Novel multiparametric imaging strategies for precision radiation therapy

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

vorgelegt von
René Mario Winter
aus Hamburg

Tübingen
2019

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

Tag der mündlichen Qualifikation:	05.03.2020
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Daniela Thorwarth
2. Berichterstatter:	Prof. Dr. Dr. Fritz Schick

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. Distortion correction of diffusion-weighted magnetic resonance imaging of the head and neck in radiotherapy position

published in Acta Oncologica 2017,**56**:1659–1663.

René M. Winter	first author, conception of data analysis and execution of the study, data acqusition and analysis, writing of the manuscript
Holger Schmidt	advice and support in PET/MR data acquisition, critical discussion of results
Sara Leibfarth	advice and support in preprocessing and data analysis
Kerstin Zwirner	patient recruitment, quality assessment of anatomical regions of interest
Stefan Welz	patient recruitment
Nina F. Schwenzer	supervision of PET/MR data acquisition
Christian la Fougère	supervision of PET/MR data acquisition
Konstantin Nikolaou	supervision of PET/MR data acquisition
Sergios Gatidis	advice and support in PET/MR data acquisition, critical discussion of results
Daniel Zips	critical discussion of results, proof-reading of the manuscript
Daniela Thorwarth	study conception, supervision of the project

II. Assessment of image quality of a radiotherapy-specific hardware solution for PET/MRI in head and neck cancer patients

published in Radiotherapy and Oncology 2018,**128**:485–491.

René M. Winter	first author, conception of data analysis and execution of the study, data acquisition and analysis, writing of the manuscript
Sara Leibfarth	advice and support in preprocessing and data analysis
Holger Schmidt	advice and support in PET/MR data acquisition
Kerstin Zwirner	manual delineation of tumor regions
David Mönnich	critical discussion of results
Stefan Welz	patient recruitment
Nina F. Schwenzer	supervision of PET/MR data acquisition
Christian la Fougère	supervision of PET/MR data acquisition
Konstantin Nikolaou	supervision of PET/MR data acquisition
Sergios Gatidis	manual delineation of tumor structures, advice and support in PET/MR data acquisition
Daniel Zips	critical discussion of results, proof-reading of the manuscript
Daniela Thorwarth	study conception, supervision of the project

III. Machine learning for the prediction of dynamic FMISO-PET voxel information by multi-parametric MRI

manuscript in preparation

René M. Winter	first author, conception of data analysis, execution of the study, data acquisition, preprocessing and analysis, writing of the manuscript
Sara Leibfarth	advice and support in preprocessing and conception of data analysis
Simon Boeke	animal handling and care, support in PET/MR data acquisition, manual delineation of tumor and muscle regions, critical discussion of results
Marcel Krueger	advice and support in PET/MR data acquisition and in PET image reconstruction
Pamela Mena-Romano	support in data preprocessing
Efe Cumhur Sezgin	support in animal handling and care, and in PET/MR data acquisition
Gregory Bowden	[^{18}F]-FMISO tracer synthesis
Jonathan Cotton	[^{18}F]-FMISO tracer synthesis
Bernd Pichler	supervision of PET/MR data acquisition
Daniel Zips	critical discussion of results, proof-reading of the manuscript
Daniela Thorwarth	study conception, supervision of the project

List of abbreviations

ADC	apparent diffusion coefficient
ANN	artificial neural network
APDs	avalanche photodiodes
ACF	attenuation correction factor
AP	anteroposterior
ASSD	average symmetric surface distance
AUC	area-under-the-curve
BOLD MRI	blood oxygen level dependent MRI
CNR	contrast-to-noise ratio
CT	computed tomography
DCE-MRI	dynamic contrast-enhanced MRI
DOGC	distance between geometric centers
DSI	Dice similarity index
DW-MRI	diffusion-weighted MRI
DWI	diffusion-weighted MRI
EPI	echo-planar imaging (MR sequence)
EES	extracellular-extravascular space
FAZA	^{18}F-fluoroazomycin arabinoside
fBV	fractional blood volume
FDG	^{18}F-fluorodeoxyglucose
FMISO	^{18}F-fluoromisonidazole
FWHM	full width at half maximum
GBDT	gradient boosted decision trees
GTV	gross tumor volume
HN	head and neck
HNC	head-and-neck cancer
HNSCC	head-and-neck squamous cell carcinoma

HU	Hounsfield unit
HX4	^{18}F-fluortanidazole
intEPI	EPI with integrated shim and frequency adjustment
ITK	Insight Segmentation and Registration Toolkit
IMRT	intensity modulated RT
IVIM	intra-voxel incoherent motion
kVp	peak kilovoltage
LAC	linear attenuation coefficient
Linac	linear accelerator for radiotherapy
LOR	line of response (PET)
LSO	lutetium oxyorthosilicate
MLC	multi leaf collimator
MRI	magnetic resonance imaging
NIST	US National Institute for Standards and Technology
OAR	organ at risk
OE MRI	oxygen-enhanced MRI
OSEM	ordered subset expectation maximization (PET reconstruction)
p.i.	post injection
PA	posteroanterior
PC	principal component
PCA	principal component analysis
PET	positron emission tomography
PET/MRI	PET and MRI combined in one system
RF	radiofrequency
ROC	receiver operating characteristic
ROC-AUC	ROC area-under-the-curve
ROI	region of interest
RPED	reversed phase-encode direction
rsEPI	readout-segmented multi-shot EPI
RT	radiation therapy
RTP	radiation therapy planning
RVD	relative volume difference
SD	standard deviation
SNR	signal-to-noise ratio

ST	soft tissue
STIR	short-TI inversion recovery (MR sequence)
SUV	standardized uptake value (PET)
T_1	longitudinal relaxation time (MRI)
T_2	transverse relaxation time (MRI)
T2w	T2-weighted (MRI)
TBR	tumor-to-blood ratio (PET)
TE	echo time (MRI)
TMR	tumor-to-muscle ratio (PET)
TOF	time-of-flight (PET)
TOLD MRI	tumor oxygenation level dependent MRI
TR	repetition time (MRI)
TSE	turbo spin echo (MR sequence)
UTE	ultra-short echo time (MR sequence)
VIBE	volumetric interpolated breath-hold examination
Z	atomic number
ZTE	zero-echo-time (MR sequence)

Acknowledgements

First, I would like to thank Daniela Thorwarth for making it possible to work on such an exciting research project. My heartfelt thanks, for your continuous guidance and support over these years. Your valuable advice and unlimited optimism was simply inspiring and always impressed and motivated me to persue and complete this work.

I would further like to thank Fritz Schick, Josef Jochum and Reinhold Kleiner for their kind support and their commitment in my dissertation committee. Many thanks also to Daniel Zips for the many valuable discussions of findings and very helpful feedbacks on manuscripts and conference abstracts.

To Marcel Krueger, Sandro Aidone and colleagues at the Werner Siemens Imaging Center for their substantial support with the preclinical imaging experiments. And to Sergios Gatidis, Holger Schmidt and Carsten Groeper for their great help with the clinical PET/MR examinations and phantom experiments. All of you played a major part in the completion of this work and therefore, thank you!

My special thanks to Sara Leibfarth. Without your support this work would simply not have been possible. Thank you for teaching me so many things, and for always giving a helping hand or some valuable advice. My special thanks also to Simon Boeke for the support, the valuable input and not least, for making the countless hours spent together in the imaging lab such a fun time.

Finally, many thanks to all my colleagues in the research group I was fortunate to meet during these years! Thank you for all the fun we had over coffee & cake or on trail running sessions on Friday afternoons; on trips to conferences, when exploring the city centers of Vienna and Barcelona or when visiting the Duomo in Milano. My time as PhD student would not have been the same without you all and will always be a very special memory for me.

Lastly, I would like to express my deepest gratitude to my friends and family for their love and support, and for giving me the strength to complete this work.

Contents

1 Introduction

Cancer is a major public-health problem worldwide and according to the most recent report on global cancer statistics (GLOBOCAN 2018) produced by the International Agency for Research on Cancer, incidence and mortality are rapidly increasing [1]. For 2018, 18.1 million new cancer cases and 9.6 million cancer deaths were estimated to have occurred worldwide. These findings clearly indicate the need for research on more effective cancer treatments.

Besides surgery and chemotherapy, an essential mode of cancer treatment today is radiotherapy (RT). It has been estimated from a comprehensive review of evidence based clinical guidelines that approximately 50% of patients with cancer would benefit from receiving RT at least once during the course of their illness [2]. RT plays an important role in curative but also in palliative treatment in patients and can be indicated as a sole therapy or in combination with surgery or concurrent chemotherapy.

The basic principle of RT is to deliver a high dose of ionizing radiation to the tumor tissue in order to induce cellular death by DNA damage, while minimizing the dose to surrounding normal tissue for the prevention of adverse side-effects. Typically, photon beams are used for treatment which are generated by 4–25 MV linear accelerators (Linacs). Photon energies range from less than 100 keV up to several MeV, depending on the maximum energy of the electrons accelerated by the Linac. Most DNA damage arises from ionizations caused by secondary electrons that originate from ionizations of other molecules, mainly water molecules, caused by the incident photons. These are mainly due to Compton scattering which is the dominating interaction of photons with matter at photon energies of 100 keV–10 MeV.

Research in radiobiology has contributed to a better understanding of cancer biology and response of normal and malignant tissue to radiation exposure and led to different concepts of fractionated treatment which are commonly applied

in clinical practice [3]. By delivering the radiation dose to the tumor in multiple fractions over several days or weeks, side effects, in particular, can be reduced as normal cells have a higher capability of DNA repair compared to cancer cells [4].

In addition, technological advances have strongly contributed to improvements in the treatment of cancer with RT. These include, in particular, improvements in the precision of dose delivery. Using modern techniques such as intensity modulated RT (IMRT), radiation fields can be applied from different angles, and field sizes and shapes can be adjusted to match the geometry of the tumor structure e.g., by using multileaf collimators. Thus, the radiation exposure of surrounding healthy tissue can be reduced while a high therapeutic dose can be delivered that is highly confined to the tumor [5].

At present, IMRT with photon radiation is standard of care, but also electrons, protons and heavier charged particles such as carbon ions can nowadays be used for treatment. These particles have different physical properties which result in different depth-dose profiles in tissue compared to photons. These can be exploited to improve the treatment of specific tumor sites. Electrons deposit their energy at low depth, which is beneficial for the treatment of near-surface or superficial cancer like skin cancer. The dose deposition of protons and ions, on the other hand, is characterized by slowly increasing deposition with a pronounced peak in higher depth (Bragg peak). This curve profile can be exploited to reduce the integral dose received by healthy tissue or to better spare specific organs at risk located closely to the target structure. The indication of expensive proton or particle therapy, however, depends on the tumor location and clinical trials are currently ongoing to identify the patients who are most likely to benefit from these elaborate technologies [6].

Such high precision treatment modalities require 3D high resolution imaging information on the tumor and surrounding tissue for precise treatment planning. Treatment planning comprises the delineation of the tumor and organs at risk and the simulation of dose distributions in these regions to determine the optimal delivery parameters of a prescribed dose. Primarily, x-ray computed tomography (CT) images are used for treatment planning. CT images provide spatially accurate anatomical information to identify the regions of interest and, since CT numbers reflect the distribution of the photon linear attenuation coefficients, can

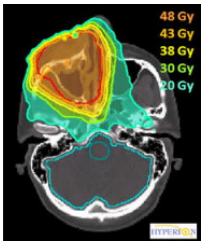

Figure 1.1: IMRT treatment plan on a CT in axial view of a 41-year-old patient with sinus cancer. The plan shows dose isocontours with a total dose of 60 Gy delivered to the tumor target volume (red contour), while brain stem (dark green) and brain (light blue) are best possibly spared from radiation dose. This optimized dose distribution is achieved by irradiating the tumor from multiple directions using multileaf collimators and intensity modulation. The figure is reprinted and adapted by permission of Springer Nature from reference [7].

be converted to electron density maps needed for dose calculation. A typical IMRT treatment plan overlayed to a planning CT is shown for a head and neck cancer patient in Fig. 1.1.

However, despite the advanced treatment concepts and the highly developed technologies for radiation dose delivery and treatment planning available today, different tumor types, such as head and neck cancer, are still associated with high rates of recurrence. About 50% of patients with locally advanced head and neck cancer experience loco-regional recurrence within three years of follow-up [8]. Moreover, RT delivered to the head and neck region is still linked to a high rate of acute and late side effects.

Modern medical imaging techniques like positron emission tomography (PET) and magnetic resonance imaging (MRI) have the potential to help identify such patients with poor response and to serve as a basis for improvements in RT treatment. For example, it has been shown that functional MR image-derived parameters such as the apparent diffusion coefficient (ADC) derived from diffusion-weighted (DW) MRI has high potential for prediction of treatment response. DW-MRI is a technique sensitive to the motion of water molecules in the cellular environment and the derived parameter ADC is an effective diffusion coefficient of the water molecules

being restricted in free motion by the biological environment, such as by hydrophobic cellular membranes. It has been shown that head and neck tumors with certain characteristics such as high stromal content, low cell density, and micro-necrosis are associated to treatment resistance and poor outcome, and that these characteristics are linked to less diffusion restriction of water molecules [9, 10]. This, in turn, is reflected by a higher ADC and different studies have found that high ADC is a valuable parameter for identifying patients with poor response [9, 11, 12].

Furthermore, dynamic contrast-enhanced (DCE) MRI has been shown to be valuable for predicting treatment outcome or local disease control [13, 14]. DCE-MRI reflects the spatio-temporal distribution of a paramagnetic contrast agent in tissue and allows to extract physiological information on perfusion and microvasculature of tissue e.g., by pharmacokinetic modeling of the acquired time-series data [15]. Here, higher vascular-related parameters seem to be related to better response to therapy [9].

On the other hand, using dedicated tracers labeled with different positron emitting radioisotopes, PET allows to image different physiopathological processes within the tumor. One example for such processes is increased glucose metabolic activity which can be detected using ^{18}F-labeled Fluorodeoxyglucose (FDG). Another example is cellular hypoxia, i.e., oxygen deprived tissue regions. Such regions are associated to a higher radioresistance and can be detected by means of tracers such as ^{18}F-Fluoromisonidazole (FMISO) accumulating in areas of low oxygen partial pressure. ^{18}F-FDG PET [16], as well as ^{18}F-FMISO-PET [17, 18] have been shown to have potential to identify head and neck cancer patients with treatment resistance and poor outcome.

Beyond the prediction of treatment response, MRI and PET based methods have the potential to improve RT treatment in many different respects. MRI allows different mechanisms to determine image contrast, such as weighting of the signal intensity by tissue type dependent relaxation times T2 or T1, and offers superior soft tissue contrast as compared to CT. Therefore, it may provide more precise information on tumor localization and spread and, thus, increase the delineation accuracy of tumor volumes [19]. In patients with head and neck cancer, contrast-enhanced T1-weighted MRI may further contribute to a higher confidence in determining the exact tumor extent and infiltration of crucial adjacent struc-

tures [20]. Using FDG-PET for imaging glucose metabolic activity has been shown to enhance the accuracy of RT target definition or staging [21, 22].

In addition, PET and MR imaging play a pivotal role in the development of new personalized treatment strategies for cancer patients with poor prognosis including the subset of patients with high-risk head and neck cancer [23–25]. Different concepts have been proposed to adapt a treatment of a patient individually. As mentioned, imaging features may help identifying subgroups of patients that are more likely to respond favorably or poorly to a given treatment. This information may then serve to tailor the treatment e.g., by prescribing a lower or higher dose to the entire tumor. On the other hand, based on specific imaging information indicating the local risk of relapse, subvolumes of the tumor may be targeted by a nonuniform dose distribution within the tumor [26]. Either a distinct subvolume of the tumor could be boosted (dose escalation, dose painting), or the dose could be prescribed on voxel level according to imaging derived probability maps of radioresponsiveness (dose painting by numbers) [25, 27].

In head and neck cancer, different imaging tools are currently being investigated as candidates for such strategies, most notably ^{18}F-FDG PET [28,29], as a surrogate for tumor burden, and PET with tracers such as ^{18}F-FMISO [30], as a surrogate for cellular hypoxia. In addition to PET, DW-MRI and DCE-MRI have been shown to be valuable for adapting RT treatment, e.g. by dose painting [31, 32].

An exemplary multiparametric data set of a patient with head and neck cancer comprising functional PET, DW- and DCE-MRI information which could potentially be used for treatment adaptation, is provided in Fig. 1.2.

The recent development of combined PET/MRI as a hybrid system has allowed to unite the advantages of both modalities within a single system [33]: the high soft tissue contrast and the wide range of functional imaging options of MRI, and the molecular information of PET. Combined PET/MRI as an integrated system may, thus, play a key role in further development and implementation of imaging based RT treatment adaptation [7, 23, 34, 35]. The accurate spatial integration of the imaging data into RT planning, however, requires patient examination in RT treatment position [19, 36, 37]. The adaptation of a PET/MRI examination to treatment position, on the other hand, is challenging as it requires RT patient im-

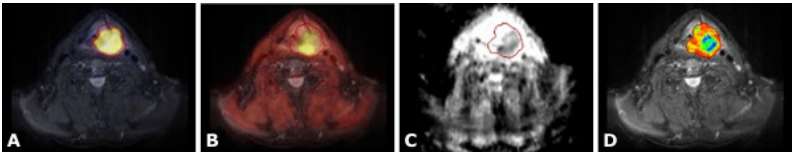

Figure 1.2: Multiparametric functional imaging of a 57-year-old male patient with head and neck cancer. A: ^{18}F FDG-PET overlayed to anatomical T2-weighted MRI. B: Combined ^{18}F FMISO PET/MRI acquired 4 h p.i.. C: ADC map derived from diffusion-weighted MRI and D: Perfusion map showing the distribution of the parameter K_{trans} derived via kinetic analysis from dynamic contrast-enhanced MRI inside the tumour volume. The radiotherapy gross tumour volume is outlined in each image slice. The figure is reprinted under the terms of the Creative Commons Attribution–NonCommercial 4.0 Unported License from reference [23].

mobilization equipment to be combined with the MRI hardware i.e., in particular, with the radiofrequency receiver coils. In contrast to stand-alone MRI systems, using a specially tailored coil setup for hybrid PET/MRI further requires appropriate PET photon attenuation correction with respect to the setup [38].

In the context of this thesis, a recently proposed RT-specific prototype setup [39] was further developed and systematically evaluated for attenuation corrected PET, and MR image quality within a clinical study. In addition to clinical feasibility, the aim was to assess whether the image quality of the prototype setup would meet the demands to accurately integrate PET/MRI data sets into RT planning.

Besides consistent patient positioning for precise data integration, high geometric accuracy of the image data is a further essential prerequisite for using medical images in RT planning [7, 19, 24]. This is a critical issue especially in DW-MRI which is often performed based on echo-planar imaging (EPI) sequences. EPI offers favorably short acquisition times reducing bulk motion-induced phase perturbations [40]. However, such EPI techniques are sensitive to magnetic field inhomogeneities leading to image distortions and signal loss. This especially applies to the head and neck region, where unfavorable geometry and susceptibility changes at air–tissue interfaces can lead to severe geometric distortions of DW-MR images up to a few centimeters [41, 42].

As high geometric fidelity is crucial for concepts of biologically adapted RT planning, a second focus in this work was on the assessment of the geometric accuracy of DW-MRI in head and neck, and on the definition of an appropriate method for distortion correction.

Multimodality imaging in general, and simultaneous PET/MRI in particular, have been shown to be of clinical value in many respects related to the treatment of cancer, such as target volume delineation in RT planning as outlined above, staging, the detection of distant metastases, monitoring of response to therapy or follow up [43–45]. Nonetheless, a primary role of multimodality imaging to help improve outcome in patients treated with RT may revert to characterizing the primary tumor and its microenvironment [26, 46].

One important feature to determine on such cellular level is hypoxia. Tumor hypoxia is one of the main factors for treatment resistance and poor prognosis, and has been shown to be present in many solid tumor types including head and neck cancer [18, 47–50]. PET imaging using different hypoxia sensitive tracers such as ^{18}F-FMISO is one means of noninvasive detection of hypoxia and has been explored in patient [18, 51] as well as in small animal studies [52]. Clinical studies, in particular, have shown that imaging of tumor hypoxia has high prognostic value with respect to RT response, when image information is derived from time series data [17, 30, 53–55] as such dynamic data includes both information on early temporal tracer distribution in the tumor tissue related to vascular properties as well as information on late retention effects [17, 30, 53–58].

Unfortunately, in many clinics the availability of hypoxia PET is limited and dynamic imaging protocols are extensive. In head and neck cancer patients, for instance, such protocols often consist of multiple scans from the time of injection up to 4 hours p.i. [17, 53]. Therefore, it would be an advantage if similar information on tumor hypoxia could be obtained by more easily available, functional MRI using brief protocols that are readily translatable into routine clinical practice.

In the third part of this thesis, a dedicated method was therefore developed to predict information on tumor hypoxia derived from dynamic FMISO-PET solely by functional MRI. The method is based on the training of different machine learning models using parameters derived from DW- and DCE-MRI as predictor variables.

2 Objective and outline

It is the aim of this work to contribute methods which leverage the integration of multiparametric PET and MR imaging into RT for future biologically individualized treatment of head and neck cancer patients. The methods presented in Part I and II of this thesis focus on different prerequisites for precise image integration into RT treatment planning. The method presented in Part III takes aim at advancing the use of multiparametric MRI as a tool for treatment adaptation strategies targeting hypoxic, radioresistant subregions within a tumor. In the following, a brief outline of each part is presented.

- **Part I** (Chapter 4) addresses the extension, implementation and characterization of a RT prototype setup designed for combined PET/MRI examinations of head and neck cancer patients in RT treatment position. The motivation is to thereby facilitate precise integration of the imaging data into treatment planning. A prototype setup consisting of a flat table top and MR coil holders for flexible body coils was extended by a dedicated add-on to allow for patient fixation with RT positioning tools. A first objective was to assess the feasibility of applying the experimental setup for patient examination in the context of a clinical study, using calculated attenuation maps of the hardware in addition to MR-based attenuation maps of the patient for retrospective PET photon attenuation correction. A second objective was to systematically evaluate the prototype setup for PET and MR image quality from the angle of integrating the image data into RT treatment planning. Therefore, a suitable evaluation method was developed based on quantitative measures for pairwise similarity analysis of RT target structures defined repeatedly on images acquired with the experimental setup and images acquired with a diagnostic reference setup.

- **Part II** (Chapter 5) deals with the geometric accuracy of DWI as part of a RT specific PET/MRI protocol for the head and neck region. Geometric accuracy is a critical issue in DWI in this anatomic region, but a further essential prerequisite for using the images in RT planning. Therefore, the focus was (i) on the assessment of geometric accuracy, (ii) on the definition of an appropriate method for distortion correction and (iii) on the stability of determining the diffusion coefficient ADC. Correction of susceptibility induced distortions was implemented based on repeated DWI data acquisition with reversed phase encoding directions, a method originally proposed for DWI in the brain [59]. To assess the geometric accuracy of original and corrected images in the critical head and neck area, an evaluation method was defined based on quantitative measures for pairwise similarity analysis of anatomical structures repeatedly delineated on original/distortion corrected images and an anatomical T2 weighted MRI reference image.

- **Part III** (Chapter 6) aimed at finding a set of parameters derived from clinically applicable MRI protocols that might serve as a tool for tumor characterization and as a basis for future treatment adaptation. The strategy to pursue this aim was to make use of supervised learning of imaging information on a known biological factor for radiation treatment resistance i.e., tumor hypoxia measured by dynamic FMISO-PET. Different machine learning models were trained to learn that information based on parameters derived from DWI and DCE-MRI used as input variables. The dynamic FMISO-PET/MRI data which formed the basis for model development was measured in small animal models of human head and neck cancer and a final, best performing model was tested in an independent cohort.

3 Background

3.1 PET photon attenuation correction methods for combined PET/MRI

3.1.1 Principles and physical aspects of PET imaging

Using dedicated pharmaceuticals labeled with different positron emitting radionuclides (e.g., ^{18}F, ^{64}Cu, ^{68}Ga,), PET allows to image different physiopathological processes within a patient (or an animal model). Such processes of interest may be glucose metabolic activity of tumor tissue, which can be detected using the ^{18}F-labeled glucose analog FDG, or tumor hypoxia, which can be traced using ^{18}F-Fluoroazomycin arabinoside (FAZA) [60], ^{18}F-Fluortanidazole (HX4) [61] or the most extensively studied hypoxia specific radiotracer ^{18}F-FMISO [18]. In order to localize the functional information, PET is commonly combined with the anatomical imaging modality CT or, more recently, MRI.

To conduct a PET scan, the radiotracer is administered to the patient by intravenous injection. An uptake time between injection and scanning is generally necessary for adequate image contrast and mostly depends on the retention mechanism of the tracer, the tissue or organ of interest and the half-life of the radionuclide. In PET, radionuclides are chosen that decay by positron emission. In this work, ^{18}F was used which features a half-life of 109.77 min and a probability of positron decay of 96.9% (3.1% electron capture). The positron decay equation for ^{18}F reads:

$$^{18}\mathrm{F} \longrightarrow {}^{18}\mathrm{O} + e^{+} + \nu_{\mathrm{e}} + E\,, \tag{3.1}$$

where ^{18}O denotes a stable nonradioactive isotope of oxygen, e^{+} the emitted positron, ν_{e} an electron neutrino and E excess energy partitioned as kinetic energy between the positron and the electron neutrino with varying amounts. Annihila-

tion of the emitted positron with an electron of the surrounding tissue leads to a pair of 511 keV photons being emitted in opposing directions. These photons can be externally detected by a ring-shaped PET system typically consisting of scintillation crystals coupled to photomultiplier tubes. For PET integrated into MRI, magnetic field-insensitive readout electronics such as avalanche photodiodes (APDs) are used instead of photomultiplier tubes which are used in conventional stand-alone PET or PET/CT systems [62].

From the detector signals, the density distribution of the positron emitting radionuclide in the patient and hence, the concentration of the tracer molecule, can mathematically be reconstructed. The reconstruction relies on coincidence detection of two photons and on associating the detection with an annihilation event having occurred in an undetermined position along a straight line (line of response, LOR) connecting the two detector elements which recorded the pair of photons. The coincidence window is typically in the order of a few nanoseconds. PET image reconstruction is then mainly based on iterative algorithms such as filtered back projection (FBP) or ordered subset expectation maximization (OSEM) to recover the original activity distribution from the full set of the acquired LOR projection data (the recorded counts along the LORs) [63].

In Figure 3.1, the iterative process of image reconstruction with FBP is demonstrated for an exemplary FDG-PET image slice of a patient's head (patient data from the study in Part I/Chapter 4). The LOR projection data is typically stored in sinogram format (Figure 3.1 A), with matrix row data corresponding to the subset of LORs at a certain angle, and columns indicating the distance of an LOR to the center of the projection plane (image plane). Summation of successive back projections of the sinogram matrix row data on an image plane of same dimension (3.1 B/E/H) and normalization to the number of iterations then allows to recover the actual PET image (Figure 3.1 I).

Advanced PET systems are capable to further enhance spatial resolution and signal-to-noise ratio (SNR) by making use of measuring the time-of-flight (TOF) difference between two photons to better approximate the localization of the annihilation event along an LOR [64]. However, TOF availability is limited for integrated PET/MRI, as here, mostly timing characteristics of avalanche photodiodes are too slow [65].

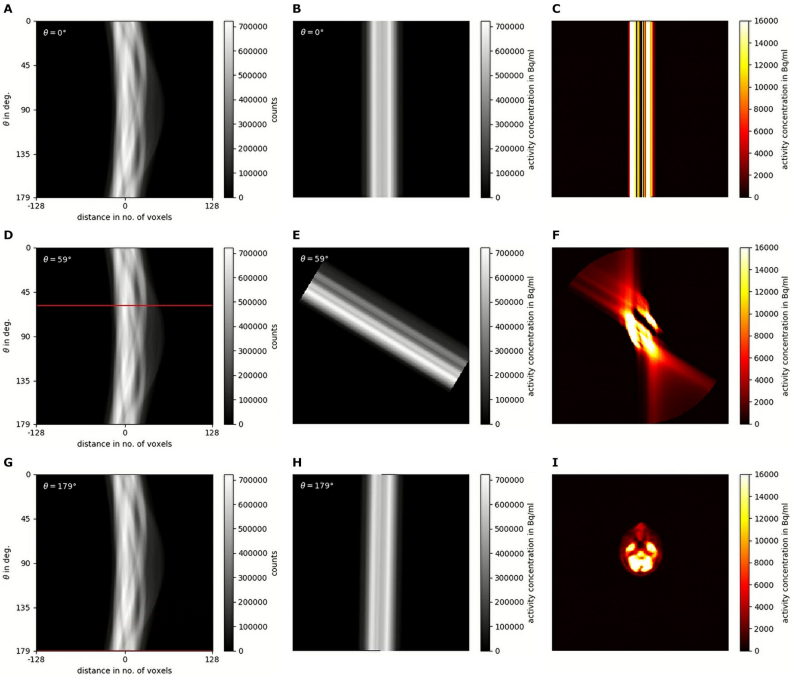

Figure 3.1: PET image reconstruction based on iterative back projection of the acquired raw data collected in sinogram format (A). Each data point in the sinogram image corresponds to the number of recorded counts along a specific LOR, which is defined by its angle and distance to the image center. The images B/E/H in the second column of the figure show the back projection of the sinogram matrix row data indicated as red lines in A/D/G, for three different angles. The third column of the figure shows the reconstructed image at the respective angle i.e., the normalized summation of backprojections after 1, 60 and a total of 180 iterations (C/F/I).

For image analysis and integration of images into RT planning, the spatial resolution of an image plays an important role. Often, the PET image grid is used as reference grid, e.g., for resampling MR image data in the context of joint PET/MR image analysis. This is opportune, as often resolution in PET is lower compared to MR images, especially inplane. With a modern clinical PET system a spatial resolution of approximately 4 mm (full width at half maximum, FWHM) can be achieved [66]. The average spatial resolution of the PET system implemented in the clinical PET/MRI scanner used in this work, the Siemens mMR Biograph, has been determined as 4.3 mm [67]. With current small animal PET inserts for combined preclinical PET/MRI, a higher spatial resolution of approximately 1–2 mm can be achieved [62, 68]. The spatial resolution of a PET image is limited by several factors with the size of the detector element being the most dominant one [69]. Further factors mainly include the positron range before annihilation, and acollinearity of the annihilation photons. A fundamental limit for the spatial resolution Γ (in mm FWHM) can be determined as:

$$\Gamma = \sqrt{(d/2)^2 + s^2 + (0.0044R)^2}\,, \tag{3.2}$$

where d denotes the width of a detector element, s the positron range and R the radius of the detector ring system [69].

