Aus der

Universitätsklinik für Radioonkologie mit Poliklinik Tübingen
Sektion Biomedizinische Physik

Quantitative MR imaging for functional characterization in MR-guided radiotherapy of head and neck cancer

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Humanwissenschaften

der Medizinischen Fakultät
der Eberhard Karls Universität
zu Tübingen

vorgelegt von
Habrich, Jonas Norbert
2025

Dekan:	Professor Dr. B. Pichler
1. Berichterstatter:	Professorin Dr. D. Thorwarth
2. Berichterstatter:	Professor Dr. K. Scheffler
Tag der Disputation:	11.04.2025

Contents

Figures

Tables

Abbreviations

^{18}F-FDG	^{18}F-fluorodeoxyglucose
^{18}F-FMISO	^{18}F-fluoromisonidazole
ADC	apparent diffusion coefficient
CT	computed tomography
DCE	dynamic contrast-enhanced
DFG	German Research Council
DNA	deoxyribonucleic acid
DPBC	dose painting by contours
DPBN	dose painting by numbers
DTI	diffusion tensor imaging
DW	diffusion-weighted
EPI	echo-planar imaging
ERE	electron return effect
GS	Gleason Score
GTV	gross tumor volume
HNC	head and neck cancer
HNSCC	head and neck squamous cell carcinoma
HPV	human papillomavirus

HRS	high-risk subvolume
IB	imaging biomarker
ICC	intraclass correlation coefficient
IGRT	image guided radiotherapy
IMRT	intensity modulated radiotherapy
IVIM	intravoxel incoherent motion
MRgRT	magnetic resonance-guided radiotherapy
MRI	magnetic resonance imaging
MS	multi-shot
OAR	organs-at-risk
OE	oxygen-enhanced
PET	positron emission tomography
PSMA	prostate-specific membrane antigen
PVP	polyvinylpyrrolidone
QIB	quantitative imaging biomarker
QIBA	Quantitative Imaging Biomarkers Alliance
RC	repeatability coefficient
RDC	reproducibility coefficient
relRC	relative repeatability coefficient
RF	radiofrequency
SE	spin-echo
SNR	signal-to-noise ratio
SPECT	single photon emission computed tomography

SPLICE	split acquisition of fast spin echo signals
TSE	turbo-spin-echo
VMAT	volumetric intensity modulated arc therapy
wCV	within-subject coefficient of variation
wSD	within-subject standard deviation

1 Introduction, Aims and Objectives of the thesis

1.1 Introduction

Cancer is one of the most prominent and deadliest diseases for humanity with an estimated incidence of 20 million cases in the year 2022 alone [24]. However, due to the increasing life expectancy, cancer will become more important in the foreseeable future with around 35 million new cases of cancer predicted for the year 2035 [24]. Therefore, great efforts are undertaken in order to improve the diagnostics and therapy of cancer resulting in earlier diagnosis and better prognosis for many patients.

One of the cornerstones in the treatment of cancer is radiotherapy which can be used as a standalone technique for curative or palliative approaches, but also in combination with other treatment modalities such as surgery, chemotherapy or immunotherapy. In the last years, radiotherapy has been undergoing a transition to more and more focused therapy approaches. Beginning with simple treatment plans consisting of two opposing fields over the usage of three-dimensional (3D) conformal treatment plans to intensity modulated radiotherapy (IMRT) and volumetric intensity modulated arc therapy (VMAT), patient outcome undergoing radiotherapy has majorly improved [89, 173]. The integration of different imaging modalities like computed tomography (CT) or magnetic resonance imaging (MRI) not only into radiotherapy treatment planning, but also into treatment adaptation just before radiation delivery, known as image guided radiotherapy (IGRT), marked another huge step towards the best possible patient treatment. Nevertheless, different tumor sites like pancreas [225] or the head and neck region [229] still present a challenge for radiotherapy with 29-40% locoregional recurrence rates, indicating that further improvement of radiotherapy is indispensable.

One of the latest strategies to even further enhance radiotherapy outcome is the personalization of radiotherapy with different aspects being investigated. For once, the usage of biomarkers for radiotherapy in order to gather information about the probable prognosis of patients for their prescribed treatment as well as the stratification into different risk groups is desired, enabling the application of adapted treatment strategies based on the patient's risk status [178]. Clinically already established biomarkers are the prostate-specific membrane antigen (PSMA) level, evaluated through positron emission tomography (PET), and the Gleason Score (GS) derived from tissue biopsies to stratify patients with prostate cancer into low-, intermediate- and high-risk patients and adapting their radiotherapy treatment strategy thereon [42]. For head and neck cancer (HNC), the human papillomavirus (HPV) p16 status as well as the patient's smoking history have been identified as prognostic factors for radiotherapy outcome [7, 111].

In recent years the focus of finding and establishing biomarkers for radiotherapy shifted towards quantitative imaging biomarkers (QIBs) derived from multimodal imaging like PET or MRI due to major improvements in image quality and scanner availability. Additionally, imaging QIBs provide 3D spatial resolution which is an important prerequisite for future radiotherapy treatment individualization. While anatomical CT and MRI were already clinically used in radiotherapy treatment planning, an increasing number of studies have shown that functional imaging, which can characterize biological and physiological processes of the human body, might provide information about treatment outcome and early-response assessment [6, 66]. Different PET tracers like ^{18}F-fluoromisonidazole (^{18}F-FMISO) or ^{18}F-fluorodeoxyglucose (^{18}F-FDG) yielded promising results in terms of predicting tumor response and patient outcome in various tumor sites [4, 78, 171, 201, 209], but due to additional radiation doses for the patients and huge logistic efforts needed for PET imaging, the spotlight turned towards QIBs derived from functional MRI. Here, a wide variety of imaging techniques and thereof derived QIBs are available. Dynamic contrast-enhanced (DCE)-MRI evaluates the dispersion of an injected contrast medium and allows the quantification of tumor vascularity [66, 158] which has been correlated to patient response and survival [55, 72]. Another QIB is ΔR_2* derived from oxygen-enhanced (OE)-MRI, determined by providing a patient with 100% oxygen and measuring the change in tumoral longitudinal relaxation time T_1 [157], but its prognostic value has yet to be shown in larger patient studies. One of the most promising QIBs is the apparent diffusion coefficient (ADC) derived from diffusion-weighted (DW)-MRI because the measured signal is independent from a scanner's magnetic field

strength, no contrast medium is needed and measurement times are relatively short [66]. Through the acquisition of two opposing gradient pulses, the motion of water molecules can be quantified which is a surrogate for cell density or size [66, 158] and has been correlated to patient outcome and early-response in different tumor sites [68, 75, 109, 148].

Although MRI acquisitions are part of the routine radiotherapy treatment planning process, additional scans over the course of radiotherapy still demand high logistic efforts to monitor patient response to radiotherapy with MRI-based QIBs. Here, the clinical introduction of hybrid MR-Linacs, combining a conventional electron linear accelerator and a MRI scanner, featured a paradigm shift. MR-Linacs do not only allow for daily adaptation of a patient's radiotherapy treatment plan to positioning errors or anatomical shifts, but also the acquisition of daily quantitative MRI is now easily feasible [175, 176]. However, the combination of two high technologically developed systems implies the drawback that only a limited hardware for the MRI system is available (cf. Section 1.7). Therefore, all quantitative imaging techniques and thereof derived QIBs must be technically validated before QIBs can be clinically introduced. In the Department of Radiation Oncology Tübingen an 1.5 T MR-Linac was installed and first patient treatments started in 2018. One of the main research areas of the department is the identification and integration of QIBs from functional imaging into radiotherapy to personalize patient treatment and potentially improve radiotherapy outcome in the future. In this context the acquisition of ADC values derived by DW-MRI on the 1.5 Tesla MR-Linac should be technically validated as part of a research grant by the German Research Council (DFG).

1.2 Head and neck cancer

In the year 2020, approximately 870.000 people were diagnosed with HNC worldwide [195], with 90% of all cases presenting as head and neck squamous cell carcinoma (HNSCC) [125]. Although other tumor sites present with considerably higher incidences, HNC remains of high medical interest as the incidence of HNC is projected to increase by 40% in the year 2040 [128]. Additionally, HNC patients show moderate 5-year survival rates of 68% for tumors in the oral cavity and pharynx and 61% in the larynx [186] while prognosis for patients with hypopharyngeal tumors is even worse with 25% relative survival [53]. Late diagnosis, resistance to different

therapies, frequent recurrent metastasis and relapses are causing the poor prognosis of HNC patients [16]. Additionally, various risk factors play a pivotal role for the development and progression of HNC including tobacco and alcohol abuse, oral health and HPV status [127].

Different treatment strategies are available for HNC patients, but for early stage carcinomas (stage I and II) surgery or radiotherapy are the preferred options. However, about 60% of patients present with locally advanced HNC (stage III and IV) displaying large primary tumors and metastatic regional lymph nodes [28]. The latter suffer from a high risk of local recurrences as well as distant metastasis, resulting in a 5-year overall survival below 50% [23]. While surgery is usually reserved for small primary tumors, chemoradiotherapy remains the standard treatment option for many patients [28]. However, radiotherapy in the head and neck area is challenging because target volumes are mostly large and can change throughout radiotherapy, many relevant organs-at-risk (OAR) like small muscles, salivary glands or mucosa must be taken into account and substantial patient weight loss over the course of radiotherapy is possible [108]. Hence, biomarkers allowing stratification of patients in different risk groups and the possibility of treatment adaptation are highly needed to individualize the radiotherapy of HNC patients.

1.3 Biomarkers

The introduction of biomarkers and their translation into standard clinical practice has revolutionized all medical fields and majorly improved patient care by identification of diseased people, staging of diseases, the indication of a disease's response or the prediction of a patient's response to a certain treatment [58]. In general, a biomarker is defined as a "characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention." [58]. Biomarkers can be divided into molecular markers like genes or proteins and in biomarkers derived from one or more imaging modalities, known as an imaging biomarker (IB). In the treatment of HNC patients, the identification of a patient's HPV status as a prognostic molecular biomarker was one of the most important discoveries in the last years. Deoxyribonucleic acid (DNA) testing on HPV status or p16 immunohistochemistry are used to differentiate patients with and without HPV infection (HPV^+ / HPV^-)

[210] as different studies have demonstrated the significant stratification potential of HPV status for oropharyngeal HNC patients regarding survival prognosis. Here, the landmark study by Fakhry et al. [45] in 2008 determined that HPV^+ oropharyngeal or laryngeal patients had a 95% overall survival compared to only 62% in HPV^- patients. However, further research is required to discover additional biomarkers, allowing the definition of new risk-subgroups and potential treatment intensification or de-escalation [28].

Although all medical fields highly profit from the steadily increasing number of IBs being discovered, evaluated and introduced into clinical practice, IBs play a pivotal role in the field of oncology. Due to the fact that every tumor has a unique microstructure [36], it is inevitable that the same treatment will not result in the same outcome for every patient. This further strengthens the need to discover possibilities for patient stratification and treatment personalization options in the field of oncology. Especially for radiation therapy, IBs have great potential to enhance treatment personalization because of major improvements in the hardware and software components of all imaging modalities in the last decades as well as the fact that CT and MRI are being routinely used for radiotherapy to define target structures and OAR while also providing tissue electron densities to calculate radiotherapy treatment plans. Additionally, radiotherapy highly profits from IBs because of their three dimensional resolution which is crucial for the quality of radiotherapy treatment plans.

Although few IBs are currently used in standard clinical care and therefore actually guide treatment decisions like the circumferential resection margin status from MRI for rectal cancer [197] or the bone scan index from single photon emission computed tomography (SPECT) for prostate cancer [179], most IBs are not yet part of routine clinical workflows. This problem arises because an IB must undergo a time-, cost- and labor-intensive process between the discovery of the IB and transition and usage in routine medical practice. This process also differs substantially between specimen-derived biomarkers and IBs for which reason O‘Connor et al. [156] published the imaging biomarker roadmap (cf. Figure 1) describing the process every IB must traverse to be employed in standard medical care.

According to this roadmap [156], the aim for every IB is the application in clinical routine and guidance of patient management. To achieve this goal, each IB has to pass a series of domains while three tracks occur in parallel during this process. Initially, every biomarker starts in the "Discovery" domain where an IB is discov-

ered in specific studies searching for biomarkers or through unintended results from different fields of science like physics, chemistry or engineering. Afterwards, first studies evaluate the newly discovered IB in animal, in vitro or human studies.

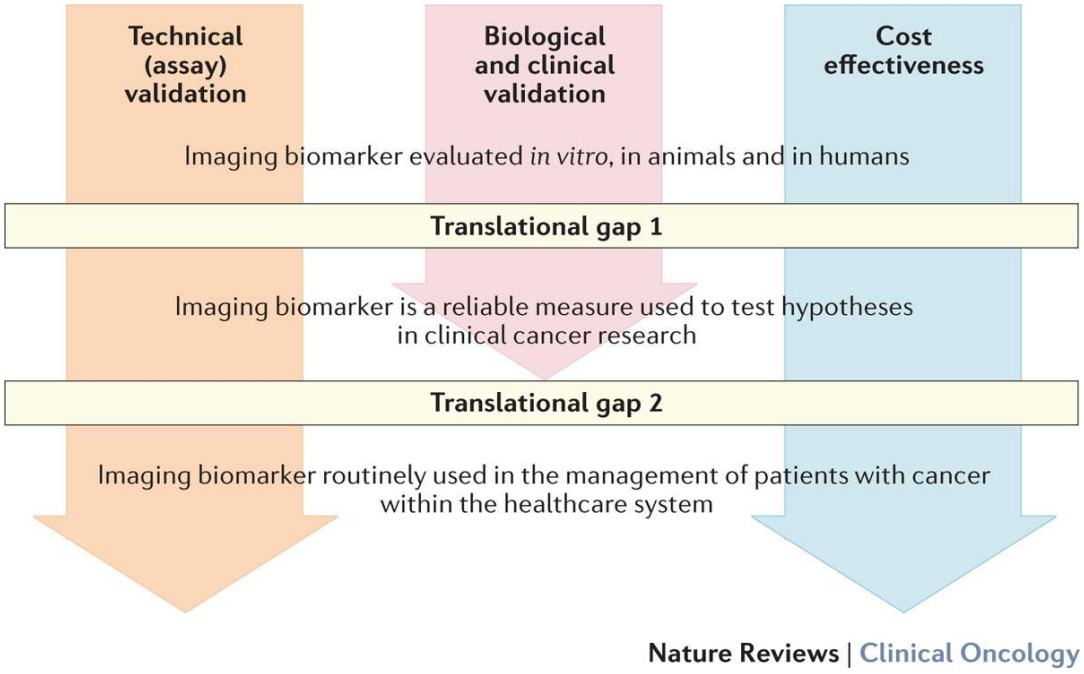

Figure 1: The imaging biomarker roadmap proposed by O´Connor et al. [156] showing three different tracks a biomarker must go through in order to cross both translational gaps and become used in routine clinical practice. Figure from O´Connor et al. (DOI: 10.1038/nrclinonc.2016.162) is licensed under the CC BY 4.0 license.

With these studies, the IB advances into the second domain "Validation" where, beginning with single center studies and going forward to international multicenter studies, the IB is technically and biologically validated while always considering the cost-effectiveness of its acquisition and evaluation. Technical validation aims at making an IB internationally measurable with comparable results. Therefore, studies determining the precision, bias and availability of the biomarker are essential prerequisites to decide about possible applications of the IB (cf. Section 1.4). In contrast to specimen-derived biomarkers, where the technical validation is often completed early in the roadmap, for IBs technical validation may be evaluated until late into the third domain due to the complex setup of human imaging studies. In parallel to the technical validation, also the biological validation of the IB is performed. Here, the IB has to demonstrate a correlation to early response of patients,

treatment outcome or tumor biology. Identification of the prognostic or predictive potential of an IB enables decisions about treatment strategies and personalization. A prognostic biomarker provides information about the outcome of patients to their disease independently of the treatment, which can be used to stratify patients into different treatment options. Predictive biomarkers on the other hand are able to predict the outcome of a specific treatment for the patient [162]. During this second domain, the IB has to take the first major step towards the clinical routine applications by demonstrating reliability and enabling investigations of medical research hypotheses, thus crossing the first translational gap.

In the last domain "Qualification", the IB has to prove fitting for certain purposes and is linked to clinical endpoints. Therefore, large prospective clinical trials are mandatory to show the usefulness of an IB for medical decision-making and not only the correlation to patient outcome. By performing all necessary steps of this domain, the IB crosses the second translational gap and may be used in routine patient management.

1.4 Technical validation of biomarkers

Multimodal imaging has emerged as one of the most important tools in nearly all fields of medicine by qualitative inspection of these images in order to secure patient diagnoses or monitor treatment response. But with the aim of more personalized medicine in mind, the research focus has shifted to also quantitative evaluation of these images, therefore searching for QIBs. While the qualitative image evaluation is often relatively straight forward, but dependent on the user experience, the quantitative evaluation is often preceded by critical analysis of the whole acquisition and evaluation process of the QIB. Because in order to adapt the diagnostic or therapeutic patient workflow, technical limits of QIBs must be determined to ensure safe and validated clinical decision. QIB acquisitions suffer from different uncertainties during the whole analytical process, beginning with the problem of a heterogeneous patient collective. Furthermore, a variety of scanners and evaluation softwares are available for measuring the QIB [177] (cf. Figure 2). All underlying uncertainties must be quantified to determine the reliability and reproducibility of the QIB which is known as technical validation of QIBs. Because of the rising interest in performing these kind of technical validation studies, the Quantitative Imaging

Biomarkers Alliance (QIBA) has been formed in the year 2007, providing guidelines for standardized processes, experiments and measurement variables needed to fulfill technical validation [92].

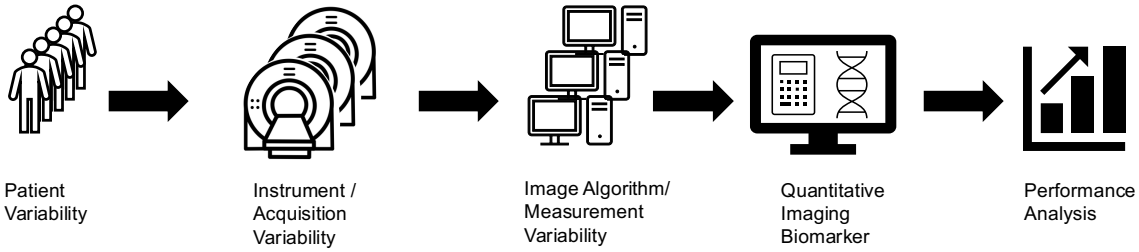

Figure 2: Shown is the analytical process every quantitative imaging biomarker undergoes for its evaluation and the uncertainties during the process. Figure based on [177]

An optimal technically validated QIB can be measured at any timepoint anywhere in the world with comparable results, but considering all the uncertainties in the analytical process, no QIB can achieve perfect validation. Therefore, the technical limits of the QIB have to be determined before decisions about the utility can be made. An important aspect when inspecting and using a QIB is that the technical validation provides no information whether there is a correlation to patient biology or clinical outcome. For a complete technical validation of a QIB, its availability, bias and precision have to be evaluated [156].

Availability of a QIB is a prominent, but often forgotten, aspect in the technical validation process. To ensure broad application of the QIB, its acquisition and analysis must be easily executable and well-tolerated by the patients. Otherwise, the biomarker can be accurate and promising for improved patient outcome, but will never be translated into routine clinical practice [156]. QIBs derived from PET have the advantage of targeting specific metabolic properties of the human body through a broad spectrum of available tracers that visualize and thereof quantify for example hypoxia [182]. However, hypoxia PET imaging suffers from additional radiation doses for the patient as well as high logistic effort for the imaging schedule, while also many short-lived tracers are only available to exclusive institutions owning cyclotrons. Therefore, QIBs derived from MRI moved into the focus of QIB research because MRI is broadly available and inherits no additional radiation dose for patients.

Another important QIB characteristic is the bias which estimates the systematic measurement error between the true underlying value and the average of all performed measurements [92]. This information is crucial for all possible applications of the QIB by ensuring that either the correct value is quantified or a constant bias of the measurement can be counterbalanced after the acquisition. Furthermore, QIBs displaying a variable bias might be unsuitable for further biological and technical studies to establish new clinical biomarkers.
In order to determine the bias of a QIB, large clinical patient studies are pursued. However, these studies are often not feasible because usually the true, to be measured, value is unknown and bias can consequently not be quantified. Therefore, most studies evaluating technical bias of a QIB perform phantom measurements where the true value is known, but have the downside that the measurement settings with phantoms are idealized and the bias will be underestimated in contrast to the real bias in a heterogeneous patient cohort [156]. For diffusion-weighted MRI acquisition and thereof derived ADC value validation, CaliberMRI offers a suitable diffusion phantom. It contains 13 vials with traceable aqueous solution of polyvinylpyrrolidone (PVP) ranging from 0% to 50% PVP which correspond to specific ADC values in the range of 128-1127 $\cdot 10^{-6} \frac{\mathrm{mm}^2}{\mathrm{s}}$ [91] (cf. Figure 3). Therewith, ADC bias can be evaluated for different sequences on various MRI scanners.

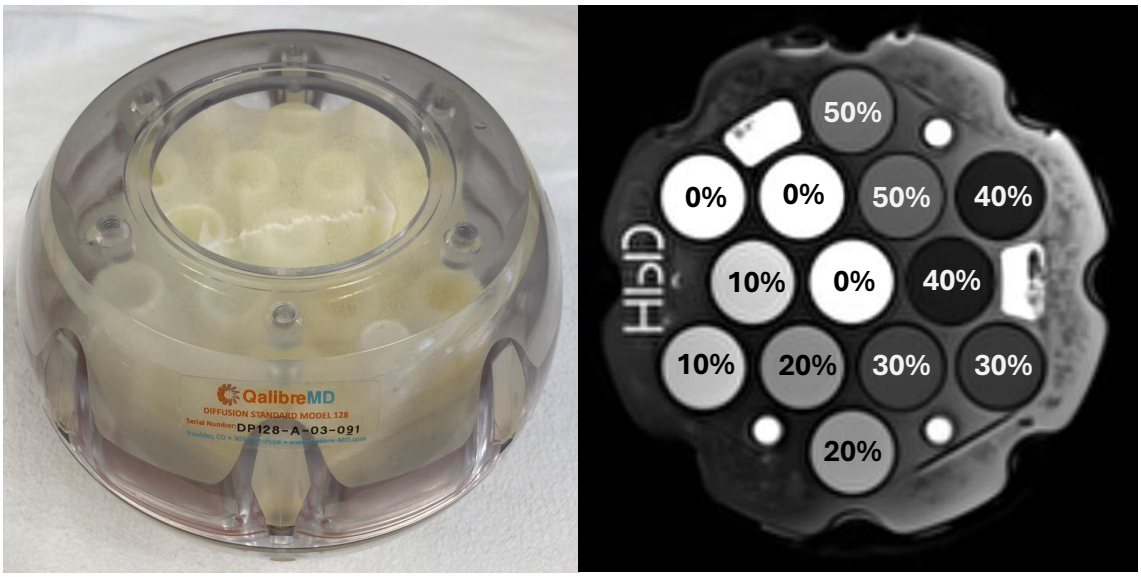

Figure 3: Shown on the left is the diffusion-weighted standard model 128 phantom from CaliberMRI and on the right an axial T2-weighted MRI scan of the phantom with all 13 vials labeled with their according PVP concentration.

The other fundamental characteristic of every QIB is its precision. It "is the closeness of agreement between measured quantity values obtained by replicate measurements

on the same or similar experimental units under specified conditions" [92]. Per definition, a distinction is made between repeatability and reproducibility which together specify precision. When multiple measurements of a QIB are made under nearly constant conditions like operator, scanner or imaging sequence, repeatability can be estimated. Quantification of repeatability is possible by calculation of the repeatability coefficient (RC) [177]

$$RC = 2.77 \cdot wSD = 2.77 \cdot \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} \sigma_i^2} \tag{1}$$

with σ_i^2 as the variance for every of the N subjects with repeated measurements and the within-subject standard deviation (wSD). The RC provides a measure for the least significant difference between two measurements under the same conditions with 95% confidence. It can be applied to decide if the difference between two values, recorded longitudinally at different times, results from true change or is caused by measurement uncertainty [185]. When the magnitude of a QIB influences repeatability, the variance for every subject is divided by the mean over every patient μ_i^2. Thereof, the within-subject coefficient of variation (wCV) and relative repeatability coefficient (relRC) can be calculated [177]:

$$relRC = 2.77 \cdot wCV = 2.77 \cdot \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} \frac{\sigma_i^2}{\mu_i^2}}. \tag{2}$$

The interpretation and application of the wCV and relRC is analog to their absolute equivalents with the wCV quantifying the extent of relative measurement variation and the relRC its 95% confidence estimate, thereby providing the opportunity of deciding whether a relative change is caused by measurement uncertainty or real change.

In routine clinical cancer workflows, diagnostic CT and MRI is often performed in the department of radiology, while for a later following radiotherapy specially dedicated planning CT and MRI are acquired. Hence, differences in scanners, operators or imaging protocols are inevitable which introduce additional variability between two QIB measurements. For the quantification of this added uncertainty, the reproducibility of a QIB must be evaluated to make results under variable measurement conditions interpretable. Therefore, the same study concepts and similar evaluation metrics as for repeatability analyses are appropriate with the difference of actively varying the measurement conditions in these studies [185].

1.5 Magnetic resonance imaging

One of the fundamental properties of subatomic particles is their angular momentum, called spin. Nuclei with an even number of neutrons or protons show no externally measurable spin, but nuclei with an odd number do. For charged particles, the spin creates a magnetic field. This property led to the groundbreaking idea of measuring the magnetic field of nuclei in the human body and use the gathered information for a non-invasive imaging method [170]. This concept is known as MRI and was first presented by Lauterbur [112] and Mansfield [137] in 1973.

In general, magnetic moments are randomly aligned, but when exposed to an external magnetic field B_0, parallel or antiparallel alignment to B_0 occurs. As the least energy requiring state is preferred, a slightly higher number of nuclei align themselves in parallel to the external magnetic field. This excess nuclei create a net magnetization M_Z along B_0 which is the basis for subsequently measurable signal generation. Due to the extremely small number of contributing nuclei, MRI is mainly measured from frequently occurring nuclei in the human body, most often protons of hydrogen atoms.
Moreover, all spins rotate around their respective magnetic axis which is called precession [33]. Depending on the external magnetic field strength and the gyromagnetic ratio γ, every atom features precession at its specific larmor frequency ω:

$$\omega = \delta \cdot B_0. \tag{3}$$

For protons at a commonly used magnetic field strength of 1.5 Tesla, the larmor frequency winds up as 63.87 MHz [170].

The longitudinal magnetization M_Z is not directly measurable because of the much stronger overlaying static magnetic field M_0. However, by applying a radiofrequency (RF) excitation pulse at the proton larmor frequency ω, the magnetization can be tilted up to 180°. These excited protons preceed in-phase with each other, creating a net magnetization M_T in the transverse plane which induces a current in nearly placed receive coils. After switching the RF pulse off, two relaxation processes begin. For one, spins loose energy through the interactions with surrounding molecules and thereby start to fall back into their preferred lowest energy state by flipping back along the longitudinal axis. Therefore, the longitudinal magnetization M_0, which has been decreased through the RF pulse, steadily increases back to its original

value described as time- and tissue-dependent T1 relaxation [212] (cf. Figure 4):

$$M_Z(t) = M_Z(0) \cdot (1 - e^{-\frac{t}{T1}}). \tag{4}$$

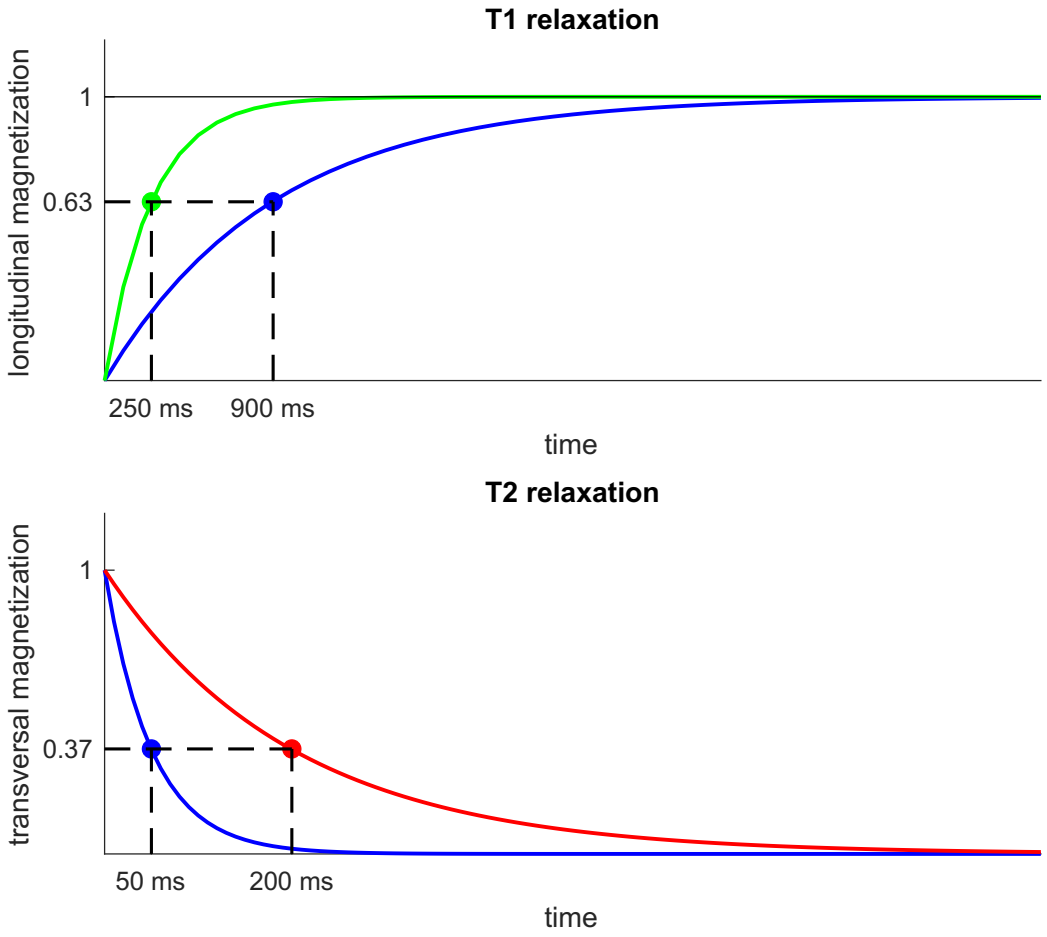

Figure 4: Exemplary longitudinal T1 relaxation is shown in the top figure for fat (green; T1: 250 ms [25]) and muscle tissue (blue; T1: 900 ms [12]) for a magnetic field strength of 1.5 Tesla at human body temperature. Bottom figure displays the transversal T2 relaxation process for muscle tissue (blue; T2: 50 ms [12]) and blood (red; T2: 200 ms [12]).

