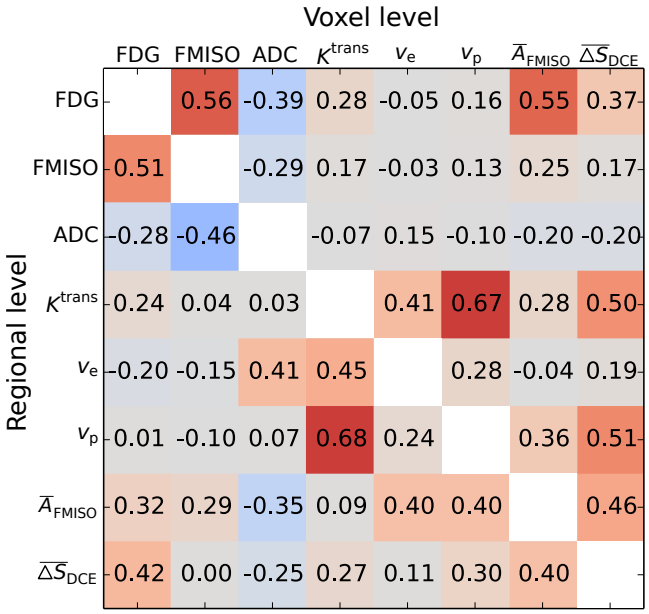

Figure 6.3: Correlation matrix showing for each pair of functional information the median Spearman correlation coefficients obtained over all respective patient datasets. The upper right triangle shows coefficients derived on the voxel level, whereas the lower left triangle shows the coefficients derived on the regional level.

0.61, $N = 7$), $\overline{A}_{FMISO}/\overline{\Delta S}_{DCE}$ ($r = 0.40$, range: -0.09 – 0.61, $N = 5$), and ADC/FDG ($r = -0.28$, range: -0.98 – 0.62, $N = 10$).

The inter-patient variation of Spearman correlation coefficients for both voxel and regional analysis are shown in Figure 6.4 for the pairs of highest median voxel correlations. Moreover, correlation coefficients are shown for each patient individually in Figure 6.5.

6.4 Discussion

While weak correlations were observed between some functional imaging modalities in the median over all patient datasets, distinct correlations were only present on an individual basis. This applies both to the voxel-based and the regional analysis. FDG and FMISO showed the largest inter-modality median correlations in our study. However, the respective correlation coefficients varied strongly within the patient cohort. This result is in line with the findings of Rajendran et al. [142] and Thorwarth et al. [143]. Similar to Houweling et al. [145],

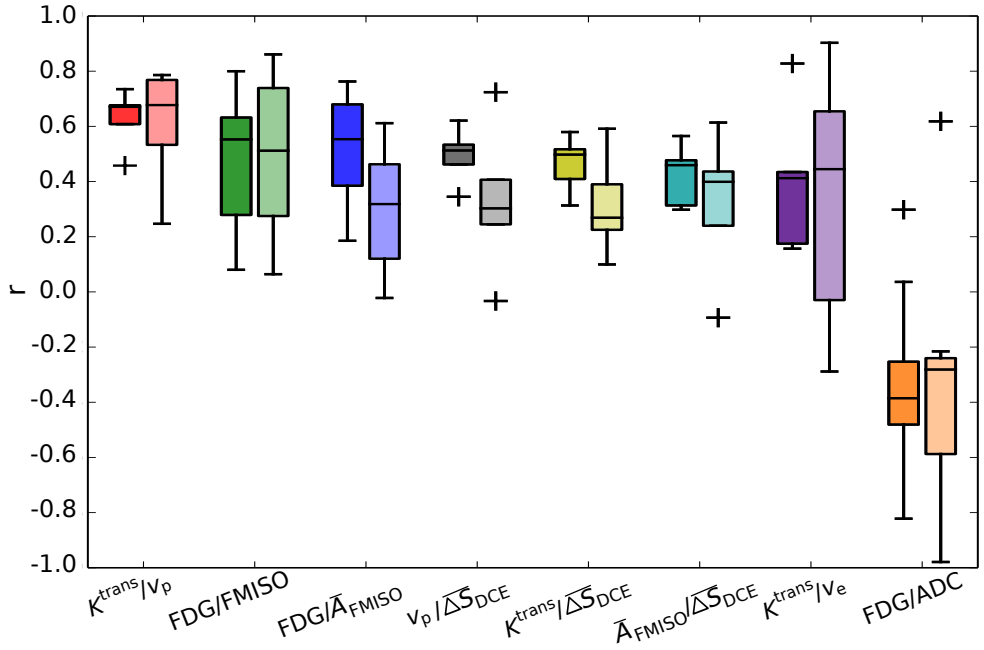

Figure 6.4: Boxplots showing the inter-patient variation of Spearman correlation coefficients for the pairs with the highest median voxel correlations according to Figure 6.3. For each functional pair, the results from voxel-based analysis (dark color, left boxes) and from regional analysis (light color, right boxes) are shown.

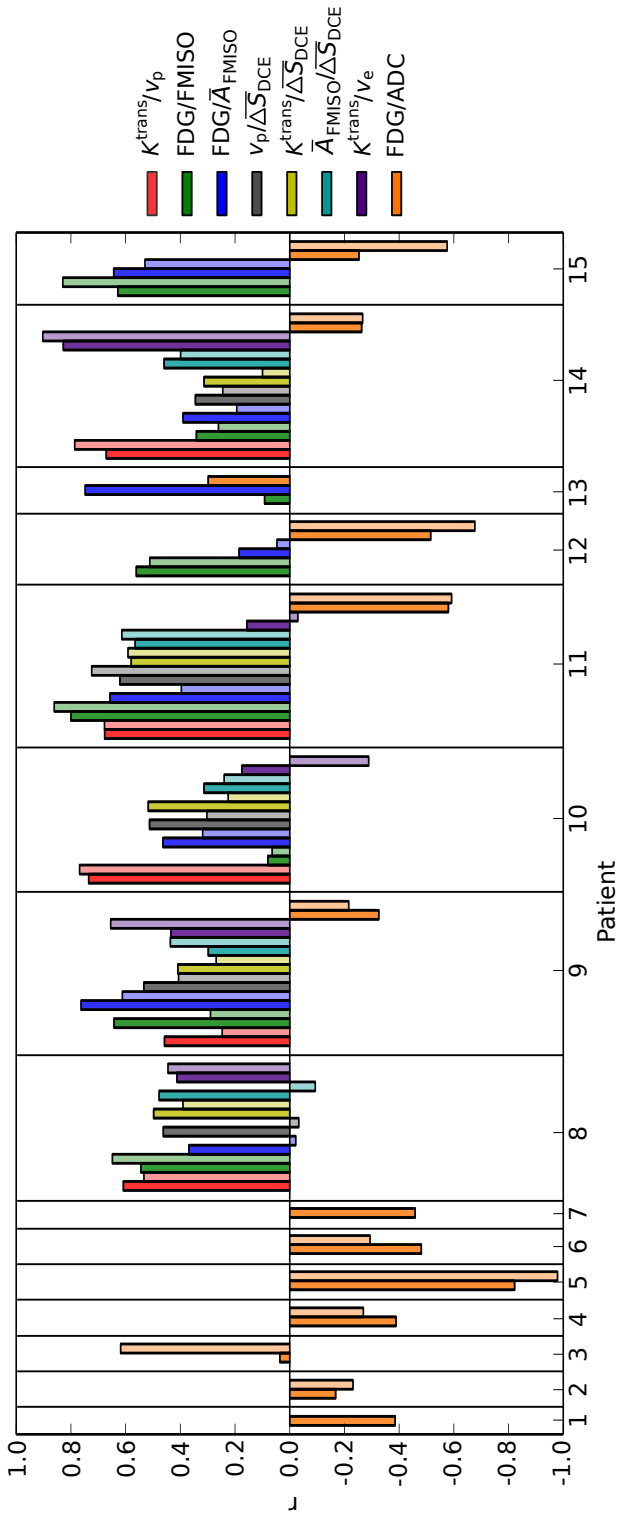

Figure 6.5: Spearman correlation coefficients obtained for the individual patients. Color encoding is the same as in Figure 6.4, with dark and light colors showing the results from voxel-based and regional analysis, respectively.

moderate negative correlations were observed between FDG and ADC, with a pronounced variability across patients. No correlations were observed between FMISO and DCE-derived parameters maps. This is different from Newbold et al. [146], Donaldson et al. [147], and Jansen et al. [148]. However, our study is not readily comparable to the results of these authors, since the first authors quantified hypoxia by means of pimonidazole staining after tumor resection and the latter performed the analysis for neck nodal metastases.