For the clinical and small animal PET/MRI systems used in this study, the detector crystal width was 4 and 1.5 mm, respectively. The positron range is dependent on the radionuclide and the density of the matter it travels through and is 0.5 mm fwhm in soft tissue for ^{18}F, which is the minimum value among different radionuclides commonly used for medical PET. The range is due to the positron being emitted from the nucleus with up to a few MeV of kinetic energy (0.64 MeV for ^{18}F), such that it travels a certain distance in the tissue before it thermalizes and annihilates with an electron. In a combined PET/MR system, the presence of a strong magnetic field of a few Tesla forces the positron to follow a helical path, which reduces its range between emission and annihilation, especially in the plane transverse to the magnetic field. However, it has been shown that this effect is not significant for ^{18}F in soft tissue at 3 T in terms of an improved resolution [70].

The acollinearity factor $0.0044R$ depends on the detector ring size, which is set to accommodate the patient or experimental animal being examined. For the

clinical and small animal PET/MRI systems used in this study, radius R was 400 and 60 mm, respectively. Acollinearity is due to the fact that the bound state of positron and electron (positronium) has some residual momentum before it decays into the pair of annihilation photons, such that the photons may not exactly be emitted at 180° with respect to the reference frame of the detector system [69].

3.1.2 Interaction of PET annihilation photons with matter

Interactions of the annihilation photons with matter is an important factor for PET image quality. Photon attenuation, scatter and random coincidence effects result in a reduction in image quality, however, these factors can partially be corrected for.

Photon attenuation can be described by the exponential law of Lambert-Beer, which reads:

$$\frac{I}{I_0} = \mathrm{e}^{-\int \mu(x)\mathrm{d}x} \tag{3.3}$$

where I_0 is the intensity of photon radiation that would be measured in the absence of attenuation, I the remaining intensity due to attenuation and $\mu(x)$ the photon energy-dependent linear attenuation coefficient varying with the material at position x along the photon path. Integration is over the photon path. In the case of PET, both annihilation photons have to escape the patient without attenuation so that a true coincident event can be recorded by two detector elements. With Equation 3.3, the probability for this scenario can easily be estimated. Thus, the probability for the two photons passing through the patient without interaction is equal to the product of their respective single probabilities and given by:

$$\mathrm{e}^{-\int_0^v \mu(x)\mathrm{d}x} \cdot \mathrm{e}^{-\int_v^d \mu(x)\mathrm{d}x} = \mathrm{e}^{-\int_0^d \mu(x)\mathrm{d}x} \tag{3.4}$$

with d the total path length in the tissue, v the distance that one photon travels from the location of the positron annihilation to escape, d minus v the distance that the other annihilation photon must travel and $\mu(x)$ the position-dependent linear attenuation coefficient at the photon energy 511 keV [71]. Thus, the probability of both photons traversing the patient without interaction is independent from the location of the annihilation event along an LOR, and is equal to the probability of a single photon traversing the total path length.

Attenuation coefficents are generally given as mass attenuation coefficients normalized to the density of the material. In data libraries provided by the US National Institute for Standards and Technology (NIST), values of mass attenuation coefficients [72] and densities [73] are listed for a wide range of materials. According to [72] and [73], the mass attenuation coefficient and the density of human soft tissue (st) are similar to the values for water, i.e., $\mu_{511\,\mathrm{keV,st}}/\rho_{\mathrm{st}} \approx 0.096\,\mathrm{cm}^2/\mathrm{g}$ and $\rho_{\mathrm{st}} \approx 1\,\mathrm{g/cm}^3$. Assuming uniform attenuation, the probability expression simplifies to $\mathrm{e}^{-\mu d}$ and with $\mu_{511\,\mathrm{keV,st}} = 0.096\,\mathrm{cm}^{-1}$, the probability of a pair of annihilation photons traversing e.g., $d = 10\,\mathrm{cm}$ of soft tissue without interaction would, thus, be just about 38%.

Most of the interactions (99.76%) are due to Compton scattering, which is the dominant type of interaction for 511 keV photons in soft tissue (cf. Figure 3.2). Moreover, Compton scattering is also the dominant type for fat tissue (99.81%), bone (99.07%) or materials of MR hardware components, which may be present in the field of view of the PET detector, since at the energy level of 511 keV, it is not much dependent from the atomic number of the material (cf. Figure 3.3 A).

Photons undergoing Compton scattering in tissue deviate from their original direction, with some loss of energy. So on the one hand, if the energy of the scattered photon falls within the energy window of the detector crystal, a coincidence count may still be recorded. However, due to the angular deviation of the scattered photon, the associated LOR will be misinterpreted. At the energy of 511 keV, scattered photons mainly propagate in forward direction, with little energy loss [74]. Thus, incorrect counts are recorded that group around correct counts, which results in image blurring. A more narrow energy window in detection would lower the rate of recorded scatter, but the relative energy resolution of detector scintillation crystals is limited to approximately 10% [75], which means that at least, scattered photons with a scatter angle $\phi \lesssim 27°$ cannot be discriminated since the energy loss associated to such angles is less than 50 keV. The PET/MRI system used in this work was characterized by a relative energy resolution of 14.5%; an energy window of lutetium oxyorthosilicate (LSO) crystals of 430–610 keV; and a scatter fraction of the recorded signal of 37.9% (as determined by NEMA, National Electrical Manufacturers Association, protocol) [67].

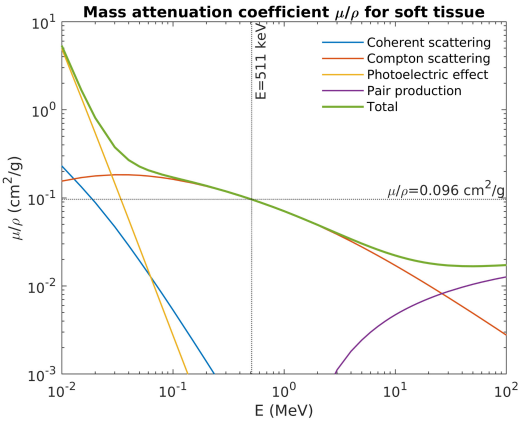

Figure 3.2: Total mass attenuation coefficient and different components (coherent scattering, Compton scattering, photoelectric absorption and pair-production in nuclear field) for soft tissue as a function of photon energy. Soft tissue data is simulated by data for water which has a similar effective atomic number ($Z \approx 7$). The total mass attenuation coefficient relative to the energy of annihilation photons detected in PET is highlighted (dotted horizontal line). Curves were generated using values obtained from reference [72].

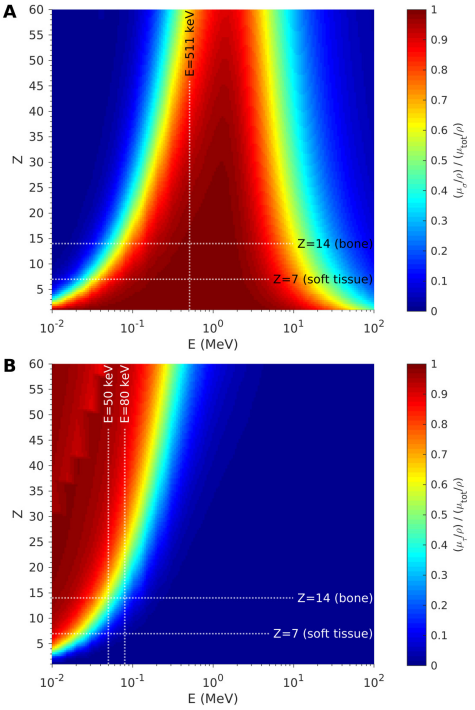

Figure 3.3: A: Ratio of Compton scattering-related attenuation to total attenuation, $(\mu_\sigma/\rho)/(\mu_{\text{total}}/\rho)$, as a function of atomic number Z and energy E of the incident photon. The effective atomic numbers associated to human soft tissue and bone are highlighted as dotted horizontal lines; the energy of annihilation photons $E = 511\,\text{keV}$ is highlighted as a dotted vertical line. At 511 keV, Compton scattering is the dominant type of interaction for human tissue, also for elements with high Z such as Calcium (Z=20) contained in bone or teeth, or Titanium (Z=22) eventually present in implants. This holds true for materials used in MR radiofrequency coils which may be placed within the PET field of view. B: Similarly, the ratio of photoelectric absorption-related attenuation to total attenuation, $(\mu_\tau/\rho)/(\mu_{\text{total}}/\rho)$, is presented as a function of atomic number and photon energy. Here, the range of effective CT energies of $E = 50$–$80\,\text{keV}$ for typical 80 kV to 140 kV x-ray spectra are highlighted. At these energies, the contribution of photoelectric absorption to total attenuation is low for low-Z associated matter like air, fat tissue and soft tissue, as compared to Compton scattering, but becomes more dominant with increasing Z, such as in bone. Surface plots were generated using tabular data from reference [72].

On the other hand, if the energy loss of scattered photons is too high, no according coincidence signal can be recorded. Such photon attenuation-related signal loss is not the same for all LORs and needs to be corrected for to prevent image artifacts due to underestimation of the activity concentration [71].

In combined PET/CT systems, the information of x-ray transmission can be used to correct the PET data for attenuation. In combined PET/MRI, this CT information is not intrinsically available and, thus, requires different methods for PET photon attenuation correction of the MRI hardware and the patient. Attenuation maps of the hardware can be determined once beforehand based on CT data acquired in a separate CT scan and can subsequently be deposited on the PET/MR system. Attenuation maps of the patient can be determined during PET/MR examination by dedicated MRI sequence techniques. Such CT- and MR-based correction methods are discussed in detail in the following three Sections.

3.1.3 Attenuation correction

PET information can be corrected for attenuation by applying an attenuation correction factor (ACF) for each LOR during image reconstruction. This correction process is illustrated in Figure 3.4 for the same exemplary PET image slice as shown before (Fig. 3.1). Here, an attenuation corrected PET sinogram (Fig. 3.4 C) is obtained by elementwise matrix multiplication of the original PET sinogram (Fig. 3.4 A) with a sinogram of the ACF (Fig. 3.4 B). The ACF is equal to the inverse of the attenuation probability along the LOR given in Equation 3.4, that is

$$\mathrm{ACF} = \exp\left[\int_{\mathrm{LOR}} \mu_{511\,\mathrm{keV}}\,(x)\,\mathrm{d}x\right]. \tag{3.5}$$

The final corrected PET image (Fig. 3.4 D) can then be obtained from the corrected sinogram data by iterative reconstruction algorithms such as FBP, as described above (Section 3.1.1).

Thus, correction with ACF just requires knowledge of the linear attenuation coefficient $\mu_{511\,\mathrm{keV}}$ at each position x in the image and conversion of such an attenuation map to a map of the ACF in a sinogram format. Such an ACF sinogram can be obtained by forward projection of the attenuation map using Equation 3.5

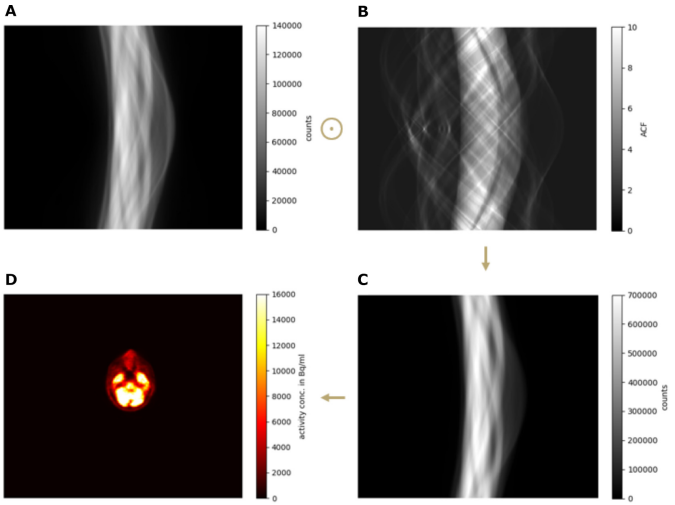

Figure 3.4: PET photon attenuation correction. Elementwise matrix multiplication of a measured PET sinogram (A) with a ACF sinogram (B) yields an attenuation corrected PET sinogram (C). D: An attenuation corrected PET image reconstructed from the corrected sinogram using FBP. The PET data refers to an image slice of a patient's head in axial view; the ACF sinogram is derived from a combined attenuation map of MRI hardware and patient as shown in Figure 3.5.

as illustrated in Figure 3.5 for an attenuation map of MRI hardware and a patient's head. Here, ACF projection profiles (Fig. 3.5 D–F) obtained from forward projection of $\mu_{511\,\mathrm{keV}}$ along LORs, i.e., along the columns of the attenuation map matrix (Fig. 3.5 G–I), are shown for three different angles. These profiles can then be collected row-wise in a data matrix (Fig. 3.5 A–C) with number of columns corresponding to the number of columns of the attenuation map, and complete iteration over 180° results in a final ACF sinogram (Fig. 3.5 C).

3.1.4 CT-based attenuation maps

For PET/MRI, an attenuation map of the MRI transmit and receive coils can be obtained by processing data from an x-ray CT scan of the equipment, as CT measures the linear attenuation coefficients averaged over the x-ray energy spectrum. The latter is dependent on the tube voltage. The effective energy of a 120 kV spectrum, for instance, is approximately 60 keV. At such energy, the linear attenuation coefficients are very different from those for 511 keV photons and vary greatly with the atomic number of the material because of the photoelectric effect which, at x-ray energies, becomes more dominant with increasing atomic number (cf. Figure 3.3 B). Therefore, the ratio of linear attenuation coefficients at x-ray and PET photon energy is not constant but depends upon the material in the individual voxel.

A common method to estimate linear attenuation coefficients at 511 keV from CT data relies on estimating the composition of the material by the measured CT number HU (Hounsfield unit). HU is the average linear attenuation coefficient at the x-ray energy normalized to the coefficient for water, $\mu_{\text{x-ray,w}}$:

$$\mathrm{HU}(x) = 1000 \cdot \frac{\mu_{\text{x-ray}}(x) - \mu_{\text{x-ray,w}}}{\mu_{\text{x-ray,w}}} \tag{3.6}$$

where HU = 0 for a voxel x corresponding to water and HU = −1000 for air, as $\mu_{\text{air}} \approx 0$.

A method introduced by Carney et al. is based on applying a bilinear transform to estimate $\mu_{511\,\mathrm{keV}}$ from measured HU [76]. The motivation behind the bilinear transform is that for matter associated with low Z and therefore low fraction of photoelectric absorption such as air (Z=7.37), skeletal muscle (Z=7.06), soft tissue

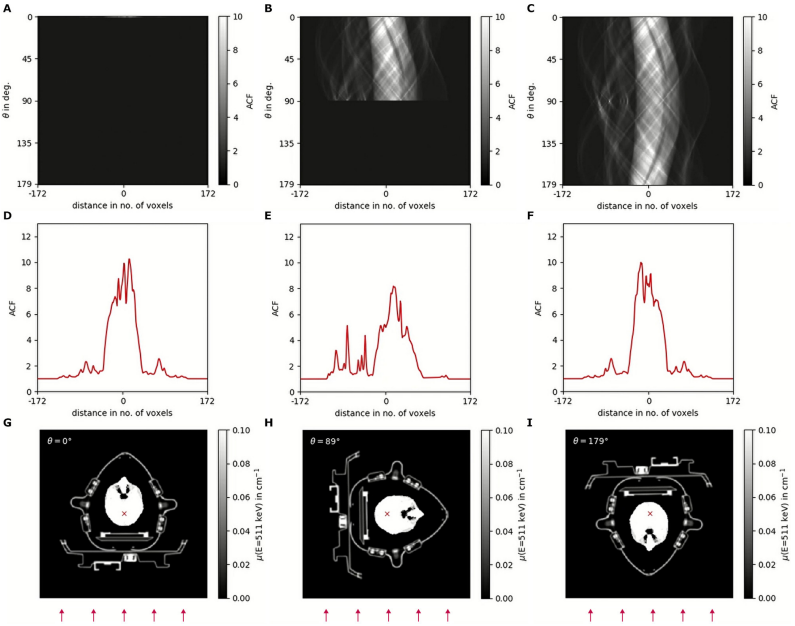

Figure 3.5: ACF sinogram (C) obtained by Equation 3.5 and iterative forward projection of a combined attenuation map for MRI hardware (MR table, overlay, coil holders and coils) and a patient's head (G, axial image slice). D–F: ACF projection profiles obtained through projection along LORs (columns of the image matrix) for three different angles; direction of projection is indicated by red arrows. A–C: ACF projection data collected row-wise for all angles in a data matrix with same number of columns as the attenuation map.

(Z=7.07) or water (Z=7.22, values from [73]; cf. Figure 3.3 B), the ratio of the total mass attenuation coefficients at CT energy and PET energy is approximately constant. For materials with higher Z such as bone, on the other hand, the contribution of photoelectric absorption at CT energy is more important, and the ratio of coefficients is much different. Terefore, a different scaling factor is required for bone-like voxels [77].

The transform proposed by Carney et al. reads:

$$\mu_{511\,\mathrm{keV}} = \begin{cases} 9.6 \cdot 10^{-5}\,\mathrm{cm}^{-1} \cdot (\mathrm{HU} + 1000) & \text{for HU} \leq \mathrm{BP} \\ a \cdot (\mathrm{HU} + 1000) + b & \text{for HU} > \mathrm{BP} \end{cases} \tag{3.7}$$

where breakpoint BP and parameters a and b (in cm^{-1}) depend on the tube voltage setting of the CT. Values of BP, a and b for various peak kilovoltage (kVp) settings are provided in [76]. The presented transform was optimized for human tissue and was defined by separated linear least-square fits to measured HUs in phantom reference tissue inserts with known 511 keV linear attenuation coefficients. kVp-dependent values for BP range from 30–52 such that Equation 3.7 yields correct linear attenuation coefficients for air (HU $= -1000$, $\mu_{511\,\mathrm{keV}} = 0$) and water (HU $= 0$, $\mu_{511\,\mathrm{keV}} = 0.096\,\mathrm{cm}^{-1}$), and approximately correct values for the different soft tissues. The linear equation for HU $>$ BP assumes that voxels are composed of a certain ratio of bone tissue to soft tissue. Thus, equation 3.7 yields adequate values for $\mu_{511\,\mathrm{keV}}$ in the patient, but also in certain hardware components, when the assumption on the composition of the material is valid. In this work, this method was used for calculating hardware attenuation maps of an RT table overlay and an add-on for mask fixation.

To estimate 511 keV linear attenuation coefficients of different hardware such as of MRI radiofrequency coils, this assumption on the composition of the material has been shown to not be valid and different values for the constants a and b have to be used. In [78], a and b were determined for an MR flexible radiofrequency coil by linear fitting of HU measured by CT plotted against $\mu_{511\,\mathrm{keV}}$ measured by a 511 keV reference transmission scan. Using these adapted values for a and b instead of values listed in [76] to determine a 511 keV attenuation map of radiofrequency coils prevents from overestimating $\mu_{511\,\mathrm{keV}}$ and, thus, from overestimating the activity

concentration in the image at respective locations. The attenuation maps used in this work for correction of a pair of flexible coils relied on this adapted transform.

3.1.5 MR-based attenuation maps

Unlike attenuation maps of MRI hardware which have to rely on pre-acquired CT or 511 keV transmission data, a 511 keV attenuation map of the patient (human μ-map) can be determined during the PET/MR examination. Different methods have been developed. Some are based on estimating attenuation information directly from the PET emission data [79], but most methods are based on MRI data [80,81]. In contrast to CT, MR based attenuation correction is challenging as MR images do not provide a direct correlation to photon attenuation. While both CT signal and 511 keV photon attenuation are mainly determined by the electron density of the material, the MR signal intensity is related to the proton density and characteristic relaxation times. Moreover, low signal intensity of bone and lung in images acquired with conventional MR sequences because of low proton density and short transverse relaxation times, makes it difficult to detect these tissue types.

Nonetheless, several attenuation correction methods have been developed to indirectly provide information on photon attenuation based on MRI [80, 81]. These methods can mostly be categorized into two categories. The first is that of atlas or template based methods [82, 83] which typically rely on a CT template, e.g. a pair of previously aligned MRI and CT images of a normal subject. This template may be transformed to the MR image of the patient being examined, or a previously trained algorithm may be used to generate a pseudo-CT image based on the MRI. The human μ-map may then be obtained by traditional scaling of the HU of the transformed template CT or pseudo-CT, as described above.

The second category is based on MRI using Dixon [84, 85], ultrashort-echo-time (UTE) [85, 86] or zero-echo-time (ZTE) techniques [87]. In these methods, the patient MR image is segmented into different tissue classes and constant attenuation values are assigned to each class. Recently, advanced methods have been proposed to assign continuous attenuation coefficients on voxel level [88, 89]. UTE and ZTE techniques, in particular, allow for identifying bone tissue.

In the first generation of PET/MRI systems like the one used in this work, mostly segmentation-based methods with discrete attenuation values are used that ignore bone, because of their straightforward implementation and fast runtime [90].

In segmentation-based methods, the tissue classes fat and water (nonfat soft tissue), are typically identified using Dixon technique [91], which is an imaging and processing technique for fat and water separation exploiting the difference in resonance frequencies between fat- and water-bound protons (chemical shift). The technique is based on the assumption that only water and fat contribute to the signal of an image. Under this assumption, and neglecting errors e.g., in magnetic field homogeneity, the signal can theoretically be expressed as:

$$S = S_w + S_f \cdot e^{i\alpha} \tag{3.8}$$

where S_w and S_f represent the magnitudes of the magnetizations at a given pixel for water and fat, respectively [92]. α is the phase angle of fat relative to that of water due to their chemical shift difference:

$$\alpha = \gamma \cdot B_0 \cdot \sigma \cdot \Delta t \tag{3.9}$$

where γ is the proton gyromagnetic ratio, σ is the chemical shift factor of fat relative to water, and B_0 is the external magnetic field of the scanner. In a typical two-point Dixon acquisition, two separate images are acquired with different Δt, representing changes in echo time (TE), which are tuned to obtain specific values of α. Thus, one image can be acquired with water and fat being in-phase ($\alpha = 0$, $e^{i\alpha} = 1$) and another being 180° out-of-phase ($\alpha = \pi$, $e^{i\alpha} = -1$):

$$S_1 = S_w + S_f \tag{3.10}$$

$$S_2 = S_w - S_f \tag{3.11}$$

By averaging the sum and differences of voxel intensities in in- and opposed-phase images S_1 and S_2, water only and fat only images S_w and S_f can be generated,

respectively:

$$\frac{1}{2}\left[S_1 + S_2\right] = \frac{1}{2}\left[(S_w + S_f) + (S_w - S_f)\right] = S_w \tag{3.12}$$

$$\frac{1}{2}\left[S_1 - S_2\right] = \frac{1}{2}\left[(S_w + S_f) - (S_w - S_f)\right] = S_f \tag{3.13}$$

The MR image segmentation approach used in this work was a 5-class segmentation of an MR image into air, lung, fat tissue, a fat–nonfat tissue mixture, and nonfat soft tissue. In this approach, segmentation is based on processing the in-phase, fat, and water images S_1, S_f, and S_w obtained with a two-point Dixon technique [90]. Bone tissue is ignored and treated as soft tissue. Air and tissue are determined by thresholding the intensity-normalized in-phase MR image. Misclassification of low-intensity bone voxels as air is reduced by a morphological image closing filter [84]. Fat and nonfat soft tissue are identified based on the Dixon fat and water only images. This is done by assigning voxels that show more than twice the intensity in the fat image as compared to the water image, to fat, and vice versa for water. Voxels with a ratio of fat and water intensities of more than 1/2 but less than 2, are interpreted as a mixture of fat and water-like soft tissue. To these voxels, an attenuation coefficient equal to the mean value of the fat and water attenuation coefficients is assigned. Lung tissue is detected as the largest connected group of low-intensity voxels [90]. The linear attenuation coefficients which are finally assigned to the five tissue classes are: $0\,\mathrm{cm}^{-1}$ for air, $0.0224\,\mathrm{cm}^{-1}$ for lung, $0.0854\,\mathrm{cm}^{-1}$ for fat tissue, $0.0927\,\mathrm{cm}^{-1}$ for the fat-nonfat tissue mixture, and $0.1\,\mathrm{cm}^{-1}$ for water-like soft tissue.

Dixon images and an attenuation map derived using this approach are presented for an exemplary patient in Figure 3.6. In addition, an attenuation map obtained by combining the Dixon technique with an atlas based technique for the integration of bone information is shown.

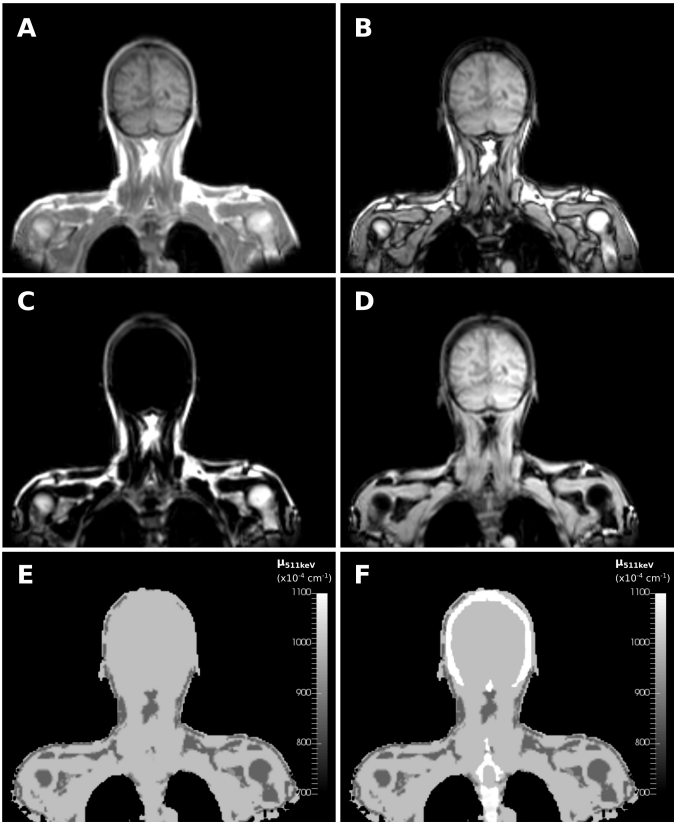

Figure 3.6: Illustration of Dixon technique for MR based photon attenuation correction of a 49-year-old male patient with head and neck cancer examined with combined PET/MRI. A: In-phase MR image acquired in typical coronal orientation. B: Opposed-phase image, same coronal slice. C–D: Fat and water images derived from in- and opposed-phase images (cf. Equations 3.12–3.13). E: 511 keV linear attenuation coefficient map obtained by 5-class Dixon segmentation [90] containing discrete values, ignoring bone. F: Dixon based attenuation map similar to E, with additional atlas based attenuation information on bone. Note, that for visualization purposes, the window level for attenuation maps in E and F is set to a subinterval of the total range of attenuation values (range E, 0–1000 $\times 10^{-4}\text{cm}^{-1}$; range F, 0–2488 $\times 10^{-4}\text{cm}^{-1}$).

3.2 Functional MRI and distortion correction methods for diffusion-weighted MR image data

3.2.1 Principles and physical aspects of MR imaging

Most imaging techniques including x-ray or CT operate with short-wavelength radiation with wavelengths being smaller than the structures to be imaged. Wavelengths of x-rays, for instance, are in the order of magnitude of around 10^{-11}–10^{-10} m). MRI, on the contrary, uses electromagnetic radiation with wavelengths in the order of 1–10 m (non-ionizing radiowaves). Despite the large wavelengths, MRI allows for image resolutions in the millimeter or submillimeter range by exploiting the intrinsic magnetic properties of atomic nuclei and their behaviour in a high magnetic field.

The nuclear spin or intrinsic angular momentum $\boldsymbol{I}$ is associated with a magnetic moment, $\boldsymbol{\mu} = \gamma \boldsymbol{I}$, that can interact with external magnetic fields. The gyromagnetic ratio γ is the ratio of magnetic moment to angular momentum and is specific to each nucleus. The nuclear spins partially align in an external static magnetic field of field strength B_0, precessing around its axis with a specific angular frequency named Larmor frequency $\omega_L = \gamma B_0$. ω_L is dependent on the nucleus type and the magnetic field strength and for protons, ω_L is 42.6 MHz per T. The spins can further interact and resonate with electromagnetic waves of a frequency corresponding to ω_L, and can thus be excited to induce a measurable highfrequent voltage in a nearby receiver coil.

Many types of nuclei exhibit a nuclear spin. The condition for a nonzero nuclear spin is fulfilled for all atomic nuclei except those with an even number of protons and at the same time, an even number of neutrons. In medical imaging, mostly hydrogen (^{1}H) nuclei (protons) are used because of high natural abundance in the human body and high gyromagnetic ratio.

The strong static magnetic field $\boldsymbol{B_0}$ in a MR tomograph is generated by a superconducting magnet coil. The field strength is typically in the order of magnitude of a few T. This yields a slight surplus of protons aligned along $\boldsymbol{B_0}$ in energetically more favorable parallel orientation as compared to antiparallely aligned protons. At 3 T and physiologic temperatures, for instance, it is a surplus of $1/10^5$. The number of excess protons generates a net macroscopic magnetization $\boldsymbol{M}$. As the precess-

ing protons absorb radiofrequency energy tuned to their precessional frequency ω_L and transition to the higher energy, anti-parallel direction, magnetization $\boldsymbol{M}$ is deflected from its alignment along $\boldsymbol{B_0}$ creating a magnetization component perpendicular to the field axis and rotating at ω_L. The rotating transverse magnetization induces a measurable high-frequent voltage in a closely positioned receiver coil.

Spatial encoding of the signal demands additional gradient coils to generate a magnetic field with field strength in the order of mT/m that superimposes on B_0, to attain a defined position-dependency of ω_L. Upon a sequence of RF pulses, magnetic field gradient settings and signal acquisitions, raw signal data can be collected in k-space and can be transformed to image data by Fourier Transform. This yields an image matrix containing complex numbers $z = |z|\exp(i\phi)$ whose magnitude values $|z|$ give the grey-scale values of the actual MR image.

Besides the spatial proton density, the tissues' T2 and T1 relaxation properties are typically exploited to create image contrast. T2 relaxation describes the process by which the transverse component of magnetization decays after a radiofrequency pulse due to dephasing of the spins. This exponential decay is characterized by the time constant T2, the time required for the transverse magnetization to fall to approximately 37% (1/e) of its initial value. T1 relaxation describes the process by which the longitudinal magnetization returns to its initial value after a radiofrequency pulse, as spins favor to return to the lower energy, parallel orientation. This exponential recovery process is characterized by the time constant T1, the time needed for recovery of approximately 63% $(1 - 1/e)$ of the initial longitudinal magnetization.

T1 and T2 relaxation processes occur at different time scales and vary between different tissue types, as they are influenced by molecular motion, size and interactions. MRI sequences can provide different contrasts depending on chosen sequence parameters which have influence on whether T1- or T2-based contrast is more pronounced. This is especially useful for the differentiation, e.g., of tumor versus adjacent normal tissue and of great relevance to radiotherapy in terms of target volume definition.