Additionally, spins in the transversal plane interact with each other and transfer energy onto each other, resulting in some spins gaining and loosing phase relative to each other, thereby becoming more and more out-of-phase. This dephasing occurs in an exponential time frame until no transversal net magnetization M_T is left [8] which is described as tissue dependent T2 relaxation [212]:

$$M_T(t) = M_T(0) \cdot e^{-\frac{t}{T2}}. \tag{5}$$

For the above explained T2-relaxation to be applicable, the static magnetic field B_0 must be perfectly homogeneous. In the human body however, tissues own different magnetic susceptibility characteristics which cause small inhomogeneity of B_0, causing additional dephasing of spins in the transversal plane and a faster T2 relaxation. This process is described by the T2* relaxation with T2' as the signal loss induced by the inhomogeneity of B_0 [170]:

$$\frac{1}{T2^*} = \frac{1}{T2'} + \frac{1}{T2}. \tag{6}$$

In order to spatially encode the acquired signal, additional gradients are applied. First, a slice-selection gradient is exhibited superimposed to B_0 to alter the larmor frequency of protons along one axis. Secondly, a phase-encoding gradient is applied which causes the protons to precess faster along another axis. After switching the gradient off, all protons still feature the same larmor frequency, but are dephased to each other, making differentiation along the second axis possible. Thirdly, a frequency-encoding gradient is exercised perpendicular to the phase-encoding gradient. Here, the precession frequencies of protons along the third axis are diversified, allowing differentiation of each spatial position.
With every RF pulse, all protons of one slice are excited together and therefore only one signal per phase-encoding step is recorded per slice. Through the acquisition of multiple signals with varying strength of the phase-encoding gradient, sufficient signal is received and stored in the frequency space (k-space). A spatial image can then be reconstructed when the data from k-space is translated by Fourier transformation [33].

1.6 Diffusion-weighted magnetic resonance imaging

Diffusion-weighted magnetic resonance imaging utilizes the motion of water molecules inside the human body to quantify the diffusivity of different tissues. While water outside the human body can move unrestricted and therefore succumbs random Brownian motion, water molecules inside the human body are restricted in their movement by cell membranes and other macromolecules [101]. Hence, anatomic regions with a high density of cells and intact cell membranes restrict the diffusion of water, while less dense tissues display a higher diffusivity indicating a correlation between the degree of water diffusion in different tissues to the underlying cell density [54, 193].

In order to measure and quantify water diffusion in vivo, Stejskal and Tanner [191] adapted a standard T2-weighted spin-echo (SE) MRI sequence with two additional gradients (cf. Figure 5). After the application of a 90° excitation RF pulse, two identical diffusion gradients G_{diff} are deployed around the 180° refocusing RF pulse. In case the excited spins are static, the first diffusive gradient initiates spin dephasing which is almost completely reverted afterwards by the second diffusive gradient. If the spins move in between both diffusive gradients, not all dephasing is compensated for and signal loss is induced, accordingly. Hence, the amount of water diffusion is proportional to the signal attenuation [11, 101].

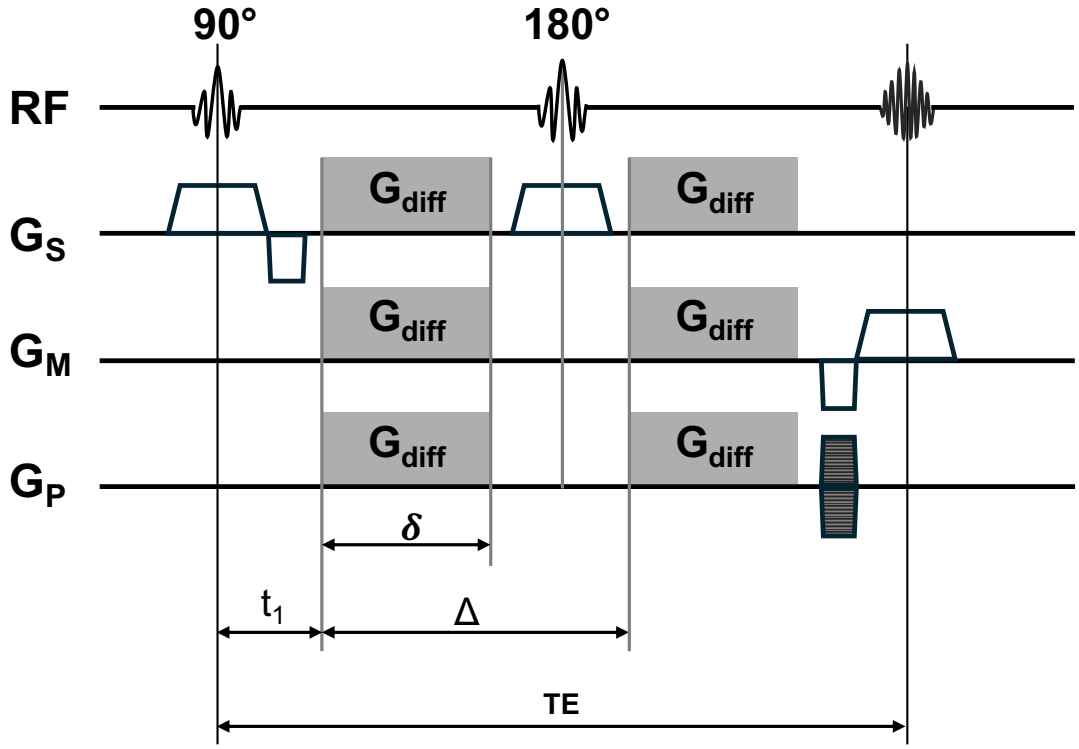

Figure 5: Basic structure of a diffusion-weighted spin-echo magnetic resonance imaging sequence with a 90° exciting radiofrequency (RF) pulse and the 180° refocusing pulse. G_{diff} describes the strength of the diffusion-weighted gradients, δ their duration, Δ the spacing between the gradients and t_1 the spacing between the 90° RF pulse and the start of the first G_{diff}. G_S represents the slice selection direction, G_M the readout direction and G_P the phase-encoding direction. Figure based on [11].

The degree of diffusion-weighting of every DW sequence can be regulated by adjusting the strength of G_{diff}, their duration δ or the spacing between both diffusive gradients Δ. These parameters can be combined into one single variable determining

the amount of diffusion-weighting, called b-value [191]:

$$b[s/mm^2] = \gamma^2 \cdot {G_{diff}}^2 \cdot \delta^2 \cdot (\Delta - \frac{\delta}{3}). \tag{7}$$

While highly diffusive water molecules like in the intravascular space already exhibit large signal loss for small b-values, other tissues display diffusive properties only at higher b-values. Hence, by acquiring multiple images with different degrees of diffusion-weighting, the diffusive property can be evaluated in vivo. Many human tissues present a heterogeneous microstructure and therefore anisotropic diffusion coefficients which can comprise clinically interesting and relevant evidence. To gather directional diffusion information, diffusion gradients in six or more directions are applied to form a diffusion tensor, known as diffusion tensor imaging (DTI) [120]. However, in radiotherapy directional information is often not required for which reason DW gradients are measured in the three orthogonal directions and one averaged isotropic image is produced [123].

When quantitative analysis of DW imaging is pursued, images with at least two different b-values must be acquired. For low b-values, the signal loss is dominated by perfusion effects while for higher b-values mostly diffusion of water molecules is measured [164]. In order to quantify the diffusivity, the simplest model to implement is the monoexponential model:

$$SI = SI_0 \cdot e^{-b \cdot ADC} \tag{8}$$

By fitting a exponentially decaying function over at least two different b-values and their acquired signal intensity SI, the ADC can be calculated. If the chosen b-values are particularly high, ADC solely quantifies the diffusion of water molecules. However, in the range of small b-values, water diffusion and perfusion overlap as the blood flow in capillaries mimics diffusion [118]. While diffusion and perfusion are different phenomena, the combined flow of blood water molecules between capillaries allows the interpretation as a pseudodiffusion effect, thus enabling the quantification of water perfusion and diffusion using a single DW sequence. This concept was introduced by Le Bihan in 1986 as intravoxel incoherent motion (IVIM) [119]. Through the acquisition of additional low b-values, the tissue water diffusion D, water diffusion coefficient in blood D_{blood}, flowing blood fraction f_{IVIM} and pseudodiffusion

coefficient D^* can be calculated using a biexponential model [118]:

$$SI = SI_0 \cdot (f_{IVIM} \cdot e^{-b \cdot (D^* + D_{blood})} + (1 - f_{IVIM}) \cdot e^{-b \cdot D}) \quad (9)$$

Here, the signal intensities present as the sum of two exponentially decaying curves, describing tissue and blood separately. IVIM allows the in vivo quantification of water perfusion by small adaptations of routinely used DW sequences as well as does not require the application of contrast agents necessary for example in DCE, thus presenting high promise in the search for new QIBs [118].

In the last years, DW-MRI moved into the focus of biomarker research for HNC patients. Different studies identified the pretreatment mean ADC value as a prognostic biomarker in HNC. Following the rationale that poor treatment response and patient outcome is associated with HPV^- tumors, low cellularity and micronecrosis [98], different studies have shown that high pretreatment mean ADC values predict poor treatment response in HNC patients [70, 95, 109, 155, 161]. Therefore, mean pretreatment ADC value as a QIB may be promising for patient stratification into intensified or de-escalated radiotherapy groups.
Furthermore, intratreatment ADC changes hold great potential as a QIB for response-adaptive radiotherapy because relative change may be more reproducible across different institutions and scanner types due to differences in acquisition protocols and subsequent delineation and analysis strategies influence absolute ADC values [98]. Based on the knowledge that mean ADC values in HNC are rising over the course of radiotherapy [98], various studies evaluated the prognostic value of relative mean ADC increase. Here, an ADC change below 14-24% was shown to be predictive for treatment failure in HNC patients with at least two years of follow-up [97, 141, 206]. However, high logistic efforts are necessary to enable multiple DW acquisitions over the course of radiotherapy to measure the relative change of mean ADC values. Here, magnetic resonance-guided radiotherapy (MRgRT) offers the possibility of acquiring daily sequential DW-MR images, providing an unique setting for the application of longitudinal QIBs in future response-adaptive MRgRT.

1.7 Magnetic resonance-guided radiotherapy

After many technological improvements over the last decades, IGRT with X-ray imaging for patient position verification remains the standard treatment option for

most patients scheduled for radiotherapy [13], allowing accurate patient positioning in order to replicate the treatment planning setting in the best way possible. However, patients are exposed to additional radiation and most tumors are not visible on X-ray images, making treatment adjustments to anatomical changes of the patient nearly impossible.
In order to overcome these problems, Raaymakers et al. [176] at the University Medical Center in Utrecht developed and published the concept of a hybrid MR-Linac which combines an 1.5 T MRI scanner with a state-of-the-art linear accelerator. Therewith, online real-time MRI with high soft tissue contrast without additional dose and daily treatment plan adaptation to the actual patient position or anatomy are feasible. Hence, higher radiation doses to tumors may be applied and OAR spared by reduced margins around target volumes and OAR. Additionally, daily functional MRI acquisition and evaluation of thereof derived QIBs is feasible without further logistic efforts. In general, two different workflows are realized on the MR-Linac. At first, the daily acquired MRI is correlated to the image and dose distribution from the reference setting. In case of high conformity between the new MRI scan and reference image, no new treatment plan optimization is necessary and only the patient positioning will be corrected. When the new anatomic image substantially differs from the reference setting, anatomical changes are counterbalanced by new delineations and a full reoptimization of the treatment plan is performed. While the latter workflow offers a more precise plan adaptation, substantially more time is required.

However, the underlying magnetic field influences the energy transport of secondary electrons inside the patient while the linear accelerator hardware disturbs the magnetic field and may cause MRI artifacts. To this date, four different systems of MR-Linacs were developed, all with differences in magnetic field strength and hardware design [34], but only the Elekta Unity system is commercially available at this point. It combines a high magnetic field strength of 1.5 Tesla with a modern 7 MV linear accelerator, providing the possibility of standard radiotherapy treatments while producing high-quality MR images (cf. Figure 6). The X-ray source is mounted on the annular gantry which enables 360° rotation around the cylindrical cryostat which compromises the coils for the static magnetic field. One of the most important hardware adaptation in the hybrid design of the MR-Linac is the split-coil design of the superconducting magnet for the magnetic field generation. Positioned in this 22 cm gap is the cryostat and other little attenuating components, accounting for the manageability of the radiation beam passing [172]. However, the necessary

compromises because of the hybrid nature of the system led to limited capacities of the MRI specifications. While clinical diagnostic scanners at the moment are capable of gradient strengths up to 80 mT/m and slew rates of 200 mT/m/ms, the MR-Linac Unity is restricted to 34 mT/m and 120 mT/m/ms (cf. Table 1).

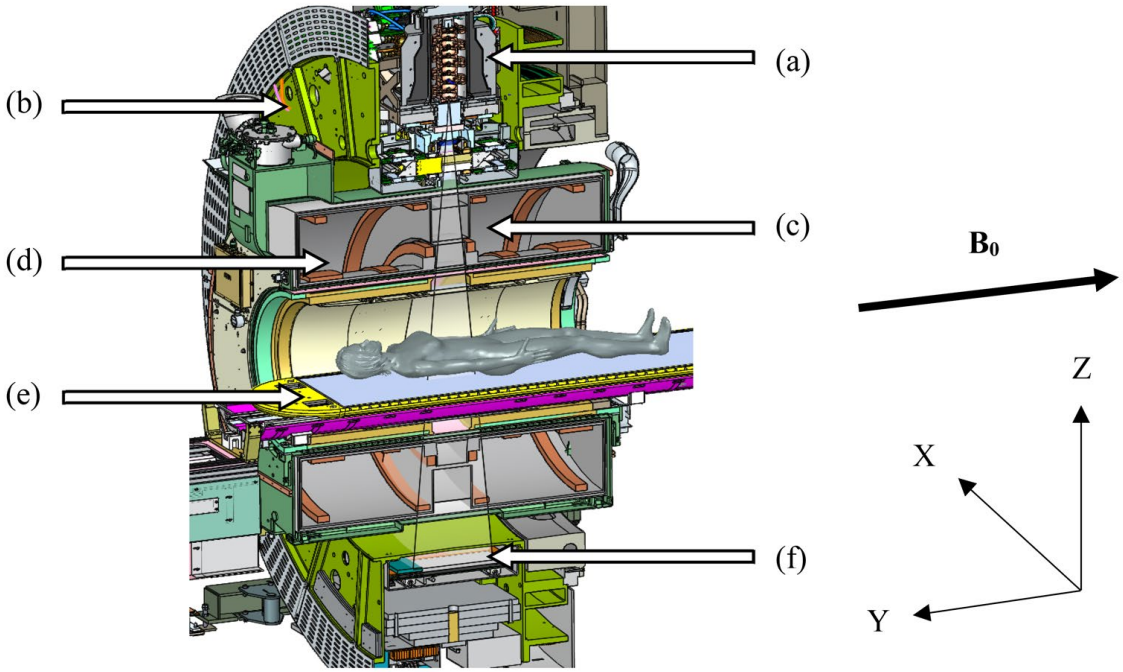

Figure 6: Design of the MR-Linac Unity. Shown is the waveguide of the linear accelerator (a), the gantry ring (b), the cryostat with the passing radiation beam (c), the coil system inside the cryostat (d), patient positioning system (e) and megavoltage imaging system (f). The direction of the static magnetic field B_0 is in the negative y-direction. Figure from Powers et al. (DOI: 10.1007/s13246-022-01113-7) is licensed under the CC BY 4.0 license.

Consequently, when the same sequence protocol is acquired on the MR-Linac and a diagnostic scanner, the underlying settings will differ. While this is less relevant for the qualitative evaluation of anatomical MR images, when DW images are quantified, different settings will lead to differences in echo times among other things. Therefore, the acquisition of a specific b-value on both scanners, will generate two different signals.

Table 1: Characteristics of the MRI system of the 1.5 Tesla MR-Linac Unity from the technical description given by the vendor and commissioning results from the multicenter study by Tijssen et al. [202].

Technical build		**Commissioning results from multi-institutional study**	
Field strength	1.5 T	Absolute field strength	63.88 MHz ± 2.5 kHz
Bore diameter	70 cm	Ghosting (echo-planar-imaging)	0.3 ± 0%
Bore length	130 cm	Flip angle accuracy	0.35 ± 0.24%
Gradient amplitude	34 mT/m	B0 homogeneity, 35 cm diameter of spherical volume	0.9 ± 0 ppm
Gradient maximum slew rate	120 mT/m/ms		
RF nominal frequency	63.87 MHz		
RF maximum frequency offset	305 kHz		

1.8 Aims and objectives

The overarching aim of this thesis was to technically validate the acquisition of DW imaging and thereof derived ADC values on a 1.5 Tesla MR-Linac. DW imaging is an established imaging technique on conventional diagnostic MRI scanners and is already routinely used in improving patient diagnosis by qualitative assessment. The introduction of MRgRT provided the opportunity to continuously acquire functional MRI and investigate the potential of new QIBs derived from e.g. DW imaging. However, the MRI hardware of these MR-Linacs is limited in contrast to the conventional gold standard MRI scanners. Therefore, the acquisition of QIBs derived from DW-MRI on hybrid MR-Linacs require full technical validation as one major cornerstone in the roadmap of every QIB.

The first objective of this thesis was to quantify the technical limit of the acquisition of mean ADC values on our 1.5 T MR-Linac. Because of lower gradient strength and slower slew rates caused by the MR-Linac hardware, adapted sequences had to be employed to measure the same diffusion properties as on conventional scanners. Here, the repeatability of DW-MRI in HNC was investigated in a clinical setting, quantifying the difference between two identical measurements. Hereby, an estimate for differentiating real pathological change from measurement uncertainty was determined.

The second objective was to evaluate the ADC differences between acquisitions on the MR-Linac and a conventional scanner. Many previous studies with promising results in terms of prognostic value of DW imaging were conducted on gold stan-

dard diagnostic MRI machines. In order to translate results from these diagnostic scanner onto the MR-Linac and vice versa, the reproducibility of DW-MRI in HNC compared to a diagnostic scanner should be evaluated. Furthermore, the bias of the implemented DW sequences on the MR-Linac was determined in a phantom setting to quantify measurement shifts from the true baseline values.

The third objective of this thesis was to evaluate the feasibility of acquiring DW images and thereof derived high-risk subvolumes (HRSs) sequentially over the course of chemoradiotherapy in HNC patients. The adaptive MRgRT workflow is more time-consuming and effortful for patients due to treatment plan adaptations and additional MRI acquisitions. This is particularly important when biologically adaptive MRgRT should be implemented in the future to allow an even better personalization of radiotherapy. Here, dose escalation to tumor subvolumes which correlate to higher radioresistancy might be one option to enhance treatment outcome for patients. Therefore, the feasibility of sequentially acquired DW-MRI and thereof derived intratumoral HRSs on the MR-Linac was assessed in HNC in the context of a clinical study.

2 Results and Discussion

2.1 Repeatability of diffusion-weighted magnetic resonance imaging in head and neck cancer at the 1.5T MR-Linac

Authors: **Jonas Habrich**, Simon Boeke, Marcel Nachbar, Konstantin Nikolaou, Fritz Schick, Cihan Gani, Daniel Zips, Daniela Thorwarth

Published in: Radiotherapy and Oncology 2022 Vol. 174 Pages 141-148

DOI: 10.1016/j.radonc.2022.07.020

The research was carried out in accordance to the approved statement by the local ethics board (No. 659/2017BO1, NCT04172753)

2.1.1 Abstract

Background and purpose: Functional information acquired through diffusion-weighted magnetic resonance imaging (DW-MRI) may be beneficial for personalized head and neck cancer (HNC) radiotherapy. Technical validation is required before DW-MRI based radiotherapy interventions can be realized clinically. The aim of this study was to assess the repeatability of apparent diffusion coefficients (ADC) derived from DW-MRI in HNC using echo-planar imaging (EPI) on a 1.5 T MR-Linac.
Material and methods: A total of eleven HNC patients underwent test/retest DW-MRI scans at least once per week during fractionated radiotherapy at the MR-Linac. An EPI DW-MRI test scan (b = 0, 150, 500 s/mm^2) was acquired before the start of adaptive MR-guided radiotherapy in addition to an identical retest scan after irradiation. Volumes-of-interest (VOI) were defined manually for parotid (PTs) and submandibular glands (SMs), gross tumor volume (GTV) and lymph nodes (LNs). Mean ADC was calculated for all VOI in all test/retest scans. Absolute/relative repeatability coefficients (RCs/relRCs) as well as intraclass correlation coefficients (ICCs) were determined for all VOIs.
Results: A total of 81 datasets were analyzed. Mean test ADC values were 1380/1416, 950/1010, 1520 and 1344 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$ for left/right SM and PT, GTV and LNs, respectively. Accordingly, RC (relRC) values were determined as 271/281 (19.4/21.8%) and 138/155 (13.3/15.2%), 457 (31.3%) and 310 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$ (23.5%). ICC resulted in 0.80/0.87, 0.97/0.94, 0.75 and 0.83 for left/right SM and PT, GTV and LNs, respectively.
Conclusion: The repeatability of ADC derived from EPI DW-MRI at the 1.5 T MR-Linac appears reasonable to be used for future biologically adapted MR-guided radiotherapy.

2.1.2 Introduction

Functional magnetic resonance imaging (MRI) has shown to be promising in terms of better target volume delineation [126] as well as providing prognostic information in regard of therapy outcome [139]. Main methodical approaches are diffusion weighted (DW) MRI and dynamic contrast enhanced (DCE) MRI, which have demonstrated potential in predicting outcome after radiotherapy [109, 228]. Especially DW-MRI has shown promising results for the prediction of therapy outcome in head and neck cancer (HNC) [97, 109, 139]. The application of DW-MRI is based on restricted Brownian motion of water molecules and therefore reflecting microanatomy of dif-

ferent tissues. During the acquisition of DW images two opposing magnetic field gradients are applied which do not affect stationary spins but dephase moving spins. Quantification of water diffusion is based on the calculation of the apparent diffusion coefficient (ADC) which may reflect tissue cellularity [215]. Before using quantitative imaging biomarkers (QIBs) derived from functional MRI in clinical practice for response assessment and treatment decisions, their quality must be validated in terms of repeatability, reproducibility, variability and accuracy [156, 185]. Several recent studies validated the acquisition of QIB such as ADC, intravoxel incoherent motion (IVIM) parameters, T1 or T2 for diagnostic scanners in different tumor sites like breast cancer, liver cancer or orbital lesions [9, 121, 146, 153]. A recent technical development in radiation therapy was the combination of a magnetic resonance tomograph with a linear accelerator (MR-Linac) [1, 175]. Apart from the possibility of daily radiotherapy plan adaptation based on anatomical MRI, MR-Linacs allow for functional MRI before and during each treatment session which might be used for an individualized treatment plan adaptation, but also for QIB assessment and extraction, to monitor and predict therapy outcome [81, 106, 200]. Since technical capabilities, especially the performance of the gradient system, are inferior compared to up-to-date clinical MRI systems and lower signal-to-noise ratio (SNR) due to suboptimal coil setup, functional MRI acquisition on MR-Linacs must be validated to guarantee robust integration into biologically adaptive MR-guided radiotherapy. Yang et al. demonstrated that the acquisition of longitudinal DW-MRI on the combined 0.35 T MRI-guided tri-cobalt 60 system is feasible and changes of ADC values over the course of radiotherapy can be monitored [223]. For the 1.5 T MR-Linac a multicenter phantom study showed that quantitative imaging was possible with respect to repeatability, reproducibility and accuracy. A bias $\pm$ limits of agreement of $0.007 \pm 0.027 \cdot 10^{-3} \frac{\mathrm{mm}^2}{\mathrm{s}}$ was reported, whereas a short-term repeatability of $0 \pm 9\%$ was found [106]. Recently, two studies by Lawrence et al. [114] and Kooreman et al. [104] demonstrated technical validation of DW-MRI sequences on a MR-Linac. The first study [114] evaluated repeatability in tumors of the central nervous system and found ADC bias for white matter, grey matter, gross tumor volume (GTV) and clinical target volume (CTV) in the range of -0.08 to -0.1 $\mu\mathrm{m}^2/\mathrm{ms}$ whereas the bias in the cerebrospinal fluid was considerably higher with -0.5 $\mu\mathrm{m}^2/\mathrm{ms}$. Kooreman et al. [104] determined repeatability coefficients (RCs) of IVIM parameters in patients with prostate cancer. They found RCs of 0.09 in the non-cancerous prostate and $0.44 \cdot 10^{-3} \frac{\mathrm{mm}^2}{\mathrm{s}}$ in the tumor. However, the head-and-neck area represents a challenging body part for the acquisition of QIB because of motion artefacts caused by respiration or swallowing and susceptibility artifacts

through air/tissue or bone/tissue boundaries [49]. This might lead to higher variations of QIB in the head-and-neck region in contrast to other areas like pelvis or abdomen. To date, no assessment of DW-MRI repeatability has been performed in HNC on a hybrid MR-Linac.
Consequently, the aim of this study was to assess the repeatability of an echo-planar (EPI) based DW-MRI sequence in patients with HNC at the 1.5 T MR-Linac in a test/retest experiment.

2.1.3 Material and methods

Patients
Between May 2020 and April 2021 eleven patients (median age: 64 y, range 38–77 y) with locally advanced HNC were treated within a prospective clinical trial with 35 fractions of 2 Gy at the 1.5 T MR-Linac (Unity, Elekta AB, Sweden). The trial was approved by the local ethical committee (no. 659/2017BO1, NCT04172753). All patients gave written informed consent.

Imaging protocol
All MRI scans were acquired on the 1.5 T MR-Linac. The study protocol consisted of two identical DW-MRI acquisitions with b-values of 0, 150 and 500 s/mm^2 as well as a T1- and a T2-weighted MRI. Examination of the patients was performed in RT position using a thermoplastic mask and the radiolucent whole-body receive coils of the MR-Linac [32]. Single-shot spin-echo echo-planar imaging (SS-SE-EPI) was used to acquire diffusion weighted images one to two times per week in patients concomitant to seven weeks of radiotherapy. During selected treatment fractions, a test scan was obtained before starting radiation treatment, i.e. during plan adaptation, whereas the retest scan was also acquired before or during irradiation on the same day with a time delay of approximately 10 min, without patient repositioning between both scans. Details about all MRI sequences and parameters are listed in Table 2, while the timepoints of all image acquisitions can be found in Supplementary Table S1.

Image processing and analysis
All b-value test images were rigidly registered to the first fraction's test image for each patient using the open-source toolkit elastix (Version 5.0.1) with a transformation matrix consisting of translations only [99]. Additionally, b-value retest images

were rigidly registered to their corresponding test image. All registrations were visually checked and manually corrected if necessary. ADC maps were derived by an in-house written python script (python 3.8.10) based on the registered diffusion weighted images with b-values 150 and 500 s/mm^2 using the mono-exponential model [191]

$$SI = SI_0 \cdot e^{-b \cdot ADC}.$$

Here, SI describes the signal intensity and SI_0 the calculated signal intensity at b = 0 s/mm^2. Volumes of interest (VOIs) were delineated by a board-certified radiation oncologist on the b500-image for each test image using the open-source software 3D Slicer (Version 4.10) [47] and then propagated to the ADC maps. The T2 image was used to check for plausibility of the contours. VOIs were defined for the left and right submandibular gland, left and right parotid gland, GTV and all suspicious lymph nodes (LNs). Due to a limited craniocaudal scan size, absolute volumes of normal tissues varied between different treatment fractions in case VOIs were not fully covered.

Table 2: Details of the MRI sequences.

parameter	**DW MRI**	**T1w MRI**	**T2w MRI**
Sequence	SS-SE-EPI	3D GR	3D SE
Voxel size [mm^3]	2.78 x 2.78 x 4	0.59 x 0.59 x 1.2	0.68 x 0.68 x 1.1
Field of view [mm^3]	400 x 400 x 100	520 x 520 x 250	520 x 520 x 297
TR/TE [ms]	4811/68	13/4.5	2100/375
Flip angle [°]	90	27	90
Bandwidth [Hz/px]	1930	333	459
b-values (averages)	0 (3), 150 (5), 500 (8) s/mm^2	—	—

The mean test ADC value was calculated for each VOI. Voxels with negative ADC values or ADC $> 3.1 \cdot 10^{-3} \frac{\text{mm}^2}{\text{s}}$ were considered as fitting failures and were therefore excluded from further analysis.

Statistical analysis

ADC data was further analyzed using Matlab 2020a (MathWorks, Natick, MA, USA). Repeatability of the ADC acquisition was assessed by calculation of the RC and the relative repeatability coefficient (relRC) [159]. RC was defined as

$$RC = 1.96 \cdot \sqrt{2} \cdot wSD = 1.96 \cdot \sqrt{2} \cdot \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} \sigma_i^2}$$

with N being the number of patients, wSD as the within-subject standard deviation and σ_i^2 the within-subject variance. Accordingly, the relRC was defined as

$$relRC = 1.96 \cdot \sqrt{2} \cdot wCV = 1.96 \cdot \sqrt{2} \cdot \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} \frac{\sigma_i^2}{\mu_i^2}}$$

where μ_i defines the mean ADC value of each patient and wCV the within-subject coefficient of variation. ADC differences from two measurements, which are higher than the RC or relRC, are caused with 95% confidence by real changes of the measured tissue and not only by measurement uncertainties [159]. Additionally, Bland-Altman analysis was performed including the assessment of limits of agreement (LoA), defined as mean difference between test and retest ± RC [177] to display the repeatability of ADC measurement using the MR-Linac in the different VOIs.

The correlation between the relRC and volume was tested calculating Spearman's rank correlation coefficient using Matlab2020a ($\alpha = 0.05$). Furthermore, intra-class correlation coefficients (ICCs) based on a single-rating, absolute agreement and two-way mixed effects model and their 95% confidence intervals (CIs) were calculated using SPSS statistical package version 28.0.0.0 (SPSS Inc, Chicago, Illinois) [103].

2.1.4 Results

A total of 87 imaging datasets from eleven patients (median age 62 years, range 38–77 years; two women, nine men) were available, six of which had to be excluded from analysis due to image artifacts or missing retest images. The mean time (range) between test and retest measurement was 9.2 min (3–14 min) with all test images acquired before the daily irradiation and nine retest images acquired before and 72 during irradiation, respectively. Detailed information about patient data is given in Table 3.

Table 3: Patient characteristics

Patient	Sex	Age [y]	Primary tumor site	TNM-classification[a]	p16-status	UICC stage	Median (range) time between test/retest [min]	# test/retest acquisitions
P01	m	69	Hypopharyngeal	T2N1M0	-	III	4 (3-12)	7
P02	w	57	Oropharyngeal	T3N1M0	-	III	7 (3-10)	6
P03	m	77	Oropharyngeal	T3N1M0	+	II	9 (9-13)	7
P04	m	60	Oropharyngeal	T3N2M0	+	II	10 (9-10)	7
P05	m	64	Oropharyngeal	T2N1M0	+	I	10 (9-10)	6
P06	m	58	Oropharyngeal	T3N2bM0	-	IVA	10 (9-12)	7
P07	m	38	Oropharyngeal	T4N2M0	+	III	9 (9-14)	5
P08	m	76	Oropharyngeal	T3N2M0	+	II	10 (9-12)	7
P09	m	76	Oropharyngeal	T2N2M0	+	II	10 (9-13)	9
P10	m	58	Oropharyngeal/ Hypopharyngeal/ laryngeal	T4aN2cM0	-	IVA	10 (9-10)	12
P11	w	71	Supraglottic laryngeal SCC	T2N2bM0	-	IVA	9 (9-10)	8

[a]Tumor classification was done according to TNM classification 8th edition, 2018.

Mean ± standard deviation (SD) VOI size was 7.6 ± 1.3 cm^3 and 6.9 ± 1.5 cm^3 for left and right submandibular gland, 13.8 ± 8.0 cm^3 and 14.5 ± 7.5 cm^3 for left and right parotid gland, 15.8 ± 13.6 cm^3 for GTV, and 5.6 ± 5.2 cm^3 for LNs, respectively. Mean test ADC over all patients and all fractions for the left submandibular gland was 1380 ± 203 and 1416 ± 284 for the right submandibular gland, 950 ± 258 for the left parotid gland and 1010 ± 233 for the right parotid gland, 1520 ± 295 for the GTV and 1344 ± 286 $\cdot 10^{-6} \frac{mm^2}{s}$ for the LNs, respectively. Individual ADC values for all patients and every VOI can be found in Supplementary Table S2. In Fig. 7 a representative example of patient #9 is presented, showing test/retest diffusion weighted as well as anatomical MRI data of one fraction.
Absolute RCs were determined to be 271/281, 138/155, 457 and 310 $\cdot 10^{-6} \frac{mm^2}{s}$ for left/right submandibular gland, left/right parotid gland, GTV and LNs, respectively. Repeatability of ADC, as determined by Bland-Altman analysis, resulted in bias or mean ADC difference ± LoA of -76.8 ± 270.6 for left and -42.5 ± 281 for right submandibular gland, -39.6 ± 138 for left and -27.1 ± 154.8 for right parotid gland, -83.6 ± 456.5 for GTV, and -41.8 ± 310.2 $\cdot 10^{-6} \frac{mm^2}{s}$ for LNs (cf. Fig. 8).