We also found a moderate correlation between $\overline{A}_{\mathrm{FMISO}}$ and $\overline{\Delta S}_{\mathrm{DCE}}$. This indicates that they may be measures of similar physiological parameters. However, according to this first analysis the observed correlation does not seem to be sufficient to infer the early FMISO information during wash-in from the DCE data. Instead of using $\overline{\Delta S}_{\mathrm{DCE}}$ maps directly for correlation analysis, they could have also been converted to maps of contrast agent concentration using native T1 maps derived from the VIBE acquisitions. While this might have a slight impact on correlations quantified with the Spearman coefficient due to the dependency of the relation between signal enhancement and concentration on native T1, the conversion to concentration maps would introduce an additional source of error due to uncertainties in native T1 derivation.

For ADC and v_{e} maps, correlations may be expected as ADC is commonly related to the fraction of EES, and v_{e} is interpreted as the fraction of EES itself. However, in our study weak correlations are only observed on a regional level. One explanation of missing correlations could be that DCE parameter maps in regions with low vascularization are not reliable due to the weak delivery of contrast agent. However, correlation analysis between ADC and v_{e} should be performed with further datasets to provide more representative results.

Some of the highest correlations were found between the DCE-based maps K^{trans}, v_{e}, v_{p}, as well as $\overline{\Delta S}_{\mathrm{DCE}}$. This may be either due to inherently correlated parameter estimates in the extended Tofts model used for data analysis, or due to biological relations between the respective parameters.

The determination of multimodal parameter correlations may be substantially compromised by different factors, such as geometrical inaccuracies associated with imaging techniques and image registration, as well as interpolation errors and statistical uncertainties. Geometrical distortions are particularly present in the ADC maps, which were acquired using EPI sequences. For future acquisitions, ideally sequences which are less prone to geometrical distortions should be used in combination with a method for geometrical distortion correction [151, 152]. Also, since no patient positioning system was used during image acquisition in the combined PET/MR examinations, movement of the patients during image acquisition cannot be excluded a priori. Hardware so-

lutions for effective patient immobilization are currently being developed [141]. Finally, geometrical uncertainties are associated with images that were transferred to the PET/MR datasets by deformable registration, which may lead to a reduction of absolute correlation values [153]. An independent analysis of the errors introduced by the different factors is not possible with realistic patient data. In order to account for geometrical uncertainties, a correlation analysis on a regional level was added to the voxel analysis. Such a regional analysis is more robust with respect to geometrical uncertainties, interpolation errors and image noise, whereas averaging may underestimate existing correlations, and additional statistical uncertainties may be introduced. Both increases and decreases in correlation coefficients compared to the voxel-based analysis were observed. However, similar inter-patient distributions were observed (cf. Figure 6.4). As a main result of our study we found large variations of correlation coefficients between patients, which most probably can not be explained by the present inaccuracies alone.

DCE parameter maps were derived with the extended Tofts model. However, model parameters may be misinterpreted for some physiological conditions such as highly vascularized tissues with intermediate flow [154]. For other conditions, the model may not fit the data accurately. Other models with fewer assumptions like the four parameter two-compartment exchange model (2CXM) could be used instead if data quality is sufficient in terms of temporal resolution, signal-to-noise ratio and artifacts [92].

Only a limited number of patients was available in this study, especially with respect to DCE-MRI data. Further evaluation should be performed when more patient data is available.

The results in this study extend the correlation analyses performed in previous studies by considering a comprehensive set of functional data. The present results suggest that the different functional datasets derived from DCE-MRI, DW-MRI, FDG-PET and FMISO-PET provide complementary information. Since all these imaging methods were proven to be prognostic for treatment outcome [14–19, 41, 43, 44], this suggests that each method may be of separate value for the adaptation of treatment strategies. However, only pairwise correlations have been analyzed so far. It appears interesting to elaborate if the information from one functional imaging method could be deduced from a combination of several other functional imaging methods. Such an analysis could in the future be performed with machine learning approaches [155] when more patient datasets are available. On the other hand, one may obtain more coherent correlation results if only subgroups of HN tumors are analyzed, for example patients with equal tumor localization, size and staging.

Analyses exploring a potential redundancy between functional PET/MR data may be of value for RT and other treatment modalities due to several reasons. Firstly, using redundant image data and parameter maps for biologically adapted treatments would unnecessarily increase the number of parameters to be adapted with respect to improved outcome. Thus, the correlation analysis performed in the present study constitutes a first step towards the integration of functional imaging into treatment individualization. Before biologically adapted treatments can be used clinically, a number of additional steps are required, such as the correlation of functional parameters to treatment outcome and a thorough regional failure analysis. Further research is needed to clarify which parameter combination provides accurate information about locoregional control probability. Secondly, functional imaging data may concatenate multiple physiological parameters, and interpretation is not always straightforward. A more detailed understanding of functional images and the parameter maps obtained by post-processing models is necessary [154,156]. Exploring a potential inter-dependency between different datasets may support the interpretation of functional imaging data. Moreover, present or missing correlations between different datasets could potentially also be associated with biological evidence related to treatment response of individual patients. A more comprehensive picture of these issues would allow for a knowledge-driven treatment adaptation, which would then need to be validated in clinical trials.

6.5 Conclusion

Multiparameteric PET/MR provides a substantial amount of different functional imaging data, which may be highly beneficial for cancer treatment adaptation. The results of our study suggest that the associated datasets provide complementary information, and thus could all be of separate value for defining treatment adaption strategies, as well as for treatment response assessment and follow-up. Results of this correlation study might in the future contribute to the design of individually adapted treatment approaches based on multiparametric functional PET/MR imaging.

Compliance with ethical standards

Funding Parts of this project were supported by the European Research Council (ERC), starting Grant No. 335367.

Conflict of Interest None.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

7 Concluding discussion and outlook

Concluding discussion

Radiation treatment with IMRT as well as advanced imaging methods have become available in the last years, potentially providing a basis for major advancements in RT treatment. In this respect, combined PET/MR provides a plethora of co-registered multimodal imaging information for tumor delineation and characterization. In this thesis, several image processing algorithms and methods for integrating PET/MR data into RT have been developed and evaluated. All studies were performed on a limited number of patient datasets, and can thus be considered primarily as a methodological development, for which further evaluation and potential adjustment of parameters and methods should be pursued in future. In the following, conclusions from the performed studies are summarized and limitations as well as potentials are discussed.

The accuracy of the algorithm developed for deformably registering MR and CT datasets (*Part I*) was found to be in the order of half the voxel size of the images. It may therefore be of sufficient accuracy for the integration of PET/MR data into the RT planning process. However, for clinical application, a further evaluation on more patient datasets is necessary.