3.2.2 Dynamic contrast-enhanced and diffusion-weighted imaging

Besides proton density, T1 or T2 weightings, which basically provide anatomical information, different weightings and image contrasts can be obtained, e.g., by using dedicated exogenous contrast agents (contrast-enhanced MRI), or by using specific sequences of radiofrequency pulses and field gradient settings to sensitize MRI for the diffusion of water molecules (diffusion-weighted (DW) MRI). By the analysis of images obtained with such methods, different functional information on a tumor and insights into the tumor microenvironment can be obtained. These may help to retrieve prognostic or predictive information on treatment outcome, or may assist in a better evaluation of treatment efficacy during treatment.

In T1-weighted contrast-enhanced MRI, image contrast can be modified by a strongly paramagnetic contrast agent administered to the patient intravenously. The paramagnetic agent, typically gadolinium based, alters the relaxation properties of tissue molecules in its vicinity. Predominantly, T1 relaxation time of nearby hydrogen protons is shortened, which results in enhanced signal intensity in T1 weighted images. DCE-MRI relates to the measurement of the T1 changes in tissue over time by recording time-series image data after a bolus administration of the contrast agent. Following the injection, the contrast agent molecules are circulated to organs and tissue through blood supply and may pass through the vascular endothelium via diffusion into the extracellular, extravascular space of the tissue. This leads to local concentration dependent contrast enhancement which is, throughout the time course, eventually lowered through transport of the agent back out of the tissue. Observing and quantifying the time course of such contrast enhancement then allows to relate to different physiological properties of the tissue. These include blood flow, blood vessel permeability and different tissue volume fractions [15]. For such analysis, different pharmacokinetic models have been proposed [93, 94], but also alternative, data-driven approaches may be used to extract parameters of relevance to RT, such as principal component analysis (PCA). PCA of dynamic image data is addressed in Section 3.4.1.

DW-MRI, on the other hand, allows to visualize local diffusion of water molecules in the tissue. Diffusion, i.e., Brownian motion, is thermally driven and influenced by the cellular environment of the water molecules, such as hydrophobic cellular membranes. Thus, DW-MRI allows to relate to the cellular architecture of the

tissue, such as to cell density, organization of the tissue in terms of shape and size of the intercellular space, and to the integrity of cellular membranes [95]. Technically, diffusion-weighting can be achieved by an additional pair of gradients of the same magnitude and duration but with reversed sign applied between the radiofrequency excitation and signal collection. This way, stationary spins are not influenced by the gradients, in contrast to diffusing spins, for which dephasing occurs at a degree related to the local diffusion. This is, because for non-stationary or diffusing spins, the position-dependent phase change received through the first gradient is not recompensated by the second gradient. For an ensemble of diffusing spins, this results in a reduction in signal that is higher than for an ensemble of spins with few mobility. And the higher the diffusion-weighting for the image acquisition, the higher the signal reduction in regions of diffusing spins.

From a serial acquisition of images with the same sequence parameter settings but different strengths of diffusion weighting (b-value setting), a physical parameter map of the apparent diffusion coefficient ADC can be derived, e.g., for tumor tissue, by modeling the diffusion-related signal decay. ADC represents an effective diffusion coefficient of water molecules, which are restricted in free motion by the cellular environment, such as by the cellular membranes [95]. The b-value, which is varied to control the diffusion-weighted contrast, is a MR sequence parameter that summarizes the settings of the magnitude G of the diffusion sensitizing gradients, the temporal duration δ of a gradient, as well as the time gap Δ between the two gradients:

$$b = \gamma^2 G^2 \delta^2 (\Delta - \delta/3) . \tag{3.14}$$

Different models can be used to derive information on diffusion from a series of b-value images. A typical model to extract the ADC from such DW-MRI data assumes a mono-exponential signal decay with increasing b-value:

$$S = S_0 \exp(-ADC \cdot b) \tag{3.15}$$

where S denotes the measured signal intensity of an image voxel and S_0 a secondary fit parameter describing the signal intensity at $b = 0$. An example for mapping the ADC in a tumor is given in Fig. 3.7.

Due to its robustness, the mono-exponential model is often chosen over more complex models, such as intra-voxel incoherent motion (IVIM) [96, 97]. IVIM is a bi-exponential model, that allows for more advanced separation of perfusion and diffusion effects, but poses challenges in image acquisition and quality, and is less robust than the mono-exponential model especially for analysis in head-and-neck tumors at a voxel level [98].

In general, several factors should be considered at the time of image acquisition in order to optimize DW-MR image quality for robust data analysis. Mainly, additional averages should be performed to compensate for the reduction in SNR at high b-values. Moreover, parallel imaging, short TE and fat suppression are recommended [95]. Breath-hold techniques or respiratory triggering may be favorable to reduce motion induced artifacts. Advanced shimming techniques help to reduce magnetic field inhomogeneity and related image distortions [41]. The latter is a critical issue especially in DW-MRI of the head and neck and, therefore, it is discussed in the next section.

3.2.3 Image distortion and distortion correction

DW-MRI is often performed based on echo-planar imaging (EPI) sequences. This is a technique, that offers short acquisition times through rapid data collection in k-space. In single-shot EPI, for instance, the total of k-space is filled by a continuous echo train after a single radiofrequency excitation pulse. Such k-space filling is achieved through a series of alternating readout and phase-encoding gradients leading to zigzag-like k-space sampling trajectories. Thus, the technique is extremely fast in terms of image acquisition and therefore, less sensitive for motion artifacts [40].

However, this sampling and encoding approach makes EPI very sensitive to magnetic field inhomogeneities. Apart from scanner imperfections and eddy-currents arising e.g., from strong diffusion-encoding gradients, B_0 field inhomogeneities are particularly severe near boundaries between materials or tissues with large susceptibility differences [99]. In contrast to conventional imaging techniques, the field-sensitive direction in EPI is the phase-encoding direction. In the frequency-encoding direction, strong gradients with a large receiver bandwidth are applied for rapid sampling of k-space. In comparison, the bandwidth in the phase-encoding

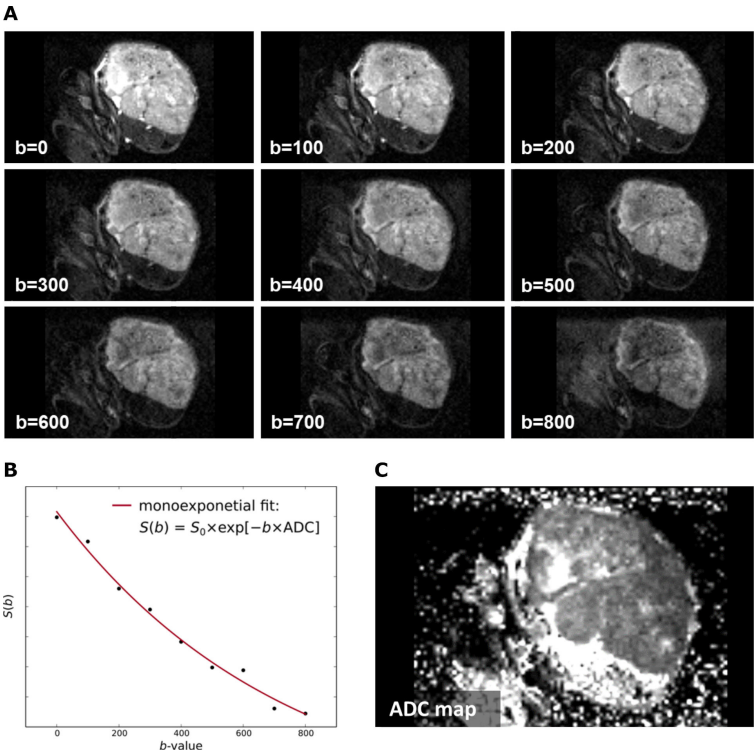

Figure 3.7: Illustration of ADC mapping based on a mono-exponential model for DW-MRI data of a xenografted head-and-neck tumor. A: DW-MR image series acquired with various diffusion-weightings controlled by the b-value (given in s/mm^2), presented for one tumor slice (axial view). B: A mono-exponential fit to the diffusion related signal decay for a tumor voxel using (3.15) to extract the parameter ADC. C: ADC map obtained by modeling the DW-MRI data for each image voxel.

direction is very low due to the continuous encoding scheme. In our patient study, for instance, the bandwidth per pixel in phase-encoding direction was just about 20 Hz, versus 1736 Hz in frequency-encoding direction. The low pixel bandwidth makes the phase-encoding direction much more sensitive to small offsets in precession frequency from magnetic field inhomogeneities accumulated during the long echo-train readout. For example, given a bandwidth in phase-encoding direction of 20 Hz/pixel, a local frequency offset Δf_B of, e.g., 100 Hz induced by B_0 static field inhomogeneity or magnetic susceptibility differences, would result in an unwanted pixel displacement in phase-encoding direction of 5 pixels.

Especially in the head-and-neck region of a patient, distortions can be severe due to the complex geometry of chin, jaw and shoulders and magnetic susceptibility differences at interfaces of air and soft tissue, or soft tissue and bone. Since high geometric accuracy is essential for RT planning purposes, such distortions need to be corrected for, before image derived information, such as information on tumoral ADC, can be used for treatment adaptation strategies.

A typical method for distortion correction is field mapping, i.e., estimating the magnetic field inhomogeneity $\Delta B_0(\boldsymbol{r})$ from the phase information of two (distortion unsusceptible) gradient echo images acquired with different echo time (TE) [99]. The phase map of each complex image data set z can be obtained by $\phi = \frac{Im(z)}{Re(z)}$. In gradient echo imaging, the resonant nuclear spins accumulate transverse phase during the acquisition in dependence on the local magnetic field $B(\boldsymbol{r}) = B_0(\boldsymbol{r}) + \Delta B_0(\boldsymbol{r})$ they experience, such that the resulting complex image will acquire a phase

$$\phi(\boldsymbol{r}) = \omega(\boldsymbol{r}) \cdot \mathrm{TE} = \gamma B(\boldsymbol{r}) \cdot \mathrm{TE}\,. \tag{3.16}$$

In the case of repeated imaging with different TE, where $\Delta\mathrm{TE} = \mathrm{TE}_2 - \mathrm{TE}_1 > 0$, the magnetic field inhomogeneity can be determined from the phase difference $\Delta\phi(\boldsymbol{r}) = \phi_2(\boldsymbol{r}) - \phi_1(\boldsymbol{r})$:

$$\Delta\phi(\boldsymbol{r}) = \gamma B(\boldsymbol{r}) \cdot \Delta\mathrm{TE} \tag{3.17}$$

$$\Rightarrow\ \Delta\phi(\boldsymbol{r}) = \gamma\left[B_0(\boldsymbol{r}) + \Delta B_0(\boldsymbol{r})\right] \cdot \Delta\mathrm{TE} \tag{3.18}$$

$$\Rightarrow\ \Delta B_0(\boldsymbol{r}) = \frac{\Delta\phi(\boldsymbol{r})}{\gamma\Delta\mathrm{TE}} - B_0(\boldsymbol{r})\,. \tag{3.19}$$

Then, knowing ΔB_0, the pixel shift in phase encoding direction of an EPI image can be calculated. ΔB_0 can be expressed as a frequency deviation $\Delta f_B = \frac{1}{2\pi}\Delta\omega = \frac{1}{2\pi}\gamma\Delta B_0$, such that the pixel shift will be given by:

$$u = \frac{\Delta f_B}{\Delta f_{pp}} \tag{3.20}$$

where Δf_{pp} is the bandwidth per pixel in phase encoding direction [100].

A downside of such field mapping approaches is that they require additional scan time and are susceptible for errors related to patient movements or to phase wrapping. Moreover, using a field map to correct for distortions in a single distorted EPI image would not be able to properly correct local areas in which the information of several voxels have been mapped into a single voxel (signal compression), where it would lead to a loss of detail in the distortion-corrected image [59].

To overcome these drawbacks, alternative methods for field map estimation and distortion correction have been proposed. The idea of one such technique proposed for DW-MRI in the brain, is to repeatedly acquire DW-MRI data with reversed phase-encode direction (RPED), such as to obtain pairs of images with equal magnitude distortions in opposing directions [59, 101]. From one of these image pairs, preferably from the zero or low b-value image pair which has higher SNR, the underlying magnetic field map can be estimated using an image registration approach and applied to the sets of repeatedly acquired b-value images to obtain unwarped, distortion-reduced data. The process is illustrated in Figure 3.8. The principle of estimating the field map reverts to finding the transformation field, that, when applied to the two images, would maximize the similarity of the unwarped images. This approach is a parametrized approach based on the representation of the unknown transform by cosine basis functions and based on minimizing a similarity metric in the form of the sum-of-squared differences between the unwarped images. Application of the final transform to the pair of images is based on least-squares restoration.

The advantage of RPED-based correction compared to classic field map based correction of a single EPI image is, that information loss in areas of signal compression in one image can be compensated by information from corresponding areas of signal expansion in the other image acquired with inverse phase encoding direc-

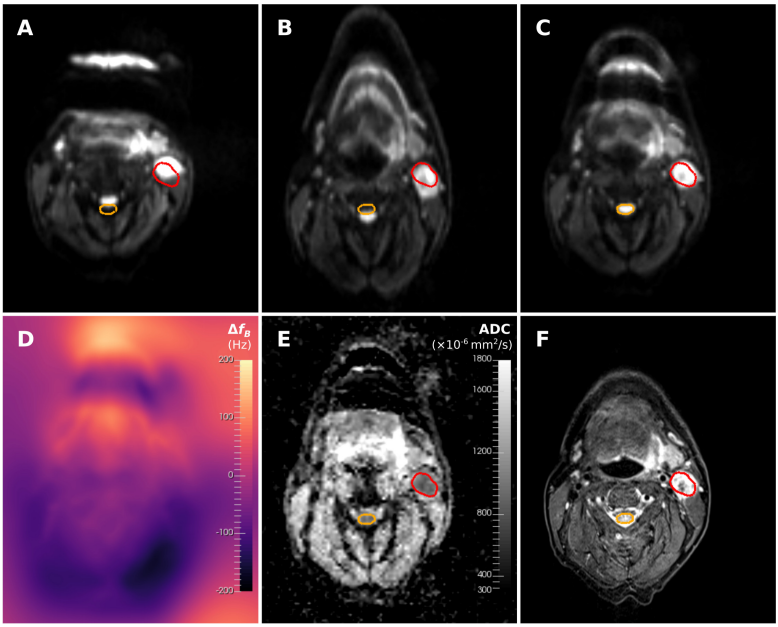

Figure 3.8: Illustration of distortion and correction for DW-MRI data of a 55-year-old male patient with head and neck cancer examined with combined ^{18}F-FDG PET/MRI. Original DW-MR image (b=150 s/mm^2) with phase-encoding direction anteroposterior (A) and posteroanterior (B) acquired in axial orientation. C: Distortion corrected image with RPED technique. D: RPED technique based field map displaying B_0 field inhomogeneity in terms of frequency shift. E: Distortion corrected ADC map. F: T2w STIR (short-TI inversion recovery) as geometric reference image. Two structures of interest are shown: a diseased lymph node (red, PET-defined) and normal spinal cord (orange, delineated on the T2w STIR). These structures are superimposed on DW-MR images and the ADC map to highlight geometric (in-)accuracy. Note a minor artifact remaining in the chin region.

tion [102]. This advantage of the method, however, can only be fully exploited, if all b-value images are acquired repeatedly with RPED.

A recent study has shown that the RPED correction technique outperforms different methods including the classic field mapping approach based on repeated gradient echo imaging with different echo times. This is mainly due to its additional advantage of lost information recovery in regions of signal compression [102].

Since high geometric fidelity is essential for RT purposes, RPED technique was implemented in this work for distortion correction of DW-MRI. It was evaluated whether this method presented for neuroimaging could be translated to combined PET/MRI data acquisition with RT imaging setup in patients with head-and-neck cancer.

3.3 Measures for image quality assessment

For different analyses performed in this work, the geometrical or quantitative accuracy of images was assessed by image comparison to a reference image. This approach was central to the comparison between distortion corrected DW-MR and anatomical MR reference images (discussed above), as well as to the comparison between attenuation corrected PET images acquired with RT and diagnostic setup, an issue addressed earlier in Section 3.1.4.

Methods for image quality assessment included pairwise comparison of radiooncological structures of interest that were defined as binary masks on two images. This strategy was based on the assumption, that images acquired with experimental and reference technique can be considered as equivalent, if manual or automatic, threshold-based contouring on the images yields structures with similar shape, size and location. Therefore, the following volume and distance based measures were used to assess the similarity of two structures, or voxel sets, A and B:

- Dice similarity index (DSI), a frequently used evaluation measure in medical image analysis. The measure quantifies the match of two sets A and B by normalizing the size of their intersection to the average of their sizes.

$$\mathrm{DSI}(A,B) = \frac{|A \cap B|}{\frac{1}{2}\left(|A| + |B|\right)} \tag{3.21}$$

Thus, DSI = 1 for a perfect match and DSI = 0 for a total mismatch.

- Relative volume difference (RVD) of a set A compared to a reference set B:

$$\mathrm{RVD}(A,B) = \frac{|A| - |B|}{|B|} \tag{3.22}$$

- Average symmetric surface distance (ASSD), the average of all the distances from points x on the boundary of A (L_A) to the boundary of B (L_B), and from points y on L_B to L_A, respectively:

$$\mathrm{ASSD}(A,B) = \frac{1}{|L_A| + |L_B|} \left(\sum_{x \in L_A} d(x, L_B) + \sum_{y \in L_B} d(y, L_A) \right) \tag{3.23}$$

  where the distance measure for a point x from a set of points L is defined as $d(x,L) = \min_{y \in L} d(x,y)$, with $d(x,y)$ the Euclidean distance between the points.
- Distance between geometric centers (DOGC) of two structures A and B:

$$\mathrm{DOGC}(A,B) = d(c_A, c_B) \tag{3.24}$$

  with $c_A = \frac{1}{|A|} \sum_{x \in A} x$, the central voxel of A, and $c_B = \frac{1}{|B|} \sum_{y \in B} y$, the central voxel of B.

3.4 Machine learning techniques

In the context of the study on predictive modeling of FMISO-PET derived hypoxia information by MRI derived image parameters, different machine learning techniques were implemented in the analysis. Machine learning refers to a wide-ranging set of techniques for understanding relationships or structure from data. Such techniques can be grouped into two categories: unsupervised and supervised learning techniques.

Unsupervised learning is a type of learning that can be beneficial in finding unknown patterns in data observations without supervising output or pre-assignment of data labels. Typical methods for unsupervised learning include principal component and cluster analysis (**Sections 3.4.1, 3.4.2**). These techniques were used to

extract a reduced set of parameters describing high-dimensional time-series image data, and to group similar time series data together in order to identify characteristic patterns.

Supervised learning techniques, on the other hand, describe learning strategies that involve building a model for predicting or estimating an output based on observed input parameters. The model is build by inferring a function from exemplary input-output paired data (labeled training data), which maps the input to the known output. Ideally, a well trained model is then able to generalize and predict the output of new, previously unseen data (independent test data). As types of supervised techniques, gradient boosted decision trees (GBDT, **3.4.3**) and artificial neural networks (ANN, **3.4.4**) played a central role in this work to find a prediction model that allows for estimating the hypoxia PET information from different MRI derived image parameters.

3.4.1 Principal component analysis

Principal component analysis (PCA) was used to derive parameter maps from time-series image data at a voxel level. PCA is a data-driven technique based on transforming data in terms of observations of a set of potentially correlated variables, into a set of values of linearly uncorrelated variables. Its design makes it possible to subsequently use the method for dimensionality reduction or noise filtering of the data.

PCA transforms the data from the original variable space to a new coordinate system whose coordinates are found in an iterative approach:

1. The first coordinate is chosen as the one when the variance of the data projections on the coordinate is greatest among all possible coordinate choices;
2. Similarly, the subsequent coordinate is defined by producing greatest variance of data projections on a new coordinate orthogonal to the previous one. This step is repeated until the number of new coordinates corresponds to the number of original variables.

This way, the new coordinates, the principal components (PCs), form a new set of basis vectors to represent the data. For full representation, the number of PCs has to be equal to the number of original variables. But in general, most of the variation in the data will be accounted for by a lower number of PCs [103].

In the case of PCA applied to time-series signal data from image voxels, the total number of PCs generated corresponds to the number of original time points. Let p denote the total number of time points or PCs, respectively, and $\boldsymbol{s}_i = (s_{i1}, \ldots s_{ip})$, the vector representing the signal time-series data of the ith voxel at the various p time-points. Let further $\mathbf{pc}_j = (\mathrm{pc}_{1j}, \ldots \mathrm{pc}_{pj})^{\mathrm{T}}$, with $\|\mathbf{pc}_j\| = 1$, denote the jth principal component, with T the transpose. Then these new basis vectors map each data vector $\boldsymbol{s}$ to a new vector of principal component projection coefficients $(c_{i1}, \ldots c_{ip})$ with

$$c_{ij} = \boldsymbol{s}_i \cdot \mathbf{pc}_j \quad \text{for } j = 1, \ldots p \tag{3.25}$$

with c_{ij}, the projection coefficient corresponding to the jth principal component.

The total set of projection coefficients, that is $n \times p$ matrix $\mathbf{C} = (c_{ij})$, with n, the total number of data (voxels), is a unique and full representation of the data in principal component space. This allows for the signal data $\boldsymbol{s}_i$ of a voxel to be reconstructed as linear combination of the PCs:

$$\boldsymbol{s}_i = \sum_{j=1}^{p} c_{ij} \mathbf{pc}_j^{\mathrm{T}} \tag{3.26}$$

with vector $\mathbf{pc}_j^{\mathrm{T}}$ the transpose of $\mathbf{pc}_j$. It is a main characteristic of the method that the data variance captured by the PCs successively decreases from the first to the last PC. Therefore, typically the linear combination can be truncated at the first PCs such that dimensionality is reduced and the data filtered. The process of PCA for analysis of dynamic image data is presented in Fig. 3.9.

3.4.2 K-means clustering

K-means clustering is an iterative algorithm designed to partition a number of n observations into a chosen number of k clusters, with each observation being assigned to the cluster with the nearest center. The aim is to find a data partitioning that minimizes the sum of the whithin-cluster, sum-of-squares point-to-cluster-center distances:

$$\arg\min_{C} \sum_{i=1}^{k} \sum_{x \in C_i} d^2(x, c_i) \tag{3.27}$$

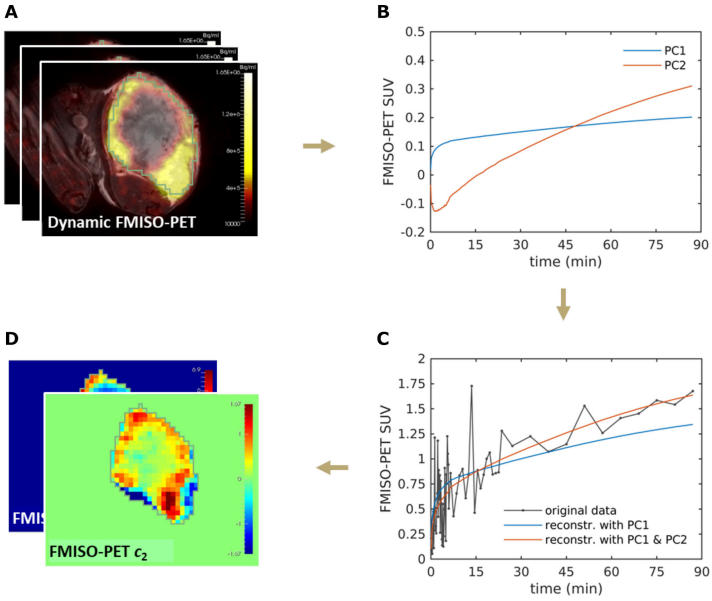

Figure 3.9: Illustration of PCA performed on FMISO-PET time-series data (dynamic data) acquired in xenografted head-and-neck tumors. A: Dynamic FMISO-PET data for an exemplary tumor. B: First two principal components (PC1 and PC2) obtained by PCA on normalized FMISO-PET data; the results refer to the analysis performed on the total of image voxels of a number of tumors. C: Measured dynamic FMISO-PET data for an exemplary tumor voxel (black data) shown together with the data reconstructed using (3.26) truncated at $p = 1$ (blue data) and $p = 2$ (red data), respectively. D: Parameter maps of the projection coefficients $\boldsymbol{c}_{j=1}$ and $\boldsymbol{c}_{j=2}$ for the same tumor as presented in A.

where x is a data point from a set of n observations $\{x_1, x_2, \dots x_n\}$, $C = \{C_1, C_2, \dots C_k\}$ the total set of k clusters, $c_i = \frac{1}{|C_i|} \sum_{x \in C_i} x$, the center (mean) of cluster C_i and $d^2(x, c_i)$ the squared Euclidean distance between point x and cluster center c_i.

Clusters satisfying (3.27) are found by an iterative approach [104]:

1. Initialization: to randomly choose k points as initial cluster center locations;
2. Assignment: to calculate point-to-cluster-center distances of all observations to each cluster center and to assign each observation to the cluster with the nearest center;
3. Update: to calculate the mean of the observations in each cluster to obtain k new center locations.
4. Iteration: to repeat steps 2–3 until assignment of data points to clusters does not change or a predefined maximum number of iterations is reached.

The k-means algorithm can converge to a local minimum. However, multiple replicates can be performed with different starting points. Thus, in practice the solution yielding the lowest minimum (lowest total sum of distances among all replicates) is chosen as final result.

3.4.3 Gradient boosted decision trees

In the third study presented in this work, the prediction problem was formulated as a classification problem of a qualitative, discrete response variable (the hypoxia class type) and different MRI derived image parameters served as predictor variables. To address the prediction problem, the supervised learning technique gradient boosted decision trees (GBDT) was chosen as a candidate model for its capability to approximate complicated nonlinear functions and to handle a multi-class response variable.

GBDT belongs to the family of decision tree ensemble based learners. A decision tree is based on segmenting the space of all predictor variable values into a number of non-overlapping subregions, and assign a constant or a label to each subregion. Let the subregions be denoted R_j, $j = 1, \dots J$ with γ_i the assigned constant. The predictive rule for a tree model T and an observation x would then be [105]:

$$x \in R_j \rightarrow T(x) = \gamma_j \,. \tag{3.28}$$

For a regression tree, for instance, the constant γ_i assigned to the subregion could be the mean response associated the observations contained in the subregion; for a classification tree, it could be the class mode, i.e., the most commonly occurring class of training observations in the respective subregion.

Now let some training data consist of p predictor variables and a response for each of N observations, i.e., (x_i, y_i) for $i = 1, \dots N$, with $x_i = (x_{i1}, \dots x_{ip})$. Then a decision tree can be more generally expressed as:

$$T(x; \Theta) = \sum_{j=1}^{J} \gamma_j I(x \in R_j) \tag{3.29}$$

with $I(x \in R_j)$, a binary indicator function taking on the value 1 for $x \in R_j$ and 0, otherwise; and $\Theta = \{R_j, \gamma_j\}_1^J$ the set of parameters defining the tree [105].

To build (or grow) a decision tree relates to finding the optimal R_j and γ_j which define the partitioning of the data space. This is thus an optimization problem with respect to Θ, as Θ summarizes these parameters:

$$\hat{\Theta} = \arg\min_{\Theta} \sum_{j=1}^{J} \sum_{x_i \in R_j} L(y_i, \gamma_j)\,. \tag{3.30}$$

To approximately solve this optimization problem typically a top-down recursive splitting algorithm is used, so as to find the best splits that result in the best R_j [105]. As a criterion to evaluate the quality of a split, some loss function $L(y_i, \gamma_j)$ is used. For regression, typically minimization of the sum of squares, $\sum(y_i - \gamma_j)^2$, serves as loss function which results in $\bar{y_j}$, the mean of the y_i in region R_j, as the best estimation of γ_j. The loss function takes on small values if the variance of y_i in the regions is low. It is thus a measure for purity of the subregions (node purity). High node purity is preferable as it indicates that the predicted response for an observation is likely to be close to its true value.

For classification, where the response is a discrete class label k of classes $k = 1 \dots K$, a different loss function is needed as criterion, such as cross entropy or log loss:

$$D(y_i, \hat{p}_{jk}) = -\sum_{k=1}^{K} \hat{p}_{jk} \log \hat{p}_{jk} \tag{3.31}$$

with $\hat{p}_{jk} = \frac{1}{N_j} \sum_{x_i \in R_j} I(y_i = k)$ the fraction of training observations in the jth region that belong to the kth class; and N_j the number of training observations in R_j [106]. Function D takes on small values for $\hat{p}_{jk}$ close to either zero, or one. This is the case when regions mainly contain observations belonging to the same class [107]. D is thus a measure for purity of the subregions. Here, high purity (small D) is preferable as it indicates a high probability of predicting the majority class in the region; it is thus a good loss function to minimize in order to obtain good predicitive performance for a classification tree.

Similar to other tree ensemble based learners, like random forest, GBDT involves combining multiple decision trees into a single model. This way, a higher prediction accuracy and generalization capacity can be achieved as compared to a single tree model. In contrast to random forest, GBDT is an additive model and trees are grown sequentially, such that each tree uses information from the previously grown trees to step by step slowly minimize the prediction error. Moreover, GBDT involves fitting the trees to modified versions of the input data instead of fitting the input data directly.

A boosted tree model can basically be described as a sum of trees (base learners) [105]:

$$f_M(x) = \sum_{m=1}^{M} T(x; \Theta_m) \tag{3.32}$$

which are built in a forward stagewise procedure, such that at step m, a tree T is added to the current model:

$$f_m(x) = f_{m-1}(x) + T(x; \Theta_m)\,. \tag{3.33}$$

This reverts to finding the tree T, parametrized by Θ_m, that best possibly improves f_{m-1} with respect to some loss function to provide a better model f_m that is closer to y. This means that, at each step m one has to solve:

$$\hat{\Theta}_m = \arg\min_{\Theta_m} \sum_{i=1}^{N} L(y_i, f_{m-1}(x_i) + T(x_i; \Theta_m))\,. \tag{3.34}$$

The solution tree is the one that maximally reduces (3.34). It would thus be the tree that corresponds to the negative gradient of the loss function L. If the loss function was squared error loss, i.e., $L(y_i, f_{m-1}) = \frac{1}{2}(y_i - f_{m-1}(x_i))^2$, then the gradient would be $y_i - f_{m-1}(x_i)$, which are the current residuals of the previous model f_{m-1}, and the solution tree could be approximated by a regression tree that best predicts these residuals. The resulting constants in each region R_{jm} would just be the mean of the residuals in that region. In other words, instead of fitting the actual data, the trees are fitted to the residuals resulting from the prediction of the previous model.