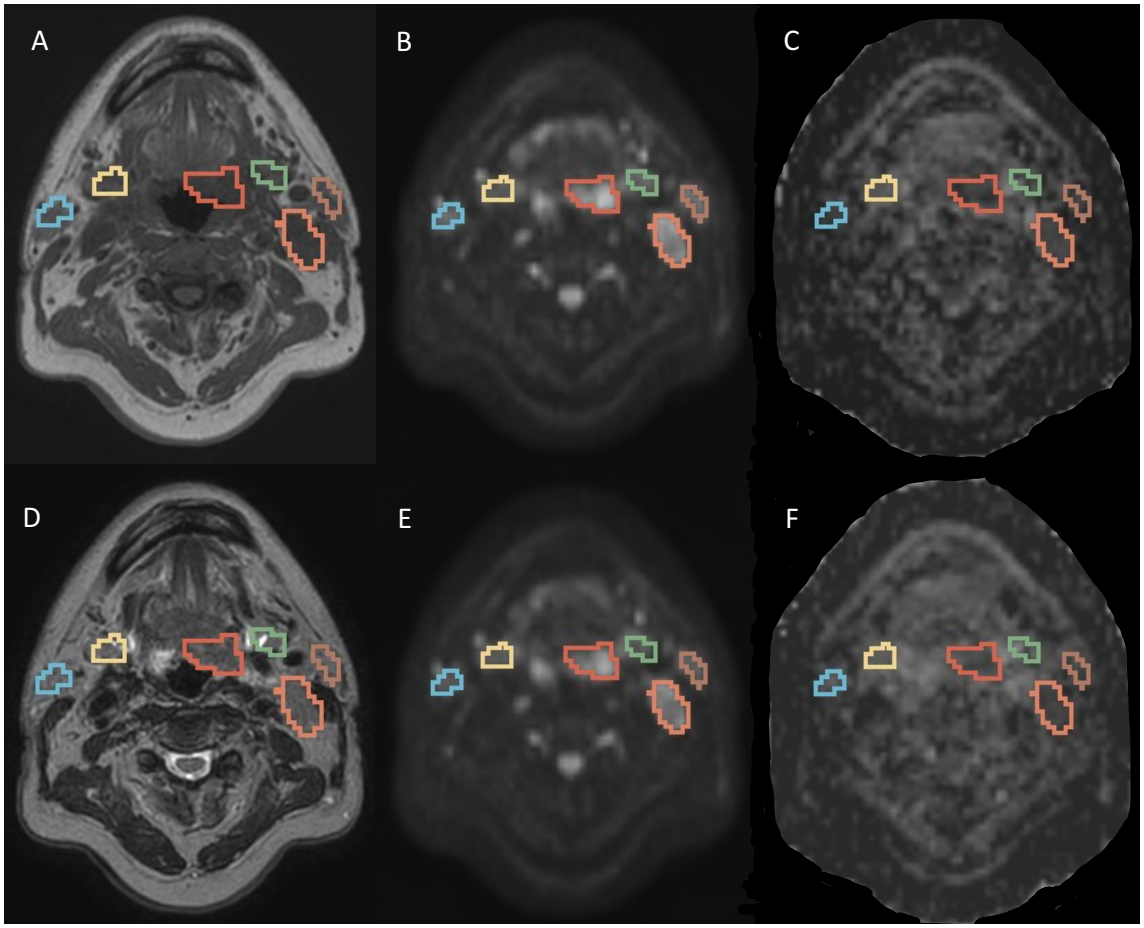

Figure 7: Representative images of the study protocol sequences for patient #9 at fraction 3 with T1w image (A), b500 test image (B), test ADC map (C), T2w image (D), b500 retest image (E) and retest ADC map (F). Blue and brown contour describe the right and left parotid gland, yellow and green contour describe the right and left submandibular gland and red as well as orange contour describe the GTV and conspicuous lymph node.

The relRC, calculated over all 81 fractions, resulted in 19.4% for left and 21.8% for right submandibular gland, 13.3% for left and 15.2% for right parotid gland, 31.3% for the GTV and 23.5% for the LNs, respectively. In addition, Fig. 9 shows the patient-individual variation of relRC for all VOIs where repeatability was calculated for each patient separately. The right submandibular gland showed two outliers with relative repeatability values over 35%.

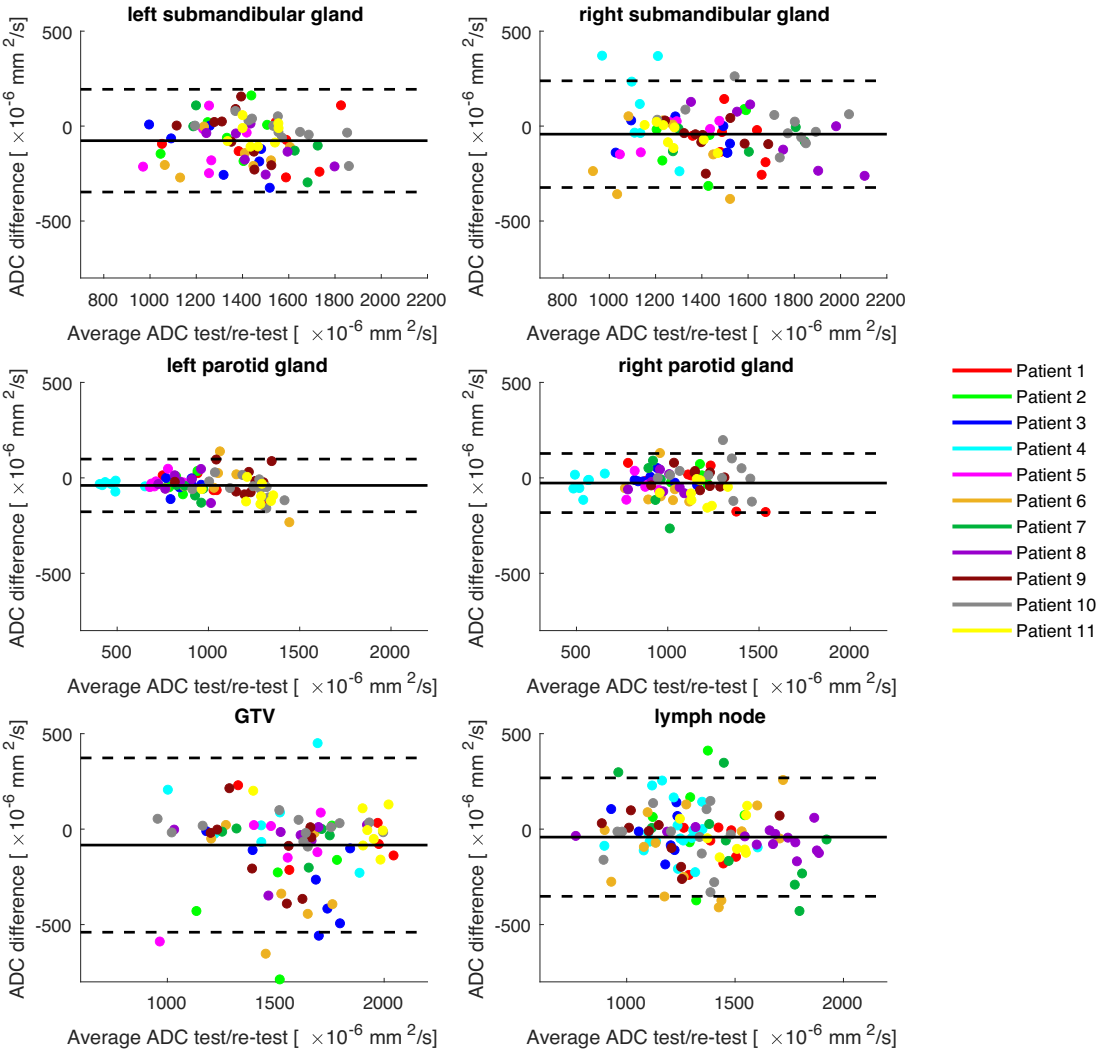

Figure 8: Bland-Altman analysis for submandibular glands, parotid glands, GTV and lymph nodes. Solid lines indicate the mean ADC difference whereas the solid lines represent the limits of agreement as the mean ADC difference ± RC.

Fig. 10 presents a graphical representation of relRC as a function of absolute VOI size. Overall, a significant correlation between relRC and volume was found (Spearman's Rho 0.27, p = 0.025), indicating better repeatability in larger volumes.
ICC values (95% CI) for left/right submandibular gland resulted in 0.80/0.87 (0.53–0.90/0.80–0.92), 0.97/0.94 (0.89–0.98/0.91–0.97) for left/right parotid gland, 0.75 (0.61–0.84) for the GTV and in 0.83 (0.75–0.88) for the LNs.

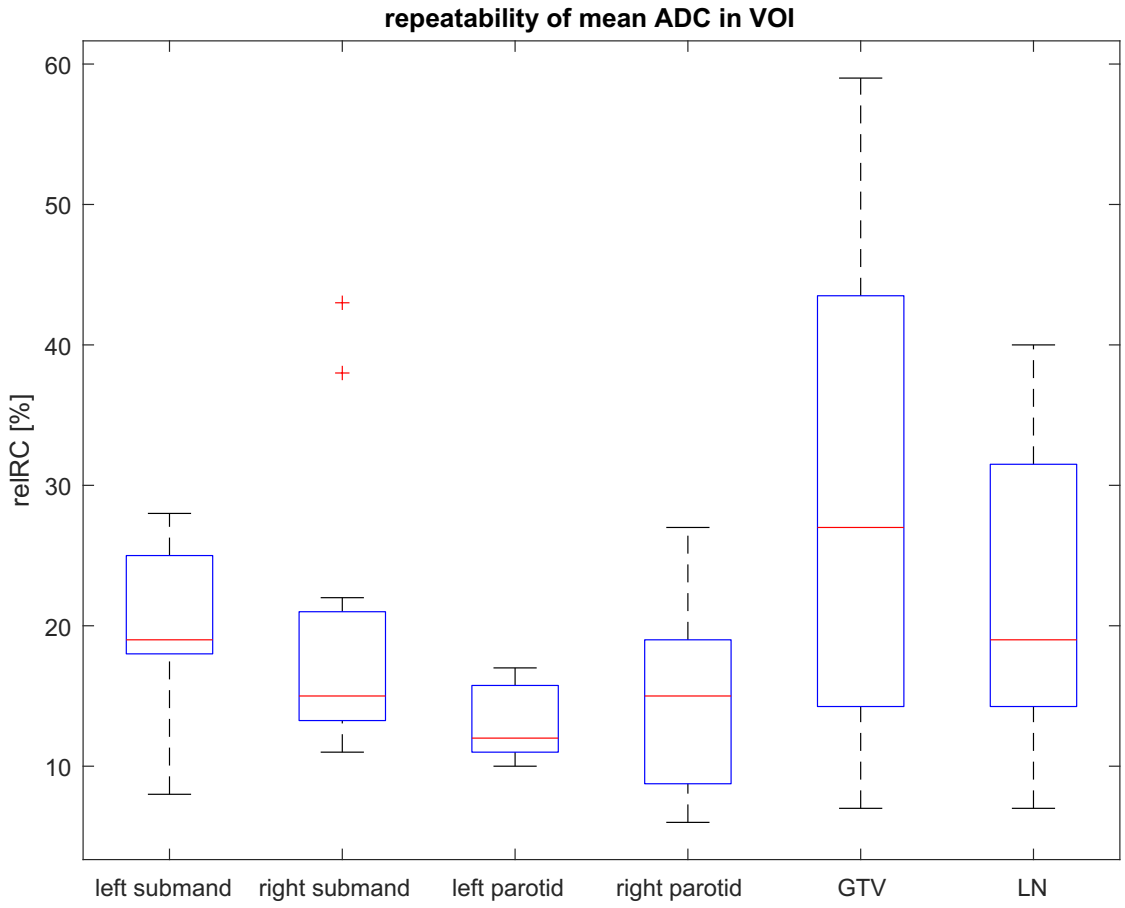

Figure 9: Repeatability of VOIs analyzed for every patient separately.

2.1.5 Discussion

In this study we assessed repeatability of DW-MRI measurements using EPI on a 1.5 T MR-Linac in patients with HNC during radiotherapy. For normal tissues as well as target volumes mean relative repeatability was demonstrated to be between 13% and 31%, depending on the VOI size, anatomic region and potentially also signal changes during treatment. In general, the mean ADC values determined in this study agreed with previously reported values. For the submandibular glands, mean ADC values in the range of 1140 - 1350 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$ were published by earlier studies [39, 132, 199], whereas for the parotid glands, values between 850 and 1150 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$ were found by other groups [39, 132]. Meanwhile, our mean ADC values for submandibular and parotid glands resulted in 1403 and 987 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$. Various other studies determined pretreatment ADC values in HNC in the range 920–1220 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$ [27, 49, 97, 152, 208]. Therefore, our results seem to be corroborated by previously reported mean ADC values, especially when considering measurement

uncertainties and expected ADC increase during radiotherapy [97, 167].
So far, only few studies investigated the repeatability of DW-MRI on MR-Linacs. In a phantom study, Kooreman et al. [106] evaluated repeatability of the ADC value by Bland-Altman analysis with a bias of 0 and LoA of 9%. Another study investigated the repeatability of IVIM parameters in prostate cancer [104]. Here, RCs of 0.09 in non-cancerous prostate and 0.44 $\cdot 10^{-3} \frac{\mathrm{mm}^2}{\mathrm{s}}$ in tumor were reported. Lawrence et al. [114] evaluated repeatability for brain tumors at the 1.5 T MR-Linac. They found within-subject coefficient of variation values of 1.1% in the GTV and 0.9% in the clinical target volume whereas we found a wCV of 11.3% for GTV in HNC. These differences may be caused by motion and susceptibility artefacts, which might be more pronounced in HNC compared to measurements of the brain [117].

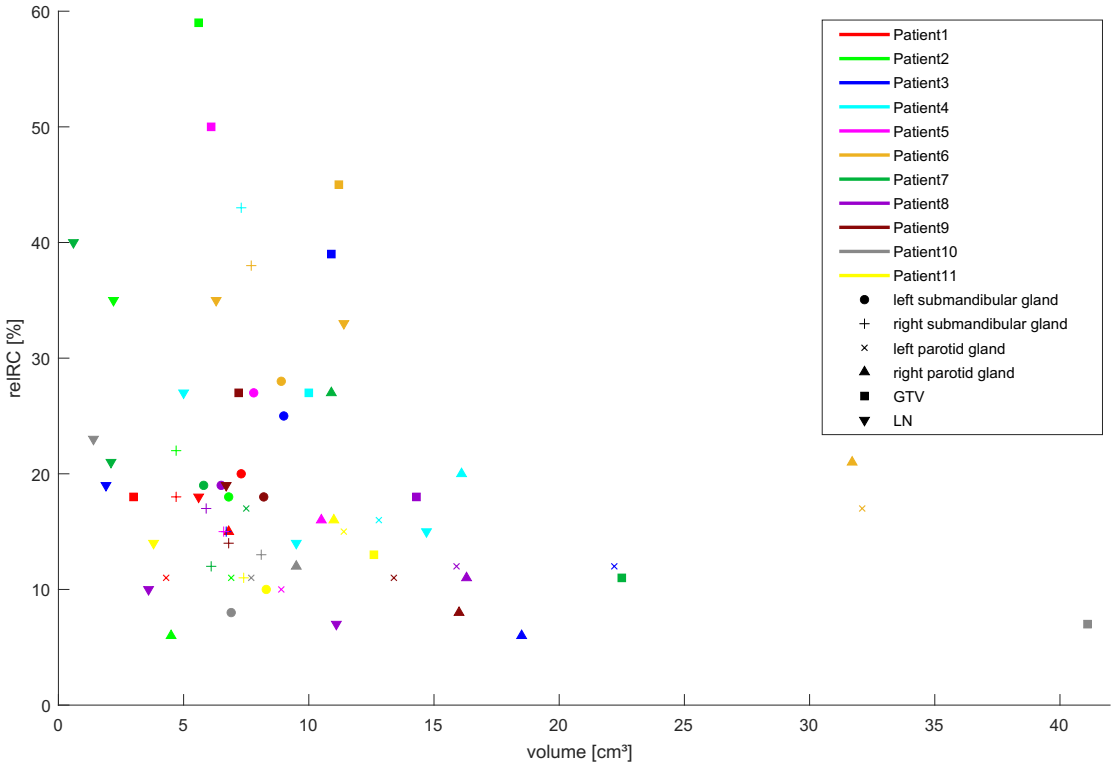

Figure 10: Repeatability of all VOIs for every patient related to the mean volume of the VOI over therapy.

For diagnostic scanners several studies investigating the technical validation of DW-MRI were performed previously. Shukla-Dave et al. [185] summarized repeatability coefficients for three tumor entities. For brain tumors a relRC of 11%, 26% for liver cancer and 47% for prostate cancer were reported. Repeatability analysis of functional MRI parameters in HNC are quite rare. Paudyal et al. [166] reported a pretreatment wCV of 2.38% for head and neck squamous cell carcinoma in nine

patients, imaged at a 3 T scanner.
Variations of ADC values might potentially also be caused by the influence of irradiation. While the first diffusion-weighted MRI in this study was acquired during adaptation of the daily plan, the second scan was performed during or after irradiation. This may have led to changes in anatomy and microstructure of tumor and surrounding tissue. However, some studies evaluated that the tumor microstructure changes only in one to two weeks [131, 138, 224]. Moreover, Lis et al. [130] reported changes in perfusion parameters one hour after single irradiation of spine metastasis with a high dose of 24 Gy. Therefore, it appears unlikely that ADC values change within minutes after irradiation.
Another potential influence on the mean ADC value could be delineation uncertainties, but several studies showed good to excellent interobserver reproducibility of mean ADC values of whole tumors [110, 116].
While SS-EPI is mostly used in DW-MRI, this technique suffers from several shortcomings [165]. EPI-sequences are prone to geometric distortion and signal loss because of susceptibility artefacts, for example at air-tissue boundaries which are especially a problem in the head and neck area, or inhomogeneities of the static magnetic field. These issues may have also affected image acquisition in our study and therefore the calculation of the ADC value and the repeatability. Potential alternatives for DW-MRI acquisition are dedicated methods for distortion correction of EPI [217, 218] or the use of different sequences, such as turbo-spin-echo (TSE) or SPLICE [183]. TSE-based DW-MRI reduces susceptibility artefacts by radio frequency refocusing pulses, whereas respective sequences need longer acquisition times and have lower SNR [165]. In recent years TSE sequences have been developed that have improved SNR, minimize distortion artefacts and may thus be superior to EPI DW-MRI [76, 147, 165]. Although the MR-Linac hardware design differs from that of diagnostic scanners regarding field strength, slew rate and gradient strength [200], accordingly adapted TSE-sequences might provide improvements for DW-MRI on the MR-Linac. The EPI sequence used in this study used three b-values and was adapted to the recommendations of the MR-Linac consortium with respect to DW-MRI [105]. Nevertheless, studies evaluating dedicated TSE-sequences may be needed in the future to further optimize DW-MRI acquisition on hybrid MR-Linacs.
A major limitation of our study was the missing repositioning of patients between the two diffusion weighted acquisitions. In general, studies investigating the repeatability of functional imaging had a short break of approximately 10 min between the two acquisitions with the patient being fully repositioned on the scanner couch [86, 121, 133]. Some studies performed the second scan even several days after the first

scan [189, 213]. Therefore, it was possible to include the uncertainty caused by patient repositioning and other biological factors in the calculation of the ADC value. However, our workflow on the MR-Linac did not allow for patient repositioning as image acquisitions, treatment plan adaptation and irradiation of the patient had to be done in one single session of 30–45 min. Moreover, ADC values may change during the course of radiotherapy, wherefor a break of days might induce biological changes of the ADC.

Another limitation of our study was associated with potential uncertainties and variations in ADC calculation caused by misregistration of the b-value images. Since all test- and retest images acquired for one patient were registered to a common frame of reference, minor registration errors may have led to ADC discrepancies and therefore higher repeatability coefficients. This might be especially problematic at tumor boundaries where a large signal decrease can occur within a few millimeters. This might also be a reason for moderate to good ICC values determined for GTVs and LNs with 0.75 and 0.83 while the glandular tissues showed good to excellent repeatability with ICCs up to 0.97. This further strengthens that RC and relRC values calculated for normal tissues showed better repeatability compared to target volumes. Moreover, lower repeatability expressed by higher RC/relRC values in tumor tissue may also be due to higher levels if tissue heterogeneity compared to functionally intact organs as well as changes in absolute ADC caused by fractionated radiation treatment.

In general, a formal classification of repeatability values does not exist. Consequently, there are no clear cut-off values for the interpretation of RC values in terms of imaging quality assessment. Furthermore, clinical acceptability highly depends on the required sequence implementation and related QIB. If a certain level of RC can be considered good, strongly depends on the observed QIB change and foreseen intervention. Seen the hardware limitations of hybrid MR-Linacs, the overall repeatability of ADC values, measured with an MR-Linac, can be considered good if comparable values as reported for diagnostic scanners are obtained. Shukla-Dave et al. summarized relRC values for tumors in the brain, liver and prostate, measured on diagnostic scanners, which ranged from 11% to 47% [185], which seem comparable to relRCs determined in our study.

Considering the repeatability values determined in this study, ADC-based patient stratification as well as assessment of treatment response enduring the course of radiotherapy might be realized. According to several studies, ADC changes of 30–40% can be expected 2–3 weeks after the start of therapy [97, 167, 206]. Consequently, we hypothesize that with respect to our observed ADC changes and measured relRCs

over the course of radiotherapy, real ADC changes and constitutive interventions as well as prognostication of treatment outcome seem to be feasible in week 2–3 of radiotherapy. However, before ADC-based interventions can be clinically implemented, validation studies on sequential ADC changes in relation to the reported RCs must be carried out. Also, with further developments of DW sequences regarding SNR and improved geometric distortions, TSE-based sequences such as SPLICE, may provide a promising basis for future adaptive radiotherapy.
To confirm the repeatability of ADC measurements using DW-MRI at 1.5 T MR-Linacs, independent validation in the context of multi-center studies is required in addition to investigations regarding the potential of ADC-based radiotherapy interventions in HNC.

2.1.6 Conclusion

In this study, repeatability of mean ADC measured by echo-planar based DW-MRI on a 1.5 T MR-Linac in normal tissue and HNC was assessed. Mean relative repeatability coefficients of ADC ranged between 13 and 31% depending on tissue type and absolute volume, demonstrating the limits for discriminating pathological ADC changes from measurement uncertainties. Consequently, future biologically adapted MR-guided radiotherapy interventions can be designed with respect to these limits. Nevertheless, further research including external validation and improvement of DW sequences is necessary.

Financial support

This study received funding by the German Research Council (DFG Grants No. ZI 736/2-1, PAK 997/1: TH1528/6-1 ,SCHI498/14-1, NI707/7-1).

Conflict of interest disclosure statement

The Department of Radiation Oncology Tübingen receives financial and technical support by Elekta, Philips, Siemens, Dr. Sennewald Medizintechnik, Kaiku Health, TheraPanacea, PTW and ITV in the context of research cooperation's.
The Department of Diagnostic and Interventional Radiology, University of Tübingen reports institutional collaborations, consultation and grant supported by Siemens

Healthineers, Erlangen, Germany and Bayer Healthcare, Leverkusen, Germany outside this study.

Supplementary Material

The supplementary material of the publication is depicted in the supplementary material of the dissertation at section 9.1

2.2 Reproducibility of diffusion-weighted magnetic resonance imaging in head and neck cancer assessed on a 1.5 T MR-Linac and comparison to parallel measurements on a 3 T diagnostic scanner

Authors: **Jonas Habrich**, Simon Boeke, Victor Fritz, Elisa Koerner, Konstantin Nikolaou, Fritz Schick, Cihan Gani, Daniel Zips, Daniela Thorwarth

Published in: Radiotherapy and Oncology 2024 Vol. 191 Pages 110046

DOI: 10.1016/j.radonc.2023.110046

The research was carried out in accordance to the approved statement by the local ethics board (No. 659/2017BO1, NCT04172753)

2.2.1 Abstract

Background and purpose: Before quantitative imaging biomarkers (QIBs) acquired with magnetic resonance imaging (MRI) can be used for interventional trials in radiotherapy (RT), technical validation of these QIBs is necessary. The aim of this study was to assess the reproducibility of apparent diffusion coefficient (ADC) values, derived from diffusion-weighted (DW) MRI, in head and neck cancer using a 1.5 T MR-Linac (MRL) by comparison to a 3 T diagnostic scanner (DS).
Material and methods: DW-MRIs were acquired on MRL and DS for 15 head and neck cancer patients before RT and in week 2 and rigidly registered to the planning computed tomography. Mean ADC values were calculated for submandibular (SG) and parotid (PG) glands as well as target volumes (TV, gross tumor volume and lymph nodes), which were delineated based on computed tomography. Mean absolute ADC differences as well as within-subject coefficient of variation (wCV) and intraclass correlation coefficients (ICCs) were calculated for all volumes of interest.
Results: A total of 23 datasets were analyzed. Mean ADC difference (DS-MRL) for SG, PG and TV resulted in 142, 254 and 93 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$. wCVs/ICCs, comparing MRL and DS, were determined as 13.7 %/0.26, 24.4 %/0.23 and 16.1 %/0.73 for SG, PG and TV, respectively.
Conclusion: ADC values, measured on the 1.5 T MRL, showed reasonable reproducibility with an ADC underestimation in contrast to the DS. This ADC shift must be validated in further experiments and considered for future translation of QIB candidates from DS to MRL for response adaptive RT.

2.2.2 Introduction

Hybrid MR-Linacs (MRLs), which combine magnetic resonance imaging (MRI) and a linear accelerator, were clinically introduced in recent years [1, 48, 151, 175]. This innovative radiotherapy (RT) technology enables the acquisition of anatomical MRI before and during RT for daily online treatment adaptation, with the potential to reduce doses to organs at risk (OAR) [29, 63, 87]. Furthermore, MRLs not only allow for the acquisition of longitudinal anatomical but also functional MRI during the course of RT [85, 107].
While anatomical imaging, especially MRI, provides information about the spread of a solid tumor, functional MRI can assess the microenvironment of the tumor describing biological tumor properties like cell density, hypoxia, oxygenation or permeability [35, 43, 74]. A variety of functional MRI techniques and thereof derived

quantitative imaging biomarkers (QIBs) have been described with respect to their potential for RT, each trying to quantify different properties of the tumor [93, 149, 157]. One of the most prominent functional MRI techniques is diffusion weighted (DW) MRI [149]. Here, two opposing magnetic field gradients are applied which dephase moving spins and hence signal loss is induced. By measuring with at least two different b-values, which depend on the gradient timing and strength, thereby determining the amount of diffusion weighting and signal loss, the apparent diffusion coefficient (ADC) can be calculated, which correlates inversely with tumor cellularity [215].

QIBs derived from functional MRI have been shown to be promising in terms of outcome prediction of RT with many studies focusing on QIBs from especially DW and DCE MRI [19, 64, 109, 134, 139, 228]. However, most of these studies were performed on diagnostic scanners (DSs), while recent studies using MRLs were focusing on the technical validation of the QIB acquisition by assessing repeatability, reproducibility, accuracy and variability [156, 185]. While repeatability describes the precision of repeated measurements with measurement conditions, which should be as similar as possible, reproducibility assesses repeatable measurement precision under varying scanning settings like different scanners or imaging protocols [185]. Thorough technical validation is a prerequisite before integration of QIBs into the clinical RT workflow as the technical performance of modern DSs is still superior to the MRL's.

At the moment, two MRL systems are commercially available, one with a field strength of 0.35 T [100] and one with 1.5 T [216]. Due to the recent clinical implementation of MRLs, studies concerning the technical validation of MRLs in phantoms as well as patients are still limited. Yang et al. [223] evaluated the feasibility of longitudinal DW-MRI in a phantom study and showed that the ADC value can be measured with an error smaller than 5 %. Another phantom study on the 0.35 T MRL also showed that DW-MRI is feasible, but the measured ADC values were significantly lower than on two other DSs [124]. For the 1.5 T MRL, a phantom study by Kooreman et al. [106] investigated bias and limits of agreement of ADC among other quantitative imaging parameters. For tumors of the central nervous system the study by Lawrence et al. [114] investigated the repeatability of ADC values as well as the differences to a DS. In head and neck cancer (HNC) McDonald et al. [143] determined repeatability of three DW sequences on the 1.5 T MRL and a 1.5 T DS in primary tumors, lymph nodes and parotid glands (PGs), while reproducibility by comparison of the MRL to the DS was only assessed for PGs, but not for tumors, in ten patients resulting in a mean ADC difference of -310 $\cdot 10^{-6} \frac{\mathrm{mm}^2}{\mathrm{s}}$ between the

echo-planar imaging (EPI) sequences of the DS and the MRL. In addition, our group evaluated the repeatability of ADC measurements on the 1.5 T MRL [61]. While repeatability of ADC values in HNC tumors on the MRL was recently investigated, to date the reproducibility of ADC values between the MRL and a DS in HNC was only evaluated for organs at risk, but not for tumors.
Therefore, the aim of this study was to evaluate the reproducibility of EPI DW-MRI in head and neck cancer between a 1.5 T MRL and DS for glandular tissues as well as target volumes (TVs).

2.2.3 Material and methods

Patients
Between August 2018 and August 2021, 15 patients with locally advanced HNC were treated on the 1.5 T MRL (Unity, Elekta AB, Sweden) within a prospective clinical trial consisting of 35 fractions in six to seven weeks with concurrent chemotherapy and additional DW imaging on a DS. The study was approved by the local ethical committee (no.659/2017BO1, NCT04172753) and all patients gave written informed consent.

Imaging protocol
The study protocol consisted of DW-MRI measurements on the MRL as well as on a 3 T DS (Magnetom Vida, Siemens Healthineers, Germany). Image acquisition was scheduled on both scanners twice, once before the start of fractionated RT and optional after two weeks of RT. The maximum allowed time difference between the corresponding pairs of MRL and DS measurements was limited to five days before and 3 days in week two of RT to minimize potential biases by therapy-induced ADC changes. The spin-echo (SE) EPI DW-MRI sequence for the DS, which is used routinely for treatment planning of HNC patients and was developed similar to the exemplary head and neck DWI sequence published by Shukla-Dave et al. [185], consisted of 8 b-values from 0 up to 1000 s/mm^2. On the MRL, a SE EPI DW-MRI sequence had been implemented initially, which consisted of b-values 0, 200, 500 and 800 s/mm^2 (EPI4b). Starting from patient #7, the DW-MRI sequence was adapted according to the recommendations of the MR-Linac consortium regarding acquisitions of ADC values on the 1.5 T MRL Unity [105]. Consequently, the new SE-EPI DW-MRI sequence acquired b-values 0, 150 and 500 s/mm^2 (EPI3b). Detailed information about all three sequences used for DW-MRI

acquisition in this study is listed in Table 4. All examinations were acquired in treatment position using the individual thermoplastic RT mask of each patient. For the MRL, the system's radiolucent whole-body receive coils were used while on the DS the standard 18-channel body receive coils were utilized.