Deformable registration algorithms and their validation are in general associated with difficulties. Deformable registration is an ill-posed problem [114]. In homogeneous regions with low structural information, there is little evidence about the 'true' spatial correspondence. Tumors often appear homogeneous in anatomical images, while functional imaging methods may reveal a biological heterogeneity. Thus, if functional imaging information should be co-transformed to the reference image for local tumor characterization, the registration accuracy in the tumor region is of utmost importance. In the developed algorithm, a heuristic regularization term is introduced to ensure smooth deformation fields, especially in regions with limited anatomical information, and suppress physically implausible deformations.

In some publications also an attempt is made to include more knowledge into deformable registration approaches. In the work of Brock et al. [157, 158],

information of biomechanical behavior of tissue is included in the registration algorithm. However, biomechanical parameters are not exactly known for the various tissues and are assumed to be homogeneous and isotropic within the approach. Also, segmentations of different tissue types have to be available beforehand for the application of this concept.

Multi-institution studies may be the best way for a comprehensive comparison of different registration algorithms [159]. In future, a comprehensive dataset of MR and CT images of HNC patients, with pre-defined annotations for quantitative evaluation could be established for this purpose.

Registration of PET/MR data to the planning CT could be facilitated by imaging the patient in RT-specific position in the PET/MR scanner. However, particular hardware solutions for this purpose are not yet commercially available. One such solution is currently being developed and evaluated with respect to image quality [141,160]. Even with the possibility of applying positioning aids in PET/MR it is likely that slight deformations between PET/MR and CT remain, and deformable registration may still be favorable over rigid registration alone.

Further potential applications of deformable registration in RT are the assessment of local treatment response, as well as failure analysis in case that recurrences are discovered in follow-up imaging. In these cases changes in patient geometry have to be accounted for, which only can be realized with deformable algorithms.

The co-segmentation algorithm for the derivation of tumor volumes from the combination of FDG-PET and T2-weighted MR images (*Part II*) yielded an accuracy that lies within variations between manual contours drawn by three independent observers. The study also showed that the combination of PET and MRI information for tumor volume definition reduced inter-observer variations compared to MRI alone. This supports the assumption that multimodal imaging improves tumor delineation. Variability of tumor delineation may be further reduced by including additional imaging methods, for example further MR sequences. This hypothesis should be evaluated in future studies.

Since the algorithm has a modular design, it is easy to include additional modalities into the co-segmentation process. However, it could also be evaluated in future if, instead of a linear combination of single-modality tumor probability maps, a more complex combination of different imaging information would be more appropriate for the segmentation process. From an image processing point of view the integration of quantitative imaging data, such as quantitative parameter maps from functional imaging, or maps of MR relaxation times T1 and T2, into the segmentation algorithm is most straightforward.

The developed algorithm should in future ideally be validated on histological tissue sections, which provide a ground truth for the local classification into tumor and normal tissue. As for registration, multi-institution studies would provide a basis to compare the accuracy and robustness of different segmentation approaches.

Multiparametric functional imaging may allow for a comprehensive biological tumor characterization and the individual adaptation of treatment strategies. As a first step towards such concepts, pairwise correlation analyses of FDG-PET, FMISO-PET, DCE-MRI and DW-MRI data were motivated and performed (*Part III*). The results revealed distinct correlations for individual patients only, while high inter-patient variabilities with weak correlations in the median were obtained.

Since no general correlations between the functional datasets were observed, each method may be of separate value for treatment adaptations. However, general relations could still be present in the combination of three or more datasets. Machine learning approaches could be used to address this question in more detail. Results of such studies may in future contribute to the integration of functional PET/MR data into individualized therapy concepts. Additional evaluations, such as the analysis of the relationship between complementary functional imaging data and treatment outcome, will have to be performed in order to finally define biologically adapted treatment strategies.

Outlook

Combined PET/MR is highly promising for RT as it provides comprehensive anatomical, functional and molecular information. At least in the near future CT will remain an important imaging modality for RT due to its value for RT dose calculation and easy implementation in the RT treatment room as a cone beam CT device. However, MR-based treatment planning (cf. Section 3.3.2), and MR image guidance in the treatment room are currently being developed. For the latter, prototype MR-linac systems have already been developed and first systems are commercially available [31, 161, 162].

It is not clear yet if combined PET/MR will be established in clinical practice as it is the case for PET/CT. The advantages over the usage of PET/CT in combination with stand-alone MR scanners are intrinsic co-registration of PET/MR data, the simplified logistics, and the saving of the CT radiation if CT could be omitted in the treatment planning process. If combined PET/MR does not

prevail in future the general concepts and methods of this thesis remain valid, but would have to be applied to PET and MR data from separate devices. In this case, an accurate previous registration of PET and MR datasets would a prerequisite. PET/MR imaging is promising in all steps of the RT workflow that require imaging data (cf. Figure 3.2). Further research is necessary before the full potential of PET/MR imaging for RT can be exploited.

With the development towards multimodality imaging, automatic image processing will become more important. A large amount of different datasets is not manageable by a human observer alone. Moreover, with automatic algorithms a higher degree of standardization in treatment decision, planning and evaluation can be achieved. Additional image processing approaches for local characterization of tumors, such as texture analysis, are recently being evaluated for RT purposes [163, 164]. With the increase of imaging data, as well as of datasets derived from the original images, machine learning approaches are likely to become more important as a methodology for dimensionality reduction and the relation of datasets to clinical endpoints.

The benefit of PET/MR in RT tumor volume delineation, biologically adapted dose prescription, and response assessment has finally to be assessed in clinical studies. Individualized treatment concepts will hopefully lead to a reduction of the number of local treatment failures and a better patient outcome.

Bibliography

[1] Delaney G, Jacob S, Featherstone C, Barton M. The role of radiotherapy in cancer treatment. Cancer 2005; 104:1129–1137.

[2] Freund L. Grundriss der gesamten Radiotherapie für praktische Ärzte. Urban & Schwarzenberg Wien 1903.

[3] Lenmalm F. Förhandlingar vid Svenska Lakare-Sallskapets Sammankomster ar19899. Isaac Markus Stockholm 1900; 205–209.

[4] Coutard H. Principles of X-ray therapy of malignant diseases. Lancet 1934; 224:1–8.

[5] Bamberg M, Molls M, Sack H (Eds): *Radioonkologie: Grundlagen.* W. Zuckerschwerdt Verlag 2009.

[6] Fuks Z, Horwich A. Clinical and technical aspects of conformal therapy. Radiother Oncol 1993; 29:219.

[7] Webb S. The physical basis of IMRT and inverse planning. Br J Radiol 2003; 76:678–689.

[8] Fippel M. Fast Monte Carlo dose calculation for photon beams based on the VMC electron algorithm. Med Phys 1999; 26:1466–1475.

[9] Cozzi L, Fogliata A, Lomax A, Bolsi A. A treatment planning comparison of 3D conformal therapy, intensity modulated photon therapy and proton therapy for treatment of advanced head and neck tumours. Radiother Oncol 2001; 61:287–297.

[10] Schulz-Ertner D, Tsujii H. Particle radiation therapy using proton and heavier ion beams. J Clin Oncol 2007; 25:953–964.

[11] Budach V, Stuschke M, Budach W, Baumann M, Geismar D, Grabenbauer G, Lammert I, Jahnke K, Stueben G, Herrmann T, Bamberg M, Wust P, Hinkelbein W, Wernecke KD. Hyperfractionated accelerated chemoradiation with concurrent fluorouracil-mitomycin is more effective than dose-escalated hyperfractionated accelerated radiation therapy alone in locally

advanced head and neck cancer: Final results of the radiotherapy cooperative clinical trials group of the German Cancer Society 95-06 Prospective Randomized Trial. J Clin Oncol 2005; 23:1125–1135.

[12] Herk MV. Errors and margins in radiotherapy. Sem Radiat Oncol 2004; 14:52–64.

[13] Ling C, Humm J, Larson S, Amols H, Fuks Z, Leibel S, Koutcher JA. Towards multidimensional radiotherapy (MD-CRT): biological imaging and biological conformality. Int J Radiat Oncol Biol Phys 2000; 47:551–560.