In general, an approximate solution for (3.34) is given by fitting a new tree at each stage m to the negative gradient of the respective loss function with respect to the predictions from the previous model (gradient boosting approach). In GBDT classification, typically multinomial deviance is used as loss function and, in order to solve (3.34), regression trees are fitted to the pseudo-residuals obtained by the negative gradient of the deviance function, such that the final output of f_M are probabilities for observations to belong to a class k.

Similar to other machine learning methods, different hyper-parameters can be tuned in a cross validation setting where the training data is repeatedly split into subpartitions used for repeated training and testing of different versions of a model type. For instance, boosting-related hyper-parameterscan be tuned, such as the number of iterations, i.e., the number of trees M. A different category of hyper-parameters is tree shape related, and influence the subregios R_{jm} and J_m (the number of subregions R_{jm}) of the mth tree. An example is setting a maximum limit of observations N_J to be allowed in a subregion.

Both boosting and tree related hyper-parameters can be tuned and tested in a cross validation setting to find an improved version of the model with increased predictive performance and better generalization capacity. To assess the predictive performance of the model, score functions like area-under-the-curve of the receiver operating characteristic (ROC-AUC) can be used.

3.4.4 Artificial neural networks

As an alternative model to GBDT, an artificial neural network (ANN) was chosen to approach the prediction problem in Part III. Similar to GBDT, it is a supervised

learning technique capable of approximating complicated nonlinear functions and to handle a multiclass response variable.

Different types of ANN exist. Here, a multilayer perceptron was used, a type of ANN consisting of at least three layers (input, hidden and output layer) of connected units (nodes, or neurons). Each unit of a layer is connected to every unit of the next layer such that an ANN can be graphically represented as a network connecting the input data X in the input layer via one or multiple hidden layers to the output layer (Fig. 3.10). For a single hidden layer network and for the case of multiclass classification of K classes, the output layer consists of K units with the kth unit modeling the probability of class k. The response variable Y_k is coded as 0 – 1 for the kth class. The hidden layer consists of M units, where the mth unit, Z_m, is computed by applying a non-linear activation function to a linear combination of the input X:

$$Z_m = \sigma(\alpha_{0m} + \alpha_m X),\ m = 1, \dots M\,, \tag{3.35}$$

with σ the activation function, e.g. a sigmoid function, α_{0m} an intercept parameter, and α_m the weights from the input units to the mth hidden unit [108]. The hidden units owe their name to the fact that the Z_m they are creating, are not directly observed but further processed. That is, the response Y_k is modeled as a function of linear combinations of the Z_m:

$$f_k(X) = g_k(H),\ k = 1, \dots K, \tag{3.36}$$

with $H = (H_1, \dots H_K)$ and

$$H_k = \beta_{0k} + \beta_k Z,\ k = 1, \dots K\,, \tag{3.37}$$

where $Z = (Z_1, \dots Z_M)$, and β_{m0}, β_m are further intercept and weight parameters [108]. g_k is typically chosen as a function like the softmax function, so as to obtain class-membership probabilities.

During training of the ANN, the set of model parameters $\{\alpha_{m0}, \alpha_m, \beta_{m0}, \beta_m\}$ is optimized with respect to a loss function. As a loss function for classification, typically the cross-entropy function is chosen, similar to (3.31) used for GBDT. As optimization algorithm, gradient descent is normally chosen. The gradient of the

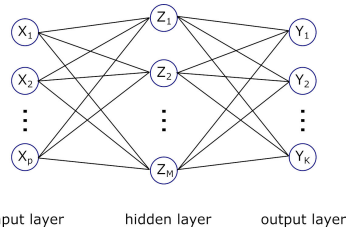

Figure 3.10: Network diagram for a single hidden layer neural network (ANN) used for classification. Each unit in the hidden layer transforms the values from the previous layer (here the input layer) with a linear transform, followed by a non-linear activation function as described by (3.35). The output layer finally receives the values from the previous hidden layer to transform them into some output values, e.g., class-membership probabilities, as described by (3.36) and (3.37).

loss function with respect to the model parameters can be derived using the chain rule, which translates into the backprojection algorithm. Using backprojection, the model parameters are iteratively updated to improve the network's predictive performance.

Similar to GBDT and other machine learning methods, different hyper-parameters can be tuned in a cross validation setting to test different versions of a model type. For ANN, for instance, the hidden layer size and number of layers can be increased to increase the flexibility of the model to closer fit the model to the training data. To avoid overfitting, the model can further be regularized e.g., by penalizing weights with large magnitudes.

4 Part I: Assessment of image quality of a radiotherapy-specific hardware solution for PET/MRI in head and neck cancer patients

René M. Winter[1*], Sara Leibfarth[1], Holger Schmidt[2], Kerstin Zwirner[3], David Mönnich[1,5], Stefan Welz[3], Nina F. Schwenzer[2], Christian la Fougère[4,5], Konstantin Nikolaou[2,5], Sergios Gatidis[2], Daniel Zips[3,5], Daniela Thorwarth[1,5]

[1] Section for Biomedical Physics, Department of Radiation Oncology, University Hospital and Medical Faculty, Eberhard Karls University Tübingen, Tübingen, Germany
[2] Department of Diagnostic and Interventional Radiology, University Hospital and Medical Faculty, Eberhard Karls University Tübingen, Tübingen, Germany
[3] Department of Radiation Oncology, University Hospital and Medical Faculty, Eberhard Karls University Tübingen, Tübingen, Germany
[4] Department of Nuclear Medicine, University Hospital and Medical Faculty, Eberhard Karls University Tübingen, Tübingen, Germany
[5] German Cancer Consortium (DKTK), partner site Tübingen; and German Cancer Research Center (DKFZ), Heidelberg, Germany
[*] Corresponding author at: University Hospital Tübingen, Department of Radiation Oncology, Section for Biomedical Physics, Hoppe-Seyler-Str. 3, 72076 Tübingen, Germany. E-mail address: rene.winter@med.uni-tuebingen.de

published in
Radiotherapy and Oncology 2018,**128**:485–491.

Abstract

Background and purpose: Functional PET/MRI has great potential to improve radiotherapy planning (RTP). However, data integration requires imaging with radiotherapy-specific patient positioning. Here, we investigated the feasibility and image quality of radiotherapy-customized PET/MRI in head-and-neck cancer (HNC) patients using a dedicated hardware setup.

Material and Methods: Ten HNC patients were examined with simultaneous PET/MRI before treatment, with radiotherapy and diagnostic scan setup, respectively. We tested feasibility of radiotherapyspecific patient positioning and compared the image quality between both setups by pairwise image analysis of ^{18}F-FDG-PET, T1/T2-weighted and diffusion-weighted MRI. For image quality assessment, similarity measures including average symmetric surface distance (ASSD) of PET and MR-based tumor contours, MR signal-to-noise ratio (SNR) and mean apparent diffusion coefficient (ADC) value were used.

Results: PET/MRI in radiotherapy position was feasible – all patients were successfully examined. ASSD (median/range) of PET and MR contours was 0.6 (0.4 — 1.2) and 0.9 (0.5 -– 1.3) mm, respectively. For T2-weighted MRI, a reduced SNR of −26.2% (−39.0 – −11.7%) was observed with radiotherapy setup. No significant difference in mean ADC was found.

Conclusion: Simultaneous PET/MRI in HNC patients using radiotherapy positioning aids is clinically feasible. Though SNR was reduced, the image quality obtained with a radiotherapy setup meets RTP requirements and the data can thus be used for personalized RTP.

4.1 Introduction

Within the scope of radiotherapy (RT) planning, imaging information is primarily used for precise delineation of target volumes and calculation of an optimal radiation dose distribution. In current clinical practice, these two steps of the RT workflow are most commonly based on computed tomography (CT) data, since CT offers high geometric fidelity and provides the electron density information of the tissue [109]. However, in the current era of precision radiation oncology [110], the additional integration of data from different imaging modalities such as (functional) magnetic resonance imaging (MRI) or positron emission tomography (PET) has great potential to improve and individualize RT planning [23–25]. First, T2- or T1-weighted (contrast-enhanced) MRI shows superior soft tissue contrast as compared to CT. Therefore, it may provide more precise information on tumor localization and spread and thus increase the delineation accuracy of target volumes [19,20,111]. In addition, functional information assessed with PET [16, 17], diffusion-weighted (DW) MRI [11, 12, 31] or dynamic contrast-enhanced (DCE) MRI [13, 14, 32] has been associated with RT outcome in different tumor sites including head-and-neck cancer (HNC). Its integration may thus be a promising strategy for adapting RT planning for patients individually [26, 27, 112]. The combination of PET and MRI as a hybrid system now offers spatially and temporally coregistered anatomical and functional image data and may therefore become a key technology for individual therapy adaptation [7, 23, 34, 35, 113].

The integration of combined PET/MRI into RT, however, requires patient examination in RT position for accurate image alignment with the RT planning CT [19,36,37,114,115]. Especially for HNC, it has been shown that the accuracy of rigid or deformable registration algorithms is strongly improved when patient images are acquired in RT position for both CT and MRI [114, 115]. The adaptation to treatment position, on the other hand, is challenging as it requires RT immobilization equipment to be combined with the MRI hardware, in particular with the radiofrequency coils used for signal reception. For stand-alone MRI systems, dedicated RT coil setups have been presented [116, 117]. However, these setups cannot readily be transferred to a combined PET/MRI system since they do not meet the demands of hardware PET attenuation correction, i.e., foremost, a fixed and reproducible positioning of each hardware device for the usage of predefined

attenuation maps [38]. For combined PET/MRI of HNC, an initial RT-specific solution has recently been proposed by Paulus et al. [39]. It comprises a flat table top and MR coil holders for flexible body coils.

In the present study, we upgraded the initial setup with a dedicated add-on, designed and manufactured in-house, for the use of a RT mask fixation system. Besides feasibility assessment of patient imaging in RT treatment position, the aim of the study was to systematically evaluate the image quality of the customized setup in a clinical setting in comparison with a diagnostic setup. Our hypotheses were, that (i) following attenuation correction, good agreement between PET data with RT and diagnostic setup is achievable, that (ii) MR image quality with RT setup is inferior but still sufficient for RT planning applications, and that (iii) potentially reduced MR image quality does not adversely affect the stability of DW-MRI in terms of the mean apparent diffusion coefficient (ADC).

4.2 Material and Methods

Study design and imaging protocol

During the pilot phase of a prospective clinical trial (NCT -02666885), ten patients with loco-regionally advanced head-and-neck squamous cell-carcinoma (HNSCC) of the oro- or hypopharynx were examined before the start of multimodal treatment (surgery and adjuvant RT) with simultaneous PET/MRI (Biograph mMR, Siemens Healthcare GmbH, Erlangen, Germany). The imaging protocol included ^{18}F-FDG PET, T2-weighted (T2w) MRI using turbo spin echo (TSE) technique, T1w MRI after contrast agent administration (gadolinium) using volumetric interpolated breath-hold examination (VIBE) technique and DW-MRI (b = 150 and 800 s/mm^2). Further protocol details are given in Table 4.1. PET and MR sequence parameters are listed in Supplementary Tables 4.2 and 4.3, respectively. Following intravenous injection of ^{18}F-FDG, two consecutive scans were performed for each patient with RT-specific and diagnostic setup. The RT setup consisted of a flat MR table top and a pair of C-shaped coil holders (Qfix, Avondale, PA, USA) for 6-channel flexible body matrix coils, as introduced in [39]. In addition, an in-house designed add-on was mounted onto the MR table top which allows for patient fixation with a thermoplastic RT mask (ITV, Innsbruck, Austria). The

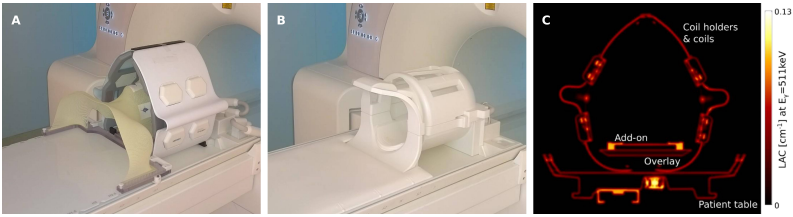

Figure 4.1: Coil setups for RT specific (A) and diagnostic imaging (B) of the head-and-neck on the Biograph mMR. Setup A consists of a flat table top, C-shaped coil holders and an add-on for patient fixation with a thermoplastic mask. For setup A, a μ-map of the hardware components (C) was generated and used for attenuation correction during PET image reconstruction.

diagnostic setup consisted of the state-of-the-art 16-channel head-and-neck coil of the mMR system. Both hardware setups are depicted in Fig. 4.1. Hearing protection was ensured using earplugs. Scan limits were infraclavicular level to skull base. For both setups, the first element of the spine array coil was activated to improve signal acquisition in the neck region.

Data processing

Since objects located within the field of view of the PET detector may lead to attenuation and scattering of the PET photons before signal detection, reliable quantification requires data correction. Therefore, attenuation correction was performed during PET data reconstruction using attenuation maps (μ-maps) of both patient and hardware. The individual human μ-map was acquired based on a Dixon fat-water separation technique [84]. For the RT scan, a hardware μ-map was used that was created from CT images of the hardware components by bilinear transformation of CT Hounsfield units into linear attenuation coefficients (LAC) at the characteristic PET photon energy level of $E_\gamma = 511\,\text{keV}$, as described in [39, 76]. For the diagnostic scan, default vendor-supplied attenuation correction of the hardware was applied.

DW-MRI was performed using echo-planar imaging (EPI) technique. Since EPI is sensitive to B0-field inhomogeneities and susceptibility changes especially in the head-and-neck region which can lead to image distortions and signal loss, a method

Table 4.1: Patient characteristics and imaging protocol.

Patient	Tumor site	No. of lesion ROIs	Att. corr. FDG-PET	MR-based μ-map	T2w MRI (TSE)	T1w MRI (VIBE)	Diffusion-weighted MRI				
							b=150 (AP)	b=150 (PA)	b=800 (AP)	b=800 (PA)	ADC d.c.
#01	OP	2	✓	✓	(✓)[c]	✓	✓	✓	(✓)[b]	(✓)[b]	(✓)[b]
#02	OP/HP/VAL	3[a]	✓	✓	✓	✓	✓	✓	✓	✓	✓
#03	HP	2	✓	✓	✓	✓	✓	✓	✓	✓	✓
#04	OP	2[a]	✓	✓	✓	✓	✓	✓	✓	✓	✓
#05	OP	1	✓	✓	✓	✓	✓	✓	✓	✓	✓
#06	OP/MC	2	✓	✓	✓	✓	✓	✓	✓	✓	✓
#07	BT/UV	2	✓	✓	✓	✓	✓	✓	✓	✓	✓
#08	HP	2	✓	✓	✓	✓	✓	✓	✓	✓	✓
#09	OP	1	✓	✓	✓	✓	✓	✓	✓	✓	✓
#10	HP	1	✓	✓	✓	✓	(✓)[c]	(✓)[c]	(✓)[c]	(✓)[c]	(✓)[c]

Abbreviations: OP – oropharynx; HP – hypopharynx; VAL – vallecula; MC – mouth cavity; BT – base of tongue; UV – uvula; ROI – region of interest; att. corr. – attenuation corrected; TSE – turbo spin echo; VIBE – volume interpolated breath hold; AP – anteroposterior; PA – posteroanterior; d.c. – distortion corrected.

[a] One FDG avid ROI was no RT target; one MR ROI was only weakly PET positive and not considered in PET analysis.

[b] Error in MR sequence parameter settings for scan with diagnostic setup.

[c] Strong geometric distortions in DW-MRI with RT setup.

for distortion correction was applied based on repeated data collection with reversed phase- encoding directions (RPED) as described in [118].

Image analysis

The evaluation of PET and MR image quality was performed patient-by-patient by systematic image comparison between RT and diagnostic setup. PET images were compared by estimating the similarity of corresponding regions of interest (ROI). To better account for the difference in positioning between the two examinations, rigid image registration was applied locally with a binary mask using elastix [119]. ROIs were defined in primary tumors and FDG-positive lymph nodes by creating operator-independent 3D threshold contours that comprised voxels with PET activity concentrations greater than or equal to 50% of the local maximum [120]. For pairwise ROI comparison ($n = 17$), the following similarity measures were calculated in MATLAB R2017b (The MathWorks, Inc., Natick, Massachusetts, United States): Dice similarity index (DSI), relative volume difference (RVD), average symmetric surface distance (ASSD) and Euclidean distance of geometric centers (DOGC).

To evaluate the impact of the RT-tailored coil configuration on MR-based PET attenuation correction, human μ-maps were compared. Rigid registration was applied to the pair of μ-maps in two steps, i.e., for head and neck separately. Nearest neighbor interpolation was chosen for final resampling to preserve discrete μ-map values. For each one of the four tissue classes present in the RT μ-map, its relative fraction and the corresponding mean attenuation coefficient in the reference μ-map were determined.

For MR image quality assessment signal- and contrast-to-noise ratios (SNR, CNR) were calculated. In T2w and distortion corrected DW-MR images, SNR was determined in four anatomical ROIs defined manually in the submandibular glands (left, right) and spinal cord at positions C1–2 and C4–5. In addition, SNR of lesion and CNR of lesion versus adjacent tissue were determined for T2w MRI based on the PET-derived ROIs. Image noise was estimated as the standard deviation (SD) of the signal intensity in a background region [121]. For T1w MRI direct quantitative comparison was not feasible as images were acquired at different times after a single contrast agent administration. To further investigate if MR image

quality with RT setup would allow for accurate delineation, RT target structures were contoured manually by a radiation oncologist in training and a board-certified radiologist (KZ, SG) on MR images from both scans using information of T2w and contrast-enhanced T1w MRI. Rigid image registration allowed for the calculation of similarity measures between the ROI pairs (DSI, RVD, ASSD and DOGC).

To assess the stability of DW-MRI, ADC maps were derived from the distortion corrected *b*-value images. Mean ADC values were compared between RT and diagnostic setups in the lesions based on the PET-derived ROIs. A variability or repeatability coefficient was calculated as the SD of ADC percentage change multiplied by 1.96 [122].

To assess the difference (i) in SNR and CNR in T2w and DW-MRI and (ii) in ADC values between the scan setups, statistical analysis was performed using a Wilcoxon signed rank test (MATLAB R2017b). In either case, a *p*-value below .05 was considered statistically significant.

4.3 Results

All patients were successfully examined with simultaneous ^{18}F-FDG PET, anatomical and DW-MRI in diagnostic as well as in RT setup using a dedicated hardware solution for RT-specific patient positioning. However, one DW-MRI dataset with RT setup presented strong distortion artifacts, probably due to patient swallowing, which could not be corrected using RPED; one DW-MRI dataset with diagnostic setup was incomplete due to wrong protocol settings.

Fig. 4.2 shows an example for high ROI agreement between RT and diagnostic PET scan. Relative to the measurement with diagnostic setup, the analysis of ROI similarity yielded a cohort median DSI of 0.88 (range: 0.69–0.94) and RVD of −1% (−40-–24%). Similarly, median ASSD and DOGC were found to be 0.6 mm (0.4–1.2 mm) and 0.9 mm (0.4–3.8 mm), respectively.

Within regions where the human μ-map with RT setup identified soft tissue, fat, the intermediate class between soft tissue and fat, and air with respective LACs of 1000, 854, 927 and $0 \times 10^{-4}\,\mathrm{cm}^{-1}$, the following median (range) values were found in the μ-map with diagnostic setup: 989 (898–996), 865 (812–923), 935 (888–976) and 780 (50–993) $\times 10^{-4}\,\mathrm{cm}^{-1}$, respectively. Mean relative fractions

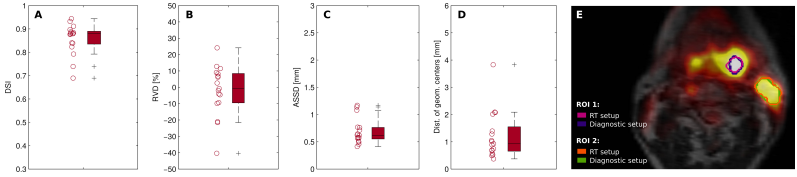

Figure 4.2: Comparison of PET-derived ROIs in FDG-positive tumor and lymph node regions between images acquired with RT and diagnostic scan setup. Boxplots present results of four different similarity measures for all patient ROIs. A–D: Dice similarity index (DSI), relative volume difference (RVD), average symmetric surface distance (ASSD) and the distance of geometric centers (DOGC). E: ROIs in two lesions are shown exemplarily for one patient (#06) on a fused PET/MR image in axial view.

of the four tissue classes in RT and diagnostic μ-map ($\pm$SD) were determined as $77.9 \pm 15.6/75.1 \pm 25.7\%$, $13.7 \pm 13.4/13.9 \pm 19.2\%$, $7.2 \pm 2.7/8.8 \pm 6.7\%$ and $1.2 \pm 1.4/2.3 \pm 1.5\%$, respectively. We refer to Supplementary Fig. 4.6 for data plots and an exemplary μ-map.

Exemplary images of T2w MRI with RT and diagnostic setup are presented in Fig. 4.3. ROI-based analysis of T2w and distortion corrected DW-MRI with b-values of 150 and 800 s/mm^2 resulted in a cohort median difference in SNR of -26.2% (-39.0–-11.7%), -37.9% (-66.7–17.9%) and -31.4% (-65.9–20.9%), respectively. Similarly, a relative difference of SNR in lesion of -32.2% (-39.3–2.8%) and of CNR in lesion versus adjacent tissue of -31.3% (-44.7–10.6%) was found between the coil setups for T2w MRI (Fig. 4.4). Differences in SNR and CNR were found to be significant for both T2w and DW-MRI ($p < .001$).

However, high similarity was found for MR delineated contours. Median DSI and RVD were 0.85 (0.68-–0.89) and 0% (-18—50%), respectively. ASSD and DOGC were 0.9 mm (0.5—1.3 mm) and 1.4 mm (0.3—4.0 mm) (Fig. 4.5).

No significant difference was found between ADC values generated with RT and diagnostic setup. For the lesions, a cohort median difference in ADC of -1.7% (-25.5—24.1%) (p = n.s.) was determined (Supplementary Fig. 4.7). The repeatability coefficient was 17.6%.

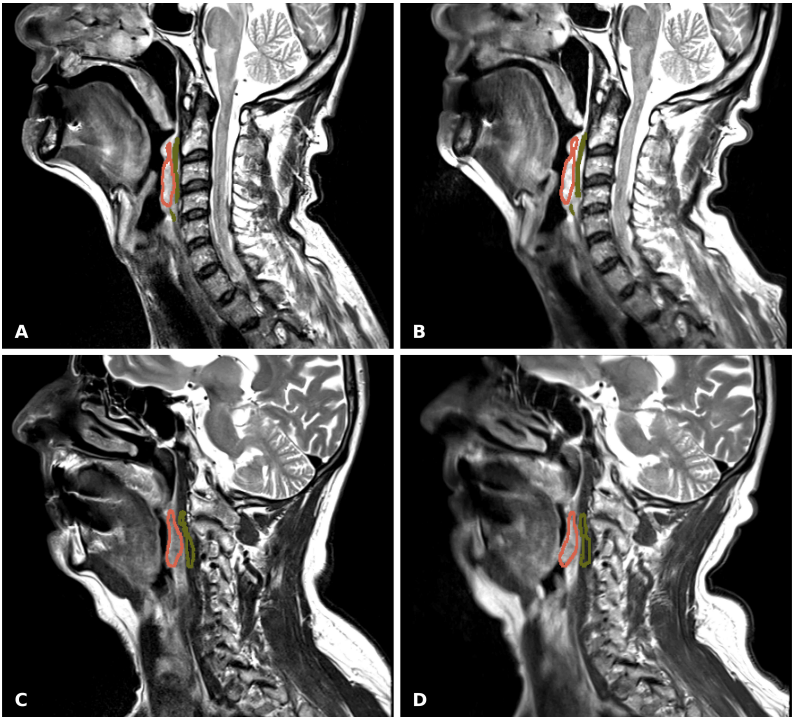

Figure 4.3: T2w MRI using turbo spin echo (TSE) technique of two patients (A, B: #05; C, D: #09) acquired with RT specific (A, C) and diagnostic scan setup (B, D). Note the difference in patient positioning between both scans. ROIs are shown for tumor (red) and adjacent tissue (green). The difference in tumor SNR and CNR in (A) relative to (B) was −21.8% and −21.1%, respectively (good correspondence). Similarly, −30.7% and −31.5% were found for (C) relative to (D) (average correspondence).

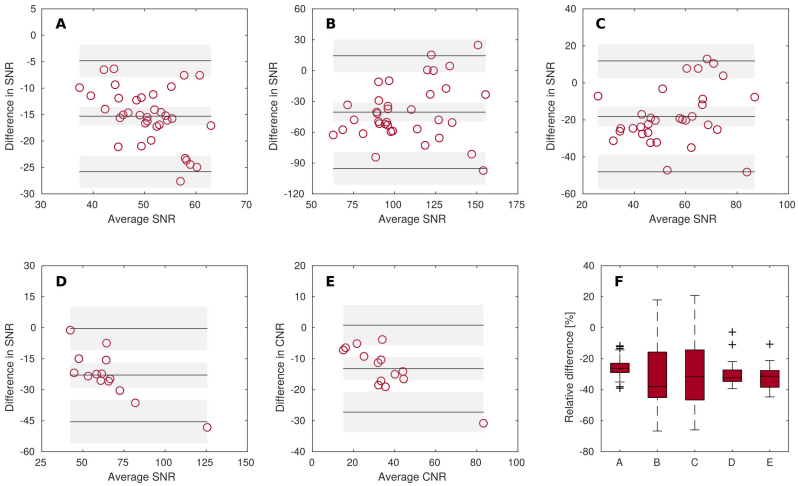

Figure 4.4: Bland Altman plots (A–E) for SNR and CNR for measurements with RT and diagnostic setup. Graphs show difference between two measurements plotted against their average. Solid lines indicate the mean of differences and the limits of agreement (mean $\pm 1.96 \times$ SD). Corresponding confidence intervals are shown as gray-shaded areas. A–C: SNR in anatomical ROIs for T2w and distortion-corrected DW-MRI with b-values of 150 and 800 s/mm^2, respectively. D, E: SNR in lesion and CNR in lesion vs. adjacent tissue for T2w MRI, respectively. F: Boxplot for the relative difference of SNR and CNR between both scan setups for A–E, accordingly.

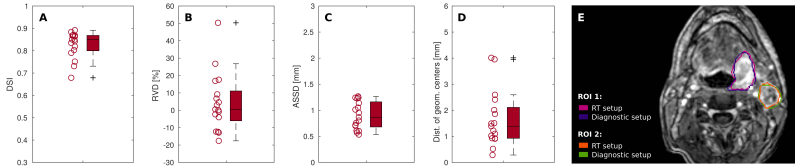

Figure 4.5: Pairwise comparison of MR delineated target structures between RT and diagnostic scans. Boxplots present results of four different similarity measures for all contours. A–D: Dice similarity index (DSI), relative volume difference (RVD), average symmetric surface distance (ASSD) and the distance of geometric centers (DOGC). E: ROIs in two lesions are shown exemplarily for one patient (#06) on a T1w contrast-enhanced MR image in axial view.

4.4 Discussion

In personalized RT, treatment adaptation based on PET/MR information could be of particular relevance. However, data integration requires precise alignment with the RT planning CT. Advanced strategies like deformable image registration are available and may yield reasonable results [123], but multimodal alignment is more precise if both examinations are conducted in RT position, especially for head-and-neck [115]. The purpose of this study was to assess feasibility and image quality of a dedicated hardware solution for PET/MRI in treatment position. Image quality assessment, in particular, was based on the pairwise comparison of FDG-PET and MR contours, MRI-derived SNR and mean ADC values between RT and diagnostic setup.

Image analysis of the first ten patients recruited in this clinical trial demonstrated the clinical feasibility of functional PET/MR examination in RT-specific position using a customized hardware setup and dedicated positioning aids – all patients were successfully examined. Although patients reported that mask fixation felt rather tight, the setup was well tolerated and no examination had to be interrupted or aborted.

Certain components of the RT setup used in this study have been presented earlier [39]. Results of phantom-based PET analysis indicated that correct hardware component attenuation correction is feasible. However, the initial setup did not yet allow for patient examination in actual RT position but needed a modified tabletop to allow for the use of head-and-neck immobilization equipment. In the present study, this add-on was designed and a CT-based attenuation map was generated for attenuation correction.

ROI-based analysis showed that PET images acquired with RT and diagnostic setup could be rated as equivalent with regards to target volume definition, as high volume agreement (high DSI/RVD; low ASSD/DOGC) was found. In particular, the cohort maximum ASSD of 1.17 mm was less than the PET voxel size of 2.8 mm. However, ASSD rather represents a low estimate of residual volume mismatch. Besides that, results may be regarded as conservative estimates as registration uncertainty is included. Direct quantitative PET comparison was not practicable due to variations in physiological tracer uptake between consecutive examinations.

In this study, hardware component μ-maps were used in offline PET data reconstruction toolkit RetroRecon. A method for more automated hardware component attenuation correction has recently been proposed [124] and may further simplify the clinical workflow, in particular if the setup is to be extended to other anatomical regions.

A fat-water separating Dixon sequence was used in this study for generation of human μ-maps. As the flexible coils of the RT setup had a greater distance to the head-and-neck, which comes along with lower SNR, the Dixon-based μ-map was verified toward correct tissue segmentation. Very good agreement of LACs between both scan setups was found except for air segmentation. However this difference is rather negligible since in the head-and-neck region, the fraction of voxels assigned to air is very low (1.2 ± 1.4 vs. $2.3\pm1.5\%$ for RT and diagnostic setup, respectively). Moreover, the discordance in air detection is likely caused by the variation of air pockets in the oral or pharyngeal region between the two examinations, rather than by deficiencies in the μ-map generation with RT coil setup.

The results are relevant not only for correct PET quantification but also for future RT planning based on PET/MRI as the sole modality, because MR Dixon sequences are an attractive approach to generate substitute or pseudo CTs [125]. For this purpose, correct tissue classification is crucial for dosimetric accuracy, but we do not expect significant differences for pseudo CTs between the setups as for Dixon-based μ-maps only minor differences were observed, as discussed above.

SNR was measured in T2w and DW-MRI. Of note is that the method for noise estimation was chosen for simplicity and its frequent use while potentially more accurate methods exist especially when parallel imaging techniques are used. However, such approaches may require e.g. specific sequence modification for additional acquisition of noise only data without radiofrequency pulse excitation [126], multiple repeated image acquisition for pixel-by-pixel noise SD or repeated acquisition for a noise estimate based on pixel-by-pixel difference [127].

Reduced SNR and CNR were observed as compared to diagnostic imaging. Yet, the image quality seems to be sufficient for RT planning applications as good agreement was found between target structures delineated on MRI. The level of agreement should be assessed against the level of variability of repeated MR delineation since manual delineation is open to both inter- and intra-observer variation [128].

Recently, these two types of variation were quantified for two head-and-neck specialists by a mean DSI of 0.80 and 0.86, respectively [129]. Our results seem to be in the same order (DSI = 0.83 ± 0.06 (mean $\pm$ SD)) and thus support the conclusion that MR image quality with RT setup appears suitable for RT planning requirements.