Table 4: Details of the DW-MRI sequences.

parameter	**EPI3b**	**EPI4b**	**Vida**
Sequence type	SS-SE-EPI[1]	SS-SE-EPI[1]	SS-SE-EPI[1]
Acquisition voxel size [mm^3]	3.03 x 2.99 x 4	3.03 x 2.99 x 4.8	2.96 x 3.29 x 6
Matrix size	132 x 134 x 25	132 x 134 x 42	128 x 76 x 24
Field of view [mm^3]	400 x 400 x 100	400 x 400 x 202	380 x 250 x 144
TR/TE [ms]	4811/68	10392/107	10800/44
Flip angle [°]	90	90	90
Bandwidth	11.197/19.4	11.180/19.4	2298
Philips: WFS[2](pix)/BW[3] (Hz/px)			
Siemens: BW[3](Hz)			
Parallel imaging technique (factor)	SENSE[4] (2)	SENSE[4] (2)	GRAPPA[5] (2)
Phase-encoding direction	Anterior-posterior	Anterior-posterior	Anterior-posterior
Fat suppression	SPAIR[6]	SPAIR[6]	—
b-values [s/mm^2] (averages)	0 (3), 150 (5), 500 (8)	0 (2), 200 (3), 500 (4), 800 (6)	0 (1), 20 (2), 40 (3), 80 (4), 120 (5), 200 (6), 500 (7), 1000 (10)
Time [min]	3:32	7:17	7:05

[1]single-shot spin-echo echo-planar-imaging

[2]water-fat shift

[3]bandwidth

[4]Sensitivity Encoding

[5]Generalized Autocalibrating Partial Parallel Acquisition

[6]Spectral Attenuated Inversion Recovery

Image processing and analysis

DW-MRIs from both scanners were rigidly registered to the computed tomography (CT) images of each patient acquired before the start of RT for treatment planning using the open-source toolkit elastix (version 5.0.1) [99]. The registration included translations only while no rotation was applied. Registrations were manually corrected after visual inspection, if necessary. ADC maps were calculated using an in-house written python script (version 3.8.10) for the registered DW-MRIs using the mono-exponential model [191]

$$SI = SI_0 \cdot e^{-b \cdot ADC}.$$

with SI describing the signal intensity and SI_0 the calculated signal intensity for b-value 0 s/mm^2. For DW-MRIs acquired on the DS, b-values of 200, 500 and 1000

s/mm^2 were used for ADC calculation, while for the EPI4b and EPI3b acquired on the MRL only 200 and 500 s/mm^2 as well as 150 and 500 s/mm^2 were used, respectively. Due to sufficient signal-to-noise ratios no noise correction was applied (cf. Supplementary Table S3). Comparability of EPI3b and EPI4b was tested previously in-house with phantom measurements and in patients, resulting in comparable mean ADC values for both MRL sequences (cf. Supplementary Figures S2 and S3).
Volumes of interest (VOIs) including submandibular glands (SG) and PGs as well as TV, combining the GTV and all conspicuous lymph nodes (LNs), were extracted from each patient's structure set created for RT planning and visually checked by a board-certified radiation oncologist using the open-source software 3D Slicer (version 4.10). Subsequently, all VOIs were adjusted to the highest b-value image of each registered DW-MRI. LNs smaller than 1 cm^3 were excluded from the analysis. Whenever normal tissues were not fully covered due to limited cranio-caudal field of view (FOV) coverage, absolute volumes of SG and PG glands were reduced in all images of this patient. OAR coverages for all timepoints and acquisitions are given in Supplementary Table S4. For further analysis, mean ADC was calculated for all VOIs.

Statistical analysis
Reproducibility of mean ADC values was assessed by calculation of the within-subject coefficient of variation (wCV) according to guidelines of the Quantitative Imaging Biomarker Alliance (QIBA) [185]. wCV was defined as

$$wCV = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} \frac{\sigma_i^2}{\mu_i^2}}$$

with N being the number of datasets, σ_i^2 the within-subject variance between MRL and DS ADC of each dataset and μ_i the mean ADC for every corresponding pair of ADC values from both scanners.
Furthermore, Bland-Altman analysis with calculation of mean differences between the DS and the MRL where the bias was defined as the difference between DS and MRL (DS-MRL) was performed. Limits ofagreement (LoA), defined as bias $\pm$ 1.96 $\cdot$ standard deviation (SD) [40], were calculated in Matlab 2020a (MathWorks, Natick, MA, USA) showing the 95 % confidence interval of the mean ADC difference. One sample Wilcoxon Signed-Rank test was performed with SPSS (version 28.0.0.0, SPSS Inc, Chicago, Illinois) to test for significant differences of ADC biases from 0 mm^2/s ($\alpha = 0.05$).

To check for a constant offset of the mean ADC difference, linear regression of ADC values from DS and MRL was determined via Matlab including calculation of coefficient of determination R^2. SPSS was used to calculate intraclass correlation coefficients (ICCs) and their 95 % confidence intervals (CI) for all VOIs based on absolute agreement, single measurements and the two-way mixed effects model [103].

Table 5: Patient and disease characteristics.

Patient	Sex	Age [y]	Primary tumor site	TNM-classification[1]	p16-status	Days between MRL and DS measurements (pretreatment/week 2)	DW sequence of the MRL
P01	m	67	Hypopharyngeal	T4N1M0	-	0/-[2]	EPI4b
P02	w	56	Oropharyngeal	T3N2cM0	+	4/3	EPI4b
P03	w	66	Hypopharyngeal	T3N2bM0	-	0/0	EPI4b
P04	m	63	Oropharyngeal	T2N2bM0	-	2/-	EPI4b
P05	m	59	Oropharyngeal	T3N2cM0	+	3/-	EPI4b
P06	w	60	Oropharyngeal	T3N1M0	+	1/-	EPI4b
P07	m	69	Hypopharyngeal	T2N1M0	-	-/3	EPI3b
P08	w	57	Oropharyngeal	T3N2cM0	-	1/-	EPI3b
P09	m	77	Oropharyngeal	T3N1M0	+	5/0	EPI3b
P10	m	64	Oropharyngeal	T3N0M0	+	0/0	EPI3b
P11	m	38	Oropharyngeal	T4N2aM0	+	2/0	EPI3b
P12	m	76	Epi-/Oropharyngeal	T3N2cM0	+	1/-	EPI3b
P13	m	58	Oropharyngeal/ Hypopharyngeal/ laryngeal	T4aN2bM0	-	3/1	EPI3b
P14	w	71	Laryngeal	T2N2bM0	-	1/3	EPI3b
P15	m	53	Oropharyngeal	T2N1M0	+	1/0	EPI3b

[1]Tumor classification was done according to TNM classification 8th edition, 2018.

[2]no corresponding imaging dataset from the MRL and DS was available.

2.2.4 Results

Overall, seven of the possible 30 datasets (corresponding images from DS and MRL) had to be excluded from analysis due to missing image acquisitions or noisy images, resulting in a total of 23 datasets (14 pretreatment, nine week 2) for analysis. Median time between DW-MRIs from the MRL and DS was one day (range: 0–5 days) before RT and zero days (range: 0–3 days) in week 2 of RT. Detailed information about the analyzed patients and the clinical features are provided in Table 5. Rep-

resentative images of calculated ADC maps and corresponding anatomical images for both timepoints of one patient are displayed in Fig. 11.

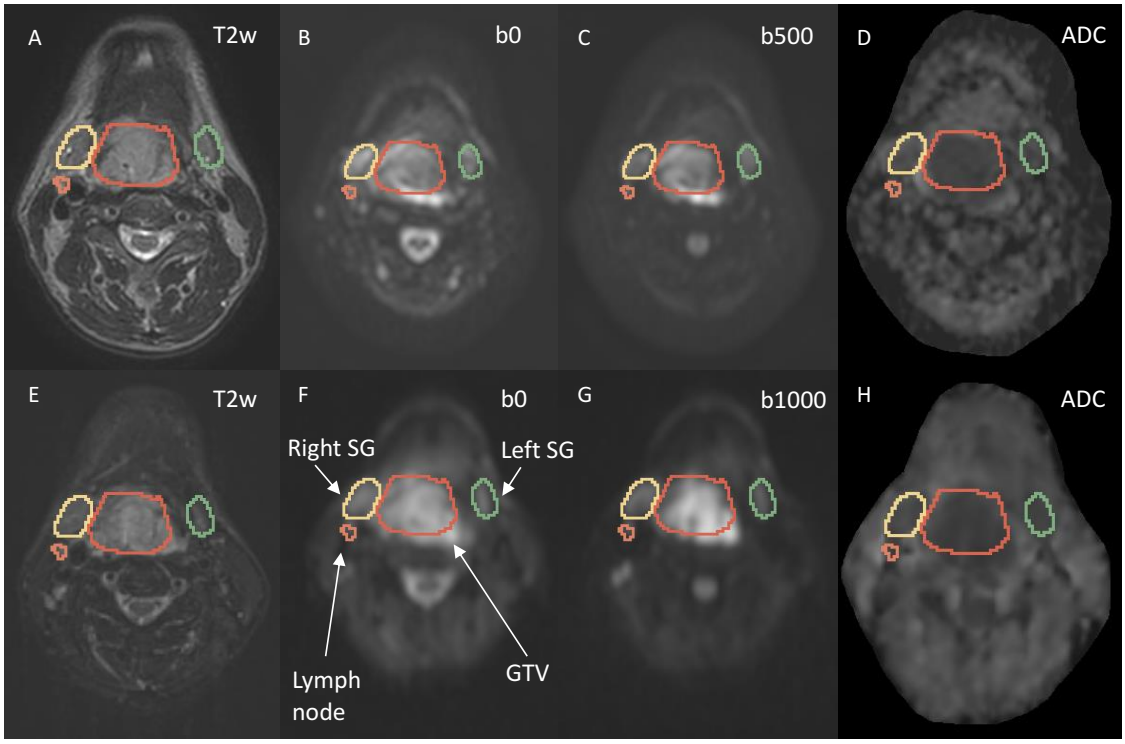

Figure 11: Exemplary pretreatment images of the acquired DW sequences of the MR-Linac (top row) and the diagnostic scanner (bottom row) of patient #13 with T2w images (A and E), b0 images (B and F), b500 image (C), b1000 image (G) and ADC maps (D and H). The green and yellow contour describe the right and left submandibular gland while the red and orange contour present the GTV and lymph node.

For SG, mean volume ± SD, evaluated from the contoured VOIs on the DW images, was 8.5 ± 2.9 cm^3 while mean volume for PG resulted in 19.1 ± 6.7 cm^3. TV had a mean size of 13.4 ± 15.0 cm^3.

Mean ADC values of normal tissues as well as target volumes measured on the MRL resulted in 1120 ± 165 and 796 ± 185 $\cdot 10^{-6} \frac{mm^2}{s}$ for the SG and PG whereas TV had a mean ADC value of 1163 ± 366 $\cdot 10^{-6} \frac{mm^2}{s}$. In contrast, mean ADC values from the DS were 1262 ± 142 and 1051 ± 125 $\cdot 10^{-6} \frac{mm^2}{s}$ for the SG and PG as well as 1256 ± 357 $\cdot 10^{-6} \frac{mm^2}{s}$ for the TV, respectively. Comparison of these ADC values in Bland-Altman analysis resulted in bias ± LoA of 142 ± 340 for SG, 254 ± 299 for PG and 93 ± 493 $\cdot 10^{-6} \frac{mm^2}{s}$ for TV (cf. Fig. 12). All ADC biases were significantly different from 0 mm^2/s ($p < 0.001$).

Reproducibility analysis revealed wCVs of 13.7, 24.4 and 16.1 % for SG, PG and TV, respectively.

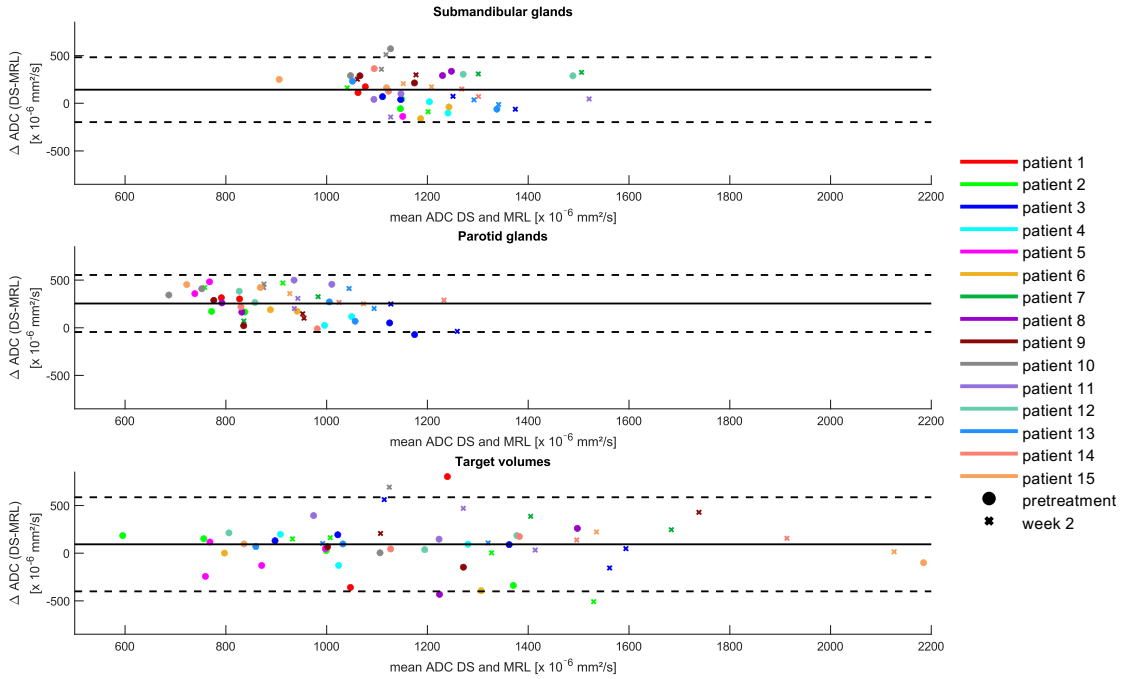

Figure 12: Bland-Altman analysis for submandibular glands, parotid glands as well as target volumes (GTV and lymph nodes) with solid lines as the mean ADC difference (bias) and the dashed lines as the limits of agreement defined as the mean ADC difference ± 1.96 · SD.

Linear regression analysis for all VOIs combined showed R^2 of 0.54. Results of regression analysis are further visualized in Fig. 13. Calculation of ICC values (95 % CI) yielded 0.26 (-0.42–0.52) for SG, 0.23 (-0.1–0.55) for PG, 0.73 (0.56–0.84) for TV and 0.62 (0.24–0.80) for all VOIs combined. Fig. 14 visualizes ADC values acquired using DS and MRL for eight patients with both measurement time points, separated in pretreatment and week 2.

2.2.5 Discussion

The aim of this study was to assess the reproducibility of DW-MRI in HNC by comparing the ADC values acquired on a diagnostic 3 T MRI scanner and a 1.5 T MRL. The comparison of mean ADC values of normal tissues as well as target volumes demonstrated mean ADC biases of 142 and 254 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$ for SG and PG, while target volumes differed by 93 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$.

Several recent studies investigated DW-MRI on DSs and evaluated thereof derived mean ADC values. A mean ADC value in the range of 1140-1350 $\cdot 10^{-6} \frac{mm^2}{s}$ was reported for SG [39, 46, 102, 132, 198], while for PG values between 850 and 1150 $\cdot 10^{-6} \frac{mm^2}{s}$ were published [39, 46, 132]. In this study, the mean ADC value of the SG was 1262 $\cdot 10^{-6} \frac{mm^2}{s}$ and 1051 $\cdot 10^{-6} \frac{mm^2}{s}$ for the PG, measured on the DS. Mean ADC values of the TV measured on the DS resulted in 1256 $\cdot 10^{-6} \frac{mm^2}{s}$, whereas other studies reported mean pretreatment ADC values of HNC in the range of 920-1220 $\cdot 10^{-6} \frac{mm^2}{s}$ [27, 49, 97, 152, 208]. When considering the expected increase of mean TV ADC during the course of radiotherapy [97, 167], our calculated ADC values from the DS seem to be in line with reported ADC values for normal tissues as well as target volumes. On the other hand, there is only one published study reporting ADC values in HNC, acquired on a MRL [143]. Here, mean pretreatment ADC values of 1140, 1330 and 1380 $\cdot 10^{-6} \frac{mm^2}{s}$ were reported for PG, GTV and lymph nodes with an EPI-based sequence, with similar ADC values in target volumes and higher ADC in the parotid glands compared to our results. However, comparison of ADC values across different studies is quite challenging due to differences in evaluated b-values which influence the calculated ADC value [31, 90, 169]. In our study, we used three b-values (200, 500 and 1000 s/mm^2) on the DS and two b-values on the MRL (150/200 and 500 s/mm^2). On the other hand, the previously cited studies analyzed a wide range of b-value combinations with many studies only evaluating two b-values with 0 s/mm^2 as the low value and a high b-value in the range of 600–1000 s/mm^2 [46, 49, 132, 152, 208]. Other studies evaluated six b-values in the range of 0–1000 s/mm^2 but starting and ending at different b-values [39, 97]. Therefore, our calculated ADC values could differ from other studies.

Since hybrid MRLs were only clinically introduced in recent years, studies performing technical validation of functional MRI on MRLs, especially DW-MRI, are still rare. Before reproducibility of functional MRI can be investigated, the repeatability of scanners and sequences should be assessed. While repeatability describes the precision of repeated measurements, performed with the same machine, procedure and under the same operating conditions, reproducibility quantifies the accordance of repeated measurements with a variation of individual parameters [185].

For the 1.5 T MRL, repeatability has been evaluated in several recent studies [61, 104, 106, 143]. Only a few studies investigated the reproducibility of ADC values measured on the MRL by comparison to DSs. The first study by Almansour et al. [5] focused on a qualitative comparison of ADC maps from a 1.5 T MRL and a 3 T DS in prostate cancer. However, they reported that the mean ADC values from the MRL were slightly higher, but the differences were not statistically significant.

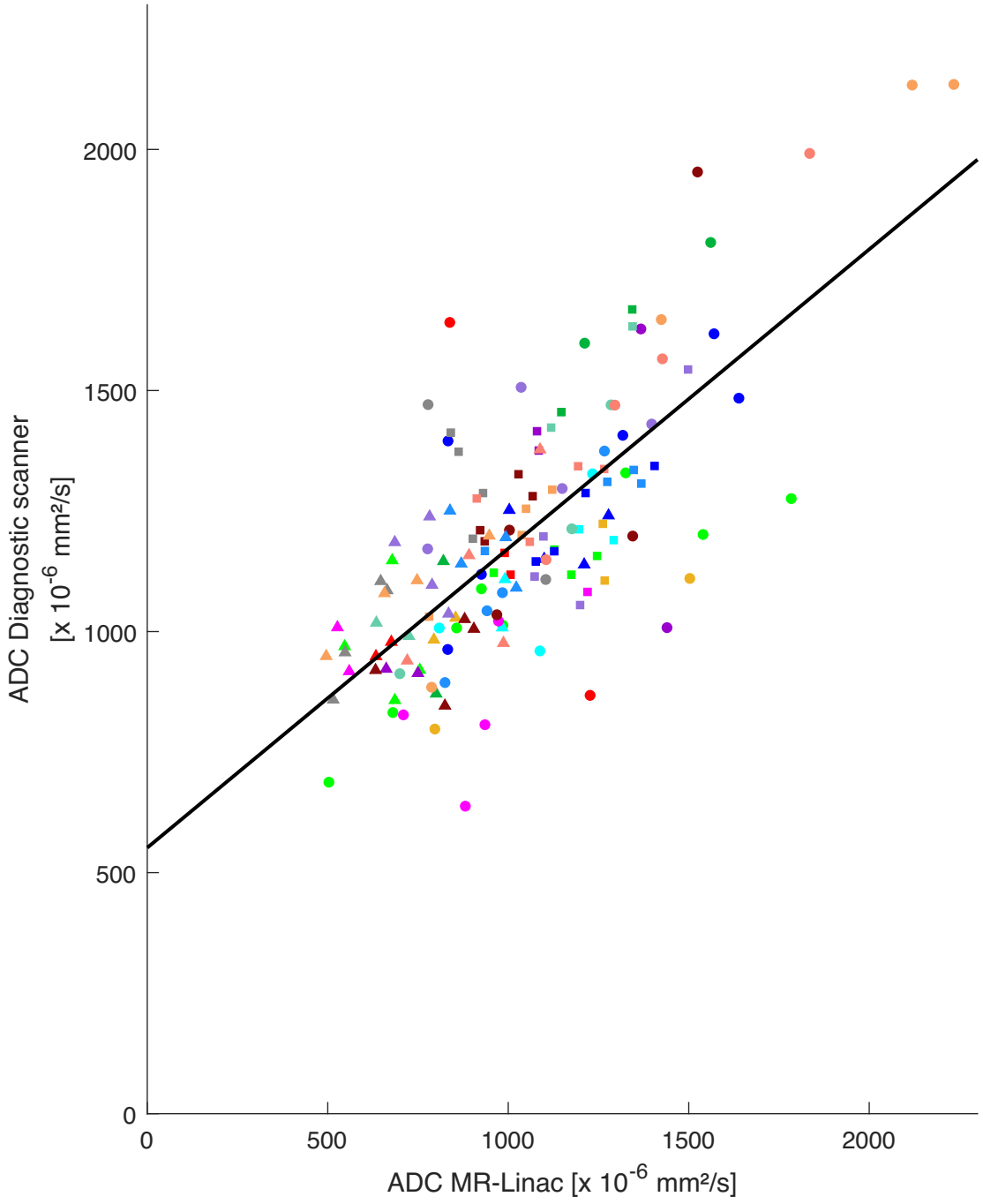

Figure 13: Scatterplot of the ADC values from the diagnostic scanner against the MR-Linac. Linear regression of all VOIs combined is displayed with the solid line. Submandibular glands are shown as squares, parotid glands as triangles and target volumes as circles. Each patient is represented by an individual color according to Fig. 12.

On the other hand, Lawrence et al. determined reproducibility in central nervous

system tumors between a 1.5 T MRL and a 1.5 T DS [114]. They described a bias (MRL-DS) of -50 to -100 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$ for grey matter, white matter and the GTV, while the cerebrospinal fluid differed by about -500 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$. McDonald et al. [143] measured ADC values in HNC patients and volunteers on a 1.5 T MRL and a 1.5 T DS. Here, a mean ADC difference of 340 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$ was determined for the PG comparing an EPI sequence of the DS with the MRL while patients were not imaged on the DS. These results corroborate our calculated biases with all studies reporting an overall underestimation of mean ADC values of the MRL.

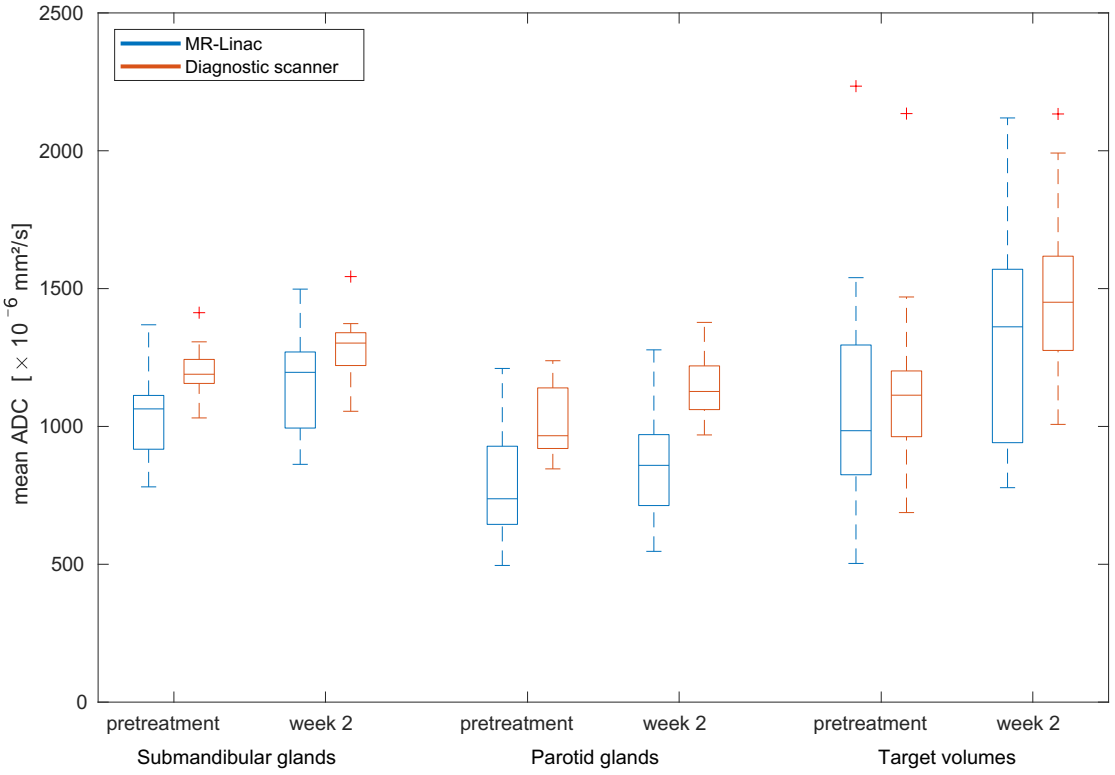

Figure 14: Mean ADC values of VOIs only for eight patients with pretreatment and week 2 measurements.

Theoretically, diffusion and therefore the ADC value are independent of a scanners magnetic field strength or the type of hardware used for the experiment [135, 160], which has been confirmed in different phantom studies [136, 227]. In contrast, several studies evaluated the ADC dependency on the magnetic field strength in vivo in various anatomical sides as well as tumors with varying results. Some studies found no differences of mean ADC values between two magnetic field strengths [41, 144] while other studies determined significant ADC differences [83, 96, 181, 194]. A possible explanation for ADC variation might be differences in echo times (TEs)

in the DW sequences. TEs of both MRL sequences used in our study were longer compared with the DSs TE. For tissues, consisting of multiple water compartments, mean ADC values can vary depending on TE because compartments with short T2 relaxation times contribute relatively less to the measured signal for long acquisition TE. Since HNC tumors are known to have spatially varying T2 relaxation times [18] and glandular tissues can possibly compromise small blood vessels, mean ADC values might have varied depending on the magnetic field strength. Further ADC uncertainties could result from differences in the acquisition protocols of both scanners, including different b-values or mixing times, as well as random image noise.
A major limitation of our study was the uncertainties and errors of the mean ADC calculation caused by misregistration. All acquired DW images of both scanners were rigidly registered to the CT used for RT planning. Therefore, small errors during the registration could lead to variations in the calculated ADC value. This problem is particularly important in the head and neck area, which contains many tissue-air boundaries, possibly resulting in large ADC differences and high reproducibility and biases, respectively. This uncertainty, resulting from misregistration, combined with the potential T2-time dependency of tissue ADC values, might be a possible explanation for the higher ADC differences in both glandular tissues and their low ICC values.
Yet another limitation of the presented study that could have affected ADC and reproducibility calculation was that both scanners used an EPI-based sequence for the acquisition of DW images. Susceptibility artifacts, caused by geometric distortions or signal loss at air-tissue boundaries, are notably prominent in the head and neck region and induce signal loss in the DW images. Furthermore, the magnetic field strength dependency of geometric distortions in DW images may cause ADC variations between the 1.5 T MRL and the 3 T DS. This effect would even be worse if the phase-encoding directions between the scanners differed. Improvement might possibly be achieved by using correction methods for geometric distortion [217, 218] or by using other types of MRI sequences such as turbo-spin-echo (TSE) or SPLICE [183]. Here, susceptibility artifacts can be reduced, but for sufficient signal-to-noise ratio (SNR) longer acquisition times are needed. Nevertheless, when TSE- or SPLICE-based DW sequences are suitably designed for usage on the MRL, they may provide comparable or improved results. McDonald et al. [143] already showed that TSE- and SPLICE-based DW sequences achieve comparable repeatability coefficients in HNC patients and better SNRs than a standard EPI-based sequence. However, all three sequences presented significant ADC biases when compared with each other.
In addition, our study included only a small number of 15 patients. Other studies,

ideally from other institutions, are therefore necessary to confirm the presented results.

2.2.6 Conclusion

In this study reproducibility of ADC values, derived from DW-MRI, in HNC patients was assessed by comparing mean ADC values of a 1.5 T MRL and a 3 T DS. Overall, a small but significant ADC bias was found, indicating an underestimation of ADC values, acquired on the MRL. When results, generated on a DS are translated to MRLs or vice versa, this shift must be considered. Before clinical implementation of response adaptive MR-guided RT based on DW-MRI, further validation studies including more patients as well as external validation with different (diagnostic) scanners are required.

Conflict of interest disclosure statement

The Department of Radiation Oncology Tübingen receives financial and technical support by Elekta, Philips, Dr. Sennewald Medizintechnik, Kaiku Health, TheraPanacea, PTW Freiburg and ITV in the context of research cooperations.
The Department of Diagnostic and Interventional Radiology, University of Tübingen reports institutional collaborations, consultation and grants supported Siemens Healthineers, Erlangen, Germany and Bayer Healthcare, Leverkusen, Germany outside this study.

CRediT authorship contribution statement

Jonas Habrich: Methodology, Software, Formal analysis, Investigation, Writing – original draft, Visualization. **Simon Boeke:** Resources, Investigation, Writing – review & editing. **Victor Fritz:** Resources, Writing – review & editing. **Elisa Koerner:** Investigation, Writing – review & editing. **Konstantin Nikolaou:** Conceptualization, Funding acquisition, Supervision, Writing – review & editing. **Fritz Schick:** Conceptualization, Funding acquisition, Supervision, Writing – review & editing. **Cihan Gani:** Resources, Writing – review & editing. **Daniel Zips:** Conceptualization, Funding acquisition, Writing – review & editing. **Daniela Thorwarth:** Conceptualization, Methodology, Validation, Funding acquisition, Supervi-

sion, Writing – review & editing.

Declaration of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study received funding by the German Research Council (DFG Grants ZI736/2-1 and PAK 997/1: TH1528/6-1, SCHI498/14-1, NI707/7-1). We also thank Dr. David Mönnich and Dr. Marcel Nachbar for their help and support on the MR-Linac.

Supplementary Material

The supplementary material of the publication is depicted in the supplementary material of the dissertation at section 9.2

2.3 Longitudinal assessment of diffusion-weighted imaging during magnetic resonance-guided radiotherapy in head and neck cancer

Authors: Simon Boeke[#], **Jonas Habrich**[#], Sarah Kübler, Jessica Boldt, Fritz Schick, Konstantin Nikolaou, Jens Kübler, Cihan Gani, Maximilian Niyazi, Daniel Zips, Daniela Thorwarth

[#] contributed equally

Published in: Radiation Oncology 2025 Vol. 20(1) Pages 15

DOI: 10.1186/s13014-025-02589-9

The research was carried out in accordance to the approved statement by the local ethics board (No. 659/2017BO1, NCT04172753)

2.3.1 Abstract

Background: For radiotherapy of head and neck cancer (HNC) magnetic resonance imaging (MRI) plays a pivotal role due to its high soft tissue contrast. Moreover, it offers the potential to acquire functional information through diffusion weighted imaging (DWI) with the potential to personalize treatment. The aim of this study was to acquire repetitive DWI during the course of online adaptive radiotherapy on an 1.5T MR-linear accelerator (MR-Linac) for HNC patients and to investigate temporal changes of apparent diffusion coefficient (ADC) values of the tumor and subvolume levels.