[14] Lambrecht M, Calster BV, Vandecaveye V, Keyzer FD, Roebben I, Hermans R, Nuyts S. Integrating pretreatment diffusion weighted MRI into a multivariable prognostic model for head and neck squamous cell carcinoma. Radiother Oncol 2014; 110:429–434.

[15] Andersen EK, Hole KH, Lund KV, Sundfør K, Kristensen GB, Lyng H, Malinen E. Pharmacokinetic parameters derived from dynamic contrast enhanced MRI of cervical cancers predict chemoradiotherapy outcome. Radiother Oncol 2013; 107:117–122.

[16] Halle C, Andersen E, Lando M, Aarnes EK, Hasvold G, Holden M, Syljuåsen RG, Sundfør K, Kristensen GB, Holm R, Malinen E, Lyng H. Hypoxia-induced gene expression in chemoradioresistant cervical cancer revealed by dynamic contrast-enhanced MRI. Cancer Res 2012; 72:5285–5295.

[17] Pak K, Cheon GJ, Nam HY, Kim SJ, Kang KW, Chung JK, Kim EE, Lee DS. Prognostic value of metabolic tumor volume and total lesion glycolysis in head and neck cancer: A systematic review and meta-analysis. J Nucl Med 2014; 55:884–890.

[18] Rischin D, Hicks RJ, Fisher R, Binns D, Corry J, Porceddu S, Peters LJ. Prognostic significance of [18F]-misonidazole positron emission tomography–detected tumor hypoxia in patients with advanced head and neck cancer randomly assigned to chemoradiation with or without tirapazamine: A substudy of Trans-Tasman Radiation Oncology Group Study 98.02. J Clin Oncol 2006; 24:2098–2104.

[19] Zips D, Zöphel K, Abolmaali N, Perrin R, Abramyuk A, Haase R, Appold S, Steinbach J, Kotzerke J, Baumann M. Exploratory prospective trial of hypoxia-specific PET imaging during radiochemotherapy in patients with locally advanced head-and-neck cancer. Radiother Oncol 2012; 105:21–28.

[20] Astner ST, Dobrei-Ciuchendea M, Essler M, Bundschuh RA, Sai H, Schwaiger M, Molls M, Weber WA, Grosu AL. Effect of 11C-methionine-positron emission tomography on gross tumor volume delineation in stereotactic radiotherapy of skull base meningiomas. Int J Radiat Oncol Biol Phys 2008; 72:1161–1167.

[21] Nestle U, Kremp S, Schaefer-Schuler A, Sebastian-Welsch C, Hellwig D, Rübe C, Kirsch CM. Comparison of different methods for delineation of 18F-FDG PET–positive tissue for target volume definition in radiotherapy of patients with non–small cell lung cancer. J Nucl Med 2005; 46:1342–1348.

[22] Lips IM, van der Heide UA, Haustermans K, van Lin EN, Pos F, Franken SP, Kotte AN, van Gils CH, van Vulpen M. Single blind randomized phase III trial to investigate the benefit of a focal lesion ablative microboost in prostate cancer (FLAME-trial): study protocol for a randomized controlled trial. Trials 2011; 12:255.

[23] van Elmpt W, Ruysscher DD, van der Salm A, Lakeman A, van der Stoep J, Emans D, Damen E, Öllers M, Sonke JJ, Belderbos J. The PET-boost randomised phase II dose-escalation trial in non-small cell lung cancer. Radiother Oncol 2012; 104:67–71.

[24] Welz S, Pfannenberg C, Reimold M, Reischl G, Mauz PS, Zips D, Alber M, Belka C, Thorwarth D. Hypoxia dose-escalation with chemoradiation in head and neck cancer: planned interim analysis of a randomized study. Radiother Oncol 2014; 111:S1.

[25] Thorwarth D, Leibfarth S, Mönnich D. Potential role of PET/MRI in radiotherapy treatment planning. Clin Transl Imaging 2013; 1:45–51.

[26] Reiser MF, Semmler W (Eds): *Magnetresonanztomographie*. Springer 2012.

[27] Brown RW, Cheng YCN, Haacke EM, Thompson MR, Venkatesan R (Eds): *Magnetic resonance imaging*. Wiley-Blackwell 2014.

[28] Verduijn GM, Bartels LW, Raaijmakers CP, Terhaard CH, Pameijer FA, van den Berg CA. Magnetic resonance imaging protocol optimization for delineation of gross tumor volume in hypopharyngeal and laryngeal tumors. Int J Radiat Oncol Biol Phys 2009; 74:630–636.

[29] Padhani AR, Liu G, Mu-Koh D, Chenevert TL, Thoeny HC, Takahara T, Dzik-Jurasz A, Ross BD, Cauteren MV, Collins D, Hammoud DA, Rustin

GJ, Taouli B, Choyke PL. Diffusion-weighted magnetic resonance imaging as a cancer biomarker: Consensus and recommendations. Neoplasia 2009; 11:102–125.

[30] Vandecaveye V, Keyzer FD, Poorten VV, Dirix P, Verbeken E, Nuyts S, Hermans R. Head and neck squamous cell carcinoma: Value of diffusion-weighted MR imaging for nodal staging. Radiology 2009; 251:134–146.

[31] Lagendijk JJ, Raaymakers BW, Raaijmakers AJ, Overweg J, Brown KJ, Kerkhof EM, van der Put RW, Hårdemark B, van Vulpen M, van der Heide UA. MRI/linac integration. Radiother Oncol 2008; 86:25–29.

[32] Bhagwandien R, van Ee R, Beersma R, Bakker C, Moerland M, Lagendijk J. Numerical analysis of the magnetic field for arbitrary magnetic susceptibility distributions in 2D. Magn Reson Imaging 1992; 10:299–313.

[33] Schakel T, Hoogduin JM, Terhaard CH, Philippens ME. Diffusion weighted MRI in head-and-neck cancer: Geometrical accuracy. Radiother Oncol 2013; 109:394–397.

[34] Balafar M, Ramli A, Saripan M, Mashohor S. Review of brain MRI image segmentation methods. Artif Intell Rev 2010; 33:261–274.

[35] Vovk U, Pernus F, Likar B. A review of methods for correction of intensity inhomogeneity in MRI. IEEE Trans Med Imaging 2007; 26:405–421.

[36] Hanvey S, Sadozye AH, McJury M, Glegg M, Foster J. The influence of MRI scan position on image registration accuracy, target delineation and calculated dose in prostatic radiotherapy. Br J Radiol 2012; 85:e1256–e1262.

[37] Seagroves TN, Ryan HE, Lu H, Wouters BG, Knapp M, Thibault P, Laderoute K, Johnson RS. Transcription factor HIF-1 is a necessary mediator of the Pasteur effect in mammalian cells. Mol Cell Biol 2001; 21:3436–3444.

[38] Warburg O. On the origin of cancer cells. Science 1956; 123:309–314.

[39] Vander Heiden MG, Cantley LC, Thompson CB. Understanding the Warburg effect: The metabolic requirements of cell proliferation. Science 2009; 324:1029–1033.

[40] Nestle U, Weber W, Hentschel M, Grosu AL. Biological imaging in radiation therapy: role of positron emission tomography. Phys Med Biol 2009; 54:R1.

[41] Koyasu S, Nakamoto Y, Kikuchi M, Suzuki K, Hayashida K, Itoh K, Togashi K. Prognostic value of pretreatment 18F-FDG PET/CT parameters including visual evaluation in patients with head and neck squamous cell carcinoma. Am J Roentgenol 2014; 202:851–858.

[42] Horsman MR, Mortensen LS, Petersen JB, Busk M, Overgaard J. Imaging hypoxia to improve radiotherapy outcome. Nat Rev Clin Oncol 2012; 9:674–687.