The purpose of a PET/MR examination for RT planning is rather different from a diagnostic one. Besides uniform patient positioning for precise alignment with RT planning CT, isotropic voxel size and geometric accuracy of MRI are essential [7,24]. In this protocol, isotropic voxel size was realized for T1w MRI and geometric accuracy was assessed for DW-MRI and reported earlier [118]. To improve accuracy of EPI-based DW-MRI by correction for B0-field inhomogeneities and susceptibility induced image distortions different techniques have been proposed [41, 130, 131]. Here, RPED technique was used. The level of geometric accuracy of DW-MRI after RPED correction was in the order of 1 mm [118]. Of note is that the method may correct for geometric distortions but cannot accurately account for signal loss or pile-up.

Comparison of ADC values within FDG-avid tumor and lymph node regions yielded no significant difference between both scan setups indicating that the RT setup does not adversely influence quantification accuracy of DW-MRI. However, deviations in ADC of up to 25.5% and a repeatability coefficient of 17.6% were observed. These values may seem large but correspond to baseline ADC variability in patients with HNSCC. Based on repeated measurements with a one-week interval in 16 patients, Hoang et al. have determined deviations of up to 25% and a repeatability coefficient of 15% [122]. Hence, we consider that the variation in ADC between our two measurements does not necessarily arise from the difference in imaging setup but may rather reflect the uncertainty of EPI-based DW-MRI in head-and-neck.

One question is whether the ADC variability will not compromise clinically relevant information. This especially applies to the measurement of baseline and intratreatment ADC changes to predict outcome or monitor early treatment response [132, 133]. Clinically this is appealing as it would allow to opt for alternative treatment strategies for patients with poor prognosis or nonresponders. It is essential, though, that data interpretation takes into account the high intrinsic

variability in ADC and yet, relevant ADC information was found e.g. for HNSCC nodal disease where baseline variability was less than intratreatment change [122].

A general advantage of using immobilization equipment during examination is the reduction in bulk motion artifacts in MR images. Artifacts were less pronounced in MR images acquired with RT mask fixation as compared to the diagnostic setup (data not shown). Artifacts due to swallowing, however, cannot entirely be avoided. Beyond that, the RT setup could potentially still be improved. Closer positioning of the coils to the patient would certainly improve the image quality, but reproducible positioning could become more challenging. Besides increasing the number of averages decreasing the resolution, decreasing the acceleration or reducing TE would give a gain in SNR. However, modifications at the expense of longer acquisition time should be balanced carefully against patient comfort as imaging with mask fixation is demanding. We recommend to opt for a total scan time of no longer than 30 min.

Potentially, similar detail to PET/MRI with RT setup in combination with a planning CT could be obtained by combining data from stand-alone MRI in RT position with a planning PET/CT. It may be with the prospect of direct MR planning for head-and-neck in the future that the value of combined PET/MRI with RT setup becomes most pronounced since the number of examinations could be reduced to one.

In conclusion, simultaneous PET/MR examination of HNC patients using RT positioning aids is clinically feasible. Besides good agreement of PET, the proposed setup comes with a compromise in MR image quality in terms of SNR. However, MR delineation accuracy was not adversely affected and ADC measurement with RT setup was found to be stable. The image quality obtained with RT setup therefore meets RT planning requirements and thus allows for precise integration of PET/MRI for future personalized treatment strategies.

4.5 Acknowledgements

This study was supported by the Center for Personalized Medicine (ZPM) of the Eberhard Karls University Tübingen; parts of the research leading to these results have received funding from the European Research Council under the European

Union's Seventh Framework Programme (FP/2007–2013)/ERC Grant Agreement n. 335367.

The authors thank Siemens Healthcare GmbH, Erlangen, Germany and Qfix, Avondale, PA, USA for providing PET/MRI devices.

4.6 Supplementary Material

Table 4.2: PET parameters.

Parameter	RT scan	Diagnostic scan
Radiopharmaceutical	18F-FDG	18F-FDG
Injected dose	238±8 MBq (range: 219–248 MBq)	–
Acquisition start time	84±0.4 min p.i. (range: 84–85 min)	139±5 min p.i. (range: 130–146 min)
Time difference to first scan	–	54 ± 5 min (range: 45–61 min)
Frame duration	6 min	6 min
Reconstruction method	3D OP-OSEM (2 it., 21 subs.) and 4 mm Gaussian filtering	3D OP-OSEM (2 it., 21 subs.) and 4 mm Gaussian filtering
Inplane resolution	$2.80 \times 2.80\,mm^2$	$2.80 \times 2.80\,mm^2$
Slice thickness	2.03 mm	2.03 mm
Matrix size	256×256	256×256
Number of slices	127	127

Values are listed as mean ± standard deviation. The time difference between the two scans was due to the time needed for setup change and because the second (diagnostic) scan of the head-and-neck region was part of a whole-body examination in caudocranial direction for staging of the disease and the detection of distant metastases. Abbr.: RT - radiotherapy; FDG - Fluorodeoxyglucose; p.i. - post injection; OP-OSEM - ordinary Poisson ordered subset expectation maximization; it. - iterations; subs. - subsets.

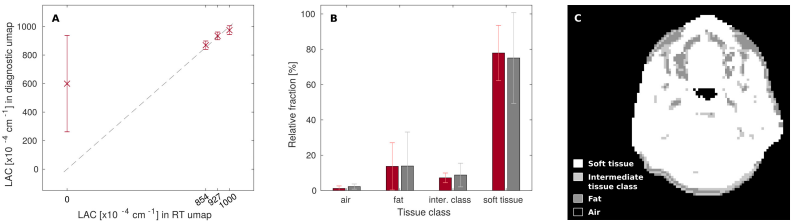

Figure 4.6: A: data plot shows linear attenuation coefficient (LAC) values (mean ± standard deviation (SD)) in MR-based patient μ-map acquired with diagnostic setup that were determined within the respective tissue class segments in the RT μ-map. B: relative fraction of each tissue class within the RT (red) and the diagnostic μ-map (gray). Error bars are shown in light color, respectively, and represent $1 \times$ SD. C: axial view of MR-based μ-map of one patient (#02, RT setup) with segmented tissue classes.

Table 4.3: MR sequence parameters.

Parameter	DW-MRI	T2w MRI	Dixon scan	T1w MRI
Inplane resolution	$2.1 \times 2.1\,mm^2$	$0.6 \times 0.6\,mm^2$	$2.6 \times 2.6\,mm^2$	$1.25 \times 1.25\,mm^2$
Slice thickness	5 mm	3 mm	2.6 mm	1.25/3.08 mm
Slice gap	1 mm	0.3 mm	–	–
Matrix size	192×156	384×384	192×79	224×192/ 320×240
Repetition time	5800 ms	7440 ms	3.6 ms	4.4/4.25 ms
Echo time	58 ms	75 ms	2.36/1.23 ms (in/opp)	1.58/1.88 ms
b-values	150 and 800 s/mm^2	–	–	–
Diffusion directionality	Isotropic	–	–	–
Turbo factor	–	14	–	–
Contrast agent	–	–	–	Gadovist
Flip angle	180 deg	160 deg	10 deg	12 deg
Number of averages	3/3 (b150/b800)	2	1	4/1
Bandwidth	1736 Hz/px	260 Hz/px	965 Hz/px	338/446 Hz/px
Field of view	$324 \times 399\,mm^2$	$240 \times 240\,mm^2$	$328 \times 500\,mm^2$	$240 \times 280/300 \times 400\,mm^2$
Number of slices	34	47	128	80/64
Parallel imaging technique	GRAPPA	GRAPPA	GRAPPA	GRAPPA
Acceleration factor	2	2	2	2
Fat suppression	None	None	–	Yes
Acquisition type	2D	2D	3D	3D

For T1w MRI, values refer to RT and diagnostic setup, respectively. Abbr.: in - in-phase; opp - opposed-phase; RT - radiotherapy scan; diagn. - diagnostic scan.

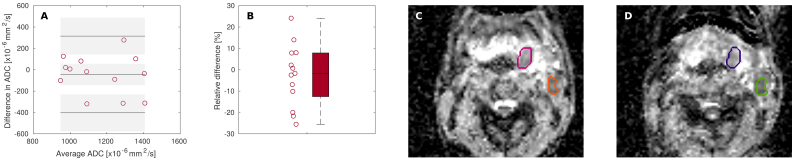

Figure 4.7: A: Bland Altman plot for mean ADC values in PET-derived ROIs measured with RT and diagnostic setup. Graph shows difference between two measurements plotted against their average. Solid lines indicate the mean of differences and the limits of agreement (mean $\pm$ 1.96 $\times$ SD). Corresponding confidence intervals are shown as gray-shaded areas. B: relative difference of ADC between both scans. C, D: ADC map of one patient (#06, axial slice) acquired with RT and diagnostic setup, respectively. ROI structures in lesions are shown in color. Note the difference in patient positioning between both scans.

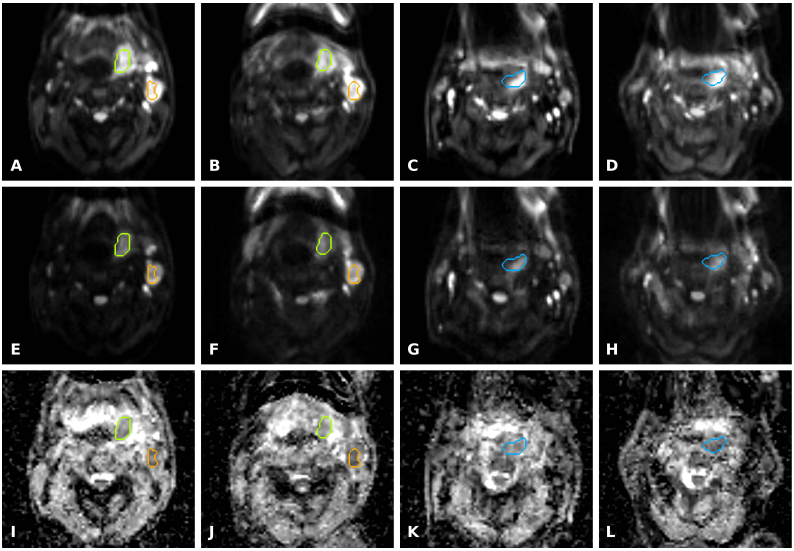

Figure 4.8: DW-MRI and ADC maps of two patients (#06 and #09) acquired with RT and diagnostic scan setups, respectively. A–D: DW-MRI with $b = 150\,\mathrm{s/mm^2}$ with RT and diagnostic setup of patient #06 (A, B) and patient #09 (C, D), respectively. E–F: DW-MRI with $b = 800\,\mathrm{s/mm^2}$ with RT and diagnostic setup of patient #06 (E, F) and #09 (G, H). I–L: Similarly, ADC maps are presented for patient #06 (I, J) and #09 (K, L). PET-derived ROIs are shown for patient #06 (green, orange) and patient #09 (blue) to estimate the difference in ADC. A mean difference in SNR between both scan setups of −24.1% and −16.6% was found for b150 and b800 images for patient #06 (good correspondence). Similarly, a difference of −29.3% and −29.6% was found for patient #09 (average correspondence). The mean difference in ADC was 7.8% and 0.7% for green and orange ROI (good correspondence) and 14% for blue ROI (average correspondence) for patient #06 and #09, respectively.

5 Part II: Distortion correction of diffusion-weighted magnetic resonance imaging of the head and neck in radiotherapy position

René M. Winter[1*], Holger Schmidt[2], Sara Leibfarth[1], Kerstin Zwirner[3], Stefan Welz[3], Nina F. Schwenzer[2], Christian la Fougère[4], Konstantin Nikolaou[2], Sergios Gatidis[2], Daniel Zips[3,5], Daniela Thorwarth[1,5]

[1] Section for Biomedical Physics, Department of Radiation Oncology, University Hospital and Medical Faculty, Eberhard Karls University Tübingen, Tübingen, Germany
[2] Department of Diagnostic and Interventional Radiology, University Hospital and Medical Faculty, Eberhard Karls University Tübingen, Tübingen, Germany
[3] Department of Radiation Oncology, University Hospital and Medical Faculty, Eberhard Karls University Tübingen, Tübingen, Germany
[4] Department of Nuclear Medicine, University Hospital and Medical Faculty, Eberhard Karls University Tübingen, Tübingen, Germany
[5] German Cancer Consortium (DKTK), partner site Tübingen; and German Cancer Research Center (DKFZ), Heidelberg, Germany
[*] Corresponding author at: University Hospital Tübingen, Department of Radiation Oncology, Section for Biomedical Physics, Hoppe-Seyler-Str. 3, 72076 Tübingen, Germany. Phone: +49 70712986061, Fax: +49 7071295920, E-mail: rene.winter@med.uni-tuebingen.de

published in
Acta Oncologica 2017,**56**:1659–1663.

5.1 Introduction

Simultaneous positron emission tomography/magnetic resonance imaging (PET/MRI) has high potential to improve radiotherapy (RT) planning. Key advantages are increased delineation accuracy and the integration of functional information that may be used for patient-individual therapy adaptation [7, 24, 25, 34, 35, 113, 134].

In particular, diffusion-weighted (DW) MRI has been associated with RT outcome in different tumor entities including head and neck [11, 12, 31]. In clinical routine, diagnostic DW-MRI is commonly performed based on echo-planar imaging (EPI) sequences due to favorably short acquisition times reducing bulk motion-induced phase perturbations [40]. However, such EPI techniques are sensitive to B0-field inhomogeneities which leads to image distortions and signal loss. This especially applies to the head and neck region [41], where unfavorable geometry and susceptibility changes can lead to high geometric mismatch of the DW- and an anatomical MR image. For future biologically adapted RT planning purposes based on DW-MRI information, however, high geometric fidelity is crucial [36,113].

PET/MRI integration into RT requires accurate image co-registration with the RT planning computed tomography (CT). Therefore, patient examination in RT-specific position is essential. Adapting PET/MRI examinations to RT position requires immobilization tools to be combined with the MRI hardware. For imaging of the head and neck, a dedicated hardware solution has recently been proposed [39]. In this study, we evaluated its clinical performance with respect to DW-MRI. Since EPI-based DW-MRI is prone to image distortions whereas RT planning demands high geometric fidelity, the focus of this investigation was (i) on the assessment of geometric accuracy, (ii) on the definition of an appropriate method for distortion correction and (iii) on the stability of the apparent diffusion coefficient (ADC).

5.2 Material and Methods

During the pilot phase of a prospective clinical trial (NCT-02666885), ten head and neck squamous cell-carcinoma (HNSCC) patients were examined with simultaneous PET/MRI (Biograph mMR, Siemens Healthineers GmBH, Erlangen, Germany) including T2-weighted (T2w) MRI and EPI-based DW-MRI ($b = 150$ and

$800\,\mathrm{s/mm^2}$), prior to the start of RT. For each patient, two scans were performed – using a RT-specific and a standard diagnostic setup (Supplementary Figure 5.3). The RT setup consisted of a flat MR table top and coil holders for 6-channel flexible body coils, as described in [39]. An in-house designed add-on was mounted onto the MR table top to ensure patient fixation with a thermoplastic RT mask (ITV, Innsbruck, Austria). MR sequence parameters were identical for both scans (Supplementary Table 1).

Distortion correction of DW-MR images was realized based on repeated data collection with reversed phase-encoding directions (RPED), i.e. with orientations anteroposterior (AP) and posteroanterior (PA), resulting in pairs of images with distortions oriented in opposing directions. From these pairs of images, the susceptibility-induced off-resonance field was estimated using a method similar to that described by Andersson et al. [59] and as implemented in open-source toolkit FSL [135]. AP and PA images were combined into a single, distortion corrected one, by finding the field that, when applied to the two images, would maximize the similarity of the unwarped images.

Geometric accuracy of distortion-corrected and uncorrected DW-MRI was derived from pairwise comparison with distortion-robust T2w-MRI. Left and right submandibular glands and the spinal cord at C1-2 and C4-5 were defined manually as regions of interest (ROIs) (i) on the original DW-MR image with $b = 150\,\mathrm{s/mm^2}$ (AP), (ii) on the corresponding distortion-corrected image, (iii) on the T2w-MR image. ROIs from DW-MRI were resampled onto the T2w-MRI grid. ROI-based similarity measures included Dice similarity index (DSI), relative volume difference (RVD), average symmetric surface distance (ASSD) and Euclidean distance of geometric centers.

To assess the quantitative fidelity of DW-MRI, ADC maps were derived from distortion-corrected b-value images and compared between scan setups by their mean values based on the pre-defined anatomical ROIs.

To evaluate (i) the benefit of RPED-technique for geometric alignment between DW-MRI and anatomical reference MRI, statistical analysis was performed using a Wilcoxon rank sum test (MATLAB R2009b). To evaluate (ii) the difference in ADC values between scan setups, statistical analysis was performed using a

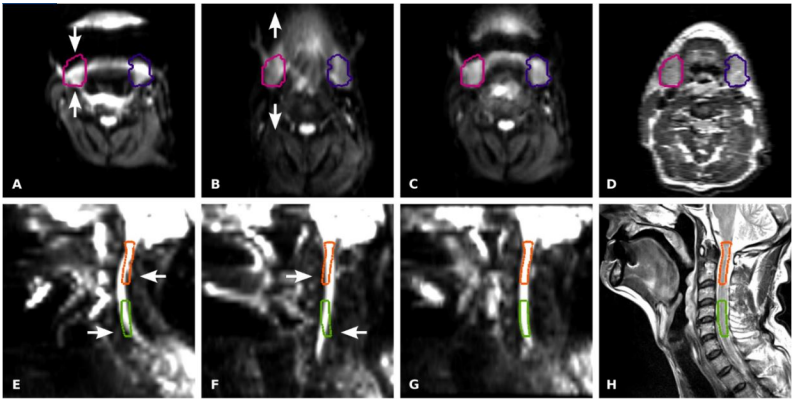

Figure 5.1: Overview of original DW-MRI ($b = 150\,\mathrm{s/mm^2}$) with phase-encoding direction anteroposterior (A, E) and posteroanterior (B, F), corresponding distortion-corrected images (C, G) and T2w-MRI (D, H) of one patient (#06). A-D: axial slice; E-H: sagittal slice. The image set was acquired with RT-specific scan setup. The four ROIs placed in submandibular glands and different sections of the spinal cord based on the T2w-MR image are shown in color and are as well displayed on the DW-MR images in order to highlight the respective geometric (in-)accuracy.

Wilcoxon signed rank test. In either case, a p-value below 0.05 was considered statistically significant.

5.3 Results

Acquisition of T2w- and DW-MRI in the diagnostic setup as well as in the RT setup was feasible in all patients. However, for one patient strong distortion artifacts and signal loss could not be corrected by using RPED, most likely due to swallowing artifacts. Figure 5.1 shows an example of original DW-MRI acquired using two different phase-encoding directions in comparison to the distortion-corrected image data as well as the anatomical T2w-MRI.

Relative to T2w-MRI, DSI (mean ± standard deviation) was $0.85 \pm 0.04/0.53 \pm 0.17$ and $0.79 \pm 0.16/0.53 \pm 0.17$ for RT-specific and diagnostic setup in distortion-corrected/uncorrected images, respectively. Similarly, RVD was found to be $-2 \pm$

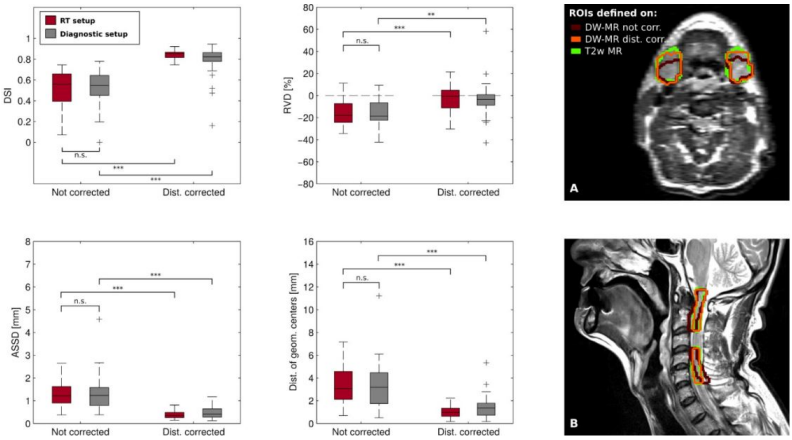

Figure 5.2: Boxplots display results of four similarity measures for inter- and intra-scan setup comparison of original, uncorrected DW-MR images ($b = 150\,\mathrm{s/mm^2}$, phase-encoding AP) and corresponding corrected images; similarity was assessed on a ROI basis with ROIs defined on the T2w-MR image serving as reference. Boxplots represent single ROI-based values over all patients. A, B: To highlight ROI similarity between the (un-)corrected DW-MR image and the T2w-MR reference image, the latter is shown for one patient (acquisition with RT setup) in two planes and ROIs that were defined on the DW-MR images are displayed together with the reference ROI.

$12/-16 \pm 12\%$ and $-3 \pm 16/-13 \pm 21\%$. ASSD was $0.4 \pm 0.2/1.3 \pm 0.5\,\mathrm{mm}$ and $0.5 \pm 0.4/1.4 \pm 0.8\,\mathrm{mm}$ and the distance of the geometric centers of the ROIs was determined as $1.0 \pm 0.5/3.3 \pm 1.6\,\mathrm{mm}$ and $1.4 \pm 1.1/3.4 \pm 2.1\,\mathrm{mm}$.

Distortion correction based on RPED yielded substantially increased geometric accuracy, demonstrated by significant improvements of the ROI-based measures DSI ($p < .001$), RVD ($p < .01$), ASSD ($p < .001$) and distance of ROI centers ($p < .001$) for RT and diagnostic scan setup alike (Figure 5.2).

The relative difference of volume-averaged ADC values in distortion-corrected maps with respect to the diagnostic setup was $1 \pm 10\%$ ($p =$ n.s., Supplementary Figure 5.4).

5.4 Discussion

For the integration of PET/MRI into RT planning, precise image fusion of PET/MRI and planning CT is mandatory. For this purpose, patient examination in RT position is essential, especially in the case of head and neck [115]. Image analysis of the first ten patients recruited in this clinical trial demonstrated the clinical feasibility of functional PET/MRI examination in RT-specific position using a customized hardware setup.

PET/MRI integration into RT also requires high geometric accuracy of the image data [36, 113]. Image analysis revealed substantial distortions in DW-MRI data. Meanwhile, the choice of coil setup did not affect the magnitude of image distortion. This was as expected, as magnetic field inhomogeneity and susceptibility artifacts are not expected to differ between coil setups. Final distortion correction with RPED-technique yielded significant improvement of geometric match between DW-MRI and anatomical image data, in particular when using the RT setup for data acquisition. Thus, DW-MRI integration into RT planning would be considerably improved when applying this correction technique.

This easy-to-implement RPED-based method for distortion correction may in the future also be relevant for other hybrid MR-applications related to RT, such as the MR-Linac or MR-hyperthermia, since technology advances in DW-MR sequence design minimizing geometric inaccuracy are generally implemented with a considerable time delay into novel, cutting-edge multimodality systems, compared to state-of-the-art MR standalone systems.

Besides RPED, different alternative techniques have been proposed for reducing artifacts in EPI-based DW-MRI. Reduction of echo spacing in readout-segmented multi-shot EPI (rsEPI), for instance, diminishes spatial inaccuracy in areas of susceptibility changes, but comes at the expense of substantially longer scan times, compared to standard single-shot EPI [130]. Another approach for the correction of geometric distortions has been developed based on the generation of magnetic field maps using data acquisition with different echo-times [131]. The main disadvantage of this technique is that field map calculation is based on multiple gradient-echo images that are acquired separately from each other and from the actual DW-MR dataset, rendering the method more sensitive to motion. Only recently, integrated 2D shim and frequency adjustment for EPI-based DW-MRI (intEPI) has been im-

plemented for the mMR system in our institution, which yielded very promising results in terms of image quality improvement and could be used in combination with RPED or other correction methods in the future, to further optimize geometric accuracy for RT purposes [41].

Although image quality using flexible coils was expected to be lower compared to a standard diagnostic head and neck coil, primarily because the distance between subject and coils is increased due to the space needed for RT positioning aids, visual assessment of images acquired with RT setup revealed rather lower levels of image artifacts since patient motion was reduced by mask fixation. Nevertheless, artifacts due to swallowing cannot be avoided by RT mask fixation.

No significant difference in corrected ADC values was found between the scan setups. Thus, data acquisition with RPED-technique using flexible radiofrequency coils in the context of RT-specific patient positioning seems not to be a limitation for quantitative DW-MRI and future multi-parametric PET/MRI schemes.

Limitations of the similarity assessment between DW-MRI and anatomical reference MRI may originate from manual ROI delineation and resampling on a common image grid. However, both steps were considered necessary for effective image comparison and calculation of similarity measures. Moreover, only one b-value ($b = 150\,\mathrm{s/mm^2}$) was used for analysis. Nevertheless, results are considered to be representative, since the susceptibility-induced field causing distortions may be, to a first approximation, considered constant for all acquired images, meaning that one set of images is internally consistent [59].

Investigation of the effects of the RT hardware setup on PET data quantification and MR image quality was beyond the scope of this work and will be analyzed in a future study.

In conclusion, PET/MRI examination of HNSCC patients in RT position using dedicated positioning aids is clinically feasible. Distortions in DW-MRI can effectively be reduced by RPED-technique. ADC acquisition with RT setup was found to be stable. The presented imaging solution thus enables more precise integration of simultaneous PET and DW-MRI into RT planning for future personalized treatment strategies.

5.5 Acknowledgements

This study was supported by the Center for Personalized Medicine (ZPM) of the Eberhard Karls University Tübingen; parts of the research leading to these results have received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007 – 2013)/ERC Grant Agreement n. 335367.

5.6 Supplementary Material

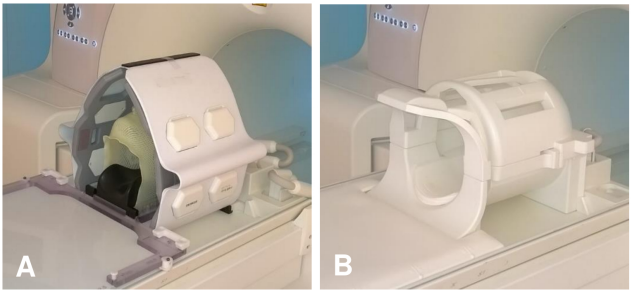

Figure 5.3: Imaging setups. A: RT-specific scan setup consisting of MR table top, coil holders for 6-channel flexible coils and table top add-on for patient fixation with RT mask. B: Diagnostic 16-channel head and neck coil that does not allow imaging in RT position.

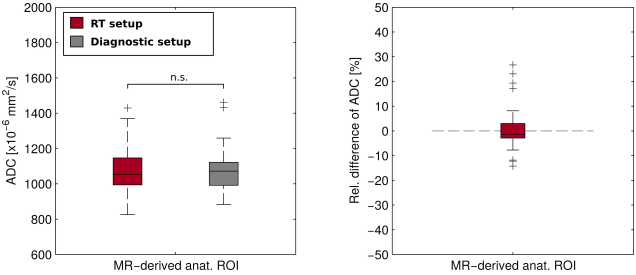

Figure 5.4: Left panel: ROI-derived mean ADC values for both scan setups. Right panel: Relative difference in ADC values between the RT-specific and the diagnostic reference setup.

Table 5.1: MR sequence parameters.

Parameter	DW-MRI	T2w-MRI
Inplane resolution	$2.1 \times 2.1\,mm^2$	$0.6 \times 0.6\,mm^2$
Slice thickness	5 mm	3 mm
Slice gap	1 mm	0.3 mm
Matrix size	192×156	384×384
Repetition time	5800 ms	7440 ms
Echo time	58 ms	75 ms
b-values	150 and $800\,s/mm^2$	–
Flip angle	180 deg	160 deg
Number of averages	3	2
Bandwidth	1736 Hz/px	260 Hz/px
Field of view	$324 \times 399\,mm^2$	$240 \times 240\,mm^2$
Number of slices	34	47
Acceleration factor	2	2
Fat suppression	None	None
Acquisition type	2D	2D

6 Part III: Machine learning for the prediction of dynamic FMISO-PET voxel information by multi-parametric MRI

René M. Winter[1*], Sara Leibfarth[1], Simon Boeke[1,2,3], Marcel Krueger[4], Pamela Mena-Romano[1], Efe Cumhur Sezgin[2], Gregory Bowden[4], Jonathan Cotton[4], Bernd Pichler[3,4], Daniel Zips[2,3], Daniela Thorwarth[1,3]

[1] Section for Biomedical Physics, Department of Radiation Oncology, University Hospital and Medical Faculty, Eberhard Karls University Tübingen, Tübingen, Germany
[2] Department of Radiation Oncology, University Hospital and Medical Faculty, Eberhard Karls University Tübingen, Tübingen, Germany
[3] German Cancer Consortium (DKTK), partner site Tübingen; and German Cancer Research Center (DKFZ), Heidelberg, Germany
[4] Werner Siemens Imaging Center, Preclinical Imaging and Radiopharmacy, Tübingen, Germany
[*] Corresponding author at: University Hospital Tübingen, Department of Radiation Oncology, Section for Biomedical Physics, Hoppe-Seyler-Str. 3, 72076 Tübingen, Germany. E-mail address: rene.winter@med.uni-tuebingen.de

manuscript in preparation

Abstract

Background and purpose: Hypoxia imaging via dynamic FMISO-PET has shown prognostic value for radiation therapy response of patients with head-and-neck squamous cell-carcinoma (HNSCC). However, in routine clinical practice FMISO-PET is subject to limited availability and extensive imaging protocols. Its replacement by MRI would be an intriguing alternative. Our hypothesis is that a multi-parametric model comprising both diffusion and perfusion related MRI parameters is able to predict dynamic FMISO-PET at a voxel level. In the present study, machine learning approaches were used to develop such model on FMISO-PET/MRI data of human HNSCC xenografts grown in mice.

Material and Methods: Combined PET/MRI data of $n = 62$ ectopic HNSCC xenografts was acquired including dynamic FMISO-PET, diffusion-weighted (DWI) and dynamic contrast-enhanced (DCE) MRI. Animal data was processed on voxel level (116 980 voxels) and split randomly into training (80%) and independent test cohorts (20%). K-means clustering was performed on noise filtered FMISO-PET time-series data to categorize each voxel into one of five different hypoxia-associated curve types. These hypoxia curve types were used as response variable for classification. Three MRI parameters served as predictors: ADC and projection coefficients c1 and c2 of the first two principal components obtained by principal component analysis of the DCE-MRI data. Gradient boosted decision trees (GBDT) and artificial neural networks (ANN) were trained and assessed by 5-fold cross-validation, and the model with best predictive performance was evaluated in the independent test cohort. ROC area-under-the-curve (AUC) served as metric to assess predictive performance.