Methods: 27 patients treated with curative RT on the 1.5T MR-Linac with at least weekly DWI in treatment position were included into this prospective analysis and divided in four risk groups (HPV-status and localisation). Tumor and lymph node volumes (GTV-P/GTV-N) were delineated on b=500 s/mm^2 images while ADC maps were calculated using b=150/200 and 500 s/mm^2 images. Absolute and relative temporal changes of mean ADC values, tumor volumes and a high-risk subvolume (HRS) defined by low ADC tumor voxels ($600<\text{ADC}<900\cdot10^{-6}\frac{\text{mm}^2}{\text{s}}$) were analyzed. Relative changes of mean ADC values, tumor volumes and HRS were statistically tested using Wilcoxon-signed-rank test.

Results: Median pretreatment ADC value for all patients resulted in $1167\cdot10^{-6}\frac{\text{mm}^2}{\text{s}}$ for GTV-P and $1002\cdot10^{-6}\frac{\text{mm}^2}{\text{s}}$ for GTV-N while absolute pretreatment tumor volume yielded 9.1 cm^3 for GTV-P and 6.0 cm^3 for GTV-N, respectively. Pretreatment HRS volumes were 1.5 cm^3 for GTV-P and 1.3 cm^3 for GTV-P and GTV-N. Median ADC values increase during 35 fractions of RT was 49% for GTV-P and 24% for GTV-N during RT. Median tumor volume decrease was 68% and 52% for GTV-P and GTV-N with a median HRS decrease of 93% and 87%. Significant differences from 0 for mean ADC were observed starting from week 1, for tumor volumes from week 2 for GTV-P and week 1 for GTV-N and for HRS in week 1 for GTV-P and week 2 for GTV-N.

Conclusion: Longitudinal DWI acquisition in HNC is feasible on a MR-Linac during the course of online adaptive MR-guided radiotherapy. Changes in ADC and volumes can be assessed, but future work needs to explore the potential for biologically guided treatment individualization.

Trial registration: NCT04172753, actual study start: 09.05.2018

Keywords: MR-guided radiotherapy; apparent diffusion coefficient; quantitative magnetic resonance imaging; head and neck cancer;

2.3.2 Background

Radiotherapy (RT) is one of the cornerstones in the treatment of head and neck cancer (HNC). Different prognostic factors exist for HNC, with the most important being human papilloma virus 16 (HPV) and smoking status in oropharyngeal cancers (OPC) and mainly tumor stage for other tumor sites [7, 111, 129]. In depth research has been done on imaging biomarkers in HNC, with partially contradictory results. Positron emission tomography either with FDG or hypoxia specific tracers were described to offer prognostic potential and have been used to escalate or de-escalate treatment by dose painting [15, 122, 145, 214, 231].
Due to its superior soft tissue contrast magnetic resonance imaging (MRI) plays a pivotal role in target volume definition and might prove beneficial for automation of these time comsuming steps in HNC [113, 211, 232]. Besides the anatomical information from MRI, it offers the possibility to acquire functional imaging such as dynamic contrast enhanced (DCE) and diffusion-weighted imaging (DWI). Several single center studies looked at different imaging biomarkers with regards to prognostic information and technical validation with contradicting results, especially in HPV associated OPC [59, 61, 62, 187, 188, 203]. Lambrecht et al. showed prognostic information of DWI acquired before radiotherapy with regards to local control [109]. Contradictory to these results, there are observations in HPV associated OPC, that DWI is not prognostic and may not be compared to non-HPV associated HNC [30, 187].
Due to the complexity and limited availability of MRI, only a few studies were reported focusing on longitudinal imaging during the course of fractionated RT [2, 149, 154, 203]. In recent years the technique of magnetic resonance linear accelerators (MR-Linac) has been introduced and first data were published for MR-guided RT for HNC [20, 142]. Initial results showed the feasibility of MR-guided RT for HNC, providing the potential to acquire longitudinal anatomical imaging with high resolution and functional imaging during the course of fractionated radiotherapy [20, 142, 200]. Moreover, in our department a potential quantitative imaging biomarker (QIB) based on a cluster of ADC values has been established in a preclinical model, suggesting a high-risk subvolume (HRS), which could be used for individualized dose prescription [19]. A validation on a clinical cohort of patients has been performed in which the retrained HRS was found to be significantly associated to outcome after primary radiochemotherapy , so this ADC-based HRS might serve as a biomarker for stratification or therapy individualization of HNC patients [219].
The aim of the present study was to explore the feasibility of online acquired, lon-

gitudinal DWI with an exploratory analysis of changes in different DWI parameters during the course of MR-guided RT in HNC patients.

2.3.3 Methods

Patient and treatment characteristics:

A total of n=28 patients with HNC, which were treated on the 1.5 T MR-Linac (Unity, Elekta AB, Sweden) between October 2018 and December 2021 have been prospectively included into this analysis within a enclosed project funded by the German Research Council. All patients gave written informed consent to be treated within this prospective trial of MR-guided adaptive RT which was approved by the local ethics committee (no. 659/2017BO1, NCT04172753). All patients were immobilized with a 5-point radiotherapy mask on a dedicated head step for the use on the 1.5 T MR-Linac in a neutral neck position [62]. The planning-CT of 2 to 3 mm slice thickness and all MRIs have been acquired in this setup as shown in detail elsewhere [20, 61, 62, 142]. Patients were instructed to limit swallowing whenever possible. Treatment consisted of fractionated RT with 60 Gy and 54 Gy in 30 fractions to the macroscopic tumor, the high risk area and the elective volume with a sequential boost to the macroscopic tumor of 10 Gy in 5 fractions, consistent with international guidelines. Concomitant weekly Cisplatin was administered. The treatment was performed on the MR-Linac with an online "adapt-to-position" workflow and offline adaptation in case of large anatomical deviations based on the treating physicians discretion [142]. Detailed patient characteristics are shown in Table 6. Follow-up was according to clinical routine with a FDG-PET/CT after 3 months in case of nodal involvement and further clinical and radiological examinations every 3 months. For OPC, risk groups have been defined according to Ang et al. with stratification based on p16-, smoking-status and TNM stage [7]. The remaining six non-OPC patients have been included into a fourth patient group.

Imaging protocol:

Imaging of all patients was performed on the 1.5 T MR-Linac except for pretreatment images of patient 25 which were acquired on a 1.5 T diagnostic scanner (Ingenia, Philips). The protocol consisted of a single-shot echo-planar imaging (SS-EPI) DWI as well as T1- and/or T2-weighted anatomical imaging. For the first eleven patients

Table 6: Patient characteristics.

Patient characteristics		**N = 27**
Age (median, range)		67 (39 - 78)
Primary tumor site	Hypopharyngeal	5
	Oropharyngeal	21
	Supraglottic Larynx	1
Primary tumor (T) stage	T1	0
	T2	9
	T3	13
	T4	5
Nodal tumor (N) stage	N0	0
	N1	9*
	N2	8+
	N2a	0
	N2b	5
	N2c	4
p16-status	Positive	15 #
	negative	12
smoking status	current	9
	former	9
	never	9

* n = 6 p16 positive

+ n = 8 p16 positive

n = 1 non-OPC

the acquired b-values of the DWI sequence were 0, 200, 500 and 800 s/mm^2 while starting with patient 12, the DWI sequence was adapted according to the recommendations of the MR-Linac consortium [105] and b-values 0, 150 and 500 s/mm^2 were acquired. For patient 2 the b-value 500 s/mm^2 was not part of the DWI sequence pretreatment and in fraction four of RT, while patient one had to be excluded from analysis due to missing DWI data. Details about the sequence parameters of both

DWI as well as the two anatomical MRI protocols are provided in Table 7. Patients were imaged in RT position either directly before the start, during beam on or after completion of RT on the MR-Linac. The imaging protocol was applied during simulation before the start of radiotherapy and sequentially during RT, approximately once per week.

Table 7: Details of the sequence parameters.

parameter	EPI3b	EPI4b	T1w MRI (T1_3D_Tra)	T2w MRI (T2_3D_Tra)
Sequence type	SS-SE-EPI[1]	SS-SE-EPI	FFE[2] 3D	MS-SE-TSE[3] 3D
Acquisition voxel size [mm^3]	3 x 3 x 4	3 x 3 x 4.8	1.2 x 1.2 x 2.4	1.2 x 1.2 x 2.2
Slice gap [mm]	0	0	-1.2	-1.1
Field of view [mm^3]	400 x 400 x 100	400 x 400 x 202	520 x 298 x 250	520 x 298 x 250
TR/TE [ms]	4811/68	10392/107	13/4.5	2100/375
Flip angle [°]	90-180	90-180	27	90-(180)
TSE factor	—	—	—	150
Water-fat shift (pix)/ bandwidth (Hz)	11.197/19.4	11.180/19.4	0.653/332.6	0.473/459.3
b-values [s/mm^2] (averages)	0 (3), 150 (5), 500 (8)	0 (2), 200 (3), 500 (4), 800 (6)	—	—
Fat suppression	SPAIR[4]	SPAIR[4]	no	no
Duration [min]	3:32	7:17	5:48	6:03

[1] single-shot spin-echo echo-planar imaging

[2] Fast Field Echo

[3] multi-shot spin-echo turbo spin-echo

[4] Spectral Attenuated Inversion Recovery

Image processing:

Volumes of interest (VOIs) defined as primary tumors (GTV-P) and all conspicuous lymph nodes (GTV-N) were delineated by a board-certified radiation oncologist (SB) on the b-500 images with the open-source software 3D Slicer (Version 4.10) taking the anatomical images as visual input to ensure the capture of the whole tumor volume. Then, ADC maps were calculated using an in-house written python script (version 3.8.10) and b-values 150/200 and 500 s/mm^2 with the mono-exponential model [191]

$$SI = SI_0 \cdot e^{-b \cdot ADC}.$$

Here SI is defined as the signal intensity and SI_0 as the calculated signal intensity at b-value 0 s/mm^2.

Furthermore, a high-risk subvolume (HRS) was defined, representing potentially radioresistant tumor subregions [19] inside both GTV-P and GTV-N, separately. This HRS was defined as all voxels inside the GTV-P or GTV-N with ADC values in the

range of 600-900·$10^{-6}\frac{mm^2}{s}$. ADC thresholds were used from a prior investigation [219] and adapted for the use on the MR-Linac to the known underestimation of ADC on the MR-Linac [62, 143, 220].

Statistics:

GTV-P and GTV-N volumes as well as mean ADC values (ADC_{mean}) and HRS volume were calculated using Matlab 2020a (MathWorks, Natick, MA, USA). Target volumes smaller than 1 cm^3 and HRS smaller than 0.2 cm^3 were excluded from further analysis. Absolute values of ADC and volumes were evaluated for all patients for baseline and every week of radiotherapy. Furthermore, linear regression as a function of treatment time regarding mean ADC values as well as volumes of GTV-P, GVT-N and HRS was fitted for all risk groups with Matlab. Additionally, relative changes of ADC and volume from baseline to every week of RT were evaluated for every risk group. One-sample Wilcoxon-signed-rank test with a significance level of 5% was used to evaluate relative changes of ADC_{mean}, absolute tumor volumes and HRS volumes in every week of radiotherapy for a significant difference from 0 using SPSS statistical package 28.0.0.0 (SPSS Inc, Chicago, Illinois).

2.3.4 Results

A total of 28 patients have been treated on the MR-Linac out of which 27 primary lesions as well as 39 lymph nodes from 27 patients were available for analysis. Median (interquartile range) number of scans per patient during treatment was 7 (1). Exemplary pretreatment anatomical as well as DW images for patient 17 including delineations of the primary tumor and calculated HRS are shown in Figure 15. At the time of analysis median follow up was 40.4 months, with five patients having died due to metastatic disease and one likely due to pneumonia after aspiration. One patient with a combined hypopharyngeal and proximal esophageal carcinoma (in initial endoscopy distant to each other) showed a recurrence in the proximal esophagus 18 months after completion of the treatment with no recurrence in the hypopharynx. No isolated local or regional recurrence have been observed in the cohort.

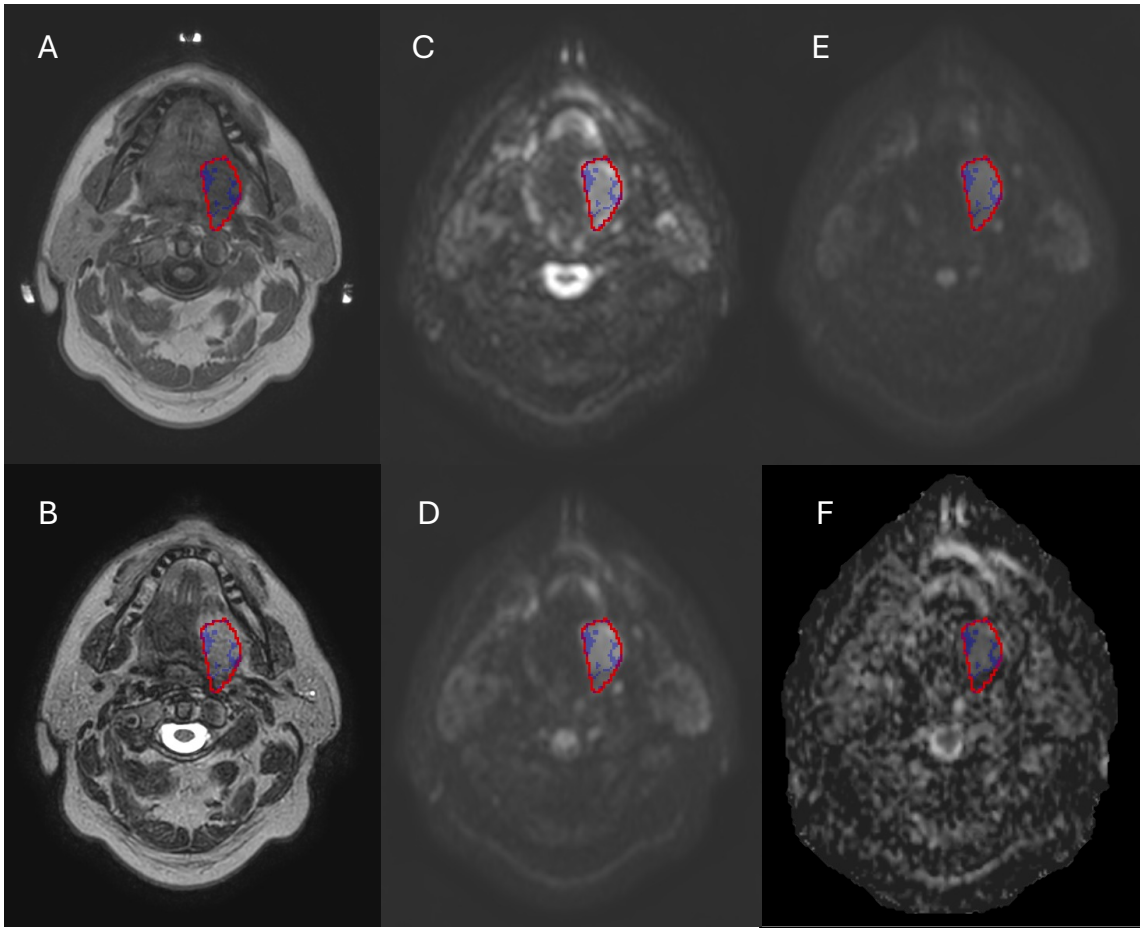

Figure 15: Exemplary pretreatment images of patient #17 including T1w (A), T2w (B), b0 image (C), b150 image (D), b500 image (E) and calculated ADC map (F) with GTV depicted in red and high-risk subvolume ($600<\text{ADC}<900 \cdot 10^{-6}\,\frac{\text{mm}^2}{\text{s}}$) in blue.

Median (interquartile range) pretreatment ADC value for all patients amounted to 1167 (282) $\cdot 10^{-6}\,\frac{\text{mm}^2}{\text{s}}$ for GTV-P and 1002 (398) $\cdot 10^{-6}\,\frac{\text{mm}^2}{\text{s}}$ for GTV-N, respectively. Over the course of radiotherapy, the ADC_{mean} of primary lesions increased to 1687 (381) $\cdot 10^{-6}\,\frac{\text{mm}^2}{\text{s}}$ and 1420 (424) $\cdot 10^{-6}\,\frac{\text{mm}^2}{\text{s}}$ of conspicuous lymph nodes in the last week of radiotherapy. The temporal changes including linear regression of ADC_{mean} for GTV-P and GTV-N over the course of radiotherapy depending on the patient's risk classification are displayed in Figures 16 and 17 for all patients.

Regarding tumor volumes, median pretreatment GTV-P volume was 9.1 (15.4) cm^3 while GTV-N displayed a median size of 6.0 (12.2) cm^3. Median primary tumor volumes decreased over the course of radiotherapy to 3.6 (3.1) cm^3 in the last week of RT. In contrast, the conspicuous lymph nodes had a median volume of 2.8 (4.4)

cm^3 in the last week of RT. For all patients, trends of absolute tumor volumes and linear regression are visualized in Figures 18 and 19 for GTV-P and GTV-N, respectively.

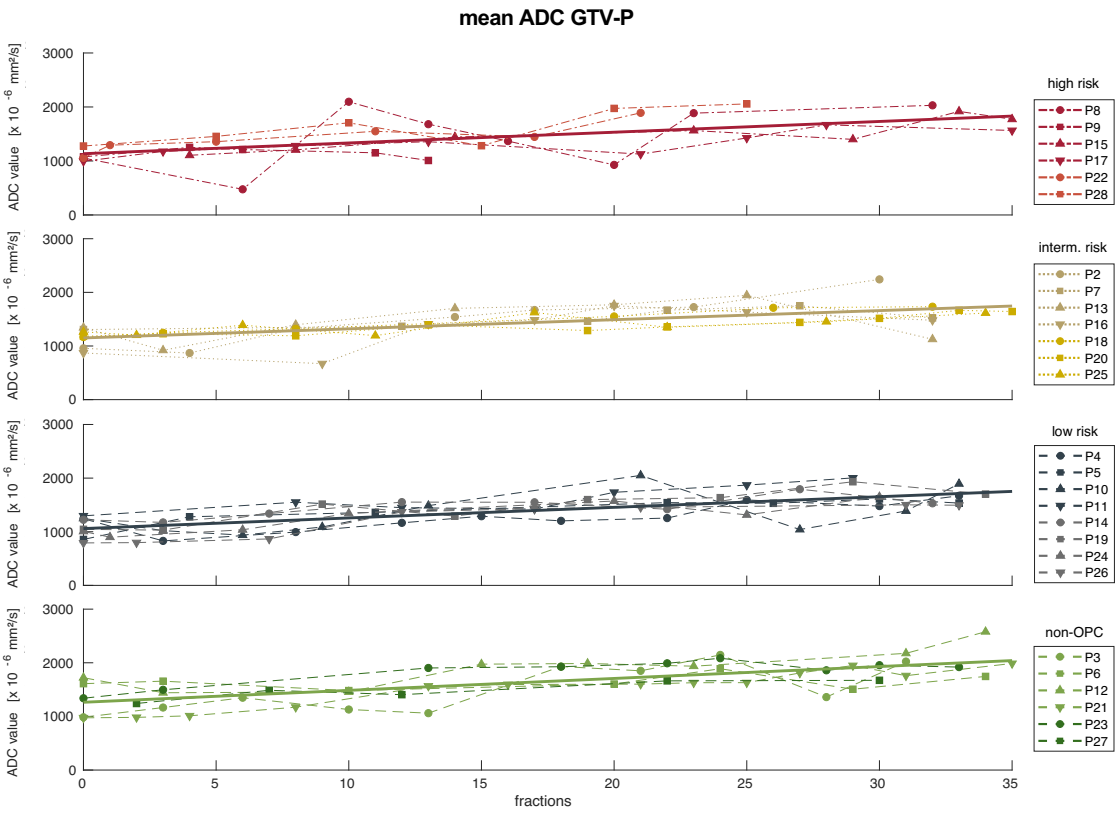

Figure 16: Mean ADC values for GTV-P over the course of radiotherapy for high (red), intermediate (gold), low risk (blue) and non-oropharyngeal (green) patients (according to Ang et al. [7]. Solid lines represent linear regression for each group of patients.

The ADC-based HRS of the GTV-P and GTV-N showed a median pretreatment size of 1.5 (3.1) cm^3 and 1.3 (2.4) cm^3, respectively. In the last week of RT, the median HRS had shrunk to 0.4 (0.0) cm^3 for GTV-P and 0.6 (0.5) cm^3 for GTV-N. Figures 20 and 21 show the trend of HRS for all analyzed patients and the linear regression. Detailed information about ADC_{mean}, absolute tumor volume and HRS volumes for every risk group pretreatment and in the last week of RT are presented in Table 8. Additionally, absolute values for ADC_{mean}, tumor volume and HRS are summarized on a weekly basis for all patients in Figure 22 and separated by risk groups in Figure 23.

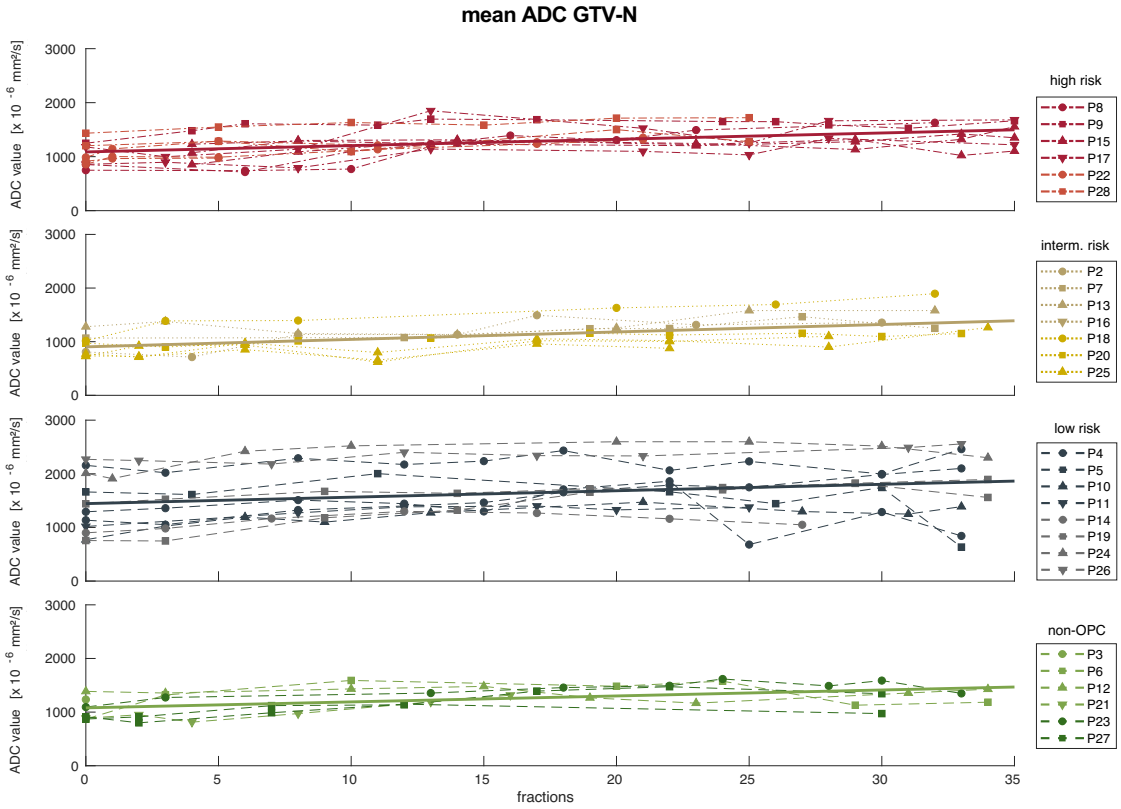

Figure 17: Mean ADC values for GTV-N over the course of radiotherapy for high (red), intermediate (gold), low risk (blue) and non-oropharyngeal (green) patients (according to Ang et al. [7]. Solid lines represent linear regression for each group of patients.

Table 8: ADC_{mean}, absolute tumor volumes and HRS volumes before the start of radiotherapy (RT), in week 7 of RT and the relative change over course of RT separated by risk groups.

Median (interquartile range) ADC_{mean} $[\cdot 10^{-6}\frac{mm^2}{s}]$	High risk		Intermediate risk		Low risk		Non-oropharyngeal	
	GTV-P	GTV-N	GTV-P	GTV-N	GTV-P	GTV-N	GTV-P	GTV-N
pretreatment	1057 (152)	952 (358)	1181 (350)	890 (309)	1134 (354)	1292 (1112)	1339 (687)	913 (364)
week 7 of RT	1845 (386)	1420 (410)	1615 (184)	1264 (537)	1532 (198)	1895 (1211)	1987 (416)	1352 (184)
Median (interquartile range) volumes [cm^3]								
pretreatment	20.5 (41.4)	11.7 (15.8)	9.1 (7.7)	3.3 (10.6)	14.0 (17.4)	4.6 (20.7)	7.2 (32.4)	6.8 (6.6)
week 7 of RT	4.0 (3.0)	5.6 (3.9)	2.8 (1.6)	1.5 (2.2)	4.2 (3.8)	3.6 (12.1)	4.0 (24.2)	2.1 (1.3)
Median (interquartile range) HRS [cm^3]								
pretreatment	4.0 (10.7)	1.9 (3.7)	1.5 (3.1)	1.3 (3.0)	2.5 (3.3)	1.0 (2.3)	1.0 (13.7)	1.3 (2.3)
week 7 of RT[a]	—	0.8 (0.6)	0.4	0.3, 0.8	0.2, 0.3, 0.6	—	0.4, 0.4	0.2, 0.4

[a] in case of three or less data points, all datapoints are presented.

Median ADC_{mean} increase to the last week of radiotherapy was 49 (52)% for GTV-P and 24 (48)% and GTV-N. GTV-P volumes showed a median decrease of 68 (24)% while GTV-N volumes decreased by 52 (62)% to the last week of RT. Furthermore, median decrease of HRS volumes yielded 93 (9)% for GTV-P and 87 (18)% for GTV-N. Significant differences ($p<0.05$) for relative change of ADC_{mean} of GTV-P and GTV-N were found starting in week 1 of RT while tumor volumes showed significant differences ($p<0.05$) beginning in week 2 for GTV-P and week 1 for GTV-N, respectively. HRS size displayed significant differences ($p<0.05$) between baseline and week 1 of RT for GTV-P and week 2 for GTV-N. Furthermore, we evaluated the relative changes of ADC_{mean}, absolute tumor volumes and HRS size depending on patient risk stratification from baseline to every week of radiotherapy (cf. Figure 24). All clinical risk categories showed an increase of ADC parameters during treatment and decrease in HRS during treatment.

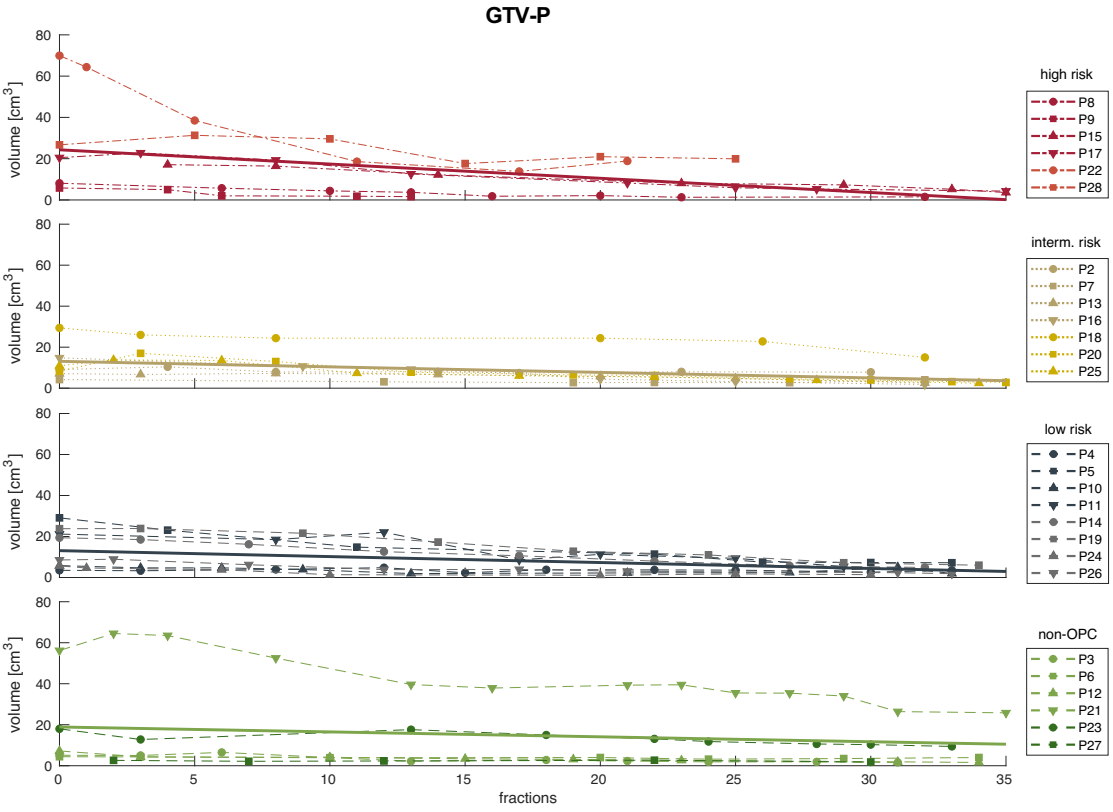

Figure 18: Volumes of GTV-P over the course of radiotherapy for high (red), intermediate (gold), low risk (blue) and non-oropharyngeal (green) patients. Solid lines represent linear regression for each group of patients.

2.3.5 Discussion

As new QIBs need to undergo technical and also clinical validation with respect to different features such as reliability and repeatability, the aim of the present analysis was to show clinical feasibility of online acquired, longitudinal DWI on a MR-Linac during fractionated RT for HNC and to characterize ADC dynamics. Previously, we reported relative repeatability coefficients (relRCs) of DWI in HNC on the MR-Linac to be 31% for GTV-P and 23% for GTV-N [61] and performed a comparison with a diagnostic scanner as a benchmark [62]. In this study, we observed ADCmean changes larger than these relRCs in weeks 3-4 for GTV-P and weeks 2-3 for GTV-N which correspond to the timepoints where significant ADC changes from 0 could be measured. Hence, week 3 seems to be a reasonable timepoint for ADC based interventions on a MR-Linac.