[43] Mortensen LS, Johansen J, Kallehauge J, Primdahl H, Busk M, Lassen P, Alsner J, Sørensen BS, Toustrup K, Jakobsen S, Petersen J, Petersen H, Theil J, Nordsmark M, Overgaard J. FAZA PET/CT hypoxia imaging in patients with squamous cell carcinoma of the head and neck treated with radiotherapy: Results from the DAHANCA 24 trial. Radiother Oncol 2012; 105:14–20.

[44] Thorwarth D, Eschmann SM, Scheiderbauer J, Paulsen F, Alber M. Kinetic analysis of dynamic 18F-fluoromisonidazole PET correlates with radiation treatment outcome in head-and-neck cancer. BMC Cancer 2005; 5:152.

[45] Sattler B, Lee JA, Lonsdale M, Coche E. PET/CT (and CT) instrumentation, image reconstruction and data transfer for radiotherapy planning. Radiother Oncol 2010; 96:288–297.

[46] Shreve PD, Anzai Y, Wahl RL. Pitfalls in oncologic diagnosis with FDG PET imaging: Physiologic and benign variants. RadioGraphics 1999; 19:61–77.

[47] Kolb A, Wehrl H, Hofmann M, Judenhofer M, Eriksson L, Ladebeck R, Lichy M, Byars L, Michel C, Schlemmer HP, Schmand M, Claussen C, Sossi V, Pichler B. Technical performance evaluation of a human brain PET/MRI system. Eur Radiol 2012; 22:1776–1788.

[48] Delso G, Fürst S, Jakoby B, Ladebeck R, Ganter C, Nekolla SG, Schwaiger M, Ziegler SI. Performance measurements of the siemens mMR integrated whole-body PET/MR scanner. J Nucl Med 2011; 52:1914–1922.

[49] Zaidi H, Guerra AD. An outlook on future design of hybrid PET/MRI systems. Med Phys 2011; 38:5667–5689.

[50] Yankeelov TE, Peterson TE, Abramson RG, Garcia-Izquierdo D, Arlinghaus LR, Li X, Atuegwu NC, Catana C, Manning HC, Fayad ZA, et al. Simultaneous PET–MRI in oncology: a solution looking for a problem? Magn Reson Imaging 2012; 30:1342–1356.

[51] Padhani A, Khan A. Diffusion-weighted (DW) and dynamic contrast-enhanced (DCE) magnetic resonance imaging (MRI) for monitoring anticancer therapy. Targeted Oncology 2010; 5:39–52.

[52] Dirix P, Vandecaveye V, De Keyzer F, Stroobants S, Hermans R, Nuyts S. Dose painting in radiotherapy for head and neck squamous cell carcinoma: Value of repeated functional imaging with 18F-FDG PET, 18F-Fluoromisonidazole PET, diffusion-weighted MRI, and dynamic contrast-enhanced MRI. J Nucl Med 2009; 50:1020–1027.

[53] Subesinghe M, Scarsbrook AF, Sourbron S, Wilson DJ, McDermott G, Speight R, Roberts N, Carey B, Forrester R, Gopal SV, et al. Alterations in anatomic and functional imaging parameters with repeated FDG PET-CT and MRI during radiotherapy for head and neck cancer: a pilot study. BMC cancer 2015; 15:137.

[54] Castadot P, Lee JA, Geets X, Grégoire V. Adaptive radiotherapy of head and neck cancer. Sem Radiat Oncol 2010; 20:84–93.

[55] Schwartz DL, Garden AS, Thomas J, Chen Y, Zhang Y, Lewin J, Chambers MS, Dong L. Adaptive radiotherapy for head-and-neck cancer: Initial clinical outcomes from a prospective trial. Int J Radiat Oncol Biol Phys 2012; 83:986–993.

[56] Mac Manus MP, Hicks RJ, Matthews JP, McKenzie A, Rischin D, Salminen EK, Ball DL. Positron emission tomography is superior to computed tomography scanning for response-assessment after radical radiotherapy or chemoradiotherapy in patients with non–small-cell lung cancer. J Clin Oncol 2003; 21:1285–1292.

[57] Pichler B, Lorenz E, Mirzoyan R, Pimpl W, Roder F, Schwaiger M, Ziegler S: Performance test of a LSO-APD PET module in a 9.4 Tesla magnet. In *IEEE Nuclear Science Symposium* 1997; 2:1237–1239.

[58] Boss A, Stegger L, Bisdas S, Kolb A, Schwenzer N, Pfister M, Claussen C, Pichler B, Pfannenberg C. Feasibility of simultaneous PET/MR imaging in the head and upper neck area. Eur Radiol 2011; 21:1439–1446.

[59] Hofmann M, Pichler B, Schölkopf B, Beyer T. Towards quantitative PET/MRI: a review of MR-based attenuation correction techniques. Eur J Nucl Med Mol Imaging 2009; 36:93–104.

[60] Carney JPJ, Townsend DW, Rappoport V, Bendriem B. Method for transforming CT images for attenuation correction in PET/CT imaging. Med Phys 2006; 33:976–983.

[61] Martinez-Möller A, Souvatzoglou M, Delso G, Bundschuh RA, Chefd'hotel C, Ziegler SI, Navab N, Schwaiger M, Nekolla SG. Tissue classification as a potential approach for attenuation correction in whole-body PET/MRI: evaluation with PET/CT data. J Nucl Med 2009; 50:520–526.

[62] Hofmann M, Bezrukov I, Mantlik F, Aschoff P, Steinke F, Beyer T, Pichler BJ, Schölkopf B. MRI-based attenuation correction for whole-body PET/MRI: Quantitative evaluation of segmentation- and atlas-based methods. J Nucl Med 2011; 52:1392–1399.

[63] Robson MD, Gatehouse PD, Bydder M, Bydder GM. Magnetic resonance: An introduction to ultrashort TE (UTE) imaging. J Comput Assist Tomo 2003; 27:825–846.

[64] Delso G, Carl M, Wiesinger F, Sacolick L, Porto M, Hüllner M, Boss A, Veit-Haibach P. Anatomic evaluation of 3-dimensional ultrashort-echo-time bone maps for PET/MR attenuation correction. J Nucl Med 2014; 55:780–785.

[65] Navalpakkam BK, Braun H, Kuwert T, Quick HH. Magnetic resonance–based attenuation correction for PET/MR hybrid imaging using continuous valued attenuation maps. Invest Radiol 2013; 48:323–332.

[66] Su KH, Hu L, Stehning C, Helle M, Qian P, Thompson CL, Pereira GC, Jordan DW, Herrmann KA, Traughber M, Muzic RF, Traughber BJ. Generation of brain pseudo-CTs using an undersampled, single-acquisition UTE-mDixon pulse sequence and unsupervised clustering. Med Phys 2015; 42:4974–4986.

[67] Johansson A, Karlsson M, Nyholm T. CT substitute derived from MRI sequences with ultrashort echo time. Med Phys 2011; 38:2708–2714.

[68] Kapanen M, Tenhunen M. T1/T2*-weighted MRI provides clinically relevant pseudo-CT density data for the pelvic bones in MRI-only based radiotherapy treatment planning. Acta Oncol 2013; 52:612–618.

[69] Uh J, Merchant TE, Li Y, Li X, Hua C. MRI-based treatment planning with pseudo CT generated through atlas registration. Med Phys 2014; 41:051711.

[70] Hofmann M, Steinke F, Scheel V, Charpiat G, Farquhar J, Aschoff P, Brady M, Schölkopf B, Pichler BJ. MRI-based attenuation correction for PET/MRI: a novel approach combining pattern recognition and atlas registration. J Nucl Med 2008; 49:1875–1883.