Results: Cross-validation on the training data yielded a similar validation fold based ROC-AUC score of 0.80 ± 0.02 (macro-averaged) for both GBDT and ANN. GBDT was favored over ANN as final model for better interpretability and yielded ROC-AUC = 0.80 in the independent test cohort. Parameter importances of ADC and DCE-MRI projection coefficients c1 and c2 in the final model were 20%, 39% and 41%, respectively.

Conclusion: A GBDT classifier for multi-parametric prediction of dynamic FMISO-PET derived hypoxia classes by DWI and DCE-MRI was trained and successfully validated on small animal data. The proposed method presents a

novel approach for voxelwise learning of functional image information. Retraining the model on patient data for clinical validation seems highly promising.

6.1 Introduction

Tumor hypoxia is an important adverse prognostic factor for treatment outcome of patients receiving radiation therapy (RT) and has been shown to be present in a wide variety of solid tumors including head-and-neck squamous cell-carcinoma (HNSCC) [18,47–50]. A well-established method to detect and assess tumor hypoxia in both preclinical and clinical settings is positron-emission-tomography (PET) with hypoxia-specific ^{18}F labeled radiotracers such as Fluoromisonidazole (FMISO), Fluoroazomycin arabinoside (FAZA) or Fluortanidazole (HX4) [18,51,60,61,136]. Clinical studies have shown that imaging of tumor hypoxia via hypoxia PET has high prognostic value for RT response of patients with HNSCC [17, 51, 60]. Planning studies investigating predictive potential and first trials adapting RT dose prescription based on PET derived hypoxia status are currently ongoing [30, 137, 138].

Hypoxia PET, in particular, has high prognostic value with respect to RT response when image information on tumor hypoxia is derived from time series data [17, 30, 53–55] as such dynamic data includes both information on early temporal tracer distribution in the tumor tissue related to vascular properties as well as on late retention effects [17, 30, 53–58]. However, in routine clinical practice dynamic hypoxia PET is subject to limited availability and extensive imaging protocols. A surrogate marker derived from functional magnetic resonance imaging (MRI) may provide a more practical and readily implementable alternative.

Under this perspective, recent studies have investigated pairwise correlations between hypoxia information and various MRI parameters [113, 139–142]. Moderate to high correlations between different FMISO-PET parameters and parameters derived from dynamic contrast-enhanced (DCE) MRI were identified for patients with HNSCC [113, 139, 140], but also high variation across patients was observed [113, 139]. Different studies found significant correlations between single DCE-MRI derived parameters and hypoxia defined by pimonidazole staining of tumor tissue from HNSCC patients [141, 142].

However, a relation between functional MRI and hypoxia PET might not be fully captured by simple pairwise correlation analysis alone. Hompland et al. presented a strategy to indirectly measure hypoxia in prostate cancer by exploiting joint information of two parameters derived from diffusion-weighted MRI (DWI) [97]. With a model including apparent diffusion coefficient (ADC) and fractional blood volume (fBV) from intravoxel incoherent motion (IVIM) modeling, good results for prediction of pimonidazole immunoscore could be obtained.

Machine learning now offers advanced tools to build prediction models that may capture very complex relationships between multimodal imaging parameters [143–145]. Therefore, we hypothesized that a machine learning model combining both diffusion and vascularity related parameters from DWI and DCE-MRI is able to predict dynamic hypoxia PET on tumor voxel level. In the present study, we developed such model on dynamic FMISO-PET/MRI data of human HNSCC xenografts grown in mice and investigated its robustness in an independent cohort.

6.2 Material and Methods

Imaging

62 ectopic HNSCC xenografts grown in mice (female immunodeficient nude mice, 4–6 weeks old, NMRI nu/nu, Charles River Laboratories) were examined with simultaneous PET/MRI (7 T MRI Bruker Biospec equipped with a PET insert [62,146]). The image protocol included dynamic FMISO-PET (90 min), T2-weighted (T2w) MRI (2D RARE), DWI (2D standard DTI) and DCE-MRI (2D FLASH). Experiments were approved according to institutional guidelines and German animal welfare regulations (35/9185.81-2/R4/16). HNSCC xenografts were derived from eight different cell-lines: (UT-SCC-45, XF354, UT-SCC-14, UT-SCC-8, FaDu, CAL-33, UT-SCC-5, SAS). 2–5 days before tumor cell injection (hind leg, subcutaneous), animals received a 4 Gy total body irradiation to suppress the residual immune system and facilitate tumor growth. During imaging experiments, animals were anesthetized using a mixture of isoflurane (1.5–2%) and air. Breathing rate was constantly monitored and body temperature regulated via water heating pads.

FMISO tracer was administered intravenously (tail vein). PET data was acquired in list mode and reconstructed into 65 time-series images with time frame setting

of $36 \times 10, 18 \times 60$ and 11×360 s using 2D-OSEM (4 iterations, 16 subsets). MR data was acquired with a mouse whole-body radiofrequency coil (Bruker). DWI was performed with nine equidistant b-values ranging from $b = 0-800\,\mathrm{s/mm^2}$. T2w MRI and DWI image acquisitions were breathing gated. DCE-MRI was acquired from 1 min before to 12.5 min after gadolinium based contrast agent injection (Gadobutrol, Gadovist, Bayer Vital GmbH, Germany) with a temporal resolution of 5.4 s. Further PET and MR acquisition parameters are listed in Supplementary Table 6.2. Complete datasets of 42 animals could be used for analysis as of 62 animals examined, five had to be excluded for small tumor size, fourteen for incomplete data acquisition and one for image artifacts.

Pre-processing

For data analysis on a common image grid, rigid image registration was applied using open source tool elastix [119]. Evaluation of the image similarity metric was constrained to a small region containing the tumor and surrounding area. The PET image with voxel size of $0.65 \times 0.65 \times 0.8\,\mathrm{mm^3}$ was chosen as fixed image. Registration was chosen over direct resampling for corrections of potential misalignments which could occur due to slight leg movements or shifts e.g., due to bladder filling. Tumor tissue was delineated on T2-weighted MRI data by a board certified radiation oncologist (SB) with open source software 3DSlicer.

Strategy for data analysis

Prediction of dynamic FMISO-PET information was defined as a multi-class classification problem of different hypoxia-associated curve types that occur within a tumor. To identify such hypoxia curve types and find a suitable set of MRI predictor variables for a dedicated classification model, animal data was processed on the voxel level (116 980 voxels). Animal data was split randomly into training (80%) and independent test cohorts (20%) to decouple model development from final validation. During data splitting, integrity of tumor structures and cell-line ratios were maintained.

Extraction of quantitative parameter maps

Principal component analysis (PCA) was used to derive parameter maps from dynamic image data on voxel level. PCA is a dedicated method for dimensionality reduction and noise filtering of data. Applied to time-series image data, it generates a set of new variables named principal components (PCs) which is of same size as the original set of variables, i.e., the set of time points, forming a new set of basis vectors for data representation. Associated projection coefficients $\boldsymbol{c}$ are, thus, a unique and full representation of the data in coefficient space. This allows for the signal S_{t_i} of a voxel at time point t_i collected in vector notation as $\boldsymbol{S} = S_{t_1}, S_{t_2}, \ldots, S_{t_T})$ to be reconstructed as linear combination of the PCs: $\boldsymbol{S} = \sum_{j=1}^{T} c_j \mathbf{PC}_j$, where c_j denotes the coefficient of that voxel corresponding to the jth principal component vector $\mathbf{PC}_j$ and T the total number of principal components equal the total number of time points. As it is characteristic for the method that data variance captured by the PCs successively decreases from the first to the last PC, the linear combination can be truncated at the first PCs to obtain dimensionality reduction and filtering.

Regarding dynamic FMISO-PET, PCA was applied after conversion of measured activity concentrations to standardized uptake values (SUV) by normalization to animal weight and injected activity. For DCE-MRI, PCA was applied after normalization of time dependent signal data S_{t_i} to relative signal increase: $\Delta S_{t_i} = (S_{t_i} - \bar{S}_0)/\bar{S}_0$, where $\bar{S}_0$ represents the mean signal before contrast agent injection. Of note is that to fully decouple final testing from training, PCA was applied to the training data only, whereas projection coefficients c_j of the test data were determined by data projection to the training PCs.

In addition, apparent diffusion coefficient (ADC) maps were derived by modeling DWI signal decay with a common mono-exponential function $S(b) = S_0 \exp(-ADC \times b)$, with b-value b denoting the diffusion weighting at image acquisition and S_0 representing DWI signal intensity at $b = 0\,\mathrm{s/mm^2}$.

Derivation of hypoxia classes associated to different types of FMISO-PET time-activity curves

To group voxels into different classes of hypoxia, FMISO-PET data represented by coefficients c1 and c2 was clustered using K-means algorithm with K = 5 clusters.

Before clustering, coefficients were normalized to zero mean and unit standard deviation. Since each data point in coefficient space represents a filtered time-activity curve of a voxel, voxels could directly be assigned to one of five different classes of cluster-specific FMISO-PET time-activity curves associated to different levels of hypoxia. These hypoxia-associated classes were used as discrete response variable for predictive modeling.

In order to keep independent testing unlinked from training, the hold out test data was not included into k-means clustering but assigned to the five hypoxia classes by nearest neighbor search of cluster centers.

Classification

Two different classification models were trained on the training data for prediction of the five hypoxia classes: gradient boosted decision trees (GBDT) and a deep learning artificial neural network (ANN) based on a multi-layer perceptron architecture. In both models, ADC, and DCE-MRI projection coefficients c1 and c2 served as predictors (features). In a first step, all possible subsets of features were tested for best predictive performance using 5-fold cross-validation. Mean area under the curve (AUC) of receiver operating characteristic (ROC) was used as score function to assess predictive performance. Calculation of ROC-AUC was adapted to the multiclass problem by binarizing the response variable according to common scheme one-vs-all.

In a second step, hyper-parameter settings of each classifier were optimized on the best performing feature subset found in step one. This was done by evaluation of dual learning curves on training and validation folds obtained by 5-fold cross-validation. Metrics used to evaluate the learning curves were mean Log Loss and ROC-AUC. The strategy for optimization involved three phases: first, to build a flexible model that slightly overfits the training data, i.e. to minimize training loss/maximize training ROC-AUC; second, to decrease the learning rate; and third, to regularize the model in order to minimize validation loss and maximize validation ROC-AUC. Minimum validation loss was used to determine the optimum number of iterations. Hyper-parameters tuned are listed in Supplementary Table 6.3.

In a third and final step, the model with best predictive performance was retrained on the full training data and tested in the independent test cohort. In

addition to ROC-AUC, a confusion matrix was used to assess the classifier's performance.

Image processing was implemented using Matlab (Release R2018b) and Python (Python 3.6.7, SciPy 0.19.1); for machine learning Python library scikit-learn (0.19.1) was used.

6.3 Results

Parameter maps

The results of PCA on the dynamic image data obtained in the training cohort are presented in Fig. 6.1. The first two PCs found for FMISO-PET and DCE-MRI data are presented in Fig. 6.1A and 6.1D, respectively. The percentage of total variance explained by the first two PCs was 79.4% and 3.3% for FMISO-PET, and 93.5% and 2.1% for DCE-MRI, respectively (Fig. 6.1B and 6.1E). In both cases, the percentage of total variance explained by PCs of lower order was $< 1\%$. Fig. 6.1C and 6.1F demonstrate the accuracy of FMISO-PET and DCE-MRI voxel curve reconstruction by truncated linear combination of PC1 and PC2.

A fused PET/MR image of an exemplary tumor and maps of FMISO-PET projection coefficients c1 and c2 relative to PC1 and PC2 are presented in Fig. 6.2A–C. Maps of all three parameters used as predictor variables for classification of FMISO-PET curve types, i.e., maps of ADC and DCE-MRI projection coefficients c1 and c2, are presented in Fig. 6.2D–F.

Hypoxia classes associated to different types of FMISO-PET time-activity curves

Results obtained in the process of finding hypoxia classes associated to characteristic FMISO-PET curve types as response variable for classification are presented in Fig. 6.2A–C. Fig. 6.2A depicts results from k-means clustering of projection coefficients c1 and c2 relative to FMISO-PET PC1 and PC2. The different curve types or hypoxia classes obtained by the clustering are presented in Fig. 6.2B. An exemplary tumor map of these hypoxia classes is shown in Fig. 6.2C. Clustering assigned 26.3%, 23.7%, 15.1%, 23.3% and 11.6% of voxels in the training cohort to hypoxia classes 1–5, respectively.

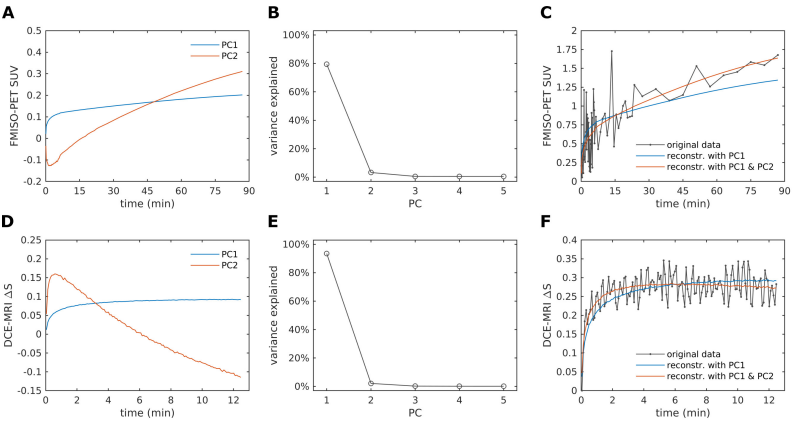

Figure 6.1: Principle component analysis. The first two principle components (PCs) are presented for FMISO-PET (A) and DCE-MRI (C). The percentage of total variance explained by the first few PCs of FMISO-PET and DCE-MRI is plotted in B and D, respectively. Exemplary voxel data of FMISO-PET (C) and DCE-MRI (F) is shown together with data reconstructions by PC1 and by truncated linear combination of PC1 and PC2.

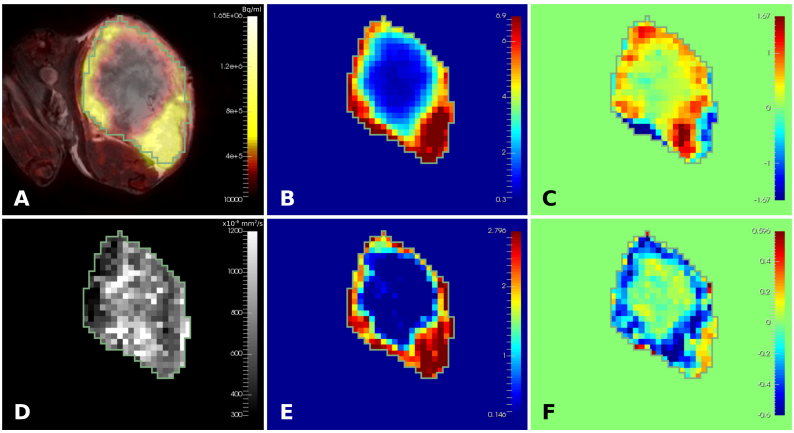

Figure 6.2: A fused image of FMISO-PET at 90 min p.i. and T2 weighted MRI is shown for a tumor in axial view in A. Parameter maps of FMISO-PET projection coefficients c1 and c2 are presented in B and C, respectively. In the bottom row, maps of the three predictor variables ADC (D), DCE-MRI projection coefficient c1 (E) and c2 (F) are presented.

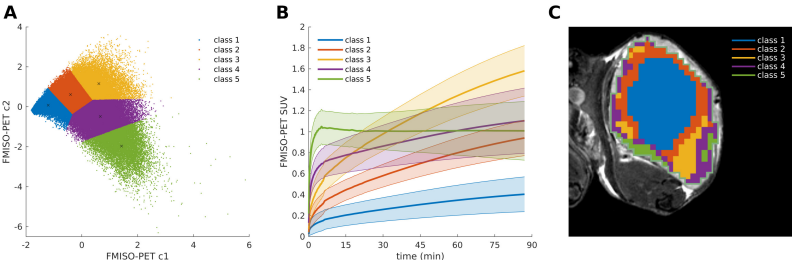

Figure 6.3: The process of finding hypoxia classes associated to FMISO-PET voxel curve types to be used as response variable for classification. A: K-means clustering of FMISO-PET voxel curve data in coefficient space of c1 and c2. B: FMISO-PET curve types obtained by clustering. Bold lines correspond to cluster centers in B; shaded areas correspond to ±1 standard deviation. C: Exemplary map of the hypoxia associated classes for a tumor from the training cohort. The same tumor is shown as in Fig. 6.2.

Table 6.1: Selection of best feature set based on the training cohort. Mean ROC-AUC averaged over the validation folds of 5-fold cross-validation was used as score to estimate predictive performance. GBDT was tested using a total number of estimators of $noe_{total} = 200$, whereas ANN was tested using a total number of iterations of $noi_{total} = 500$ and the highest scored result as well as corresponding noe and noi were identified. Feature combinations were ranked according to score. Hyper-parameter settings are listed in Supplementary Table 6.3.

Feature combination	GBDT			ANN		
	ROC-AUC ± SD	noe	rank	ROC-AUC ± SD	noi	rank
ADC	0.589 ± 0.025	14	7	0.576 ± 0.022	500	7
DCE-MRI c1	0.740 ± 0.034	188	4	0.740 ± 0.035	14	4
DCE-MRI c2	0.691 ± 0.020	102	6	0.691 ± 0.021	139	6
ADC/DCE-MRI c1	0.759 ± 0.030	104	3	0.756 ± 0.029	500	3
ADC/DCE-MRI c2	0.720 ± 0.021	159	5	0.720 ± 0.021	374	5
DCE-MRI c1/DCE-MRI c2	0.781 ± 0.024	112	2	0.782 ± 0.023	252	2
ADC/DCE-MRI c1/DCE-MRI c2	0.793 ± 0.022	162	1	0.793 ± 0.021	340	1

Abbr.: GBDT: gradient boosted decision trees, ANN: artificial neural network, ROC: receiver operating characteristic, AUC: area under the curve, SD: standard deviation, noe: number of estimators, noi: number of iterations.

Classification

Among all feature subsets evaluated on the training data using cross-validation, the set of ADC/DCE-MRI c1/DCE-MRI c2 performed best for prediction of FMISO-PET derived hypoxia classes in both GBDT and ANN. Scores are listed in Table 6.1. Exemplary learning curves obtained with that feature set and best hyper-parameter settings found during optimization are presented in Supplementary Fig. 6.8 and 6.9.

With optimized settings, GBDT and ANN yielded a similar cross-validation based ROC-AUC score of 0.80 ± 0.02, averaged over all hypoxia classes. For further evaluation, GBDT was favored over ANN as final model for better interpretability. ROC curves of all single hypoxia classes obtained with GBDT in cross-validation folds of the training cohort are presented in Fig. 6.4 along with a confusion matrix in Fig. 6.5. When retrained on the full training dataset, parameter importance of ADC, DCE-MRI c1 and DCE-MRI c2 for GBDT based classification was 20%, 39% and 41%, respectively.

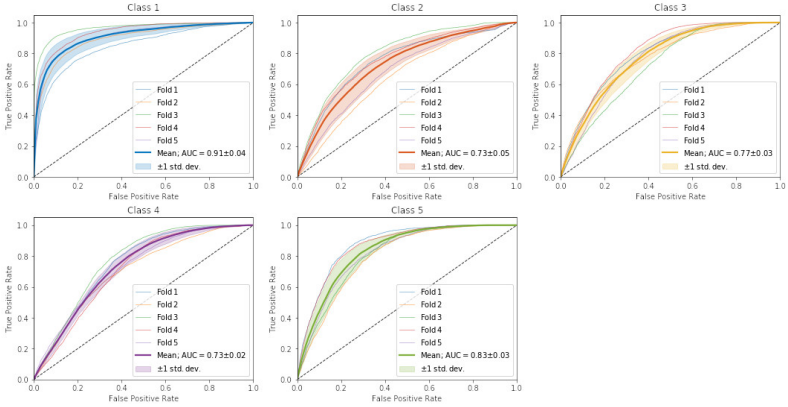

Figure 6.4: ROC curves of hypoxia associated classes 1–5 obtained with GBDT in cross-validation folds of the training data. A bold curve refers to calculated mean over the five folds (pale colored curves).

In the independent test cohort, retrained GBDT yielded a ROC-AUC score of 0.80 with single class scores ranging from 0.73 – 0.94. ROC curves and the confusion matrix obtained in that test cohort are plotted in Fig. 6.6. To visualize the partial dependence between FMISO-PET hypoxia classes and the different MRI parameters, test data was plotted in 2D parameter spaces for all three pairwise combinations of parameters and color-coded for real and predicted classes (Supplementary Fig. 6.10). Fig. 6.7 shows an exemplary tumor taken from the independent test cohort with real versus predicted map of hypoxia classes.

6.4 Discussion

The study was performed in ectopic HNSCC xenografts grown in mice that were imaged simultaneously with dynamic FMISO PET, T2w MRI, DWI and DCE-MRI. The multimodal image data was processed on voxel level which allowed for training of decision trees and neural network based classification models towards intratumoral prediction of distinct FMISO-PET derived hypoxia classes by three

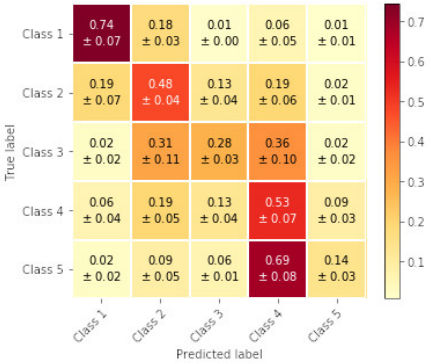

Figure 6.5: Confusion matrix of hypoxia associated classes 1–5 obtained with GBDT in cross-validation folds of the training data. Values refer to the fraction of the total number of voxels within a certain true class type i (for $i = 1, \ldots 5$) that was predicted as class type j (for $j = 1, \ldots 5$), and are given as mean ± 1 standard deviation averaged over the 5 folds. Values in the diagonal from the upper left to the lower right report the fractions of correct predictions (true positives), whereas prediction errors are represented by values outside the diagonal.

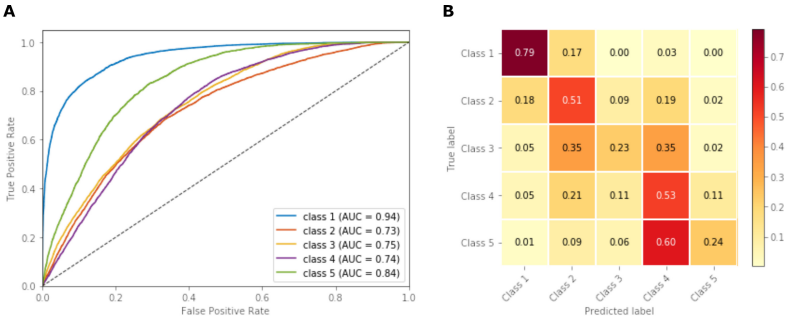

Figure 6.6: A: ROC curves and AUC scores for the five hypoxia associated classes obtained with GBDT in the independent test cohort. Confusion matrix of hypoxia associated classes 1—5 obtained with GBDT in the independent test cohort. Values refer to the fraction of the total number of voxels within a certain true class type i that was predicted as class type i (for $i = 1, \ldots 5$) that was predicted as class type j (for $j = 1, \ldots 5$). Values in the diagonal from the upper left to the lower right report the fractions of correct predictions (true positives), whereas prediction errors are represented by values outside the diagonal.

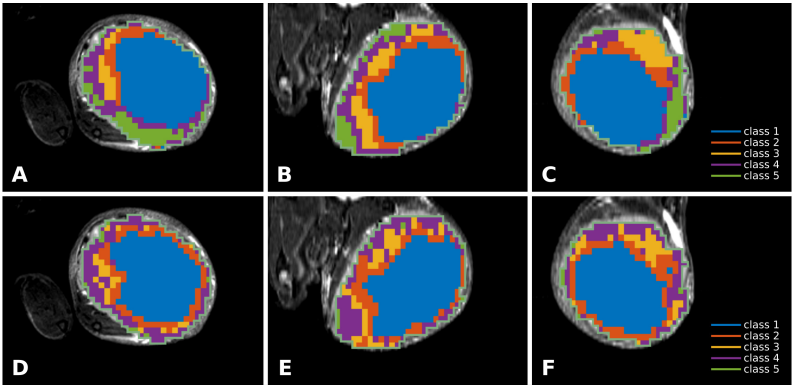

Figure 6.7: True (A–C) versus predicted map (D–F) of hypoxia associated classes for an exemplary tumor from the independent test cohort. A/D: axial view; B/E: coronal view; C/F: sagittal view.

different MRI parameters. Both models showed high predictive performance and GBDT was favored over ANN as final model for better interpretability. Robustness of GBDT was demonstrated by validation in an independent test cohort.

During model development, PCA was used as a model free approach for capturing multi-frame dynamic imaging information by a few number of quantitative parameters. For both FMISO-PET and DCE-MRI, the first two PCs explained most of the total variance of the original data and sufficed for adequate reconstruction (cf. Fig. 6.1). These findings are similar to findings of previous studies where PCA was used for parameter extraction from DCE-MRI [147–149] or dynamic PET [150, 151]. PCA was chosen over pharmacokinetic modeling as physiological parameters derived from modeling may involve several uncertainties [149, 152, 153]. These include, in particular, that model assumptions may be oversimplified to account for certain voxel dynamics such that the model does not fit adequately or model parameters may not accurately reflect the underlying physiology.

Five different hypoxia classes associated to FMISO-PET time activity curves were derived by k-means clustering, a data driven approach for automatic partitioning of data. $K = 5$ was chosen to obtain a limited number of distinct classes of time-activity curves. The curve types found show different patterns of early and late tracer uptake which may point at different levels of tissue perfusion, diffusion and hypoxia induced retention effects [57, 58]. Class 1 seems to indicate avascular, necrotic tissue as over time only slow tracer uptake occurs. Classes 2–3 are characterized by rather low early, yet pronounced late uptake which could be related to different levels of hypoxia induced retention. Classes 4 and 5 show more pronounced early and less pronounced increase or even decrease in late tracer uptake and may therefore represent rather well perfused, normoxic voxels. Mapping of the different classes yielded spatially contiguous regions mostly in an onion-shape or rim-core structure, suggesting a plausible biological basis. Overall, dynamics of derived classes seem to reflect different levels of physiological properties associated to treatment resistance but would require further biological validation for more precise characterization.

Using dynamic FMISO-PET information as response variable to predict was deliberately favored over using parameters derived from a static acquisition, such as tumor-to-muscle (TMR) or tumor-to-blood ratio (TBR) based parameters. Such

static hypoxia parameters depend on the exact time point at which images are obtained and have been shown to be associated to only limited reproducibility as definition of background activity is prone to variability [154]. Furthermore, they may mislead to a similar assessment of time-activity curves stemming from different physiologies missing out on information on vascular tissue properties [57,58]. Some studies identified static maximum TMR or TBR as prognostic measures [51, 155]. However, results of different studies demonstrated that FMISO-PET has only prognostic value for therapy outcome, or higher prognostic value, when perfusion and retention information are both taken into account [30, 53, 156].

To tackle the prediction problem in this study, classifiers were chosen from two different families of learners, i.e., decision tree ensembles and neural networks. Both supervised learning algorithms were suitable candidates since they are capable of approximating complicated nonlinear functions and capable of handling a multi-class response variable. GBDT, in particular, was opted for as it has been shown that this model most often outperforms classifiers from different families, such as K-nearest-neighbor classification, random forest or support vector machines, for a wide range of classification problems [157].

Cross-validation in the training cohort yielded ROC-AUC $\geq$ 0.73 for any of the five classes which indicates good generalization capacity of the model. ROC-AUC scores obtained in the independent test cohort were very similar to the ones obtained with cross-validation and confirm good generalization capacity and robustness of the proposed method.

Evaluation of the confusion matrices, on the other hand, revealed a certain risk of misclassification between hypoxia classes. To some extent this may be explained by uncertainties in mapping response to predictor data due to the resampling of MR image data with different spacing onto a common image grid (PET), which was necessary for voxelwise learning, but involved interpolation of signal intensities. Different spacing of MR images resulted from a tradeoff between good image quality and good coverage of the entire tumor volume. In general, it was observed that when classes were not predicted correctly, they were rather confused with adjacent than with non-adjacent ones. In particular, a low percentage of true positives ($14 \pm 3\%$ in cross-validation, 24% in the test cohort) was obtained for voxels of class 5. This might be explained by the imbalanced nature of the classification

problem with class 5 representing just 11.6% of voxels in the training cohort (12% in the test cohort). In addition, similar dependence between classes and values of MRI parameters of classes 4 and 5, as indicated by partial dependence plots in Supplementary Fig. 6.10, could be a contributing factor to misclassification. However, the fraction of false positives for class 5 was rather low (cf. off-diagonal values in column 5 of the confusion matrices shown in Fig. 6.5 and 6.6) and good ROC performance was observed. Moreover, type 4 and 5 voxels can be considered as rather well perfused, low risk voxels compared to voxels of type 1–3, as outlined above, and may be less of interest e.g., for treatment adaptation.

Different MRI techniques have been presented recently to image hypoxia such as blood oxygenation level dependent (BOLD) or oxygen-enhanced (OE) MRI, also known as tumor oxygenation level dependent (TOLD) MRI [158]. These techniques measure changes in transverse ($R2^*$) or longitudinal relaxation rates (R1), respectively, induced by inhalation of oxygen or hyperoxic gas such as carbogen. However, at present these techniques still lack sufficient technical, biological, and clinical validation to be established in routine clinical practice [158, 159]. Nevertheless, parameters derived from such techniques could easily be added and trained with the proposed model.

Recently, Hompland et al. presented a new strategy to indirectly measure hypoxia in prostate cancer by exploiting joint information of IVIM parameters derived from DWI [97]. With ADC and fractional blood volume fBV the model included a combination of oxygen consumption and supply related image parameters. This is similar to our model where joint information of ADC and projection coefficients c1 and c2 relative to PC1 and PC2 of DCE-MRI are used as input parameters. While ADC has been shown to be correlated with cell density and oxygen consumption [95], DCE-MRI time signal curves depend on temporal contrast agent distribution related to perfusion and microvasculature of tissue [15]. Therefore, high and low values of projection coefficient c2 would, in particular, translate to high and low level of early perfusion and oxygen inflow. However, a major difference is that in contrast to the model proposed by Hompland et al., our model could be directly applied to learn and predict voxel based image information since both predictor and response variables were present on voxel level. This allowed for successful training of a complex, non-linear model.