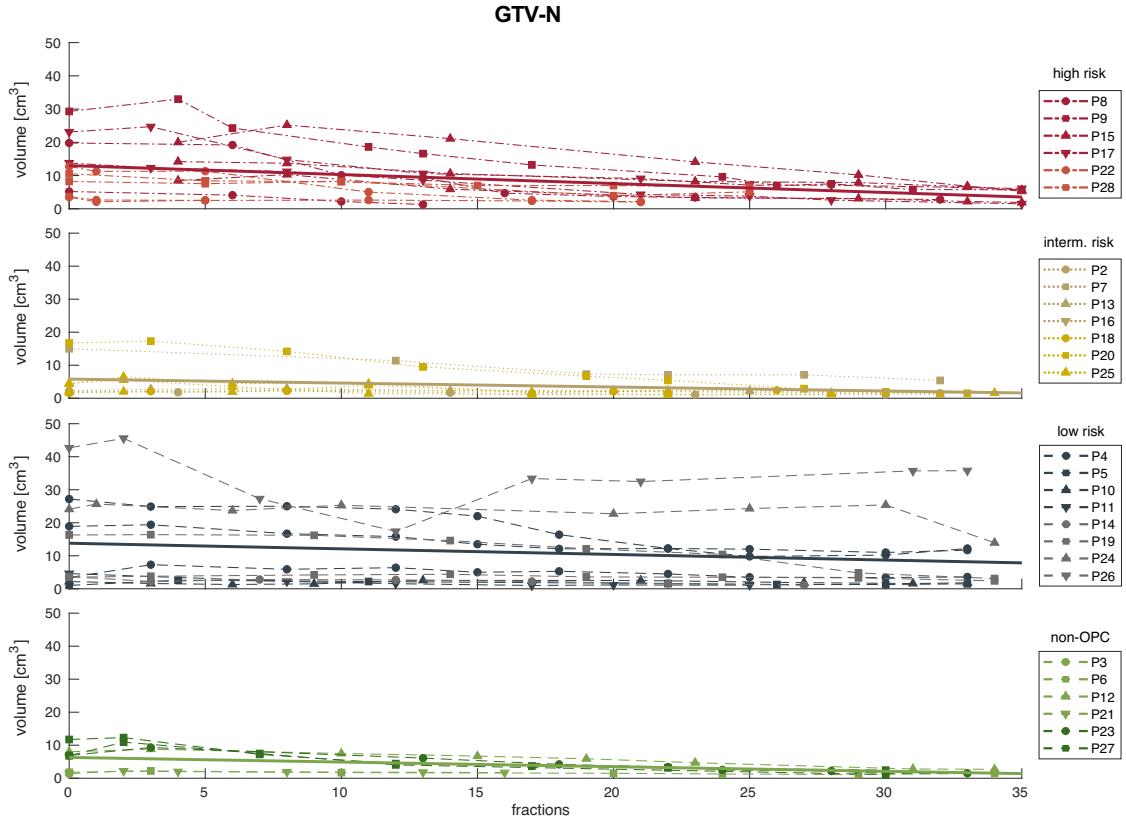

Figure 19: Volumes of GTV-N over the course of radiotherapy for high (red), intermediate (gold), low risk (blue) and non-oropharyngeal (green) patients. Solid lines represent linear regression for each group of patients.

The current patient cohort is one of the largest to date reporting longitudinal ADC during RT and one of the first to report on online acquired serial DWI. As anticipated and in congruence with previously published data, GTV-P and GTV-N

volume decreased during treatment [65, 154, 192]. El-Habashy et al. showed a decrease of GTV-P and GTV-N volumes during fractionated RT for 30 patients treated on an 1.5 T MR-Linac [60]. These volume changes built the rationale for prospective adaptive RT trials such as the MR-ADAPTOR [10, 150]. Our results also show a consistent increase in mean ADC in GTV-P and GTV-N, as reported by others in an offline setting, too [59, 154, 187, 203].

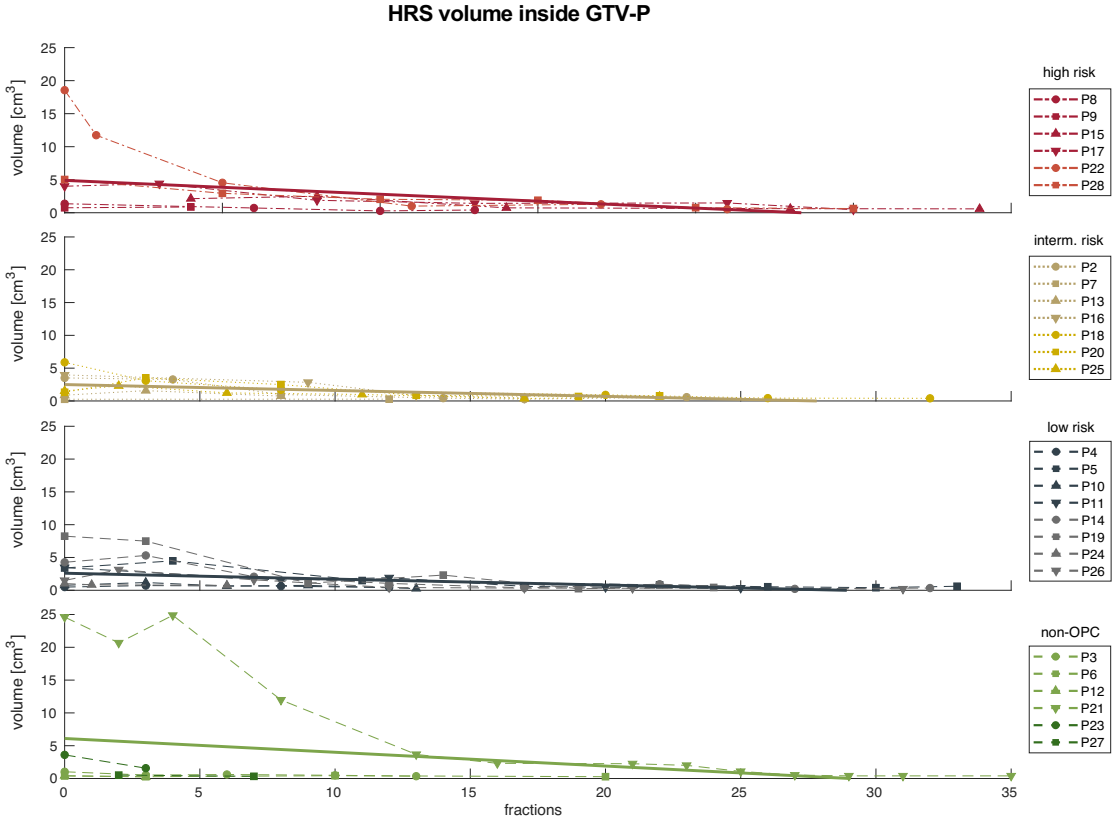

Figure 20: Trend for HRSs inside GTV-P over the course of radiotherapy for high (red), intermediate (gold), low risk (blue) and non-oropharyngeal (green) patients. Solid lines represent linear regression for each group of patients.

One major limitation is the inherent high technical uncertainty related to the acquisition and analysis of DWI. While echo-planar-imaging (EPI) for DWI has the advantage of fast image acquisition, it severely suffers from geometrical distortions caused by inhomogeneities of the magnetic field [88, 180, 204]. EPI applications in HNC patients are especially prone to geometric distortions because of many air-tissue boundaries in the respective area which cause magnetic field inhomogeneity and therefore geometric distortions. Improvements could be achieved by using image acquisition techniques like turbo-spin-echo (TSE) or split acquisition of fast spin echo signal (SPLICE), but with the downside of lower signal-to-noise ratio (SNR)

and longer acquisition times [143, 180, 183]. Furthermore, comparability of ADC values and the translation of results from other studies is complicated due to inconsistencies in selected b-values for the DWI sequences and different algorithms used for ADC calculation, different magnetic field strengths and imaging techniques which all might affect the reported ADC values [17, 44, 143]. Another limitation is the small number of patients receiving their treatment within a planned period of time, limiting the potential prognostic value of the analysis as discussed above. Therefore, the MR-Linac consortium is working on publishing consensus DWI sequences or guidelines for the 1.5 T MR-Linac to form a prospective large group of patients, scanned with a dedicated DWI sequence and to make the translation and interpretation of ADC values from different studies easier and to shed further light on the role of DWI and its prognostic value for HNC.

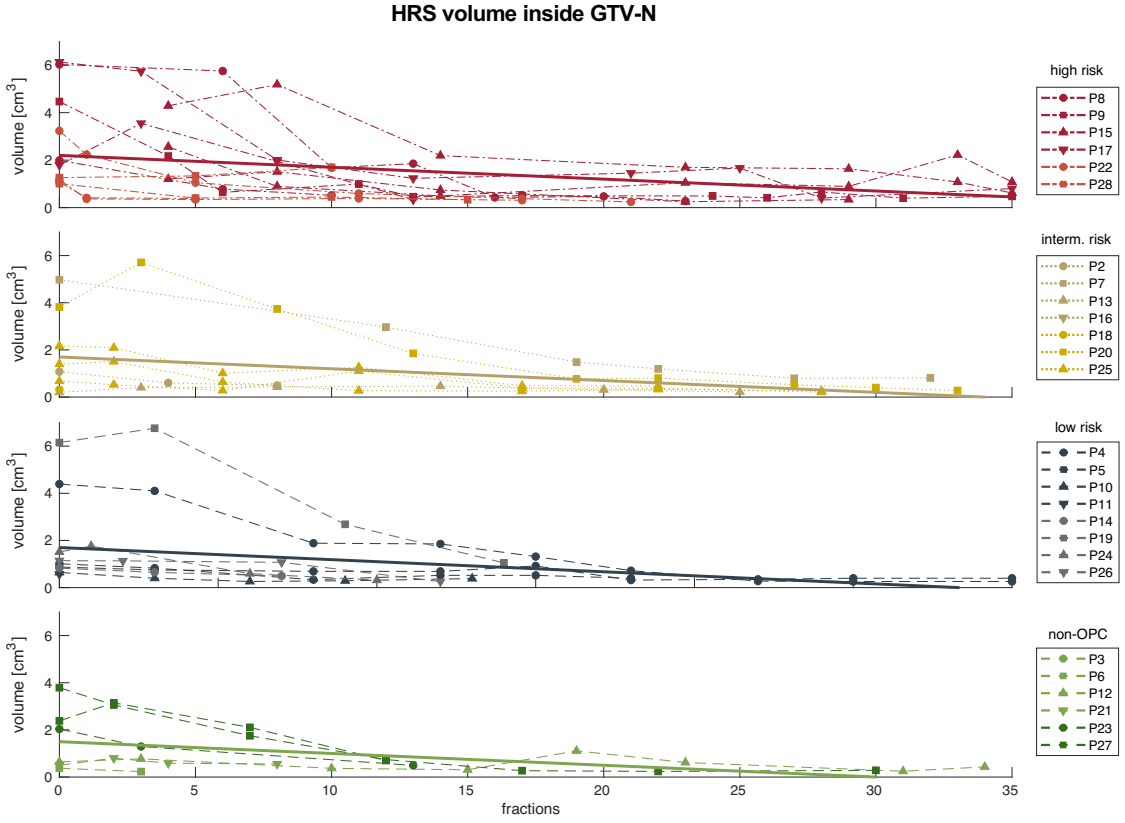

Figure 21: Trend for HRSs inside each individual GTV-N over the course of radiotherapy for high (red), intermediate (gold), low risk (blue) and non-oropharyngeal (green) patients. Solid lines represent linear regression for each group of patients.

In a preclinical trial, tumour subvolumes based on a cluster of ADC values were found to be significantly associated with radiation sensitivity and local tumor control [19]. In this clinical cohort this association was not reproduced, as the number of events was very small and moreover the reproducibility for serial DWI measurements might be worse than in a classical diagnostic setup and especially compared to a preclinical setting. In contrast to other reported data [187], we could not find a difference between OPC and non OPC or between HPV associated OPC and non-HPV associated OPC, possibly due to the small subgroups.

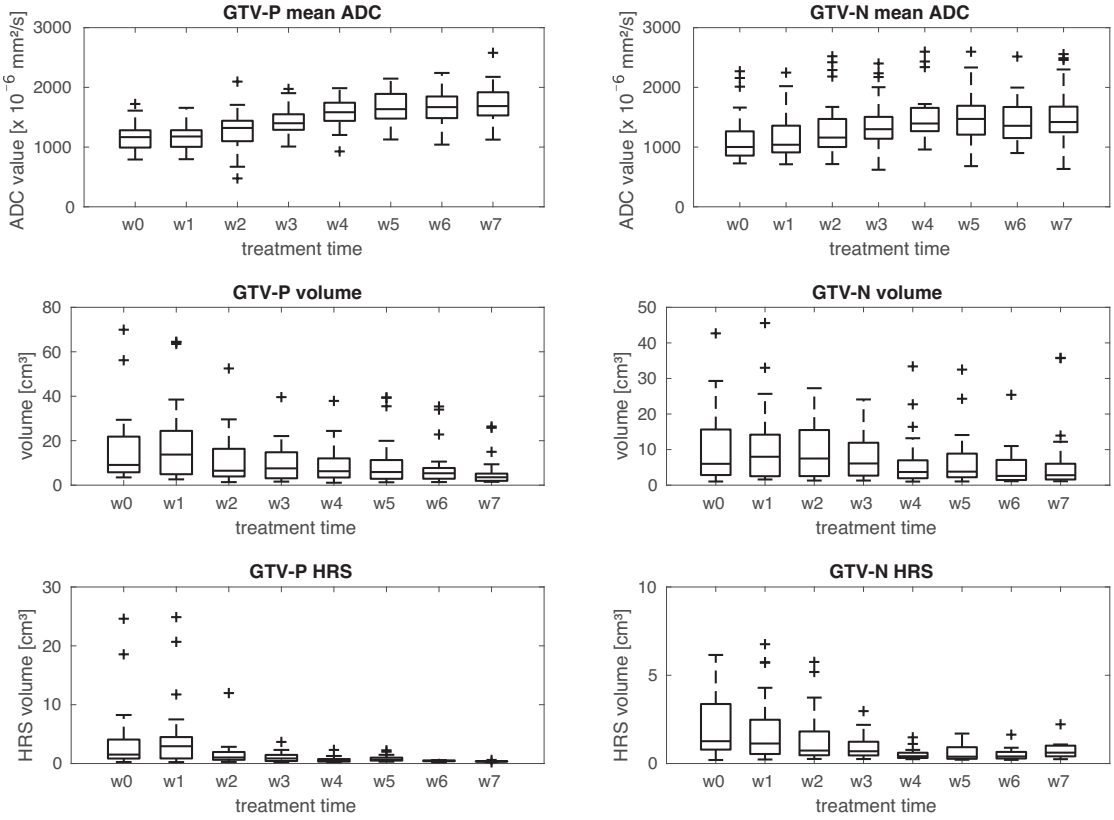

Figure 22: Boxplots of mean ADC, absolute tumor volumes and absolute HRS volumes for all patients before and in every week of radiotherapy.

With no isolated local or regional recurrence, likely due to sampling bias, no correlation of DWI with outcome in this small cohort could be generated. Moreover, the established clinical risk groups for OPC could not be differentiated by the means of mean ADC or ADC-based HRS, potentially due to the small number of patients in each subgroup. Although no obvious differences could be seen between the subgroups in OPC or non-OPC, the high risk subgroup had larger tumor volumes and showed initially a higher HRS, consistent with the potential as a QIB, whereas intermediate and low risk subgroups had already a median HRS below the proposed

threshold of 5.8 cm^3 [219]. The missing correlation with outcome seen in the present analysis, emphasizes the need of larger cohorts of patients and a group of patients suffering from less selection bias. In addition, the HRS even in the high risk and the non-OPC subgroup decreased very rapidly and were below the established threshold of 5.8 cm^3 already in week 2 for all but one of the patients (Fig. 20). This, together with the significant decrease in various parameters already early during treatment (Fig 24) emphazises the need of repeated imaging to access the dynamic change of imaging based biomarkers and needs to be taken into account for future QIB-based clinical trials. Moreover, tumor volume in some cases was very small and the resolution of the DWI sequence may affect the analysis especially in the ADC-based HRS at late time points during RT. The seen GTV volume and ADC changes early during treatment offer a further explanation for the so far inconclusive and conflicting body of literature on the potential role of DWI as a prognostic or even predictive marker in HNC [69, 71, 109, 168, 187, 188, 219]. Nevertheless, the potential of repeated online quantitative imaging and anlysis of ADC-based subvolumes may facilitate the possibility for an ADC-based dose individualization and emphasizes the consideration of adaptive approaches during QIB-based trials.

2.3.6 Conclusion

In conclusion, this study shows the feasibility of online acquired, longitudinal DWI on a tumor level as well as in ADC-based subvolumes during the course of MR-guided RT for HNC, which might build the basis of ADC-based biological individualized online adaptive RT trials. The validation of these results in a prospective multicenter study seems to be an important next step.

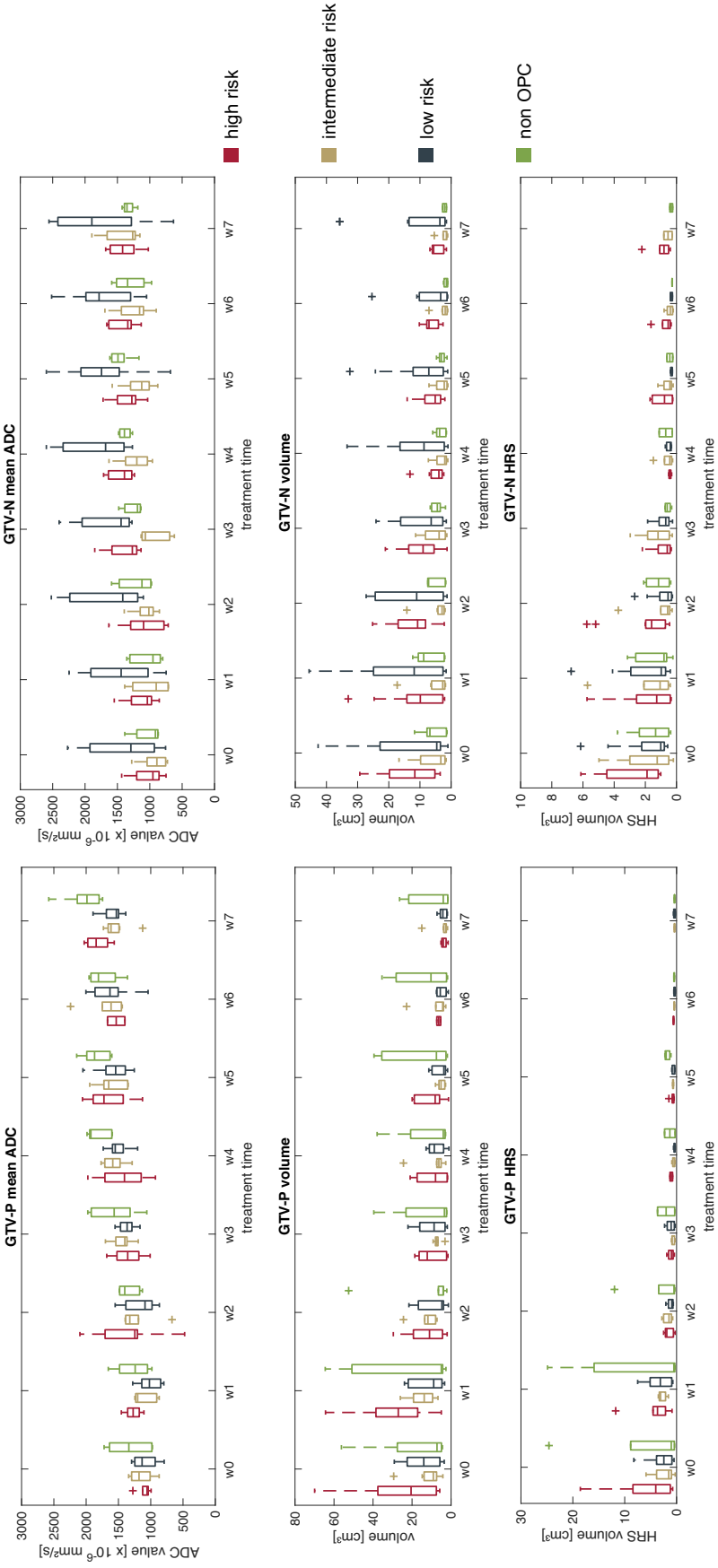

Figure 23: Boxplots of mean ADC values (top) and absolute tumor volumes (bottom) pretreatment (w0) and during radiotherapy (w1-w7) for high risk (red), intermediate risk (gold), low risk (blue) and non-oropharyngeal (green) patients.

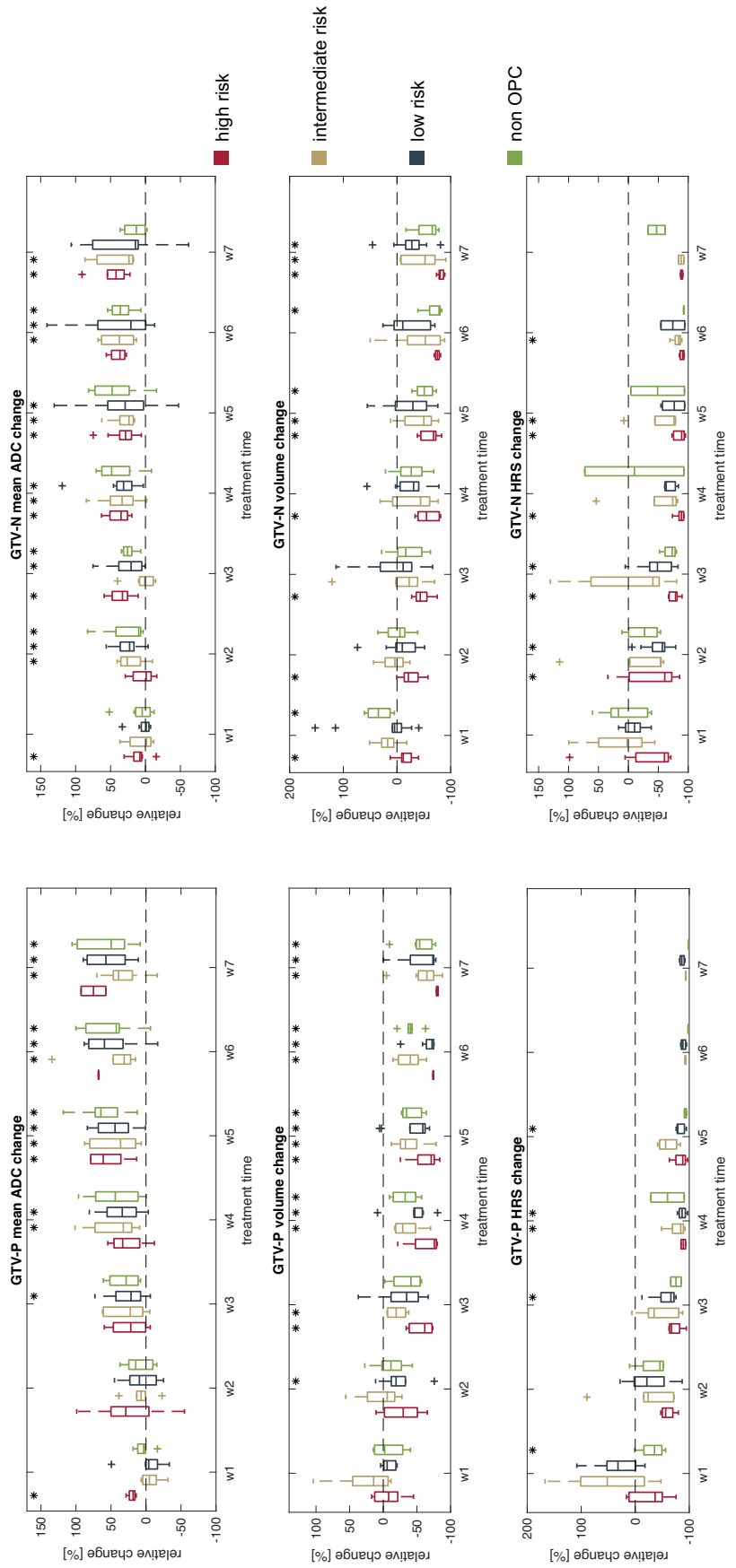

Figure 24: Boxplots of relative change of mean ADC values (top) and tumor volumes (bottom) from baseline to every week of radiotherapy (w1-w7) for high risk (red), intermediate risk (gold), low risk (blue) and non-oropharyngeal (green) patients. Outliers are shown with a plus sign. Significant different changes ($p<0.05$) from 0 are indicated by a star.

Abbreviations

ADC	apparent diffusion coefficient
CT	computed tomography
DCE	dynamic contrast enhanced
DWI	diffusion-weighted imaging
EPI	echo-planar imaging
FDG	fluorodeoxyglucose
GTV-N	conspicious lymph nodes
GTV-P	primary tumor
HNC	head and neck cancer
HPV	human papilloma virus 16
HRS	high-risk subvolume
MRI	magnetic resonance imaging
MR-Linac	magnetic resonance linear accelerator
OPC	oropharyngeal cancer
PET	positron emission tomography
QIB	quantitative imaging biomarker
relRC	relative repeatability coefficient
RT	radiotherapy
SNR	signal-to-noise ratio
SPLICE	split acquisition of fast spin echo signal
SS-EPI	single-shot echo-planar imaging
TSE	turbo-spin-echo
VOI	volume of interest

Acknowledgements

We acknowledge support from the Open Access Publication Fund of the University of Tübingen.

Author contributions

SB did a majority of the patients recruitement and data acquisition, contouring, selected analysis methods, interpreted the results and was a major contributor to wirting of the manuscript. JH selected the analysis methods, did data acquisition, solely analysed the data, interpretated the results and was a major contributor to writing the manuscript. SK helped acquiring the data. JB supported in the data acquisition. FS contributed substantially to the research concept, selection of methods and interpretation of the results. KN made major contributions to the research concept, selection of methods and interpretation of the results. JK helped with the data acquisition. CG contributed majorly to the patient recruitement and data acquisition. MN added to selection of methods, critical interpretation of results and writing of the manuscript. DZ substantially accounted for the research concept, selection of methods and interpretation of the results. DT added majorly to the research concept, selection of methods, interpretation of the results and writing of the manuscript. All authors read and approved the final manuscript.

Funding

This study received funding by the German Research Council (DFG Grants PAK 997/1: TH1528/6-1, SCHI498/14-1, NI707/7-1 and ZI 736/2-1).

Data availability

The datasets generated and analysed during the current study are not publicly available due to inclusion in an ongoing prospective registry for secondary multi-center analysis but are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The trial was approved by the local ethical committee (no. 659/2017BO1, NCT04172753) and all patients gave their written informant consent.

Consent for publication

Approval for publication of patient images was included in all patients' written informant consent.

Competing interests

The Department of Radiation Oncology Tübingen receives financial and technical support by Elekta, Philips, Dr. Sennewald Medizintechnik, Kaiku Health, TheraPanacea, PTW Freiburg, ITV and Brainlab in the context of research cooperations. The Department of Diagnostic and Interventional Radiology, University of Tübingen reports institutional collaborations, consultation and grants supported Siemens Healthineers, Erlangen, Germany and Bayer Healthcare, Leverkusen, Germany outside this study.

3 Discussion

Within the scope of this thesis, different aspects in the technical validation of diffusion-weighted magnetic resonance imaging acquired on a hybrid MR-Linac system were successfully evaluated and technical limits for mean ADC values were defined. Repeatability and reproducibility of DW-MRI were analyzed and the feasibility of longitudinal acquisition of functional MRI and thereof derived potentially radioresistant HRSs was investigated.

In the first investigation, DW MR images were acquired two times per treatment fraction in a collective of eleven head and neck cancer patients scanned approximately once per week over seven weeks of radiotherapy on the 1.5 T MR-Linac Unity. Through comparison of these two corresponding images, mean ADC differences resulting from measurement uncertainty of the imaging hardware and the protocol itself could be quantified as the repeatability on the MR-Linac. Target volumes showed relative repeatability coefficients in the range of 23 to 31% while OAR yielded relative values between 15% and 20%. Herewith, mean ADC values on the MR-Linac can be monitored and pathological changes be distinguished from measurement uncertainty.

In the second investigation, an enlarged patient collective was imaged before the start and in the second week of MRgRT on the 1.5 Tesla MR-Linac and a 3 Tesla diagnostic scanner. This approach allowed the quantification of reproducibility of mean ADC values from DW-MRI by comparison of the corresponding images from the MR-Linac and the diagnostic scanner. Results showed a significant underestimation of mean ADC values from the MR-Linac while the bias from the true underlying ADC value was shown to be minimal in a phantom experiment.

In the third project, ADC values and a HRS derived from a band of ADC values in the lower ADC range were acquired and evaluated sequentially over the course of radiotherapy for 28 HNC patients. We could show that mean ADC values consistently increase during MRgRT while tumor volumes decrease. Furthermore, a HRS as a

potential surrogate for radioresistance could be measured over the whole course of MRgRT and also showed a steady decrease in size, but with overall smaller absolute volumes.

3.1 Technical validation of diffusion-weighted magnetic resonance imaging on diagnostic and hybrid machines

Technical validation of QIBs is one of the most important steps towards their clinical implementation including evaluation of bias, repeatability and reproducibility (cf. Section 1.4). Reproducibility is often analyzed in multicenter studies and values derived from literature may be used to get an estimate for reproducibility of a QIB. On the other hand, bias and repeatability are two metrics which should be determined for every MRI scanner and sequence separately because uncertainties due to variations in scanners, imaging protocols or postprocessing strategies are per definition not included in the repeatability of a QIB. Although DW-MRI has been established on diagnostic MRI scanners and qualitatively evaluated for quite some time, studies performing technical validation of QIBs derived from DW imaging are still performed at this point because of innovations in hardware capabilities, sequence design and analysis strategies. Furthermore, with the introduction of hybrid MR-Linacs, many studies technically evaluated DW-MRI on the adapted MRI system in recent years showing the ongoing interest and demand in technical validation of DW-MRI.

Technical validation of every MRI system and corresponding DW sequence should start with phantom measurements because of controlled imaging environments and limited number of uncertainties, providing the opportunity of measuring the QIB under nearly optimal conditions. If big measurement differences are observed in these phantom measurements, further adaptations or re-development of new DW sequences are necessary before moving on with in vivo studies. Additionally, phantom studies allow bias quantification because the true underlying diffusion values are known in contrast to in vivo measurements. More than 10 years ago, Malyarenko et al. [136] evaluated mean ADC bias, repeatability and reproducibility in a multicenter study across 35 scanners from three different vendors with 1.5 T and 3 T. Using an ice-water phantom, multiple DW images with b-values up to 2000 s/mm^2

were acquired on different days in all institutes. Here, a bias smaller than 3% from the literature value was found for 95% of systems, but bias increased for off-center measurements. Intra-exam repeatability was within 1%, day-to-day repeatability within 3% and inter-system reproducibility was below 3.1% for most of the systems showing mean ADC values can be reliably measured on various diagnostic MRI systems. These results were further supported by the study of Belli et al. [14] which could also show in a multicenter phantom study that overall mean ADC bias from the nominal value was smaller than 5% and short-term repeatability was 2.6% on 26 analyzed scanners. Van Houdt et al. [82] performed a similar phantom study and determined ADC bias between -7 to 29 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$ and median relative repeatability and reproducibility coefficients of 3% and 18%, respectively.
After the clinical introduction of hybrid MR-Linacs, Kooreman et al. [106] also technically evaluated DW-MRI on the 1.5 T MR-Linac. Multiple DW images were acquired on four MR-Linac systems from four different institutions with the QIBA diffusion phantom (cf. Figure 3). Mean ADC values were accurate and median short-term repeatability and reproducibility over the whole range of ADC values, resulting in coefficients of variation of 1.3-2.2% for repeatability and 2.2% for reproducibility, respectively. Also McDonald et al. [143] performed DW-MRI using the same phantom on the 1.5 T MR-Linac with a commonly used echo-planar imaging (EPI) sequence as well as a turbo-spin-echo (TSE) and split acquisition of fast spin echo signals (SPLICE) [184] sequence. Here, a mean ADC bias between 26 and 123 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$ was found while RC and reproducibility coefficient (RDC) were below 11 and 31 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$ over all three sequences, respectively. These results are in congruence with our phantom evaluations where we obtained a mean bias below 7% for two different sequences on the 1.5 T MR-Linac and a mean bias of 2.1% on the 3 T diagnostic scanner. Furthermore, reproducibility between all pairs of the three sequences was acceptable with a mean difference over all PVP vials below 30 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$. Overall, the acquisition and subsequent ADC calculation is within reasonable limits on the diagnostic MRI scanner as well as the 1.5 T MR-Linac. However, additional in vivo studies are necessary to quantify more uncertainties introduced due to patient motion, susceptibility artifacts or considerably lower T2 times of human tissues than phantom material causing lower signal-to-noise ratio (SNR) in vivo.