[71] Paulus DH, Braun H, Aklan B, Quick HH. Simultaneous PET/MR imaging: MR-based attenuation correction of local radiofrequency surface coils. Med Phys 2012; 39:4306–4315.

[72] Meijering E, Niessen W, Pluim J, Viergever M: Quantitative comparison of sinc-approximating kernels for medical image interpolation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI'99*; Lecture Notes in Computer Science. Edited by Taylor C, Colchester A; Springer Berlin Heidelberg 1999; 1679:210–217.

[73] Holden M. A review of geometric transformations for nonrigid body registration. IEEE Trans Med Imaging 2008; 27:111–128.

[74] Collignon A, Maes F, Delaere D, Vandermeulen D, Suetens P, Marchal G: Automated multi-modality image registration based on information theory. In *Information Processing in Medical Imaging*; Dordrecht: Kluwer Academic Publishers 1995.

[75] Viola P, Wells WM: Alignment by maximization of mutual information. In *International Conference on Computer Vision*; Los Alamitos, CA: IEEE Computer Society Press 1995.

[76] Pluim J, Maintz J, Viergever M. Mutual-information-based registration of medical images: a survey. IEEE Trans Med Imaging 2003; 22:986–1004.

[77] Rueckert D, Sonoda L, Hayes C, Hill D, Leach M, Hawkes D. Nonrigid registration using free-form deformations: application to breast MR images. IEEE Trans Med Imaging 1999; 18:712–721.

[78] Klein S, Staring M, Pluim J. Evaluation of optimization methods for nonrigid medical image registration using mutual information and B-splines. IEEE Trans Image Process 2007; 16:2879–2890.

[79] Shepherd T, Teras M, Beichel RR, Boellaard R, Bruynooghe M, Dicken V, Gooding MJ, Julyan PJ, Lee JA, Lefevre S, Mix M, Naranjo V, Wu X, Zaidi H, Zeng Z, Minn H. Comparative study with new accuracy metrics for target volume contouring in PET image guided radiation therapy. IEEE Trans Med Imaging 2012; 31:2006–2024.

[80] Zaidi H, El Naqa I. PET-guided delineation of radiation therapy treatment volumes: a survey of image segmentation techniques. Eur J Nucl Med Mol Imaging 2010; 37:2165–2187.

[81] Gordillo N, Montseny E, Sobrevilla P. State of the art survey on MRI brain tumor segmentation. Magn Reson Imaging 2013; 31:1426–1438.

[82] Bauer S, Wiest R, Nolte LP, Reyes M. A survey of MRI-based medical image analysis for brain tumor studies. Phys Med Biol 2013; 58:R97.

[83] Johnson HJ, McCormick M, Ibáñez L, Consortium TIS: The ITK Software Guide. Third Edition; Kitware, Inc. 2013; http://www.itk.org/ItkSoftwareGuide.pdf.

[84] Sethian JA: *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science*. Cambridge University Press 1999.

[85] Osher S, Sethian JA. Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. J Comput Phys 1988; 79:12–49.

[86] Overgaard J, Eriksen JG, Nordsmark M, Alsner J, Horsman MR. Plasma osteopontin, hypoxia, and response to the hypoxia sensitiser nimorazole in radiotherapy of head and neck cancer: results from the DAHANCA 5 randomised double-blind placebo-controlled trial. Lancet Oncol 2005; 6:757–764.

[87] Carmeliet P, Jain RK. Principles and mechanisms of vessel normalization for cancer and other angiogenic diseases. Nat Rev Drug Discov 2011; 10:417–427.

[88] Keyes JJ. SUV: Standard uptake or silly useless value? J Nucl Med 1995; 36.

[89] Thorwarth D, Eschmann SM, Paulsen F, Alber M. A kinetic model for dynamic [18 F]-Fmiso PET data to analyse tumour hypoxia. Phys Med Biol 2005; 50:2209.

[90] Tofts, P. Modeling tracer kinetics in dynamic Gd-DTPA MR imaging. J Magn Reson Im 1997; 7:91–101.

[91] Brix G, Semmler W, Port R, Schad LR, Layer G, Lorenz WJ. Pharmacokinetic parameters in CNS Gd-DTPA enhanced MR imaging. J Comput Assist Tomo 1991; 15:621–628.

[92] Donaldson SB, West CML, Davidson SE, Carrington BM, Hutchison G, Jones AP, Sourbron SP, Buckley DL. A comparison of tracer kinetic models for T1-weighted dynamic contrast-enhanced MRI: Application in carcinoma of the cervix. Magn Reson Med 2010; 63:691–700.

[93] Koh DM, Collins DJ, Orton MR. Intravoxel incoherent motion in body diffusion-weighted MRI: Reality and challenges. Am J Roentgenol 2011; 196:1351–1361.

[94] James G, Witten D, Hastie T, Tibshirani R: *An introduction to statistical learning: with applications in R*. Springer Texts in Statistics; Springer New York 2014.

[95] Hastie T, Tibshirani R, Friedman J: *The elements of statistical learning*. Springer Series in Statistics; New York, NY, USA: Springer New York Inc. 2009.

[96] Pichler BJ, Kolb A, Nägele T, Schlemmer HP. PET/MRI: paving the way for the next generation of clinical multimodality imaging applications. J Nucl Med 2010; 51:333–336.

[97] Groenendaal G, Borren A, Moman MR, Monninkhof E, van Diest PJ, Philippens ME, van Vulpen M, van der Heide UA. Pathologic validation of a model based on diffusion-weighted imaging and dynamic contrast-enhanced magnetic resonance imaging for tumor delineation in the prostate peripheral zone. Int J Radiat Oncol Biol Phys 2012; 82:e537–e544.

[98] van der Heide UA, Houweling AC, Groenendaal G, Beets-Tan RG, Lambin P. Functional MRI for radiotherapy dose painting. Magn Reson Imaging 2012; 30:1216–1223.

[99] Thorwarth D, Geets X, Paiusco M. Physical radiotherapy treatment planning based on functional PET/CT data. Radiother Oncol 2010; 96:317–324.

[100] Andersen EKF, Kristensen GB, Lyng H, Malinen E. Pharmacokinetic analysis and k-means clustering of DCEMR images for radiotherapy outcome prediction of advanced cervical cancers. Acta Oncol 2011; 50:859–865.

[101] Osorio EMV, Hoogeman MS, Romero AM, Wielopolski P, Zolnay A, Heijmen BJM. Accurate CT/MR vessel-guided nonrigid registration of largely deformed livers. Med Phys 2012; 39:2463–2477.

[102] Dmitriev ID, Loo CE, Vogel WV, Pengel KE, Gilhuijs KGA. Fully automated deformable registration of breast DCE-MRI and PET/CT. Phys Med Biol 2013; 58:1221–1233.

[103] Lian J, Xing L, Hunjan S, Dumoulin C, Levin J, Lo A, Watkins R, Rohling K, Giaquinto R, Kim D, Spielman D, Daniel B. Mapping of the prostate in endorectal coil-based MRI/MRSI and CT: a deformable registration and validation study. Med Phys 2004; 31:3087–3094.

[104] Castadot P, Lee JA, Parraga A, Geets X, Macq B, Grégoire V. Comparison of 12 deformable registration strategies in adaptive radiation therapy for the treatment of head and neck tumors. Radiother Oncol 2008; 89:1–12.

[105] Söhn M, Birkner M, Chi Y, Wang J, Yan D, Berger B, Alber M. Model-independent, multimodality deformable image registration by local matching of anatomical features and minimization of elastic energy. Med Phys 2008; 35:866–878.

[106] du Bois d'Aische A, De Craene M, Geets X, Grégoire V, Macq B, Warfield SK. Estimation of the deformations induced by articulated bodies: Registration of the spinal column. Biomed Signal Process Control 2007; 2:16–24.