Different approaches have been presented to predict hypoxia based on easily available imaging modalities other than MRI. Even et al. used CT, ^{18}F-Fluorodeoxyglucose (FDG) PET and DCE-CT features to predict hypoxia as derived by HX4-PET in non-small cell lung cancer (NSCLC) [160]. Similarly to our study, a solely data-driven approach was used based on imaging in combination with supervised learning. The model was trained on clusters of voxels (supervoxels) to predict static TBR and correctly classified tumors as hypoxic (hypoxic volume (TBR > 1.2) $> 1\,\mathrm{cm}^3$) or nonhypoxic. However, larger hypoxic volumes were all underestimated which would limit the use of that method e.g., to adapt treatment. Beyond that, the method of grouping voxels to supervoxels increased the predictive performance in that study and could be implemented in our model to potentially further improve its performance.

Another method to predict hypoxia in head and neck cancer has been proposed by Crispin-Ortuzar et al. [161], who used a combination of FDG-PET and contrast-enhanced CT radiomics features as predictor variables. Good results were obtained with supervised learning to stratify patients based on predicted FMISO $\mathrm{TBR}_{\mathrm{max}}$. However, that model does not allow for prediction of virtual images or parameter maps such that no spatial information on tumor hypoxia can be collected and eventually be targeted for therapy adaptation.

The model we presented in this study was trained on ectopic xenograft tumors under preclinical laboratory conditions but might in principle be translated to a clinical setting. Factors like reduced breathing rates during imaging, breathing of anesthetic gas, small animal cardiovascular system, as well as ectopic tumor location may alter tracer kinetics as observed e.g., for dynamic FMISO-PET [162]. The model would therefore benefit from retraining on patient data before clinical validation.

In conclusion, a multi-parametric machine learning classifier was trained on small animal data for predictive modeling of dynamic FMISO-PET derived hypoxia classes by DWI and DCE-MRI and was successfully validated in an independent test cohort. The proposed method constitutes a novel approach for voxel-based learning of dynamic functional image information. Retraining the model on patient data for clinical validation seems highly promising.

6.5 Acknowledgements

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007–2013)/ERC Grant Agreement n. 335367.

6.6 Supplementary Material

Table 6.2: PET and MR image acquisition parameters.

	T2w MRI	DWI	DCE-MRI	FMISO-PET
Sequence	2D RARE	Diffusion weighted SE	2D FLASH	List mode acquisition, reconstruction using 2D-OSEM (4 it., 16 subs.)
Voxel size inplane	$0.14 \times 0.14\,\mathrm{mm}^2$	$0.30 \times 0.30\,\mathrm{mm}^2$	$0.40 \times 0.40\,\mathrm{mm}^2$	$0.65 \times 0.65\,\mathrm{mm}^2$
Slice thickness	0.4 mm	0.75 mm	1.0 mm	0.8 mm
Slice gap	–	0.25 mm	0.5 mm	–
Matrix size	256×256	96×76	75×60	128×128
Number of slices	60	21	13	89
TR/TE	5841/38 ms	1100/42 ms	72/1.5 ms	–
Flip angle	90 deg	90 deg	12 deg	–
Bandwidth	141 Hz/px	587 Hz/px	1420 Hz/px	–
Fat suppression	Yes	Yes	–	–
RARE factor	6	–	–	–
b-values	–	$0, 100, 200, 300, 400, 500, 600, 700, 800\,\mathrm{s/mm}^2$	–	–
Injected dose or contrast agent	–	–	0.02 mMol Gadovist	12.1 ± 2.1 MBq of ^{18}F-FMISO
Temporal resolution/framing	–	–	5.4 s	36×10 s, 18×60 s, 11×360 s
Acquisition time	–	–	$-1 - 12.5$ min	$0 - 90$ min p.i.
Breathing gated	Yes	Yes	–	–

Abbr.: RARE: rapid imaging with refocused echoes, also known as turbo spin-echo (TSE), SE: spin-echo, FLASH: fast low-angle shot, OSEM: ordered subset expectation maximization, TR/TE: repetition time/echo time.

Table 6.3: Hyper-parameters tuned included n_estimators, min_samples_split, min samples_leaf, max_features, max_depth, and learning_rate for GBDT, and hidden_layer_sizes, alpha, learning_rate and number of iterations for ANN where rectified linear unit function (relu) was used as activation function and adam algorithm as optimizer. Hyper-parameter naming corresponds to the implementation in scikit-learn.

	GBDT	ANN
Hyper-parameters used for search of best feature set	{min_samples_split: 1,000,	{hidden_layer_sizes: 10,
	min_samples_leaf: 100,	learning_rate_init: 1e-3,
	max_faetures: total no. of features in subset,	learning_rate: constant,
	max_depth: 1,	alpha: 1e-4,
	learning_rate: 0.1}	batch_size: 2% of samples in training fold}
Optimized hyper-parameters used for final model	{n_estimators: 98,	{hidden_layer_sizes: 50/50/50,
	min_samples_split: 10,000,	learning_rate_init: 1e-4,
	min_samples_leaf: 5,000,	learning_rate: constant,
	max_features: 3,	alpha: 5e-1,
	max_depth: 3,	batch_size: 2% of samples in training fold,
	learning_rate: 0.05}	max_iter: 1988}

Abbr.: GBDT: gradient boosted decision trees, ANN: artificial neural network, min_samples_split: minimum no. of samples to split an internal node, min_samples_leaf: minimum no. of samples to be at a leaf node, max_features: no. of features to consider when looking for the best split, max_depth: maximum depth of an estimator, hidden_layer_sizes: no. of neurons in each hidden layer, alpha: regularization parameter, batch_size: size of minibatches, max_iter: maximum no. of iterations.

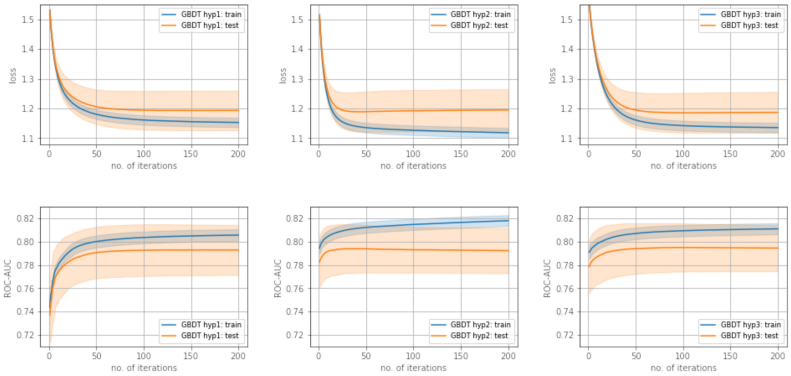

Figure 6.8: Dual learning curves of loss and ROC-AUC for GDBT. The first column shows model performance with initial hyper-parameter settings. In a first step of optimization, the maximum depth of trees was succesively increased from 1 to 3 to closer fit the model to the training data, i.e., to minimize training loss (second column). In the following, further hyper-parameters such as the maximum number of samples to be in a leaf or number of samples required to split an internal node, were tuned to regularize the model and decrease validation loss. The third column shows the learning curves for the final model with optimized settings, which are listed in detail in Supplementary Table 6.3.

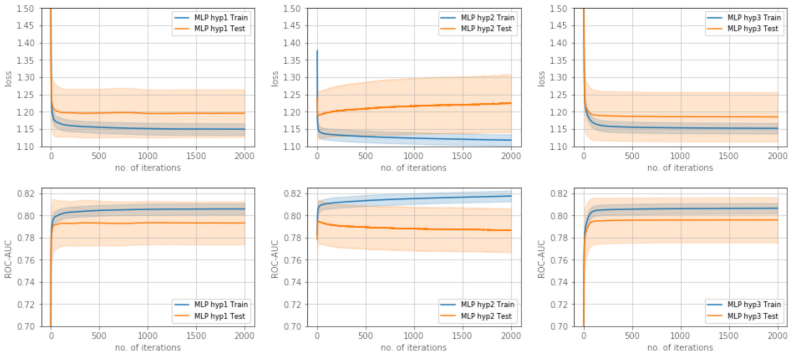

Figure 6.9: Dual learning curves of loss and ROC-AUC for ANN. The first column shows model performance with initial hyper-parameter settings. In particular, a low hidden layer size was started with, i.e., one hidden layer of 10 neurons. In a first step of optimization, the hidden layer size was successively increased to three layers of 50 neurons each in order to minimize training loss and closer fit the model to the training data (second column). In the following, L2 penalty parameter alpha was tuned to regularize the model and decrease validation loss. The third column shows the learning curves for the final model with optimized settings, which are listed in detail in Supplementary Table 6.3.

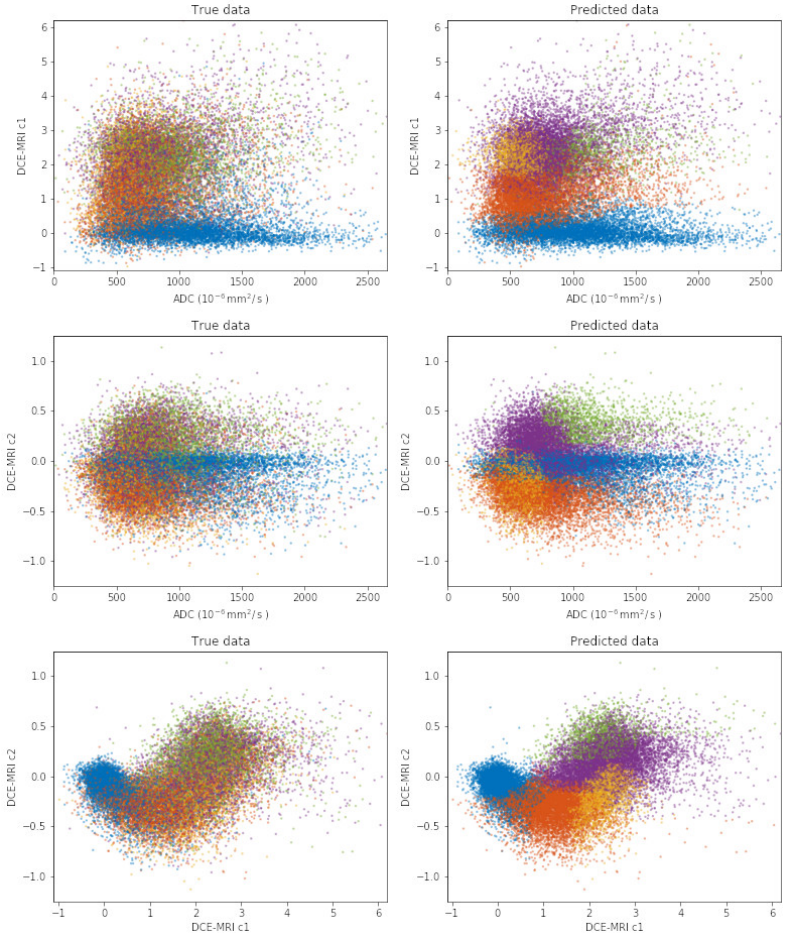

Figure 6.10: Partial dependence of FMISO-PET curve types on MRI parameters presented for the independent test cohort. Data is color coded by hypoxia class (blue: class 1, orange: class 2, yellow: class 3, purple: class 4, green: class 5). True data (left column) is presented versus predicted data (right column).

7 Concluding discussion

This work addressed three central aspects of integrating functional image data into high precision RT practice. The first was of physical-technical nature and addressed an imaging solution for combined PET/MRI of head-and-neck cancer patients in RT treatment position. The adaptation to treatment position is essential for precise image integration into RT planning, but challenging as it requires PET photon attenuation correction of the RT immobilization equipment and the MR radiofrequency coils and coil holders. Besides the image quality of attenuation corrected PET images, also the MR image quality was evaluated since the coils were used in a configuration with a greater distance to the patient so as to allow for the use of the immobilization tools. The image quality was evaluated for anatomical MRI as well as for DWI. Moreover, Dixon based attenuation maps for MR-based PET photon attenuation correction of the patient were assessed. In general, the methodology was based on a comparison of RT-specific images to diagnostic images that served as reference.

In summary, the image analysis demonstrated that PET/MRI in head-and-neck cancer patients using RT positioning aids is clinically feasible. High volume agreement was found between radiooncological structures defined on PET images acquired with RT setup and diagnostic setup, respectively. In particular, the distance measures ASSD and DOGC quantifying the discordance between the structures presented median values below the PET voxel size of 2.8 mm edge length, and below the consensus level of 2 mm accepted, for instance, as level of distortion for MR images to be used in head-and-neck treatment planning [36, 163].

A limitation was the SNR in the MR images. For T2w MRI, a reduced SNR of about -26% was observed. However, no major difference in contouring was observed between images acquired with RT, and with diagnostic reference setup: median ASSD and DOGC were, also, within the uncertainty of 2 mm, that is tolerated for MRI in RT [36, 163].

A similar reduction in SNR was observed for diffusion-weighted MR images. However, no significant difference in tumor volume-averaged ADC was found between RT-specific and diagnostic examinations. Nonetheless, further investigations should be conducted if ADC was to be processed at a voxel level.

In conclusion of the first part of this work, the image quality obtained with the RT setup meets RT requirements and the data can in general be used for future personalized therapy adaptation strategies.

The second part addressed the image quality of DWI in the head-and-neck in terms of geometric accuracy as this constitutes another essential prerequisite for using MR images in RT treatment planning. Geometric accuracy is a critical issue in EPI-based DWI, as EPI is sensitive to magnetic field inhomogeneities leading to image distortions and signal loss. This especially applies to the head-and-neck region, where unfavorable geometry and susceptibility changes at air-tissue interfaces can lead to severe geometric distortions up to a few centimeters.

An appropriate method for distortion correction was implemented based on repeated data acquisition with RPED. The results obtained in the clinical study showed, that distortions in DWI could effectively be reduced by the RPED technique. Mean ASSD and DOGC between structures repeatedly defined on distortion corrected DWI and anatomical reference images were reduced through RPED down to 0.4 and 1.0 mm, respectively. This level of remaining inaccuracy is within the uncertainty of up to 2 mm tolerated for distortions in MRI for RT of head-and-neck cancer [36, 163]. Moreover, similar to the analysis in Part I, RT mask fixation of the patient had a slightly positive effect on image quality as it helps to prevent unwanted patient movement during image acquisition.

RPED in combination with the RT imaging setup thus enables more precise integration of DWI data into RT planning for head-and-neck cancer patients and is an appropriate solution especially when sophisticated techniques for distortion reduction, such as readout-segmented, multi-shot EPI or EPI with novel integrated shim and frequency adjustment [41], are not available at a MRI system.

The third project aimed at finding a set of parameters derived from clinically readily available MRI protocols to serve as a tool for tumor characterization in

terms of treatment resistance. The strategy was to make use of supervised learning of tumor hypoxia as measured by dynamic FMISO-PET, a known indicator for radiation treatment resistance. Different machine learning models were trained to learn that information based on parameters derived from DWI and DCE-MRI. The PET/MRI data used for model development was measured in small animal models of human head and neck cancer and a final, best performing model was tested in hold out test data.

Cross-validation in the training cohort indicated good predictive performance for both model types. GBDT was favored as final model for better interpretability. Retrained GBDT yielded a mean ROC-AUC score of 0.80 in the independent test cohort confirming good generalization capacity and robustness.

As a variable to predict, dynamic FMISO-PET information was deliberately favored over different parameters derived from a static acquisition, such as tumor-to-muscle (TMR) or tumor-to-blood ratio (TBR). Some studies identified maximum TMR and TBR as prognostic measures for therapy response [51, 155]. However, results of different studies demonstrated that FMISO-PET has only prognostic value, or higher prognostic value for therapy outcome, when perfusion and retention information derived from dynamic data acquisition are both accounted for [30, 53, 156].

To address the prediction problem, GBDT was chosen as a candidate model for its capability to approximate complicated nonlinear functions and to handle a multiclass response variable. Moreover, this model most often outperforms classifiers from different families, such as K-nearest-neighbor classification, random forest or support vector machines, for a wide range of classification problems [157].

Different approaches have been presented to predict hypoxia, based on imaging modalities that are easily available in RT departments and do not rely on experimental PET tracers [160, 161]. With those models, good results were obtained in terms of classifying the tumors as hypoxic or non-hypoxic. However, in contrast to the model presented in this thesis, the methods are limited in terms of potential treatment adaptation. Larger hypoxic volumes are either underestimated in size [160], or the model does not allow for prediction of virtual images or parameter maps [161], such that no spatial information on tumor hypoxia can be collected and be used as a potential target for therapy adaptation.

To conclude, a machine learning classifier for the prediction of hypoxia-associated PET information based on multi-parametric MRI was trained and successfully tested on small animal data. The proposed method constitutes a novel approach for learning of dynamic functional image information at a voxel level. The results give new insights into the relation between dynamic hypoxia PET and functional MRI and motivate a retraining in patient data for the development of novel MRI based therapy adaptation strategies.

8 Outlook

Especially for head and neck cancer, PET and MRI may in future help improve RT in many aspects. This includes improvements in the delineation accuracy of the tumor [20, 164, 165], staging [43] or the detection of tumor recurrence after therapy [166]. In these respects, the tools presented in the first two parts of this work will certainly leverage the accurate integration of the data into RT planning if data is acquired with simultaneous PET/MRI.

Beyond that, functional image guidance will be a key factor for improving therapy outcome [110]. Here, PET/MRI will certainly play an important role in terms of characterizing the tumor and its microenvironment. Recently, a preliminary correlation analysis of functional PET and MRI parameters was conducted where the RT imaging setup and the presented methods for data correction were applied in a clinical study [167]. Similarly, more advanced multiparametric characterization of head-and-neck tumors on voxel level by means of combined PET/MRI in RT positioning could, in future, give new insights into the tumor biology for novel treatment adaptation strategies.

Regarding the RPED technique for DWI in head-and-neck, the results obtained in this work further encourage the usage of the method on a different hybrid MRI technology, the MR-Linac. With this new technology, especially with the 1.5 T MR-Linac system, daily functional imaging during RT treatment is feasible and will certainly further facilitate the development and implementation of new functional imaging based strategies for treatment adaptation. First analyses have shown that good image quality of functional MRI with a 1.5 T MR-Linac system can be achieved in phantoms [168]. However, DWI image quality in head-and-neck-cancer patients on the MR-Linac may still be compromised as compared to the phantom setting, due to the difficult patient geometry and the number of areas of susceptibility differences between different tissues present in the head and neck. Therefore, similar to DWI on a hybrid PET/MRI system, DWI on a MR-Linac system may

potentially benefit from the RPED technique for improvements in geometric accuracy. A main limitation for the implementation of RPED, however, will be the additional scan time this technique requires for repeated data acquisition.

As concerns the preclinical small animal study, the results give new insights into the relation between hypoxia PET and functional MRI and motivate a translation to a clinical setting for the development of novel MRI based therapy adaptation strategies. Therefore, a retraining of the machine learning model on patient data is required before this model can be used in clinical practice. A further confirmation of the different hypoxia associated classes found by clustering of the dynamic FMISO-PET data, e.g., with pimonidazole immunohistochemistry, could help validate the hypothesis that certain classes would be a suitable target for dose escalation. Once a biological rationale is substantiated, probability maps, obtained by the machine learning model GBDT for a voxel to belong to one of these classes, could also potentially be used for dose painting approaches.

Nonetheless, adequate measures would need to be found to help raise overall acceptance for future machine learning based treatment adaptation. Acceptance for the usage in clinical practice is challenging as it is often difficult to understand a machine learning model's process of decision making. To compensate this drawback, thorough external validation would be necessary to strengthen the evidence for a model. Notwithstanding, the issue will need to be discussed in more depth in the medical community in order for promising machine learning tools to find their way into routine clinical practice, and to improve medical image analysis along with treatment of cancer in the future.

Bibliography

[1] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians 2018; 68:394–424.

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

[3] Lacas B, Bourhis J, Overgaard J, Zhang Q, Gregoire V, Nankivell M, Zackrisson B, Szutkowski Z, Suwiński R, Poulsen M, O'Sullivan B, Corvo R, Laskar SG, Fallai C, Yamazaki H, Dobrowsky W, Cho KH, Garden AS, Langendijk JA, Viegas CMP, Hay J, Lotayef M, Parmar MKB, Auperin A, van Herpen C, Maingon P, Trotti AM, Grau C, Pignon JP, Blanchard P. Role of radiotherapy fractionation in head and neck cancers (MARCH): an updated meta-analysis. Lancet Oncol 2017; 18:1221–1237.

[4] Joiner M, Kogel Avd (Eds): Basic clinical radiobiology. London: Hodder Arnold; 4th edition 2009.

[5] Staffurth J. A Review of the Clinical Evidence for Intensity-modulated Radiotherapy. Clinical Oncology 2010; 22:643–657.

[6] Durante M, Orecchia R, Loeffler JS. Charged-particle therapy in cancer: clinical uses and future perspectives. Nature Reviews Clinical Oncology 2017; 14:483–495.

[7] Thorwarth D, Leibfarth S, Mönnich D. Potential role of PET/MRI in radiotherapy treatment planning. Clinical and Translational Imaging 2013; 1:45–51.

[8] Alterio D, Marvaso G, Ferrari A, Volpe S, Orecchia R, Jereczek-Fossa BA. Modern radiotherapy for head and neck cancer. Semin. Oncol. 2019.

[9] King AD, Thoeny HC. Functional MRI for the prediction of treatment response in head and neck squamous cell carcinoma: potential and limitations. Cancer Imaging 2016; 16.

[10] Driessen JP, Caldas-Magalhaes J, Janssen LM, Pameijer FA, Kooij N, Terhaard CHJ, Grolman W, Philippens MEP. Diffusion-weighted MR imaging in laryngeal and hypopharyngeal carcinoma: association between apparent diffusion coefficient and histologic findings. Radiology 2014; 272:456–463.

[11] Lambrecht M, Van Calster B, Vandecaveye V, De Keyzer F, Roebben I, Hermans R, Nuyts S. Integrating pretreatment diffusion weighted MRI into a multivariable prognostic model for head and neck squamous cell carcinoma. Radiotherapy and Oncology 2014; 110:429–434.

[12] Noij DP, Pouwels PJ, Ljumanovic R, Knol DL, Doornaert P, de Bree R, Castelijns JA, de Graaf P. Predictive value of diffusion-weighted imaging without and with including contrast-enhanced magnetic resonance imaging in image analysis of head and neck squamous cell carcinoma. European Journal of Radiology 2015; 84:108–116.

[13] Shukla-Dave A, Lee NY, Jansen JF, Thaler HT, Stambuk HE, Fury MG, Patel SG, Moreira AL, Sherman E, Karimi S, Wang Y, Kraus D, Shah JP, Pfister DG, Koutcher JA. Dynamic Contrast-Enhanced Magnetic Resonance Imaging as a Predictor of Outcome in Head-and-Neck Squamous Cell Carcinoma Patients With Nodal Metastases. International Journal of Radiation Oncology*Biology*Physics 2012; 82:1837–1844.

[14] Ng SH, Lin CY, Chan SC, Yen TC, Liao CT, Chang JTC, Ko SF, Wang HM, Huang SF, Lin YC, Wang JJ. Dynamic Contrast-Enhanced MR Imaging Predicts Local Control in Oropharyngeal or Hypopharyngeal Squamous Cell Carcinoma Treated with Chemoradiotherapy. PLoS ONE 2013; 8:e72230.

[15] Yankeelov TE, Gore JC. Dynamic Contrast Enhanced Magnetic Resonance Imaging in Oncology: Theory, Data Acquisition, Analysis, and Examples. Curr Med Imaging Rev 2009; 3:91–107.

[16] Min M, Lin P, Lee MT, Shon IH, Lin M, Forstner D, Bray V, Chicco A, Tieu MT, Fowler A. Prognostic role of metabolic parameters of 18f-FDG PET-CT scan performed during radiation therapy in locally advanced head and neck squamous cell carcinoma. European Journal of Nuclear Medicine and Molecular Imaging 2015; 42:1984–1994.

[17] Welz S, Mönnich D, Pfannenberg C, Nikolaou K, Reimold M, La Fougère C, Reischl G, Mauz PS, Paulsen F, Alber M, Belka C, Zips D, Thorwarth D. Prognostic value of dynamic hypoxia PET in head and neck cancer: Results from a planned interim analysis of a randomized phase II hypoxia-image guided dose escalation trial. Radiotherapy and Oncology 2017.

[18] Horsman MR, Mortensen LS, Petersen JB, Busk M, Overgaard J. Imaging hypoxia to improve radiotherapy outcome. Nature Reviews Clinical Oncology 2012; 9:674.

[19] Schmidt MA, Payne GS. Radiotherapy planning using MRI. Physics in Medicine and Biology 2015; 60:R323–R361.

[20] Kuhn FP, Hullner M, Mader CE, Kastrinidis N, Huber GF, von Schulthess GK, Kollias S, Veit-Haibach P. Contrast-Enhanced PET/MR Imaging Versus Contrast-Enhanced PET/CT in Head and Neck Cancer: How Much MR Information Is Needed? Journal of Nuclear Medicine 2014; 55:551–558.

[21] Weber WA, Schwaiger M, Avril N. Quantitative assessment of tumor metabolism using FDG-PET imaging. Nuclear Medicine and Biology 2000; 27:683–687.

[22] Bonomo P, Merlotti A, Olmetto E, Bianchi A, Desideri I, Bacigalupo A, Franco P, Franzese C, Orlandi E, Livi L, Caini S. What is the prognostic impact of FDG PET in locally advanced head and neck squamous cell carcinoma treated with concomitant chemo-radiotherapy? A systematic review and meta-analysis. Eur J Nucl Med Mol Imaging 2018; 45:2122–2138.

[23] Thorwarth D. Functional imaging for radiotherapy treatment planning: current status and future directions—a review. The British Journal of Radiology 2015; 88:20150056.

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

[25] Bentzen SM, Gregoire V. Molecular Imaging–Based Dose Painting: A Novel Paradigm for Radiation Therapy Prescription. Seminars in Radiation Oncology 2011; 21:101–110.

[26] Ling CC, Humm J, Larson S, Amols H, Fuks Z, Leibel S, Koutcher JA. Towards multidimensional radiotherapy (MD-CRT): biological imaging and biological conformality. International Journal of Radiation Oncology* Biology* Physics 2000; 47:551–560.

[27] Bentzen SM. Theragnostic imaging for radiation oncology: dose-painting by numbers. Lancet Oncol. 2005; 6:112–117.

[28] Heukelom J, Hamming O, Bartelink H, Hoebers F, Giralt J, Herlestam T, Verheij M, Brekel Mvd, Vogel W, Slevin N, Deutsch E, Sonke JJ, Lambin

P, Rasch C. Adaptive and innovative Radiation Treatment FOR improving Cancer treatment outcomE (ARTFORCE); a randomized controlled phase II trial for individualized treatment of head and neck cancer. BMC Cancer 2013; 13:84.

[29] Duprez F, De Neve W, De Gersem W, Coghe M, Madani I. Adaptive dose painting by numbers for head-and-neck cancer. Int. J. Radiat. Oncol. Biol. Phys. 2011; 80:1045–1055.

[30] Thorwarth D, Welz S, Mönnich D, Pfannenberg C, Nikolaou K, Reimold M, la Fougère C, Reischl G, Mauz PS, Paulsen F, Alber ML, Belka C, Zips D. Prospective evaluation of a tumor control probability model based on dynamic 18f-FMISO PET for head-and-neck cancer radiotherapy. J. Nucl. Med. 2019.

[31] 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. Journal of Nuclear Medicine 2009; 50:1020–1027.

[32] Wang P, Popovtzer A, Eisbruch A, Cao Y. An approach to identify, from DCE MRI, significant subvolumes of tumors related to outcomes in advanced head-and-neck cancer. Medical physics 2012; 39:5277–5285.

[33] Pichler BJ, Kolb A, Nagele T, Schlemmer HP. PET/MRI: Paving the Way for the Next Generation of Clinical Multimodality Imaging Applications. Journal of Nuclear Medicine 2010; 51:333–336.

[34] Rasmussen JH, Nørgaard M, Hansen AE, Vogelius IR, Aznar MC, Johannesen HH, Costa J, Engberg AM, Kjær A, Specht L, Fischer BM. Feasibility of Multiparametric Imaging with PET/MR in Head and Neck Squamous Cell Carcinoma. Journal of Nuclear Medicine 2017; 58:69–74.

[35] Pinker K, Andrzejewski P, Baltzer P, Polanec SH, Sturdza A, Georg D, Helbich TH, Karanikas G, Grimm C, Polterauer S, Poetter R, Wadsak W, Mitterhauser M, Georg P. Multiparametric [18f]Fluorodeoxyglucose/ [18f]Fluoromisonidazole Positron Emission Tomography/ Magnetic Resonance Imaging of Locally Advanced Cervical Cancer for the Non-Invasive Detection of Tumor Heterogeneity: A Pilot Study. PLOS ONE 2016; 11:e0155333.

[36] Paulson ES, Crijns SP, Keller BM, Wang J, Schmidt MA, Coutts G, van der Heide UA. Consensus opinion on MRI simulation for external beam radiation treatment planning. Radiotherapy and Oncology 2016; 121:187–192.

[37] Wang K, Mullins BT, Falchook AD, Lian J, He K, Shen D, Dance M, Lin W, Sills TM, Das SK, Huang BY, Chera BS. Evaluation of PET/MRI for Tumor Volume Delineation for Head and Neck Cancer. Frontiers in Oncology 2017; 7.

[38] Paulus DH, Quick HH. Hybrid Positron Emission Tomography/Magnetic Resonance Imaging: Challenges, Methods, and State of the Art of Hardware Component Attenuation Correction. Investigative radiology 2016; 51:624–634.

[39] Paulus DH, Thorwath D, Schmidt H, Quick HH. Towards integration of PET/MR hybrid imaging into radiation therapy treatment planning. Medical physics 2014; 41:072505.

[40] Skare S, Newbould RD, Clayton DB, Albers GW, Nagle S, Bammer R. Clinical multishot DW-EPI through parallel imaging with considerations of susceptibility, motion, and noise. Magnetic Resonance in Medicine 2007; 57:881–890.

[41] Gatidis S, Graf H, Weiß J, Stemmer A, Kiefer B, Nikolaou K, Notohamiprodjo M, Martirosian P. Diffusion-weighted echo planar MR imaging of the neck at 3 T using integrated shimming: comparison of MR sequence techniques for reducing artifacts caused by magnetic-field inhomogeneities. Magnetic Resonance Materials in Physics, Biology and Medicine 2017; 30:57–63.

[42] Schakel T, Hoogduin JM, Terhaard CHJ, Philippens MEP. Diffusion weighted MRI in head-and-neck cancer: geometrical accuracy. Radiother Oncol 2013; 109:394–397.

[43] Morsing A, Hildebrandt MG, Vilstrup MH, Wallenius SE, Gerke O, Petersen H, Johansen A, Andersen TL, Høilund-Carlsen PF. Hybrid PET/MRI in major cancers: a scoping review. Eur. J. Nucl. Med. Mol. Imaging 2019; 46:2138–2151.

[44] Szyszko T, Cook G. PET/CT and PET/MRI in head and neck malignancy. Clinical Radiology 2018; 73:60–69.