To this point, multiple studies have performed technical validation of DW-MRI in head and neck cancer patients on diagnostic scanners. Paudyal et al. [166] acquired Test-Retest data from nine HNC patients on a 3 Tesla diagnostic MRI scanner with ten b-values up to 2000 s/mm^2. Mean pretreatment relRC of 6.59% and 2.38% in

week 1 of radiotherapy were reported. Meanwhile, Hoang et al. [77] imaged 16 HNC patients on a 1.5 Tesla diagnostic scanner two times before the start of radiotherapy with one week between both scans with a DW imaging protocol including four b-values up to 1000 s/mm^2. Repeatability was evaluated using the intraclass correlation coefficient (ICC) and repeatability coefficient defined as 1.96 times the standard deviation of both measurements and found ICCs of 0.99 for primary tumors and 0.86 for nodal diseases while repeatability coefficients resulted in 15% for nodal diseases. For primary tumors, no repeatability coefficient was calculated due to the high number of exclusions. Additionally, Guerreiro et al. [59] determined repeatability of various OAR in the head and neck region including submandibular and parotid glands in 26 healthy volunteers using a SPLICE sequence on a 3 Tesla diagnostic scanner. RC for mean ADC ranged from 110-250 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$ while relRC values covered 13-30.5%, respectively. Furthermore, many other studies have evaluated ADC repeatability on diagnostic MRI scanners at 1.5 T and 3 T in other tumor entities like lung, rectum, liver and prostate, reporting relRC between 10-41% [22, 37, 86, 213].

In the last few years, first patient studies were published, evaluating the technical performance of DW-MRI acquired on hybrid MR-Linacs. Kooreman et al. [104] evaluated repeatability of DW-MRI in a multicenter study on the 1.5 T MR-Linac in prostate cancer patients. Although the IVIM model was used, their diffusion coefficient D was calculated like ADC in the monoexponential model and repeatability across 43 patients from three institutions resulted in RC of 440 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$ and 90 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$ for the tumor and non-cancerous prostate, respectively. Evaluation of ADC repeatability and reproducibility in brain tumors was performed in the study by Lawrence et al. [114] where DW-MRI was acquired on the 1.5 T MR-Linac and an 1.5 T diagnostic scanner. They report a tremendously low within-session RC of 33 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$ and corresponding relRC of 3%. Meanwhile, ADC values on the MR-Linac were underestimated in contrast to the diagnostic scanner with a mean bias of -100 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$. This ADC discrepancy between the MR-Linac and a diagnostic scanner is also reported by Wong et al. [220]. There, DW-MRI was also acquired on the 1.5 T MR-Linac and the same 1.5 T diagnostic scanner in 19 patients with prostate cancer. They could show either that ADC values on the diagnostic scanner were 14.5% higher than on the MR-Linac. In head and neck cancer, only McDonald et al. [143] performed technical validation of DW-MRI with different imaging techniques on the 1.5 T MR-Linac and also on a diagnostic scanner. Ten patients were imaged twice in one fraction to calculate repeatability of primary tumors, lymph nodes and

parotid glands, while ADC values were also compared with the diagnostic scanner. For ADC values, acquired with a SE-EPI sequence on the MR-Linac and calculated with b-values 150 and 500 s/mm^2, relRC resulted in 27% for primary tumors, 22% for lymph nodes and 15% for parotid glands. Using a TSE sequence, relRC presented as 21%, 30% and 16% while the SPLICE acquisition yielded 22%, 20% and 11% for the same VOIs, respectively. Furthermore, comparison of all three sequences on the MR-Linac to a benchmark SE-EPI sequence on the diagnostic scanner showed a bias (diagnostic scanner - MR-Linac) of 340, 560 and 280 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$.
These results are in line with our study, reporting a bias of 93 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$, and therefore supporting the conclusion that mean ADC values from the MR-Linac are underestimated in contrast to standard diagnostic scanners. Additionally, our relative repeatability coefficients of 31% for gross tumor volume (GTV), 24% for conspicuous lymph nodes and 15-20% for submandibular and parotid glands are not only in line with other MR-Linac studies [104, 143], but also with many studies performed on diagnostic scanners [22, 37, 59, 86, 213]. This further stresses the conclusion that mean ADC values can be acquired and calculated on a hybrid MR-Linac within limits like on benchmark diagnostic scanners, providing the basis for future studies on the prognostic and predictive value of the ADC as a QIB.

3.2 Limitations of the current magnetic resonance imaging setup

The 1.5 T MR-Linac is the first hybrid machine of its kind to combine a linear accelerator with a MRI scanner of clinical standard field strength. However, the combination of two highly complex technical machines necessitated compromises in the technical design and equipment of the MR-Linac (cf. Section 1.7) which influence the acquisition of anatomical and functional MRI. One of the main goals in further developments of MRI scanners is an increase in the static magnetic field strength B_0 as a higher field strength elevates the ratio of spins aligned parallel to the magnetic field and therefore the available SNR. For the acquisition of DW-MRI and thereof derived ADC values, the SNR plays a pivotal role as higher b-values are always accompanied by decreasing SNR because of longer gradient durations resulting in more T2 decay and signal loss. To ensure the highest SNR possible on the MR-Linac, DW gradients are applied using the maximum gradient strength and slew rates minimizing the necessary gradient durations and echo times. However, the MR-Linac is limited to 34 mT/m and 120 mT/m/ms (cf. Table 1) while clin-

ically only 15 mT/m and 65 mT/m/ms are used [105] which is considerably lower than the capability of standard diagnostic MRI scanners. This has two major effects for DW-MRI on the MR-Linac.
First of all, when the same b-value is measured on the MR-Linac and a diagnostic scanner, different settings on the machines are applied with differing gradient duration and echo times resulting in SNR differences. Hence, sequence comparison as well as propagation from the MR-Linac to a diagnostic scanner and vice versa is non trivial. As a rule of thumb, the same scanner settings from the MR-Linac result in a twofold b-value on an 1.5 T diagnostic scanner because of the smaller gradient strength and slew rate of the MR-Linac for a given diffusion gradient spacing Δ [105]. This might be a reason why measured ADC values on the MR-Linac were consistently underestimated in contrast to diagnostic scanners both in our own study [62] and many other studies [114, 143, 220].
As a second consequence, the choice of high b-values on the MR-Linac is restricted due to lower SNR. When b-values are increased, the DW gradient duration rises in parallel and therefore echo time and signal loss. Hence, at high b-values SNR may be insufficiently low and inclusion of these b-values may cause systematic errors in ADC calculation because the true image signal at the high b-value would be overshadowed by noise, resulting in a more flat-angle of the fitted signal curve and therefore lower ADC values. Kooreman et al. [105] have evaluated the SNR dependency on the b-value and corresponding echo time in prostate cancer patients and found a low SNR for b-values above 500 s/mm^2. Hence, they recommended a maximum b-value of 500 s/mm^2 for DW acquisitions on the 1.5 T MR-Linac. However, this threshold value might vary in different tumor entities as T2 relaxation times of specific tissues may deviate and therefore other cut-off b-values might also be used.

Moreover, the availability of limited RF receiving coils on the MR-Linac causes limitations in the MRI acquisitions. To maximize the measured signal and SNR, receiving coils should be placed as near as possible to the patient surface. In standard diagnostic MRI acquisitions, flexible receiving coils with a large number of channels are placed on the patient surface. The drawback of MRI with its long acquisition times can be improved by using more receive channels which allow increased parallel imaging and higher undersampling factors, reducing the scan time [230]. For the 1.5 T MR-Linac however, the design of receiving coils is restricted by two major dosimetric aspects of the coils. First, the electronics of the receiving array should not attenuate the radiation beam passing through. Standard low impedance receive coils however consist of dense materials and hence do attenuate the radiation [79],

restricting their use for MRgRT. Secondly, receive arrays placed on the patient body may act like a bolus and increase the surface dose for the patients [56]. Therefore, the anterior element of the receiving coil on the 1.5 T MR-Linac is placed inside a spacer ring and is curvely shaped allowing for the coil to be lifted above the patient while the flat posterior element is positioned 8 mm beneath the treatment table. Both elements contain four RF channels resulting in an eight channel receiving array on the MR-Linac [80]. The elevation of the anterior coil facilitates that generated secondary electrons in the coil can not contribute to the patients surface dose because they are bent off by the magnetic field known as the electron return effect (ERE) [174]. However, this coil design limits the acquired SNR on the 1.5 T MR-Linac and possibilities of parallel imaging at the moment.

Apart from these hardware limitations, the ADC as a QIB is also influenced by various other parameters. During the development of a DW sequence, b-values and their number of averages have to be chosen wisely as especially the highest b-value determines the echo time of the whole sequence and therefore the acquired SNR. Furthermore, the b-values used for the later ADC calculation highly impact the derived ADC values. Studies have shown that mean ADC values vary between different combinations of b-values chosen for ADC calculation and that the inclusion or exclusion of the b-value 0 majorly influences the mean ADC value [143, 207]. Bisgaard et al. [17] could also show in a multicenter study that mean ADC values using a full set of b-values in contrast to only b-values of 150 s/mm^2 and higher were 18% larger. This limits comparability between many published studies on quantitative DW-MRI.

Another aspect influencing mean ADC values and therefore limiting comparability and reproducibility between studies are the differences in ADC calculation algorithms. While most vendors provide built-in solutions for the calculation of ADC maps, the underlying algorithms are often not known and many studies use their own in-house algorithms instead. Here, differences arise through varying fitting methods or lower and upper bounds of ADC values which have been shown to influence mean ADC values [17, 57]. However, alignment and harmonization of used algorithms can strongly decrease these ADC variations [17].

Another restraining factor is the limited spatial accuracy of ADC values on the 1.5 T MR-Linac. Due to the split gradient coil design of the MR-Linac, eddy currents might be elevated near the iso-center, potentially leading to magnetic field variations and therefore higher ADC uncertainties near the iso-center in contrast to standard

diagnostic MRI scanners. Here, Kooreman et al. [105] determined that accurate ADC values with deviations smaller than 5% can be measured within a radius of 7 cm around the iso-center. This is of great importance in the imaging of head and neck cancer patients because of possibly large distances between the primary tumor and conspicuous lymph nodes. Hence, when the GTV is positioned in the MRI iso-center, some distant lymph nodes may be further away than 7 cm and therefore additional uncertainty for the ADC calculation in these lymph nodes is introduced.

3.3 Investigation on different magnetic resonance imaging techniques

The quality and usefulness of quantitative imaging biomarkers derived from DW-MRI strongly depend on the SNR of the acquired images. This is even more important for hybrid MR-Linacs which inherently suffer from lower image signal than diagnostic MRI scanners because of lower magnetic field strength and limited gradient capacity. Therefore, signal maximization is one of the most important objectives when DW sequences for the MR-Linacs are designed. Furthermore, sequence duration should be minimized because longer acquisition times require enhanced echo times which not only reduce SNR, but also increase the chance of patient motion and image artifacts. Additionally, DW sequences must be applicable in the context of MRgRT where more delay between anatomical image acquisition and radiation delivery causes uncertainty in dose depositions.

Therefore, the vast majority of DW sequences inherit a fast-read method to accelerate image acquisition with EPI being most commonly used [67]. In a SE-EPI sequence, rapid frequency- and phase-encoding gradients are applied after the 90° excitation RF and a 180° refocusing pulse resulting in the acquisition of all lines of k-space within one RF pulse in case of a single-shot setting [190]. Hence, EPI provides fast imaging, high SNR and relative insensitivity to macroscopic patient motion leading to its gold standard status for DW-MRI [38] on MR-Linacs. Firstly, Kooreman et al. [106] showed that ADC quantification on the 1.5 T MR-Linac using an EPI sequence is technically feasible with accurate and repeatable results in a multicenter phantom study. Afterwards, the MR-Linac consortium recommended the use of single-shot EPI sequences for DW-MRI on the MR-Linac [105]. Going forward, EPI sequences were used in many studies including our work in this thesis performing technical validation or evaluating the prognostic value of DW-MRI on

hybrid MR-Linacs [61, 62, 85, 115, 220]. However, EPI is especially prone to inhomogeneity of the static magnetic field through insufficient shimming or chemical shifts causing geometrical distortions. Particularly, the head and neck region suffers from these susceptibility artifacts because of many tissue-air interfaces with highly varying local magnetic fields [205]. This drawback impacts radiotherapy where geometrical accuracy is critical due to steep dose gradients in the target region and therefore sparing of OAR and providing tumor dose coverage.

In order to overcome the limitations from single-shot-SE-EPI imaging, different methods including correction of geometric distortions [52, 218] or the use of other imaging techniques like multi-shot (MS)-EPI, reduced field-of-view DW imaging or non-EPI techniques like TSE [222] may be applied. Particularly TSE has the potential to resolve some of the limitations of EPI as the application of additional 180° refocusing pulses can compensate the variations of Larmor frequencies introduced through susceptibility variations, therefore increasing the spatial accuracy of TSE sequences considerably in contrast to EPI [38]. Hence, several studies evaluated the use of TSE-based DW sequences showing less geometrical distortions compared to EPI sequences [3, 51, 143, 165]. However, phantom studies determined ADC differences from TSE-based sequences in regard to true ADC values [143] as well as significant differences to EPI sequences [196] which has been confirmed in patient studies [143, 226]. Moreover, the application of diffusion gradients in TSE-based sequences interferes with the equal spacing of the refocusing pulses which can lead to image artifacts [38] and the destructive interference of stimulated and spin echos can reduce SNR [3, 143]. Different approaches to overcome these problems are available, but most approaches decrease SNR even further or prolong imaging time [38], making the application of TSE-based sequences on hybrid MR-Linacs non-trivial.

A possible compromise between the benefits and limitations of EPI and TSE may be obtained by using the SPLICE read-out method as proposed by Schick in 1997 [184]. Here, spin echos and stimulated echos are acquired separately which sustains higher SNR and comparable spatial accuracy in contrast to a standard TSE-based sequence [143] moving SPLICE into the focus of DW-MRI research in radiotherapy. Schakel et al. [183] showed that geometric distortions were explicitly smaller using a SPLICE-based sequence in contrast to a standard EPI sequence in ten HNC patients while ADC values were accurately determined in an ice-water phantom for both sequences. This was supported by van der Hulst et al. [84] where SPLICE DW images were rated less distorted than EPI DW images, however lower image

contrast due to reduced SNR was observed in SPLICE. On the MR-Linac, so far only McDonald et al. [143] have evaluated DW-MRI using SPLICE. They reported that ADC values derived from a SPLICE-based sequence were repeatable and accurate while significant biases of SPLICE compared to EPI and TSE sequences were found in ten HNC patients and ten volunteers regarding primary tumors, conspicuous lymph nodes and parotid glands. However, the downside of SPLICE sequences is their long imaging time to ensure adequate SNR which is a major limiting factor in MRgRT.

Overall, EPI-based sequences still remain the gold standard for DW-MRI on hybrid MR-Linacs due to their high SNR and fast image acquisitions. TSE sequences can significantly improve spatial accuracy, but continue to struggle with sufficient SNR while image acquisition times for SPLICE impede their application on MR-Linacs as integration in the imaging and treatment protocol might be too time-consuming. However, future improvements in MR-Linac hardware as well as TSE and SPLICE protocols could potentially provide an EPI alternative for DW imaging in the future.

3.4 Perspectives on the clinical use of diffusion-weighted magnetic resonance imaging

The importance of multimodal imaging in radiotherapy has steadily increased containing diagnoses making and treatment planning. However, images are mainly used qualitatively apart from CT for dose distribution calculations of RT treatment plans while especially functional imaging has only niche applications to this date. In parallel, the wish for more distinct options of treatment personalization or response-adaptive workflows is rising, requiring more QIBs [26], but fully validated QIBs are still rare although many studies determined great potential.

Two basic approaches using QIBs are envisioned, one using mean or median values inside a VOI or on the other hand a voxelwise parameter map to steer radiation doses locally. In each case, integration of the respective QIB into the radiotherapy workflow would allow treatment personalization in form of response-adaptive radiotherapy.
Most studies evaluated mean ADC values for different organ sites and validation of this QIB is steadily published while in parallel the prognostic value of mean ADC values is shown. Over the last years, a variety of studies have shown that

the pretreatment mean ADC value is a prognostic factor in HNC patients where optimal cut-off values in the range of 860-1200 $\cdot 10^{-6} \frac{\text{mm}^2}{\text{s}}$ [70, 94, 95, 140] could optimally stratify responders from non-responders. This knowledge would allow patient stratification before the start of radiotherapy according to their likely response to treatment and implement strategies to personalize their radiotherapy by applying additional radiation boost doses for probable non-responders or perform dose de-escalation for responders. The prerequisite for this approach is a neglegible ADC bias and known reproducibility of mean ADC values across multiple institutions and scanners. In the second objective of this work, we have shown that the bias of mean ADC values on the 1.5 T MR-Linac is within reasonable limits while the underestimation of mean ADC values in contrast to diagnostic scanners has to be accounted for, when literature cut-off values should be applied.

Another possibility is monitoring of the relative change of the mean ADC value over the course of radiotherapy as the QIB. Here, Matoba et al. [141] demonstrated for HNC patients that the optimal threshold to predict locoregional failure was a fractional change of mean ADC in the primary tumor of 0.24 after three weeks of chemoradiotherapy. Furthermore, Khattab et al. [94] showed that a 33% increase in mean ADC value from baseline to week 2-3 of radiotherapy in HNC was the optimal cut-off for predicting local control while Marzi et al. [140] determined a mean ADC increase below 15.5% from baseline to mid-treatment as the optimal cut-off.
By acquiring DW-MR images not only before the start of radiotherapy, but also at least once during treatment, the relative change may be used as a QIB to implement a response-adaptive treatment approach. Radiotherapy treatment plans would then be adapted with an additional radiation boost dose when only a small mean ADC increase is determined or treatment may be de-escalated if the ADC increase is above the used threshold. However, this approach relies on the knowledge of mean ADC repeatability for discrimination whether the ADC change is large enough to not be caused by measurement uncertainty. In our first objective we evaluated that relRC of primary HNC tumors on the 1.5 T MR-Linac was 31.3% which is supported by McDonald et al. [143] where relRC resulted in 26.7% for HNC patients treated on the MR-Linac. Additionally, an ADC increase of approximately 30% was observed in week 3-4 of radiotherapy as shown in our third objective. Hence, response-adaptive MRgRT in HNC patients on the 1.5 T MR-Linac may be possible in approximately week 3 of radiotherapy.

However, translation of these findings into prospective clinical trials requires full

technical validation of mean ADC as a QIB in head and neck cancer. According to the newest QIBA profile, accuracy and precision in form of bias, linearity, repeatability and reproducibility have to be determined in phantoms and in at least 35 patients [21]. In our first and second objective we showed that mean ADC values were repeatable within 31% in 87 datasets from eleven HNC patients acquired on a hybrid MR-Linac while no relevant bias in phantom measurements was found. Though, reproducibility against a diagnostic scanner in 15 HNC patients did show an underestimation of mean ADC values on the MR-Linac, however more patient data might be needed. In order to satisfy the required patient numbers, results of different studies can be combined to give an overall estimation of these technical limits. For repeatability, Boss et al. [21] combined results from several studies reporting overall relRC of 27% for liver, 15% for breast or 8% for brain cancer.

Another potential application of functional MRI is the use of voxelwise parameters maps for intensification or de-escalation of radiotherapy treatment plans based on the radiosensitivity of tumor subvolumes identified by the parameter map. After determining biological target volumes inside the tumor based on the QIB, dose painting by numbers (DPBN) or dose painting by contours (DPBC) may be applied and treatment plans adapted accordingly [6]. Hearn et al. [73] calculated voxelwise ADC maps in ten rectal cancer patients and determined tumor subvolumes by the 40th centile ADC value as a threshold to escalate radiation dose in a DPBC strategy. Similarly in a prospective clinical study, Fu et al. [50] determined tumoral subvolumes in patients with nasopharyngeal carcinomas by voxels with ADC values smaller than the mean tumor ADC value and dose escalation was applied in contrast to a control group. A real DPBN approach was simulated by Orlandi et al. [163] prescribing every tumor voxel with an ADC value between 500 and 1500 $\cdot 10^{-6} \frac{\mathrm{mm}^2}{\mathrm{s}}$ an individual dose, but clinical treatment planning systems only allowed an thereof deviated DPBC approach.

We also showed in the third objective of this thesis that potentially radioresistant HRS in head and neck cancer can be monitored sequentially over the course of radiotherapy on the 1.5 T MR-Linac. In a prior preclinical study, the significant stratification potential of ADC clusters, representing an intratumoral HRS, was identified using a machine learning approach, being able to stratify tumors according to their level of radioresistance [19]. Based on this, the model was retrained and validated in a cohort of HNC patients where the optimal ADC cluster was identified which also significantly correlated to patient outcome [219]. Accordingly, such ADC-based HRS

may provide the basis for future studies applying dose escalation to radioresistant intratumoral HRSs. However, to reliably perform dose painting approaches, either technical validation, especially repeatability, of voxelwise ADC maps is necessary or repeatability of calculated HRS must be analyzed while determining the dosimetric consequences of the voxelwise ADC and HRS uncertainty. While for QIBs derived from functional MRI these studies are still missing, a first study has evaluated repeatability for PET-based QIBs. Wright et al. [221] determined the dosimetric differences for a DPBN approach based on the voxelwise tumor tracer uptake for two PET acquisitions before the start of RT in eight HNC patients, showing mean correlations above 0.7 when comparing both treatment plans.
However, geometric accuracy becomes increasingly important when spatially encoded dose escalation through voxelwise ADC maps is intended. In case standard EPI sequences are used, adequate registrations of the DW images to the reference image are necessary, with the need of automatic registration solutions for an online response-adaptive approach on hybrid MR-Linacs. Otherwise, TSE- or SPLICE-based sequences might provide higher geometric accuracy and may be useful for future approaches based on voxelwise ADC maps.

3.5 Conclusion

In this work, the technical foundation for the acquisition and evaluation of ADC values as a QIB derived from DW-MRI on a 1.5 T MR-Linac was laid. We determined repeatability of mean ADC values in a cohort of HNC patients by comparing a set of Test and Retest DW images acquired with a single-shot-SE-EPI sequence. Mean ADC values were repeatable within 15-20% for normal tissues and 23-31% for tumor volumes. Furthermore, we also evaluated bias and reproducibility of mean ADC values through phantom measurements on the MR-Linac and comparison of ADC values in another HNC patient cohort, imaged on the MR-Linac and a diagnostic scanner. We determined negligible ADC bias in phantoms while a significant underestimation of mean ADC values on the MR-Linac in contrast to the diagnostic scanner was found. Additionally, we showed that mean ADC values and ADC-based intratumoral HRSs can be measured sequentially over the course of radiotherapy on the 1.5 T MR-Linac in head and neck cancer patients. We demonstrated that mean ADC values were steadily increasing over the course of radiotherapy while tumor volumes and HRS were continuously decreasing.

With the technical basis presented in this thesis, the potential and limitations for using ADC as a QIB in the radiotherapy of HNC on the 1.5 T MR-Linac were determined. After the publication of several studies presenting the ADC as a prognostic QIB in radiotherapy, these technical validations enable the use of ADC in future prospective clinical studies evaluating the predictive value of DW-MRI by giving a technical framework for patient stratification based on threshold values and response-adaptive MRgRT through treatment adaptations after a certain or missing change in mean ADC. However, studies performing technical validation of QIBs and DW-MRI on hybrid MR-Linacs are still rare and most often have small patient numbers, illustrating the need for further studies in more tumor entities, larger sample sizes and in multicenter settings using harmonized DW-MRI sequences to extend the applicability of DW-MRI on hybrid MR-Linacs.

4 Summary

In recent years, quantitative imaging biomarkers (QIBs) have received an increasing amount of attention in many medical fields in order to improve patient outcome by more personalized treatments. Especially in radiotherapy where computed tomography (CT) and magnetic resonance imaging (MRI) are routinely acquired for treatment planning, great potential of QIBs is evident, but additional patient imaging is time- and cost-intensive over the course of radiotherapy.

Therefore, the newly introduced hybrid MR-Linac system, combining a linear accelerator and an MRI scanner, may provide the possibility of not only more personalized radiotherapy by daily anatomical imaging as well as treatment plan and patient positioning adaptations due to high soft tissue contrast imaging, but also the opportunity of acquiring functional MRI and thereof derived QIBs without any additional logistics. However, the technical capabilities of the 1.5 T MR-Linac are inferior compared to a standard diagnostic MRI scanner, questioning the measurement uncertainties of diffusion-weighted (DW)-MRI and thereof derived apparent diffusion coefficient (ADC) values on the MR-Linac as well as the feasibility of longitudinal DW-MRI acquisitions within MR-guided radiotherapy (MRgRT).

Therefore, the overall aim of this thesis was the technical validation of DW-MRI on the 1.5 T MR-Linac in head and neck cancer (HNC) patients, determining the inherent measurement uncertainties and deviations to a diagnostic scanner in form of repeatability and reproducibility for ADC. Additionally, feasibility of DW-MRI acquisitions sequentially during the course of radiotherapy was investigated while also the changes of mean ADC values as well as ADC-based high-risk subvolumes (HRSs) inside the tumor, potentially describing intratumoral regions with higher radioresistance, were evaluated.

In a first step, the repeatability of mean ADC values derived from DW-MRI on a hybrid MR-Linac was determined in a Test-Retest study. By evaluating repeated DW-MRI acquisitions on the 1.5 T MR-Linac from eleven HNC patients, repeata-

bility of mean ADC values in target volumes as well as normal glandular tissues was determined. Relative repeatability coefficients (relRC) resulted in about 23-31% for primary tumors and lymph nodes, while glandular tissues presented better repeatability of around 15-20%, describing threshold values for discriminating true change from measurement uncertainty. A volume-dependent deterioration of repeatability was also identified in addition.

In a second study including 15 HNC patients treated on the 1.5 T MR-Linac, DW images were acquired before the start of radiotherapy and in week two of treatment on the MR-Linac and a 3 T diagnostic MRI scanner. Reproducibility was assessed comparing the corresponding images from both scanners. We determined within-subject coefficients of variation (wCV) of 13% and 24% for target volumes and glandular tissues. Furthermore, a significant underestimation of mean ADC values from the MR-Linac in comparison to the diagnostic scanner as the gold standard was found. ADC bias was investigated in a phantom experiment where the single-shot spin-echo echo-planar-imaging sequence showed negligible bias from the true phantom values.

Lastly, we evaluated mean ADC values in a patient cohort of 28 HNC patients treated and sequentially imaged on the 1.5 T MR-Linac while also a HRS, defined by a cluster of low ADC values, was analyzed. Here, mean ADC values increased about 50% during the course of radiotherapy, while tumor volumes and HRSs decreased by 45% and 92%, respectively. Hence, sequential ADC measurement and thereof derived HRS analysis is feasible on the 1.5 T MR-Linac which provides the basis for future prospective clinical studies with ADC as a QIB or ADC-based HRS as a radioresistance surrogate for response-adaptive MRgRT.

With this work, the technical framework for the acquisition of DW-MRI and thereof derived ADC values on the hybrid 1.5 T MR-Linac was established as well as the feasibility of ADC values and intratumoral HRS acquisitions sequentially during the course of MRgRT. Therefore, future prospective studies may evaluate the predictive value of DW-MRI for more personalized radiotherapy by treatment adaptations before the start of radiotherapy or through response-adaptive MRgRT, combining the technical framework from this work with the shown prognostic value of DW-MRI.

5 Zusammenfassung

In den letzten Jahren haben quantitative bildgestützte Biomarker (QIBs) in vielen medizinischen Bereichen zunehmend an Bedeutung gewonnen, um die Behandlungsergebnisse von Patienten durch personalisierte Behandlungskonzepte zu verbessern. Vor allem in der Strahlentherapie, wo Computertomographie (CT) und Magnetresonanztomographie (MRT) routinemäßig für die Bestrahlungsplanung eingesetzt werden, herrscht großes Potential für die Anwendung von QIBs , allerdings ist die zusätzliche Bildgebung des Patienten im Verlauf der Strahlentherapie zeit- und kostenintensiv.

Das neu eingeführte hybride MR-Linac-System, das einen Linearbeschleuniger und einen MRT-Scanner kombiniert, bietet daher nicht nur die Möglichkeit einer stärker personalisierten Strahlentherapie durch die tägliche anatomische Bildgebung und darauf basierende Anpassung des Behandlungsplans und der Patientenpositionierung aufgrund des hohen Weichteilkontrast der MRT-Bildgebung, sondern auch die Möglichkeit, funktionelle MRT Bildgebung und daraus abgeleitete QIBs ohne zusätzliche Logistik zu erfassen. Allerdings ist die technische Ausstattung des 1.5 T MR-Linac im Vergleich zu einem diagnostischen Standard MRT-Scanner limitiert, was die Messunsicherheiten der diffusionsgewichteten (DW)-MRT-Bildgebung und der daraus abgeleiteten Pseudo-Diffusionskoeffizienten (ADC)-Werte auf dem MR-Linac sowie die Durchführbarkeit longitudinaler DW-MRT Akquisitionen im Rahmen der MR-geführten Strahlentherapie (MRgRT) in Frage stellt.

Daher war das übergeordnete Ziel dieser Arbeit die technische Validierung von DW-MRT Bildgebung auf dem 1.5 T MR-Linac bei Patienten mit Tumoren im Kopf und Halsbereich (HNC), wobei die inhärenten Messunsicherheiten und Abweichungen zu einem diagnostischen Scanner in Form von Wiederholbarkeit und Reproduzierbarkeit für ADC ermittelt wurden. Zusätzlich wurde die Durchführbarkeit von DW-MRT Akquisitionen über den Verlauf der Strahlentherapie untersucht, während auch die Veränderungen der mittleren ADC-Werte sowie ADC-basierte Hochrisiko-Subvolumina (HRS) innerhalb des Tumors, die potenziell intratumorale Regionen

mit höherer Strahlenresistenz beschreiben, analysiert wurden.

In einem ersten Schritt wurde in einer Test-Retest-Studie die Wiederholbarkeit der aus der DW-MRT Bildgebung auf einem hybriden MR-Linac abgeleiteten mittleren ADC-Werte ermittelt. Durch die Auswertung wiederholter DW-MRT Aufnahmen auf dem 1.5 T MR-Linac von elf HNC Patienten wurde die Wiederholbarkeit der mittleren ADC-Werte in den Zielvolumina sowie im normalen Drüsengewebe ermittelt. Die relativen Wiederholbarkeitskoeffizienten (relRC) ergaben ca. 23-31% für Primärtumore und Lymphknoten, während Drüsengewebe eine bessere Wiederholbarkeit von ca. 15-20% aufwiesen, was Schwellenwerte für die Unterscheidung zwischen echter Veränderung und Messunsicherheit beschreibt. Darüber hinaus wurde auch eine volumenabhängige Verschlechterung der Wiederholbarkeit festgestellt.

In einer zweiten Studie mit 15 HNC-Patienten, die mit dem 1.5 T MR-Linac behandelt wurden, wurden DW-Bilder vor Beginn der Strahlentherapie und in der zweiten Woche der Behandlung mit dem MR-Linac und einem diagnostischen 3 T MRT-Scanner aufgenommen. Die Reproduzierbarkeit wurde durch den Vergleich der entsprechenden Bilder von beiden Scannern bewertet. Für die Zielvolumina und das Drüsengewebe wurden Variationskoeffizienten (wCV) von 13% und 24% innerhalb der Patienten ermittelt. Außerdem wurde eine signifikante Unterschätzung der mittleren ADC-Werte des MR-Linac im Vergleich zum diagnostischen Scanner als Goldstandard festgestellt. Der systematische Fehler des ADC wurde in einem Phantomexperiment untersucht, bei dem die Single-Shot Spin-Echo echoplanare Bildgebungssequenz eine vernachlässigbare Abweichung von den wahren Phantomwerten zeigte.