[107] Mattes D, Haynor D, Vesselle H, Lewellen T, Eubank W. PET-CT image registration in the chest using free-form deformations. IEEE Trans Med Imaging 2003; 22:120–128.

[108] Klein S, van der Heide UA, Lips IM, van Vulpen M, Staring M, Pluim JPW. Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information. Med Phys 2008; 35:1407–1417.

[109] Loeckx D, Slagmolen P, Maes F, Vandermeulen D, Suetens P. Nonrigid image registration using conditional mutual information. IEEE Trans Med Imaging 2010; 29:19–29.

[110] Klein S, Staring M, Murphy K, Viergever M, Pluim J. elastix: A toolbox for intensity-based medical image registration. IEEE Trans Med Imaging 2010; 29:196–205.

[111] Klein S, Pluim JPW, Staring M, Viergever MA. Adaptive stochastic gradient descent optimisation for image registration. Int J Comput Vis 2009; 81:227–239.

[112] Christensen G, Johnson HJ. Consistent image registration. IEEE Trans Med Imaging 2001; 20:568–582.

[113] Fiorino C, Maggiulli E, Broggi S, Liberini S, Cattaneo GM, Dell'Oca I, Faggiano E, Muzio ND, Calandrino R, Rizzo G. Introducing the Jacobian-volume-histogram of deforming organs: application to parotid shrinkage evaluation. Phys Med Biol 2011; 56:3301–3312.

[114] Fischer B, Modersitzki J. Ill-posed medicine—an introduction to image registration. Inverse Probl 2008; 24:034008.

[115] Thorwarth D, Henke G, Müller AC, Reimold M, Beyer T, Boss A, Kolb A, Pichler B, Pfannenberg C. Simultaneous 68Ga-DOTATOC-PET/MRI for IMRT treatment planning for meningioma: First experience. Int J Radiat Oncol Biol Phys 2011; 81:277–283.

[116] Lorensen WE, Cline HE. Marching cubes: A high resolution 3D surface construction algorithm. SIGGRAPH Comput Graph 1987; 21:163–169.

[117] Eisbruch A. Intensity-modulated radiation therapy in the treatment of head and neck cancer. Nat Clin Prac Oncol 2005; 2:34–39.

[118] Madani I, Duthoy W, Derie C, Gersem WD, Boterberg T, Saerens M, Jacobs F, Grégoire V, Lonneux M, Vakaet L, Vanderstraeten B, Bauters W, Bonte K, Thierens H, Neve WD. Positron emission tomography-guided, focal-dose escalation using intensity-modulated radiotherapy for head and neck cancer. Int J Radiat Oncol Biol Phys 2007; 68:126–135.

[119] Thorwarth D, Eschmann SM, Paulsen F, Alber M. Hypoxia dose painting by numbers: a planning study. Int J Radiat Oncol Biol Phys 2007; 68:291–300.

[120] Rasch C, Keus R, Pameijer FA, Koops W, de Ru V, Muller S, Touw A, Bartelink H, van Herk M, Lebesque JV. The potential impact of CT-MRI matching on tumor volume delineation in advanced head and neck cancer. Int J Radiat Oncol Biol Phys 1997; 39:841–848.

[121] Rasch C, Steenbakkers R, Fitton I, Duppen JC, Nowak P, Pameijer FA, Eisbruch A, Kaanders J, Paulsen F, van Herk M. Decreased 3D observer variation with matched CT-MRI, for target delineation in nasopharynx cancer. Radiat Oncol 2010; 5:21.

[122] Emami B, Sethi A, Petruzzelli GJ. Influence of MRI on target volume delineation and IMRT planning in nasopharyngeal carcinoma. Int J Radiat Oncol Biol Phys 2003; 57:481–488.

[123] Ahmed M, Schmidt M, Sohaib A, Kong C, Burke K, Richardson C, Usher M, Brennan S, Riddell A, Davies M, Newbold K, Harrington KJ, Nutting CM. The value of magnetic resonance imaging in target volume delineation of base of tongue tumours – a study using flexible surface coils. Radiother Oncol 2010; 94:161–167.

[124] Daisne JF, Duprez T, Weynand B, Lonneux M, Hamoir M, Reychler H, Grégoire V. Tumor volume in pharyngolaryngeal squamous cell carcinoma: Comparison at CT, MR imaging, and FDG PET and validation with surgical specimen. Radiology 2004; 233:93–100.

[125] Hermans R, Feron M, Bellon E, Dupont P, den Bogaert WV, Baert AL. Laryngeal tumor volume measurements determined with CT: A study on intra- and interobserver variability. Int J Radiat Oncol Biol Phys 1998; 40:553–557.

[126] Huang W, Chan K, Zhou J. Region-based nasopharyngeal carcinoma lesion segmentation from MRI using clustering- and classification-based methods with learning. J Digit Imaging 2013; 26:472–482.

[127] Yu H, Caldwell C, Mah K, Mozeg D. Coregistered FDG PET/CT-based textural characterization of head and neck cancer for radiation treatment planning. IEEE Trans Med Imaging 2009; 28:374–383.

[128] Song Q, Bai J, Han D, Bhatia S, Sun W, Rockey W, Bayouth J, Buatti J, Wu X. Optimal co-segmentation of tumor in PET-CT images with context information. IEEE Trans Med Imaging 2013; 32:1685–1697.

[129] El Naqa I, Yang D, Apte A, Khullar D, Mutic S, Zheng J, Bradley JD, Grigsby P, Deasy JO. Concurrent multimodality image segmentation by active contours for radiotherapy treatment planning. Med Phys 2007; 34:4738–4749.

[130] Bagci U, Udupa JK, Mendhiratta N, Foster B, Xu Z, Yao J, Chen X, Mollura DJ. Joint segmentation of anatomical and functional images: Applications in quantification of lesions from PET, PET-CT, MRI-PET, and MRI-PET-CT images. Med Image Anal 2013; 17:929–945.

[131] Nyúl LG, Udupa JK. On standardizing the MR image intensity scale. Magn Reson Med 1999; 42:1072–1081.

[132] Weber WA. Quantitative analysis of PET studies. Radiother Oncol 2010; 96:308–310.

[133] Taheri S, Ong S, Chong V. Level-set segmentation of brain tumors using a threshold-based speed function. Image Vision Comput 2010; 28:26–37.

[134] Sailer SL, Rosenman JG, Soltys M, Cullip TJ, Chen J. Improving treatment planning accuracy through multimodality imaging. Int J Radiat Oncol Biol Phys 1996; 35:117–124.

[135] Dominietto M, Rudin M. Could magnetic resonance provide in vivo histology? Front Genet 2014; 4:298.

[136] Viallon M, Vargas M, Jlassi H, Lövblad K, Delavelle J. High-resolution and functional magnetic resonance imaging of the brachial plexus using an isotropic 3D T2 STIR (short term inversion recovery) SPACE sequence and diffusion tensor imaging. Eur Radiol 2008; 18:1018–1023.

[137] Knudtsen IS, van Elmpt W, Öllers M, Malinen E. Impact of PET reconstruction algorithm and threshold on dose painting of non-small cell lung cancer. Radiother Oncol 2014; 113:210–214.

[138] Bezrukov I, Mantlik F, Schmidt H, Schölkopf B, Pichler BJ. MR-based PET attenuation correction for PET/MR imaging. Semin Nucl Med 2013; 43:45–59.

[139] Sethian JA. Theory, algorithms, and applications of level set methods for propagating interfaces. Acta Numer 1996; 5:309–395.

[140] Leibfarth S, Mönnich D, Welz S, Siegel C, Schwenzer N, Schmidt H, Zips D, Thorwarth D. A strategy for multimodal deformable image registration to integrate PET/MR into radiotherapy treatment planning. Acta Oncol 2013; 52:1353–1359.