[45] Lee YZ, Ramalho J, Kessler B. PET–MR Imaging in Head and Neck. Magnetic Resonance Imaging Clinics of North America 2017; 25:315–324.

[46] Bailey DL, Pichler BJ, Gückel B, Antoch G, Barthel H, Bhujwalla ZM, Biskup S, Biswal S, Bitzer M, Boellaard R, Braren RF, Brendle C, Brindle K, Chiti A, la Fougère C, Gillies R, Goh V, Goyen M, Hacker M, Heukamp L, Knudsen GM, Krackhardt AM, Law I, Morris JC, Nikolaou K, Nuyts J, Ordonez AA, Pantel K, Quick HH, Riklund K, Sabri O, Sattler B, Troost EGC, Zaiss M, Zender L, Beyer T. Combined PET/MRI: Global Warming—Summary Report of the 6th International Workshop on PET/MRI, March 27–29, 2017, Tübingen, Germany. Mol Imaging Biol 2018; 20:4–20.

[47] Busk M, Horsman MR. Relevance of hypoxia in radiation oncology: pathophysiology, tumor biology and implications for treatment. Q J Nucl Med Mol Imaging 2013; 57:219–234.

[48] Nordsmark M, Bentzen SM, Rudat V, Brizel D, Lartigau E, Stadler P, Becker A, Adam M, Molls M, Dunst J, Terris DJ, Overgaard J. Prognostic value of tumor oxygenation in 397 head and neck tumors after primary radiation therapy. An international multi-center study. Radiotherapy and Oncology 2005; 77:18–24.

[49] Nordsmark M, Overgaard M, Overgaard J. Pretreatment oxygenation predicts radiation response in advanced squamous cell carcinoma of the head and neck. Radiother Oncol 1996; 41:31–39.

[50] Vordermark D, Horsman MR. Hypoxia as a Biomarker and for Personalized Radiation Oncology. Recent Results Cancer Res. 2016; 198:123–142.

[51] 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. Radiotherapy and Oncology 2012; 105:21–28.

[52] Peeters SGJA, Zegers CML, Lieuwes NG, van Elmpt W, Eriksson J, van Dongen GAMS, Dubois L, Lambin P. A Comparative Study of the Hypoxia PET Tracers [18f]HX4, [18f]FAZA, and [18f]FMISO in a Preclinical Tumor Model. International Journal of Radiation Oncology*Biology*Physics 2015; 91:351–359.

[53] Grkovski M, Lee NY, Schöder H, Carlin SD, Beattie BJ, Riaz N, Leeman JE, O'Donoghue JA, Humm JL. Monitoring early response to chemoradiotherapy with 18f-FMISO dynamic PET in head and neck cancer. Eur J Nucl Med Mol Imaging 2017; 44:1682–1691.

[54] Thorwarth D, Eschmann SM, Paulsen F, Alber M. A kinetic model for dynamic [18f]-Fmiso PET data to analyse tumour hypoxia. Phys. Med. Biol. 2005; 50:2209–2224.

[55] Thorwarth D, Eschmann SM, Paulsen F, Alber M. A Model of Reoxygenation Dynamics of Head-And-Neck Tumors Based on Serial 18f-Fluoromisonidazole Positron Emission Tomography Investigations. International Journal of Radiation Oncology*Biology*Physics 2007; 68:515–521.

[56] 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 Oncologica 2013; 52:1308–1313.

[57] Mönnich D, Troost EGC, Kaanders JHAM, Oyen WJG, Alber M, Thorwarth D. Modelling and simulation of [18f]fluoromisonidazole dynamics based on histology-derived microvessel maps. Phys. Med. Biol. 2011; 56:2045–2057.

[58] Kelly CJ, Brady M. A model to simulate tumour oxygenation and dynamic [18f]-Fmiso PET data. Phys. Med. Biol. 2006; 51:5859–5873.

[59] Andersson JL, Skare S, Ashburner J. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. NeuroImage 2003; 20:870–888.

[60] 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. Radiotherapy and Oncology 2012; 105:14–20.

[61] Zegers CML, Hoebers FJP, van Elmpt W, Bons JA, Öllers MC, Troost EGC, Eekers D, Balmaekers L, Arts-Pechtold M, Mottaghy FM, Lambin P. Evaluation of tumour hypoxia during radiotherapy using [18f]HX4 PET imaging and blood biomarkers in patients with head and neck cancer. Eur J Nucl Med Mol Imaging 2016; 43:2139–2146.

[62] Judenhofer MS, Wehrl HF, Newport DF, Catana C, Siegel SB, Becker M, Thielscher A, Kneilling M, Lichy MP, Eichner M, Klingel K, Reischl G, Widmaier S, Röcken M, Nutt RE, Machulla HJ, Uludag K, Cherry SR, Claussen CD, Pichler BJ. Simultaneous PET-MRI: a new approach for functional and morphological imaging. Nat. Med. 2008; 14:459–465.

[63] Hudson HM, Larkin RS. Accelerated image reconstruction using ordered subsets of projection data. IEEE Transactions on Medical Imaging 1994; 13:601–609.

[64] Conti M, Bendriem B. The new opportunities for high time resolution clinical TOF PET. Clin Transl Imaging 2019; 7:139–147.

[65] Soderlund AT, Chaal J, Tjio G, Totman JJ, Conti M, Townsend DW. Beyond 18f-FDG: Characterization of PET/CT and PET/MR Scanners for a Comprehensive Set of Positron Emitters of Growing Application—18f, 11c, 89zr, 124i, 68ga, and 90y. J Nucl Med 2015; 56:1285–1291.

[66] Jakoby BW, Bercier Y, Conti M, Casey ME, Bendriem B, Townsend DW. Physical and clinical performance of the mCT time-of-flight PET/CT scanner. Phys. Med. Biol. 2011; 56:2375–2389.

[67] 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.

[68] Hallen P, Schug D, Weissler B, Gebhardt P, Salomon A, Kiessling F, Schulz V. PET performance evaluation of the small-animal Hyperion II ^{d} PET/MRI insert based on the NEMA NU-4 standard. Biomed. Phys. Eng. Express 2018; 4:065027.

[69] Moses WW. Fundamental Limits of Spatial Resolution in PET. Nucl Instrum Methods Phys Res A 2011; 648 Supplement 1:S236–S240.

[70] Raylman RR, Hammer BE, Christensen NL. Combined MRI-PET scanner: a Monte Carlo evaluation of the improvements in PET resolution due to the effects of a static homogeneous magnetic field. IEEE Transactions on Nuclear Science 1996; 43:2406–2412.

[71] Bushberg JT, Seibert JA, Leidholdt EM, Boone JM: The Essential Physics of Medical Imaging. Lippincott Williams & Wilkins 2011.

[72] Berger M, Hubbell J, Seltzer S, Chang J, Coursey J, Sukumar R, DS Z, Olsen K. XCOM: Photon Cross Sections Database (version 1.5). [Online] Available: http://physics.nist.gov/xcom. [Accessed on 2019, September 4]. National Institute of Standards and Technology, Gaithersburg, MD 2010.

[73] Hubbell J, Seltzer S. Tables of X-Ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients (version 1.4). [Online] Available: http://physics.nist.gov/xaamdi. [Accessed on 2019, September 4]. National Institute of Standards and Technology, Gaithersburg, MD 2004.

[74] Krieger H: Grundlagen der Strahlungsphysik und des Strahlenschutzes. Vieweg+Teubner; 4 edition 2012.

[75] Guerra AD: Ionizing Radiation Detectors for Medical Imaging. World Scientific 2004. [Chapter 8.5.2 Scattering].

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

[77] Kinahan PE, Townsend DW, Beyer T, Sashin D. Attenuation correction for a combined 3d PET/CT scanner. Medical Physics 1998; 25:2046–2053.

[78] Paulus DH, Tellmann L, Quick HH. Towards improved hardware component attenuation correction in PET/MR hybrid imaging. Physics in Medicine and Biology 2013; 58:8021–8040.

[79] Boellaard R, Hofman MBM, Hoekstra OS, Lammertsma AA. Accurate PET/MR quantification using time of flight MLAA image reconstruction. Mol Imaging Biol 2014; 16:469–477.

[80] Mehranian A, Arabi H, Zaidi H. Vision 20/20: Magnetic resonance imaging-guided attenuation correction in PET/MRI: Challenges, solutions, and opportunities. Medical Physics 2017; :1130–1155.

[81] Chen Y, An H. Attenuation Correction of PET/MR Imaging. Magn Reson Imaging Clin N Am 2017; 25:245–255.

[82] 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.

[83] Sekine T, Buck A, Delso G, Voert EEGWt, Huellner M, Veit-Haibach P, Warnock G. Evaluation of Atlas-Based Attenuation Correction for Integrated PET/MR in Human Brain: Application of a Head Atlas and Comparison to True CT-Based Attenuation Correction. J Nucl Med 2016; 57:215–220.

[84] Martinez-Moller 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. Journal of Nuclear Medicine 2009; 50:520–526.

[85] Berker Y, Franke J, Salomon A, Palmowski M, Donker HCW, Temur Y, Mottaghy FM, Kuhl C, Izquierdo-Garcia D, Fayad ZA, Kiessling F, Schulz V. MRI-Based Attenuation Correction for Hybrid PET/MRI Systems: A 4-Class Tissue Segmentation Technique Using a Combined Ultrashort-Echo-Time/Dixon MRI Sequence. J Nucl Med 2012; 53:796–804.

[86] Keereman V, Fierens Y, Broux T, Deene YD, Lonneux M, Vandenberghe S. MRI-Based Attenuation Correction for PET/MRI Using Ultrashort Echo Time Sequences. J Nucl Med 2010; 51:812–818.

[87] Sekine T, Voert EEGWt, Warnock G, Buck A, Huellner M, Veit-Haibach P, Delso G. Clinical Evaluation of Zero-Echo-Time Attenuation Correction for Brain 18f-FDG PET/MRI: Comparison with Atlas Attenuation Correction. J Nucl Med 2016; 57:1927–1932.

[88] Freitag MT, Fenchel M, Bäumer P, Heußer T, Rank CM, Kachelrieß M, Paech D, Kopka K, Bickelhaupt S, Dimitrakopoulou-Strauss A, Maier-Hein K, Floca R, Ladd ME, Schlemmer HP, Maier F. Improved clinical workflow for simultaneous whole-body PET/MRI using high-resolution CAIPIRINHA-accelerated MR-based attenuation correction. European Journal of Radiology 2017; 96:12–20.

[89] Ladefoged C, Benoit D, Law I, Holm S, Hojgaard L, Hansen AE, Andersen FL. PET/MR attenuation correction in brain imaging using a continuous bone signal derived from UTE. EJNMMI Phys 2015; 2.

[90] 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.

[91] Dixon WT. Simple proton spectroscopic imaging. Radiology 1984; 153:189–194.

[92] Ma J. Dixon techniques for water and fat imaging. Journal of Magnetic Resonance Imaging 2008; 28:543–558.

[93] Tofts PS, Brix G, Buckley DL, Evelhoch JL, Henderson E, Knopp MV, Larsson HBW, Lee TY, Mayr NA, Parker GJM, Port RE, Taylor J, Weisskoff RM. Estimating kinetic parameters from dynamic contrast-enhanced t1-weighted MRI of a diffusable tracer: Standardized quantities and symbols. Journal of Magnetic Resonance Imaging 1999; 10:223–232.

[94] Brix G, Griebel J, Kiessling F, Wenz F. Tracer kinetic modelling of tumour angiogenesis based on dynamic contrast-enhanced CT and MRI measurements. Eur J Nucl Med Mol Imaging 2010; 37:30–51.

[95] Padhani AR, Liu G, Mu-Koh D, Chenevert TL, Thoeny HC, Takahara T, Dzik-Jurasz A, Ross BD, Van Cauteren M, 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.

[96] Le Bihan D, Breton E, Lallemand D, Aubin ML, Vignaud J, Laval-Jeantet M. Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. Radiology 1988; 168:497–505.

[97] Hompland T, Hole KH, Ragnum HB, Aarnes EK, Vlatkovic L, Lie AK, Patzke S, Brennhovd B, Seierstad T, Lyng H. Combined MR Imaging of Oxygen Consumption and Supply Reveals Tumor Hypoxia and Aggressiveness in Prostate Cancer Patients. Cancer Res 2018; 78:4774–4785.

[98] Leibfarth S, Winter RM, Lyng H, Zips D, Thorwarth D. Potentials and challenges of diffusion-weighted magnetic resonance imaging in radiotherapy. Clin Transl Radiat Oncol 2018; 13:29–37.

[99] Jezzard P, Balaban RS. Correction for geometric distortion in echo planar images from B0 field variations. Magnetic Resonance in Medicine 1995; 34:65–73.

[100] Jezzard P, Clare S. Sources of distortion in functional MRI data. Hum Brain Mapp 1999; 8:80–85.

[101] Holland D, Kuperman JM, Dale AM. Efficient Correction of Inhomogeneous Static Magnetic Field-Induced Distortion in Echo Planar Imaging. Neuroimage 2010; 50:175.

[102] Graham MS, Drobnjak I, Jenkinson M, Zhang H. Quantitative assessment of the susceptibility artefact and its interaction with motion in diffusion MRI. PLOS ONE 2017; 12:e0185647.

[103] Jollive IT: Principle Component Analysis. New York: Springer New York; 2 edition 2002. [Chapter 1 Introduction].

[104] The MathWorks, Inc. K-means clustering - MATLAB kmeans. [Online] Available: http://physics.nist.gov/xcom. [Accessed on 2019, November 5]. Mathworks R2019b Documentation 2019; https://www.mathworks.com/help/stats/kmeans.html#bues3lh.

[105] Hastie T, Tibshirani R, Friedman J: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition. Springer Series in Statistics; New York: Springer New York; 2 edition 2009. [Chapter 10 Boosting and Additive Trees].

[106] Hastie T, Tibshirani R, Friedman J: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition. Springer Series in Statistics; New York: Springer New York; 2 edition 2009. [Chapter 9 Additive Models, Trees, and Related Methods].

[107] James G, Witten D, Hastie T, Tibshirani R: An Introduction to Statistical Learning; *Volume 103 of* Springer Texts in Statistics. New York: Springer New York 2013. [Chapter 8 Tree-Based Methods].

[108] Hastie T, Tibshirani R, Friedman J: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition. Springer Series in Statistics; New York: Springer New York; 2 edition 2009. [Chapter 11 Neural Networks].

[109] Pereira GC, Traughber M, Muzic RF. The Role of Imaging in Radiation Therapy Planning: Past, Present, and Future. BioMed Research International 2014; 2014:1–9.

[110] Baumann M, Krause M, Overgaard J, Debus J, Bentzen SM, Daartz J, Richter C, Zips D, Bortfeld T. Radiation oncology in the era of precision medicine. Nature Reviews Cancer 2016; 16:234–249.

[111] Lagendijk JJW, Raaymakers BW, Van den Berg CAT, Moerland MA, Philippens ME, van Vulpen M. MR guidance in radiotherapy. Physics in Medicine and Biology 2014; 59:R349–R369.

[112] Thorwarth D, Eschmann SM, Paulsen F, Alber M. Hypoxia Dose Painting by Numbers: A Planning Study. International Journal of Radiation Oncology*Biology*Physics 2007; 68:291–300.

[113] Leibfarth S, Simoncic U, Mönnich D, Welz S, Schmidt H, Schwenzer N, Zips D, Thorwarth D. Analysis of pairwise correlations in multi-parametric PET/MR data for biological tumor characterization and treatment individualization strategies. European Journal of Nuclear Medicine and Molecular Imaging 2016; 43:1199–1208.

[114] Hanvey S, McJury M, Tho LM, Glegg M, Thomson M, Grose D, James A, Rizwanullah M, Paterson C, Foster J. The influence of MRI scan position on patients with oropharyngeal cancer undergoing radical radiotherapy. Radiation Oncology 2013; 8:129.

[115] Fortunati V, Verhaart RF, Verduijn GM, van der Lugt A, Angeloni F, Niessen WJ, Veenland JF, Paulides MM, van Walsum T. MRI integration into treatment planning of head and neck tumors: Can patient immobilization be avoided? Radiotherapy and Oncology 2015; 115:191–194.

[116] 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. International Journal of Radiation Oncology*Biology*Physics 2009; 74:630–636.

[117] Hanvey S, Glegg M, Foster J. Magnetic resonance imaging for radiotherapy planning of brain cancer patients using immobilization and surface coils. Physics in Medicine and Biology 2009; 54:5381–5394.

[118] Winter RM, Schmidt H, Leibfarth S, Zwirner K, Welz S, Schwenzer NF, la Fougère C, Nikolaou K, Gatidis S, Zips D, Thorwarth D. Distortion correction of diffusion-weighted magnetic resonance imaging of the head and neck in radiotherapy position. Acta Oncol 2017; :1–4.

[119] Klein S, Staring M, Murphy K, Viergever M, Pluim J. elastix: A Toolbox for Intensity-Based Medical Image Registration. IEEE Transactions on Medical Imaging 2010; 29:196–205.

[120] Lee JA. Segmentation of positron emission tomography images: Some recommendations for target delineation in radiation oncology. Radiotherapy and Oncology 2010; 96:302–307.

[121] Heverhagen JT. Noise Measurement and Estimation in MR Imaging Experiments. Radiology 2007; 245:638–639.

[122] Hoang JK, Choudhury KR, Chang J, Craciunescu OI, Yoo DS, Brizel DM. Diffusion-Weighted Imaging for Head and Neck Squamous Cell Carcinoma: Quantifying Repeatability to Understand Early Treatment-Induced Change. American Journal of Roentgenology 2014; 203:1104–1108.

[123] 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 Oncologica 2013; 52:1353–1359.

[124] Paulus DH, Oehmigen M, Grueneisen J, Umutlu L, Quick HH. Whole-body hybrid imaging concept for the integration of PET/MR into radiation therapy treatment planning. Physics in Medicine and Biology 2016; 61:3504–3520.

[125] Edmund JM, Nyholm T. A review of substitute CT generation for MRI-only radiation therapy. Radiation Oncology 2017; 12.

[126] Kellman P, McVeigh ER. Image Reconstruction in SNR Units: A General Method for SNR Measurement. Magn Reson Med 2005; 54:1439–1447.

[127] Dietrich O, Raya JG, Reeder SB, Reiser MF, Schoenberg SO. Measurement of signal-to-noise ratios in MR images: Influence of multichannel coils, parallel imaging, and reconstruction filters. J. Magn. Reson. Imaging 2007; 26:375–385.

[128] Hamilton C, Ebert M. Volumetric Uncertainty in Radiotherapy. Clinical Oncology 2005; 17:456–464.

[129] Doshi T, Wilson C, Paterson C, Lamb C, James A, MacKenzie K, Soraghan J, Petropoulakis L, Di Caterina G, Grose D. Validation of a Magnetic Resonance Imaging-based Auto-contouring Software Tool for Gross Tumour Delineation in Head and Neck Cancer Radiotherapy Planning. Clinical Oncology 2017; 29:60–67.

[130] 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. European Radiology 2014; 24:3088–3096.

[131] Windischberger C, Robinson S, Rauscher A, Barth M, Moser E. Robust field map generation using a triple-echo acquisition. Journal of Magnetic Resonance Imaging 2004; 20:730–734.

[132] Vandecaveye V, Dirix P, De Keyzer F, Op de Beeck K, Vander Poorten V, Roebben I, Nuyts S, Hermans R. Predictive value of diffusion-weighted magnetic resonance imaging during chemoradiotherapy for head and neck squamous cell carcinoma. European Radiology 2010; 20:1703–1714.

[133] King AD, Chow KK, Yu KH, Mo FKF, Yeung DKW, Yuan J, Bhatia KS, Vlantis AC, Ahuja AT. Head and Neck Squamous Cell Carcinoma: Diagnostic Performance of Diffusion-weighted MR Imaging for the Prediction of Treatment Response. Radiology 2013; 266:531–538.

[134] Leibfarth, S and Eckert, F and Welz, S and Siegel, C and Schmidt, H and Schwenzer, N and Zips, D and Thorwarth, D. Automatic delineation of tumor volumes by co-segmentation of combined PET/MR data. Physics in Medicine and Biology 2015; 60:5399.

[135] Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM. Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage 2004; 23:S208–S219.

[136] Busk M, Mortensen LS, Nordsmark M, Overgaard J, Jakobsen S, Hansen KV, Theil J, Kallehauge JF, D'Andrea FP, Steiniche T, Horsman MR. PET hypoxia imaging with FAZA: reproducibility at baseline and during fractionated radiotherapy in tumour-bearing mice. Eur J Nucl Med Mol Imaging 2013; 40:186–197.

[137] Servagi-Vernat S, Differding S, Sterpin E, Hanin FX, Labar D, Bol A, Lee JA, Grégoire V. Hypoxia-guided adaptive radiation dose escalation in head and neck carcinoma: A planning study. Acta Oncologica 2015; 54:1008–1016.

[138] Lee N, Schoder H, Beattie B, Lanning R, Riaz N, McBride S, Katabi N, Li D, Yarusi B, Chan S, Mitrani L, Zhang Z, Pfister DG, Sherman E, Baxi S, Boyle J, Morris LGT, Ganly I, Wong R, Humm J. Strategy of Using Intratreatment Hypoxia Imaging to Selectively and Safely Guide Radiation Dose De-escalation Concurrent With Chemotherapy for Locoregionally Advanced Human Papillomavirus–Related Oropharyngeal Carcinoma. International Journal of Radiation Oncology*Biology*Physics 2016; 96:9–17.

[139] Simoncic U, Leibfarth S, Welz S, Schwenzer N, Schmidt H, Reischl G, Pfannenberg C, Fougère Cl, Nikolaou K, Zips D, Thorwarth D. Comparison of DCE-MRI kinetic parameters and FMISO-PET uptake parameters in head and neck cancer patients. Med. Phys. 2017; 44:2358–2368.

[140] Jansen JFA, 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. International Journal of Radiation Oncology*Biology*Physics 2010; 77:1403–1410.

[141] Donaldson SB, Betts G, Bonington SC, Homer JJ, Slevin NJ, Kershaw LE, Valentine H, West CML, 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. International Journal of Radiation Oncology*Biology*Physics 2011; 81:1176–1183.

[142] 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. International Journal of Radiation Oncology*Biology*Physics 2009; 74:29–37.

[143] Even AJG, Reymen B, La Fontaine MD, Das M, Mottaghy FM, Belderbos JSA, De Ruysscher D, Lambin P, van Elmpt W. Clustering of multiparametric functional imaging to identify high-risk subvolumes in non-small cell lung cancer. Radiotherapy and Oncology 2017; 125:379–384.

[144] Giraud P, Giraud P, Gasnier A, El Ayachy R, Kreps S, Foy JP, Durdux C, Huguet F, Burgun A, Bibault JE. Radiomics and Machine Learning for Radiotherapy in Head and Neck Cancers. Front Oncol 2019; 9:174.

[145] Sahiner B, Pezeshk A, Hadjiiski LM, Wang X, Drukker K, Cha KH, Summers RM, Giger ML. Deep learning in medical imaging and radiation therapy. Medical Physics 2019; 46:e1–e36.

[146] Wehrl HF, Hossain M, Lankes K, Liu CC, Bezrukov I, Martirosian P, Schick F, Reischl G, Pichler BJ. Simultaneous PET-MRI reveals brain function in activated and resting state on metabolic, hemodynamic and multiple temporal scales. Nat. Med. 2013; 19:1184–1189.

[147] Featherstone AK, O'Connor JP, Little RA, Watson Y, Cheung S, Babur M, Williams KJ, Matthews JC, Parker GJ. Data-driven mapping of hypoxia-related tumor heterogeneity using DCE-MRI and OE-MRI. Magn Reson Med 2018; 79:2236–2245.

[148] Eyal E, Badikhi D, Furman-Haran E, Kelcz F, Kirshenbaum KJ, Degani H. Principal Component Analysis of breast DCE-MRI Adjusted with a Model Based Method. J Magn Reson Imaging 2009; 30:989–998.

[149] Farjam R, Tsien CI, Lawrence TS, Cao Y. DCE-MRI defined subvolumes of a brain metastatic lesion by principle component analysis and fuzzy-c-means clustering for response assessment of radiation therapy. Med Phys 2014; 41.

[150] Hapdey S, Buvat I, Carson JM, Carrasquillo JA, Whatley M, Bacharach SL. Searching for Alternatives to Full Kinetic Analysis in 18f-FDG PET: An Extension of the Simplified Kinetic Analysis Method. J Nucl Med 2011; 52:634–641.

[151] Campbell MC, Markham J, Flores H, Hartlein JM, Goate AM, Cairns NJ, Videen TO, Perlmutter JS. Principal component analysis of PiB distribution in Parkinson and Alzheimer diseases. Neurology 2013; 81:520–527.

[152] Eyal E, Degani H. Model-based and model-free parametric analysis of breast dynamic-contrast-enhanced MRI. NMR in Biomedicine 2009; 22:40–53.

[153] Joint Head and Neck Radiotherapy-MRI Development Cooperative, Elhalawani H, Ger RB, Mohamed ASR, Awan MJ, Ding Y, Li K, Fave XJ, Beers AL, Driscoll B, Hormuth Ii DA, van Houdt PJ, He R, Zhou S, Mathieu KB, Li H, Coolens C, Chung C, Bankson JA, Huang W, Wang J, Sandulache VC, Lai SY, Howell RM, Stafford RJ, Yankeelov TE, van der Heide UA, Frank SJ, Barboriak DP, Hazle JD, Court LE, Kalpathy-Cramer J, Fuller CD. Dynamic contrast-enhanced magnetic resonance imaging for head and neck cancers. Scientific Data 2018; 5:180008.

[154] Mönnich D, Welz S, Thorwarth D, Pfannenberg C, Reischl G, Mauz PS, Nikolaou K, Fougère Cl, Zips D. Robustness of quantitative hypoxia PET image analysis for predicting local tumor control. Acta Oncologica 2015; 54:1364–1369.

[155] Rajendran JG, Schwartz DL, O'Sullivan J, Peterson LM, Ng P, Scharnhorst J, Grierson JR, Krohn KA. Tumor Hypoxia Imaging with [F-18] Fluoromisonidazole Positron Emission Tomography in Head and Neck Cancer. Clin Cancer Res 2006; 12:5435–5441.

[156] 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.

[157] Olson RS, La Cava W, Mustahsan Z, Varik A, Moore JH. Data-driven advice for applying machine learning to bioinformatics problems. Pac Symp Biocomput 2018; 23:192–203.

[158] O'Connor JPB, Robinson SP, Waterton JC. Imaging tumour hypoxia with oxygen-enhanced MRI and BOLD MRI. BJR 2019; 92:20180642.

[159] O'Connor JPB, Aboagye EO, Adams JE, Aerts HJWL, Barrington SF, Beer AJ, Boellaard R, Bohndiek SE, Brady M, Brown G, Buckley DL, Chenevert TL, Clarke LP, Collette S, Cook GJ, deSouza NM, Dickson JC, Dive C, Evelhoch JL, Faivre-Finn C, Gallagher FA, Gilbert FJ, Gillies RJ, Goh V, Griffiths JR, Groves AM, Halligan S, Harris AL, Hawkes DJ, Hoekstra OS, Huang EP, Hutton BF, Jackson EF, Jayson GC, Jones A, Koh DM, Lacombe D, Lambin P, Lassau N, Leach MO, Lee TY, Leen EL, Lewis JS, Liu Y,

Lythgoe MF, Manoharan P, Maxwell RJ, Miles KA, Morgan B, Morris S, Ng T, Padhani AR, Parker GJM, Partridge M, Pathak AP, Peet AC, Punwani S, Reynolds AR, Robinson SP, Shankar LK, Sharma RA, Soloviev D, Stroobants S, Sullivan DC, Taylor SA, Tofts PS, Tozer GM, van Herk M, Walker-Samuel S, Wason J, Williams KJ, Workman P, Yankeelov TE, Brindle KM, McShane LM, Jackson A, Waterton JC. Imaging biomarker roadmap for cancer studies. Nature Reviews Clinical Oncology 2017; 14:169–186.

[160] Even AJG, Reymen B, Fontaine MDL, Das M, Jochems A, Mottaghy FM, Belderbos JSA, Ruysscher DD, Lambin P, Elmpt Wv. Predicting tumor hypoxia in non-small cell lung cancer by combining CT, FDG PET and dynamic contrast-enhanced CT. Acta Oncologica 2017; 56:1591–1596.

[161] Crispin-Ortuzar M, Apte A, Grkovski M, Oh JH, Lee NY, Schöder H, Humm JL, Deasy JO. Predicting hypoxia status using a combination of contrast-enhanced computed tomography and [18f]-Fluorodeoxyglucose positron emission tomography radiomics features. Radiotherapy and Oncology 2018; 127:36–42.

[162] Kersemans V, Cornelissen B, Hueting R, Tredwell M, Hussien K, Allen PD, Falzone N, Hill SA, Dilworth JR, Gouverneur V, Muschel RJ, Smart SC. Hypoxia Imaging Using PET and SPECT: The Effects of Anesthetic and Carrier Gas on [64cu]-ATSM, [99mtc]-HL91 and [18f]-FMISO Tumor Hypoxia Accumulation. PLoS One 2011; 6.

[163] van der Heide UA, Frantzen-Steneker M, Astreinidou E, Nowee ME, van Houdt PJ. MRI basics for radiation oncologists. Clin Transl Radiat Oncol 2019; 18:74–79.

[164] Jager EA, Ligtenberg H, Caldas-Magalhaes J, Schakel T, Philippens ME, Pameijer FA, Kasperts N, Willems SM, Terhaard CH, Raaijmakers CP. Validated guidelines for tumor delineation on magnetic resonance imaging for laryngeal and hypopharyngeal cancer. Acta Oncol 2016; 55:1305–1312.

[165] 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.

[166] Driessen JP, Peltenburg B, Philippens MEP, Huijbregts JE, Pameijer FA, de Bree R, Janssen LM, Terhaard CHJ. Prospective comparative study of MRI including diffusion-weighted images versus FDG PET-CT for the detection of recurrent head and neck squamous cell carcinomas after (chemo)radiotherapy. European Journal of Radiology 2019; 111:62–67.

[167] Zwirner K, Thorwarth D, Winter RM, Welz S, Weiss J, Schwenzer NF, Schmidt H, la Fougère C, Nikolaou K, Zips D, Gatidis S. Voxel-wise correlation of functional imaging parameters in HNSCC patients receiving PET/MRI in an irradiation setup. Strahlenther Onkol 2018; 194:719–726.

[168] Kooreman ES, van Houdt PJ, Nowee ME, van Pelt VWJ, Tijssen RHN, Paulson ES, Gurney-Champion OJ, Wang J, Koetsveld F, van Buuren LD, ter Beek LC, van der Heide UA. Feasibility and accuracy of quantitative imaging on a 1.5 T MR-linear accelerator. Radiotherapy and Oncology 2019; 133:156–162.