[141] Paulus DH, Thorwarth D, Schmidt H, Quick HH. Towards integration of PET/MR hybrid imaging into radiation therapy treatment planning. Med Phys 2014; 41:072505.

[142] Rajendran JG, Mankoff DA, O'Sullivan F, Peterson LM, Schwartz DL, Conrad EU, Spence AM, Muzi M, Farwell DG, Krohn KA. Hypoxia and glucose metabolism in malignant tumors: Evaluation by [18F]fluoromisonidazole and [18f]fluorodeoxyglucose positron emission tomography imaging. Clin Cancer Res 2004; 10:2245–2252.

[143] Thorwarth D, Eschmann SM, Holzner F, Paulsen F, Alber M. Combined uptake of [18F]FDG and [18F]FMISO correlates with radiation therapy outcome in head-and-neck cancer patients. Radiother Oncol 2006; 80:151–156.

[144] Zegers CM, van Elmpt W, Reymen B, Even AJ, Troost EG, Öllers MC, Hoebers FJ, Houben RM, Eriksson J, Windhorst AD, et al. In vivo quantification of hypoxic and metabolic status of NSCLC tumors using [18F] HX4 and [18F] FDG-PET/CT imaging. Clin Cancer Res 2014.

[145] Houweling AC, Wolf AL, Vogel WV, Hamming-Vrieze O, van Vliet-Vroegindeweij C, van de Kamer JB, van der Heide UA. FDG-PET and diffusion-weighted MRI in head-and-neck cancer patients: implications for dose painting. Radiother Oncol 2013; 106:250–254.

[146] Newbold K, Castellano I, Charles-Edwards E, Mears D, Sohaib A, Leach M, Rhys-Evans P, Clarke P, Fisher C, Harrington K, Nutting C. An exploratory study into the role of dynamic contrast-enhanced magnetic resonance imaging or perfusion computed tomography for detection of intratumoral hypoxia in head-and-neck cancer. Int J Radiat Oncol Biol Phys 2009; 74:29–37.

[147] Donaldson SB, Betts G, Bonington SC, Homer JJ, Slevin NJ, Kershaw LE, Valentine H, West CM, Buckley DL. Perfusion estimated with rapid dynamic contrast-enhanced magnetic resonance imaging correlates inversely with vascular endothelial growth factor expression and pimonidazole staining in head-and-neck cancer: A pilot study. Int J Radiat Oncol Biol Phys 2011; 81:1176–1183.

[148] Jansen JF, Schöder H, Lee NY, Wang Y, Pfister DG, Fury MG, Stambuk HE, Humm JL, Koutcher JA, Shukla-Dave A. Noninvasive assessment of tumor microenvironment using dynamic contrast-enhanced magnetic resonance imaging and 18F-fluoromisonidazole positron emission tomography imaging in neck nodal metastases. Int J Radiat Oncol Biol Phys 2010; 77:1403–1410.

[149] Mönnich D, Troost EGC, Kaanders JHAM, Oyen WJG, Alber M, Zips D, Thorwarth D. Correlation between tumor oxygenation and 18F-fluoromisonidazole PET data simulated based on microvessel images. Acta Oncol 2013; 52:1308–1313.

[150] Leibfarth S, Eckert F, Welz S, Siegel C, Schmidt H, Schwenzer N, Zips D, Thorwarth D. Automatic delineation of tumor volumes by co-segmentation of combined PET/MR data. Phys Med Biol 2015; 60:5399.

[151] Koyasu S, Iima M, Umeoka S, Morisawa N, Porter DA, Ito J, Le Bihan D, Togashi K. The clinical utility of reduced-distortion readout-segmented echo-planar imaging in the head and neck region: initial experience. Eur Radiol 2014; 24:3088–3096.

[152] Windischberger C, Robinson S, Rauscher A, Barth M, Moser E. Robust field map generation using a triple-echo acquisition. J Magn Reson Im 2004; 20:730–734.

[153] Nyflot MJ, Harari PM, Yip S, Perlman SB, Jeraj R. Correlation of PET images of metabolism, proliferation and hypoxia to characterize tumor phenotype in patients with cancer of the oropharynx. Radiother Oncol 2012; 105:36–40.

[154] Sourbron SP, Buckley DL. On the scope and interpretation of the Tofts models for DCE-MRI. Magn Reson Med 2011; 66:735–745.

[155] Ozer S, Langer DL, Liu X, Haider MA, van der Kwast TH, Evans AJ, Yang Y, Wernick MN, Yetik IS. Supervised and unsupervised methods for prostate cancer segmentation with multispectral MRI. Med Phys 2010; 37:1873–1883.

[156] Zwick S, Brix G, Tofts P, Strecker R, Kopp-Schneider A, Laue H, Semmler W, Kiessling F. Simulation-based comparison of two approaches frequently used for dynamic contrast-enhanced MRI. Eur Radiol 2010; 20:432–442.

[157] Brock KK, Sharpe MB, Dawson LA, Kim SM, Jaffray DA. Accuracy of finite element model-based multi-organ deformable image registration. Med Phys 2005; 32:1647–1659.

[158] Al-Mayah A, Moseley J, Hunter S, Velec M, Chau L, Breen S, Brock K. Biomechanical-based image registration for head and neck radiation treatment. Phys Med Biol 2010; 55:6491.

[159] Brock KK. Results of a multi-institution deformable registration accuracy study (MIDRAS). Int J Radiat Oncol Biol Phys 2010; 76:583–596.

[160] Winter R: Physical attenuation and imaging properties of a radiotherapy compatible patient positioning solution for combined PET/MR. Master's thesis; Eberhard Karls Universität Tübingen 2014.

[161] Raaijmakers AJE, Raaymakers BW, Lagendijk JJW. Magnetic-field-induced dose effects in MR-guided radiotherapy systems: dependence on the magnetic field strength. Phys Med Biol 2008; 53:909.

[162] Mutic S, Dempsey JF. The viewray system: Magnetic resonance–guided and controlled radiotherapy. Sem Radiat Oncol 2014; 24:196–199.

[163] Mahmoud-Ghoneim D, Toussaint G, Constans JM, de Certaines JD. Three dimensional texture analysis in MRI: a preliminary evaluation in gliomas. Magn Reson Imaging 2003; 21:983–987.

[164] Chicklore S, Goh V, Siddique M, Roy A, Marsden P, Cook G. Quantifying tumour heterogeneity in 18F-FDG PET/CT imaging by texture analysis. Eur J Nucl Med Mol Imaging 2013; 40:133–140.

Acknowledgements

First and foremost I wish to thank my supervisor Daniela Thorwarth for her valuable scientific and personal support during the whole process of this thesis. Without her scientific insights, kindness and continuous encouragement his thesis would not have been possible.

I also want to thank the other present and former members of the Section for Biomedical Physics and the Department of Radiation Oncology for the pleasant and cheerful atmosphere. Thanks to David, Linda, René, Alex, Urban, Vicky, Sharin, Dominik, Jens, Jenny, Anja, Valentin, Jule, Raphael, Max, Oliver, Martin, Bastian and Stefan for interesting and amusing conversations during the established after-lunch coffee breaks. Special thank to David Mönnich for sharing his distinctive kind of humor as well as his comprehensive insights into radiation therapy, and life in general, with me. Thanks to Linda Wack for uncomplicated sharing of the apartment and the nice time spent together, and to Dominik Blum and Bastian Zeeb for supporting my cake plans.

I am grateful to Holger Schmidt for taking time to answer my questions about MR sequences. Thanks to Franziska Eckert, Stefan Welz and Christine Siegel for their support with manual tumor delineations and anatomical landmarks, which were a pre-requisite for the evaluations performed in this thesis.

I am particularly grateful to Stefan, without whose support during the last years I would have never been in the position to work on this thesis. My special thank also goes to Nadia, Alex and Vicky for their attentiveness and help in difficult times.