Schließlich haben wir die mittleren ADC-Werte in einer Patientenkohorte von 28 HNC Patienten ausgewertet, die mit dem 1.5 T MR-Linac behandelt wurden und sequenzielle MRT Bildgebung erhielten, wobei auch ein HRS, definiert durch ein Band niedriger ADC-Werte, analysiert wurde. Hier stiegen die mittleren ADC-Werte im Verlauf der Strahlentherapie um etwa 50% an, während Tumorvolumen und HRS um 45% bzw. 92% abnahmen. Die sequentielle ADC-Messung und die daraus abgeleitete HRS-Analyse ist also mit dem 1.5 T MR-Linac möglich, was die Grundlage für künftige prospektive klinische Studien mit ADC als QIB oder ADC-basierten HRS als Radioresistenz-Surrogat für die reaktionsadaptive MRgRT bildet.

Mit dieser Arbeit wurde der technische Rahmen für die Aufnahme von DW-MRT

Bildgebung und daraus abgeleiteten ADC-Werten auf dem hybriden 1.5 T MR-Linac geschaffen sowie die Durchführbarkeit von ADC- und intratumoralen HRS-Messungen sequentiell im Verlauf der MRgRT. Daher könnten künftige prospektive Studien den prädiktiven Wert der DW-MRT Bildgebung für eine stärker personalisierte Strahlentherapie durch Behandlungsanpassungen vor Beginn der Strahlentherapie oder durch eine reaktionsadaptive MRgRT bewerten, wobei der technische Rahmen dieser Arbeit mit dem nachgewiesenen prognostischen Wert der DW-MRT Bildgebung kombiniert wird.

6 Publications related to the dissertation

1. Repeatability of diffusion-weighted magnetic resonance imaging in head and neck cancer at the 1.5T MR-Linac

   J. Habrich, S. Boeke, M. Nachbar, K. Nikolaou, F. Schick, C. Gani, D. Zips and D. Thorwarth

   Radiotherapy and Oncology 2022 Vol. 174 Pages 141-148

2. Recommendations for improved reproducibility of ADC derivation on behalf of the Elekta MRI-linac consortium image analysis working group

   A. Bisgaard, R Keesman, A. van Lier, C. Coolens, P. van Houdt, A. Tree, A. Wetscherek, P. Romesser, N. Tyagi, M. lo Russo, **J. Habrich**, D. Vesprini, A. Lau, S. Mook, P. Chung, L. Kerkmeijer, Z. Gouw, E. Lorenzen, U. van der Heide, T. Schytte, C. Brink and F. Mahmood

   Radiotherapy and Oncology 2023 Vol. 186 Pages 109803

3. Reproducibility of diffusion-weighted magnetic resonance imaging in head and neck cancer assessed on a 1.5 T MR-Linac and comparison to parallel measurements on a 3 T diagnostic scanner

   J. Habrich, S. Boeke, V. Fritz, E. Koerner, K. Nikolaou, F. Schick, C. Gani, D. Zips and D. Thorwarth

   Radiotherapy and Oncology 2024 Vol. 191 Pages 110046

4. Clinical validation of a prognostic preclinical magnetic resonance imaging biomarker for radiotherapy outcome in head-and-neck cancer

   R. Winter, S. Boeke, S. Leibfarth, **J. Habrich**, K. Clasen, K. Nikolaou, D.

Zips, D. Thorwarth

Radiotherapy and Oncology 2025 Vol. 204 Pages 110702

5. Longitudinal assessment of diffusion-weighted imaging during magnetic resonance-guided radiotherapy in head and neck cancer

   S. Boeke[#], **J. Habrich**[#], S. Kübler, J. Boldt, F. Schick, K. Nikolaou, J. Kübler, C. Gani, M. Niyazi, D. Zips and D. Thorwarth

   [#] contributed equally

   Radiation Oncology 2025 Vol. 20(1) Pages 15

7 References

[1] Acharya S., Fischer-Valuck B. W., Kashani R., Parikh P., Yang D., Zhao T., et al. Online Magnetic Resonance Image Guided Adaptive Radiation Therapy: First Clinical Applications. Int J Radiat Oncol Biol Phys 94 (2) (2016), pp. 394–403. DOI: 10.1016/j.ijrobp.2015.10.015.

[2] Aliotta E., Paudyal R., Diplas B., Han J., Hu Y.-C., Oh J. H., et al. Multimodality imaging parameters that predict rapid tumor regression in head and neck radiotherapy. Phys Imaging Radiat Oncol 31 (2024), p. 100603. DOI: 10.1016/j.phro.2024.100603.

[3] Aliotta E., Paudyal R., Dresner A., Shukla-Dave A., Lee N., Cerviño L., et al. Reduced-distortion diffusion weighted imaging for head and neck radiotherapy. Phys Imaging Radiat Oncol 32 (2024), p. 100653. DOI: 10.1016/j.phro.2024.100653.

[4] Allal A. S., Dulguerov P., Allaoua M., Haenggeli C.-A., El Ghazi E. A., Lehmann W., et al. Standardized uptake value of 2-[18F] fluoro-2-deoxy-D-glucose in predicting outcome in head and neck carcinomas treated by radiotherapy with or without chemotherapy. J Clin Oncol 20 (5) (2002), pp. 1398–1404. DOI: 10.1200/JCO.2002.20.5.13.

[5] Almansour H., Afat S., Fritz V., Schick F., Nachbar M., Thorwarth D., et al. Prospective Image Quality and Lesion Assessment in the Setting of MR-Guided Radiation Therapy of Prostate Cancer on an MR-Linac at 1.5 T: A Comparison to a Standard 3 T MRI. Cancers (Basel) 13 (7) (2021), p. 1533. DOI: 10.3390/cancers13071533.

[6] Alonzi R. Functional radiotherapy targeting using focused dose escalation. Clin Oncol (R Coll Radiol) 27 (10) (2015), pp. 601–617. DOI: 10.1016/j.clon.2015.06.015.

[7] Ang K. K., Harris J., Wheeler R., Weber R., Rosenthal D. I., Nguyen-Tân P. F., et al. Human papillomavirus and survival of patients with oropharyngeal cancer. N Engl J Med 363 (1) (2010), pp. 24–35.

[8] Azhar S. and Chong L. R. Clinician's guide to the basic principles of MRI. Postgrad Med J 99 (1174) (2023), pp. 894–903. DOI: 10.1136/pmj-2022-141998.

[9] Bachtiar V., Kelly M. D., Wilman H. R., Jacobs J., Newbould R., Kelly C. J., et al. Repeatability and reproducibility of multiparametric magnetic resonance imaging of the liver. PLoS One 14 (4) (2019), e0214921. DOI: 10.1371/journal.pone.0214921.

[10] Bahig H., Yuan Y., Mohamed A. S., Brock K. K., Ng S. P., Wang J., et al. Magnetic resonance-based response assessment and dose adaptation in human papilloma virus positive tumors of the oropharynx treated with radiotherapy (MR-ADAPTOR): an R-IDEAL stage 2a-2b/Bayesian phase II trial. Clin Transl Radiat Oncol 13 (2018), pp. 19–23. DOI: 10.1016/j.ctro.2018.08.003.

[11] Bammer R. Basic principles of diffusion-weighted imaging. Eur J Radiol 45 (3) (2003), pp. 169–184. DOI: 10.1016/s0720-048x(02)00303-0.

[12] Barth M. and Moser E. Proton NMR relaxation times of human blood samples at 1.5 T and implications for functional MRI. Cell Mol Biol 43 (5) (1997), pp. 783–791.

[13] Bayouth J. E., Low D. A., and Zaidi H. MRI-linac systems will replace conventional IGRT systems within 15 years. Med Phys 46 (9) (2019), pp. 3753–3756. DOI: 10.1002/mp.13657.

[14] Belli G., Busoni S., Ciccarone A., Coniglio A., Esposito M., Giannelli M., et al. Quality assurance multicenter comparison of different MR scanners for quantitative diffusion-weighted imaging. J Magn Reson Imaging 43 (1) (2016), pp. 213–219. DOI: 10.1002/jmri.24956.

[15] Berwouts D., Madani I., Duprez F., Olteanu A. L., Vercauteren T., Boterberg T., et al. Long-term outcome of 18F-Fluorodeoxyglucose-positron emission tomography-guided dose painting for head and neck cancer: matched case-control study. Head Neck 39 (11) (2017), pp. 2264–2275. DOI: 10.1002/hed.24892.

[16] Bhat G. R., Hyole R. G., and Li J. "Head and neck cancer: Current challenges and future perspectives". In: *Adv Cancer Res.* Vol. 152. Elsevier, 2021, pp. 67–102. DOI: 10.1016/bs.acr.2021.05.002.

[17] Bisgaard A. L., Keesman R., Lier A. L. van, Coolens C., Houdt P. J. van, Tree A., et al. Recommendations for improved reproducibility of ADC derivation on behalf of the Elekta MRI-linac consortium image analysis working group. Radiother Oncol 186 (2023), p. 109803. DOI: 10.1016/j.radonc.2023.109803.

[18] Bloch P., Lenkinski R. E., Loren Buhle E., Hendrix R., Bryer M., and Gillies McKenna W. The use of T2 distribution to study tumor extent and heterogeneity in head and neck cancer. Magn Reson Imaging 9 (2) (1991), pp. 205–211. DOI: 10.1016/0730-725X(91)90012-B.

[19] Boeke S., Winter R. M., Leibfarth S., Krueger M. A., Bowden G., Cotton J., et al. Machine learning identifies multi-parametric functional PET/MR imaging cluster to predict radiation resistance in preclinical head and neck cancer models. Eur J Nucl Med Mol Imaging 50 (10) (2023), pp. 3084–3096. DOI: 10.1007/s00259-023-06254-9.

[20] Boeke S., Mönnich D., Van Timmeren J. E., and Balermpas P. MR-guided radiotherapy for head and neck cancer: current developments, perspectives, and challenges. Front Oncol 11 (2021), p. 616156. DOI: 10.3389/fonc.2021.616156.

[21] Boss M. A., Malyarenko D., Partridge S., Obuchowski N., Shukla-Dave A., Winfield J. M., et al. The QIBA Profile for diffusion-weighted MRI: apparent diffusion coefficient as a quantitative imaging biomarker. Radiology 313 (1) (2024), e233055. DOI: 10.1148/radiol.233055.

[22] Boss M. A., Snyder B. S., Kim E., Flamini D., Englander S., Sundaram K. M., et al. Repeatability and Reproducibility Assessment of the Apparent Diffusion Coefficient in the Prostate: A Trial of the ECOG-ACRIN Research Group (ACRIN 6701). J Magn Reson Imaging 56 (3) (2022), pp. 668–679. DOI: 10.1002/jmri.28093.

[23] Braakhuis B., Brakenhoff R., and Leemans C. R. Treatment choice for locally advanced head and neck cancers on the basis of risk factors: biological risk factors. Ann Oncol 23 (2012), pp. x173–x177. DOI: 10.1093/annonc/mds299.

[24] Bray F., Laversanne M., Sung H., Ferlay J., Siegel R. L., Soerjomataram I., et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians 74 (3) (2024), pp. 229–263. DOI: 10.3322/caac.21834.

[25] Brown R. W., Cheng Y.-C. N., Haacke E. M., Thompson M. R., and Venkatesan R. *Magnetic resonance imaging: physical principles and sequence design*. John Wiley & Sons, 2014. DOI: 10.1002/9781118633953.

[26] Cashmore M. T., McCann A. J., Wastling S. J., McGrath C., Thornton J., and Hall M. G. Clinical quantitative MRI and the need for metrology. Br J Radiol 94 (1120) (2021), p. 20201215. DOI: 10.1259/bjr.20201215.

[27] Choi S. H., Paeng J. C., Sohn C. H., Pagsisihan J. R., Kim Y. J., Kim K. G., et al. Correlation of 18F-FDG uptake with apparent diffusion coefficient ratio measured on standard and high b value diffusion MRI in head and neck cancer. J Nucl Med 52 (7) (2011), pp. 1056–1062. DOI: 10.2967/jnumed.111.089334.

[28] Chow L. Q. Head and neck cancer. N Engl J Med 382 (1) (2020), pp. 60–72. DOI: 10.1056/NEJMra1715715.

[29] Christiansen R. L., Dysager L., Hansen C. R., Jensen H. R., Schytte T., Nyborg C. J., et al. Online adaptive radiotherapy potentially reduces toxicity for high-risk prostate cancer treatment. Radiother Oncol 167 (2022), pp. 165–171. DOI: 10.1016/j.radonc.2021.12.013.

[30] Connor S., Anjari M., Burd C., Guha A., Lei M., Guerrero-Urbano T., et al. The impact of human papilloma virus status on the prediction of head and neck cancer chemoradiotherapy outcomes using the pre-treatment apparent diffusion coefficient. Br J Radiol 95 (1130) (2022), p. 20210333. DOI: 10.1259/bjr.20210333.

[31] Corona-Villalobos C. P., Pan L., Halappa V. G., Bonekamp S., Lorenz C. H., Eng J., et al. Agreement and Reproducibility of Apparent Diffusion Coefficient Measurements of Dual-b-Value and Multi-b-Value Diffusion-Weighted Magnetic Resonance Imaging at 1.5 Tesla in Phantom and in Soft Tissues of the Abdomen. J Comput Assist Tomogr 37 (1) (2013), pp. 46–51. DOI: 10.1097/rct.0b013e3182720e07.

[32] Cuccia F., Alongi F., Belka C., Boldrini L., Hörner-Rieber J., McNair H., et al. Patient positioning and immobilization procedures for hybrid MR-Linac systems. Radiat Oncol 16 (1) (2021), p. 183. DOI: 10.1186/s13014-021-01910-6.

[33] Currie S., Hoggard N., Craven I. J., Hadjivassiliou M., and Wilkinson I. D. Understanding MRI: basic MR physics for physicians. Postgrad Med J 89 (1050) (2013), pp. 209–223. DOI: 10.1136/postgradmedj-2012-131342.

[34] Das I. J., McGee K. P., Tyagi N., and Wang H. Role and future of MRI in radiation oncology. Br J Radiol 92 (1094) (2019), p. 20180505. DOI: 10.1259/bjr.20180505.

[35] Datta A., Aznar M. C., Dubec M., Parker G. J. M., and O'Connor J. P. B. Delivering Functional Imaging on the MRI-Linac: Current Challenges and Potential Solutions. Clin Oncol (R Coll Radiol) 30 (11) (2018), pp. 702–710. DOI: 10.1016/j.clon.2018.08.005.

[36] De Visser K. E. and Joyce J. A. The evolving tumor microenvironment: From cancer initiation to metastatic outgrowth. Cancer Cell 41 (3) (2023), pp. 374–403. DOI: 10.1016/j.ccell.2023.02.016.

[37] Deckers F., De Foer B., Van Mieghem F., Botelberge T., Weytjens R., Padhani A., et al. Apparent diffusion coefficient measurements as very early predictive markers of response to chemotherapy in hepatic metastasis: a preliminary investigation of reproducibility and diagnostic value. J Magn Reson Imaging 40 (2) (2014), pp. 448–456. DOI: 10.1002/jmri.24359.

[38] Dietrich O., Biffar A., Baur-Melnyk A., and Reiser M. F. Technical aspects of MR diffusion imaging of the body. Eur J Radiol 76 (3) (2010), pp. 314–322. DOI: 10.1016/j.ejrad.2010.02.018.

[39] Dirix P., De Keyzer F., Vandecaveye V., Stroobants S., Hermans R., and Nuyts S. Diffusion-Weighted Magnetic Resonance Imaging to Evaluate Major Salivary Gland Function Before and After Radiotherapy. Int J Radiat Oncol Biol Phys 71 (5) (2008), pp. 1365–1371. DOI: 10.1016/j.ijrobp.2007.12.011.

[40] Doğan N. Bland-Altman analysis: A paradigm to understand correlation and agreement. Turk J Emerg Med 18 (4) (2018), pp. 139–141. DOI: 10.1016/j.tjem.2018.09.001.

[41] Donati O. F., Chong D., Nanz D., Boss A., Froehlich J. M., Andres E., et al. Diffusion-weighted MR imaging of upper abdominal organs: field strength and intervendor variability of apparent diffusion coefficients. Radiology 270 (2) (2014), pp. 454–463. DOI: 10.1148/radiol.13130819.

[42] Dondi F., Antonelli A., Suardi N., Treglia G., and Bertagna F. The Role of PSMA PET Imaging in the Classification of the Risk of Prostate Cancer Patients: A Systematic Review on the Insights to Guide an Active Surveillance Approach. Cancers 16 (6) (2024), p. 1122. DOI: 10.3390/cancers16061122.

[43] Dubec M. J., Buckley D. L., Berks M., Clough A., Gaffney J., Datta A., et al. First-in-human technique translation of oxygen-enhanced MRI to an MR Linac system in patients with head and neck cancer. Radiother Oncol 183 (2023), p. 109592. DOI: 10.1016/j.radonc.2023.109592.

[44] Ermongkonchai T., Khor R., Wada M., Lau E., Xing D. T., and Ng S. P. A review of diffusion-weighted magnetic resonance imaging in head and neck cancer patients for treatment evaluation and prediction of radiation-induced xerostomia. Radiat Oncol 18 (1) (2023), p. 20. DOI: 10.1186/s13014-022-02178-0.

[45] Fakhry C., Westra W. H., Li S., Cmelak A., Ridge J. A., Pinto H., et al. Improved survival of patients with human papillomavirus–positive head and neck squamous cell carcinoma in a prospective clinical trial. J Natl Cancer Inst 100 (4) (2008), pp. 261–269. DOI: 10.1093/jnci/djn011.

[46] Fan W. J., Teng F., Liu G., Zhao D. W., Li J. F., Luo Y. R., et al. Diffusion weighted imaging in submandibular gland sparing helical tomotherapy for nasopharyngeal carcinoma. Radiother Oncol 157 (2021), pp. 247–254. DOI: 10.1016/j.radonc.2021.02.004.

[47] Fedorov A., Beichel R., Kalpathy-Cramer J., Finet J., Fillion-Robin J. C., Pujol S., et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn Reson Imaging 30 (9) (2012), pp. 1323–1341. DOI: 10.1016/j.mri.2012.05.001.

[48] Finazzi T., Sörnsen de Koste J. R. van, Palacios M. A., Spoelstra F. O. B., Slotman B. J., Haasbeek C. J. A., et al. Delivery of magnetic resonance-guided single-fraction stereotactic lung radiotherapy. Phys Imaging Radiat Oncol 14 (2020), pp. 17–23. DOI: 10.1016/j.phro.2020.05.002.

[49] Fruehwald-Pallamar J., Czerny C., Mayerhoefer M. E., Halpern B. S., Eder-Czembirek C., Brunner M., et al. Functional imaging in head and neck squamous cell carcinoma: correlation of PET/CT and diffusion-weighted imaging at 3 Tesla. Eur J Nucl Med Mol Imaging 38 (6) (2011), pp. 1009–1019. DOI: 10.1007/s00259-010-1718-4.

[50] Fu S., Li Y., Han Y., Wang H., Chen Y., Yan O., et al. Diffusion-weighted magnetic resonance imaging-guided dose painting in patients with locoregionally advanced nasopharyngeal carcinoma treated with induction chemotherapy plus concurrent chemoradiotherapy: a randomized, controlled clinical

trial. Int J Radiat Oncol Biol Phys 113 (1) (2022), pp. 101–113. DOI: 10.1016/j.ijrobp.2021.12.175.

[51] Gao Y., Han F., Zhou Z., Cao M., Kaprealian T., Kamrava M., et al. Distortion-free diffusion MRI using an MRI-guided Tri-Cobalt 60 radiotherapy system: sequence verification and preliminary clinical experience. Med Phys 44 (10) (2017), pp. 5357–5366. DOI: 10.1002/mp.12465.

[52] Gao Y., Yoon S., Savjani R., Pham J., Kalbasi A., Raldow A., et al. Comparison and evaluation of distortion correction techniques on an MR-guided radiotherapy system. Med Phys 48 (2) (2021), pp. 691–702. DOI: 10.1002/mp.14634.

[53] Gatta G., Botta L., Sánchez M. J., Anderson L. A., Pierannunzio D., Licitra L., et al. Prognoses and improvement for head and neck cancers diagnosed in Europe in early 2000s: The EUROCARE-5 population-based study. Eur J Cancer 51 (15) (2015), pp. 2130–2143. DOI: j.ejca.2015.07.043.

[54] Gauvain K. M., McKinstry R. C., Mukherjee P., Perry A., Neil J. J., Kaufman B. A., et al. Evaluating pediatric brain tumor cellularity with diffusion-tensor imaging. AJR Am J Roentgenol 177 (2) (2001), pp. 449–454. DOI: 10.2214/ajr.177.2.177044.

[55] George M., Dzik-Jurasz A., Padhani A., Brown G., Tait D., Eccles S., et al. Non-invasive methods of assessing angiogenesis and their value in predicting response to treatment in colorectal cancer. Br J Surg 88 (12) (2001), pp. 1628–1636. DOI: 10.1046/j.0007-1323.2001.01947.x.

[56] Ghila A., Fallone B., and Rathee S. Influence of standard RF coil materials on surface and buildup dose from a 6 MV photon beam in magnetic field. Med Phys 43 (11) (2016), pp. 5808–5816. DOI: 10.1118/1.4963803.

[57] Ghosh A., Singh T., Singla V., Bagga R., and Khandelwal N. Comparison of absolute apparent diffusion coefficient (ADC) values in ADC maps generated across different postprocessing software: reproducibility in endometrial carcinoma. AJR Am J Roentgenol 209 (6) (2017), pp. 1312–1320. DOI: 10.2214/AJR.17.18002.

[58] Group B. D. W., Atkinson Jr A. J., Colburn W. A., DeGruttola V. G., DeMets D. L., Downing G. J., et al. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther 69 (3) (2001), pp. 89–95. DOI: 10.1067/mcp.2001.113989.

[59] Guerreiro F., Houdt P. van, Navest R., Hoekstra N., Jong M. de, Heijnen B., et al. Validation of quantitative magnetic resonance imaging techniques in head and neck healthy structures involved in the salivary and swallowing function: Accuracy and repeatability. Phys Imaging Radiat Oncol 31 (2024), p. 100608. DOI: 10.1016/j.phro.2024.100608.

[60] El-Habashy D. M., Wahid K. A., He R., McDonald B., Rigert J., Mulder S. J., et al. Longitudinal diffusion and volumetric kinetics of head and neck cancer magnetic resonance on a 1.5 T MR-Linear accelerator hybrid system: A prospective R-IDEAL Stage 2a imaging biomarker characterization/pre-qualification study. Clin Transl Radiat Oncol 42 (2023), p. 100666. DOI: 10.1016/j.ctro.2023.100666.

[61] Habrich J., Boeke S., Nachbar M., Nikolaou K., Schick F., Gani C., et al. Repeatability of diffusion-weighted magnetic resonance imaging in head and neck cancer at a 1.5 T MR-Linac. Radiother Oncol 174 (2022), pp. 141–148. DOI: 10.1016/j.radonc.2022.07.020.

[62] Habrich J., Boeke S., Fritz V., Koerner E., Nikolaou K., Schick F., et al. Reproducibility of diffusion-weighted magnetic resonance imaging in head and neck cancer assessed on a 1.5 T MR-Linac and comparison to parallel measurements on a 3 T diagnostic scanner. Radiother Oncol 191 (2024), p. 110046. DOI: 10.1016/j.radonc.2023.110046.

[63] Hall W. A., Paulson E. S., Heide U. A. van der, Fuller C. D., Raaymakers B. W., Lagendijk J. J. W., et al. The transformation of radiation oncology using real-time magnetic resonance guidance: A review. Eur J Cancer 122 (2019), pp. 42–52. DOI: 10.1016/j.ejca.2019.07.021.

[64] Halle C., Andersen E., Lando M., Aarnes E. K., Hasvold G., Holden M., et al. Hypoxia-induced gene expression in chemoradioresistant cervical cancer revealed by dynamic contrast-enhanced MRI. Cancer Res 72 (20) (2012), pp. 5285–5295. DOI: 10.1158/0008-5472.Can-12-1085.

[65] Hamming-Vrieze O., Kranen S. R. van, Heemsbergen W. D., Lange C. A., Brekel M. W. van den, Verheij M., et al. Analysis of GTV reduction during radiotherapy for oropharyngeal cancer: Implications for adaptive radiotherapy. Radiother Oncol 122 (2) (2017), pp. 224–228. DOI: 10.1016/j.radonc.2016.10.012.

[66] Harry V. N., Semple S. I., Parkin D. E., and Gilbert F. J. Use of new imaging techniques to predict tumour response to therapy. Lancet Oncol 11 (1) (2010), pp. 92–102. DOI: 10.1016/S1470-2045(09)70190-1.

[67] Hasler S. W., Bernchou U., Behrens C. P., Vogelius I. R., Bisgaard A. L., Jokivuolle M., et al. Impact of geometric correction on echo-planar imaging-based apparent diffusion coefficient maps for abdominal radiotherapy. Biomed Phys Eng Express 10 (6) (2024), p. 065010. DOI: 10.1088/2057-1976/ad7597.

[68] Hatakenaka M., Nakamura K., Yabuuchi H., Shioyama Y., Matsuo Y., Kamitani T., et al. Apparent diffusion coefficient is a prognostic factor of head and neck squamous cell carcinoma treated with radiotherapy. Jpn J Radiol 32 (2014), pp. 80–89. DOI: 10.1007/s11604-013-0272-y.

[69] Hatakenaka M., Nakamura K., Yabuuchi H., Shioyama Y., Matsuo Y., Ohnishi K., et al. Pretreatment apparent diffusion coefficient of the primary lesion correlates with local failure in head-and-neck cancer treated with chemoradiotherapy or radiotherapy. Int J Radiat Oncol Biol Phys 81 (2) (2011), pp. 339–345. DOI: 10.1016/j.ijrobp.2010.05.051.

[70] Hatakenaka M., Shioyama Y., Nakamura K., Yabuuchi H., Matsuo Y., Sunami S., et al. Apparent diffusion coefficient calculated with relatively high b-values correlates with local failure of head and neck squamous cell carcinoma treated with radiotherapy. AJNR Am J Neuroradiol 32 (10) (2011), pp. 1904–1910. DOI: 10.3174/ajnr.A2610.

[71] Hauser T., Essig M., Jensen A., Laun F. B., Münter M., Maier-Hein K. H., et al. Prediction of treatment response in head and neck carcinomas using IVIM-DWI: evaluation of lymph node metastasis. Eur J Radiol 83 (5) (2014), pp. 783–787. DOI: 10.1016/j.ejrad.2014.02.013.

[72] Hawighorst H., Weikel W., Knapstein P. G., Knopp M. V., Zuna I., Schönberg S. O., et al. Angiogenic activity of cervical carcinoma: assessment by functional magnetic resonance imaging-based parameters and a histomorphological approach in correlation with disease outcome. Clin Cancer Res 4 (10) (1998), pp. 2305–2312.

[73] Hearn N., Leppien A., O'Connor P., Cahill K., Atwell D., Vignarajah D., et al. Radiotherapy dose escalation using pre-treatment diffusion-weighted imaging in locally advanced rectal cancer: a planning study. BJR Open 6 (1) (2024), tzad001. DOI: 10.1093/bjro/tzad001.

[74] Heide U. A. van der, Houweling A. C., Groenendaal G., Beets-Tan R. G., and Lambin P. Functional MRI for radiotherapy dose painting. Magn Reson Imaging 30 (9) (2012), pp. 1216–1223. DOI: 10.1016/j.mri.2012.04.010.

[75] Hein P., Kremser C., Judmaier W., Griebel J., Rudisch A., Pfeiffer K., et al. Diffusion-weighted MRI–a new parameter for advanced rectal carcinoma? Rofo 175 (3) (2003), pp. 381–386. DOI: 10.1055/s-2003-37836.

[76] Hirata K., Nakaura T., Okuaki T., Kidoh M., Oda S., Utsunomiya D., et al. Comparison of the image quality of turbo spin echo- and echo-planar diffusion-weighted images of the oral cavity. Medicine (Baltimore) 97 (19) (2018), e0447. DOI: 10.1097/MD.0000000000010447.

[77] Hoang J. K., Choudhury K. R., Chang J., Craciunescu O. I., Yoo D. S., and Brizel D. M. Diffusion-weighted imaging for head and neck squamous cell carcinoma: quantifying repeatability to understand early treatment-induced change. AJR Am J Roentgenol 203 (5) (2014), pp. 1104–1108. DOI: 10.2214/ajr.14.12838.

[78] Hong D., Lunagomez S., Kim E. E., Lee J. H., Bresalier R. S., Swisher S. G., et al. Value of baseline positron emission tomography for predicting overall survival in patient with nonmetastatic esophageal or gastroesophageal junction carcinoma. Cancer 104 (8) (2005), pp. 1620–1626. DOI: 10.1002/cncr.21356.

[79] Hoogcarspel S. J., Crijns S. P., Lagendijk J. J., Van Vulpen M., and Raaymakers B. W. The feasibility of using a conventional flexible RF coil for an online MR-guided radiotherapy treatment. Phys Med Biol 58 (6) (2013), p. 1925. DOI: 10.1088/0031-9155/58/6/1925.

[80] Hoogcarspel S. J., Zijlema S. E., Tijssen R. H., Kerkmeijer L. G., Jürgenliemk-Schulz I. M., Lagendijk J. J., et al. Characterization of the first RF coil dedicated to 1.5 T MR guided radiotherapy. Phys Med Biol 63 (2) (2018), p. 025014. DOI: 10.1088/1361-6560/aaa303.

[81] Houdt P. J. van, Saeed H., Thorwarth D., Fuller C. D., Hall W. A., McDonald B. A., et al. Integration of quantitative imaging biomarkers in clinical trials for MR-guided radiotherapy: Conceptual guidance for multicentre studies from the MR-Linac Consortium Imaging Biomarker Working Group. Eur J Cancer 153 (2021), pp. 64–71. DOI: 10.1016/j.ejca.2021.04.041.

[82] Houdt P. J. van, Kallehauge J. F., Tanderup K., Nout R., Zaletelj M., Tadic T., et al. Phantom-based quality assurance for multicenter quantitative MRI in locally advanced cervical cancer. Radiother Oncol 153 (2020), pp. 114–121. DOI: 10.1016/j.radonc.2020.09.013.

[83] Huisman T. A., Loenneker T., Barta G., Bellemann M. E., Hennig J., Fischer J. E., et al. Quantitative diffusion tensor MR imaging of the brain: field strength related variance of apparent diffusion coefficient (ADC) and fractional anisotropy (FA) scalars. Eur Radiol 16 (8) (2006), pp. 1651–1658. DOI: 10.1007/s00330-006-0175-8.

[84] Hulst H. J. van der, Braun L., Westerink B., Agrotis G., Ter Beek L. C., Tissier R., et al. Comparison of Diffusion-weighted MRI using Singe-Shot Echo-planar Imaging (SS-EPI) and Split Acquisition of Fast Spin Echo Signal (SPLICE) Imaging, a non-EPI technique, in Tumors of the Head and Neck. AJNR Am J Neuroradiol (2024). DOI: 10.3174/ajnr.A8529.

[85] Ingle M., Blackledge M., White I., Wetscherek A., Lalondrelle S., Hafeez S., et al. Quantitative analysis of diffusion weighted imaging in rectal cancer during radiotherapy using a magnetic resonance imaging integrated linear accelerator. Phys Imaging Radiat Oncol 23 (2022), pp. 32–37. DOI: 10.1016/j.phro.2022.06.003.

[86] Intven M., Reerink O., and Philippens M. E. Repeatability of diffusion-weighted imaging in rectal cancer. J Magn Reson Imaging 40 (1) (2014), pp. 146–150. DOI: 10.1002/jmri.24337.

[87] Intven M. P. W., Mol van Otterloo S. R. de, Mook S., Doornaert P. A. H., Groot-van Breugel E. N. de, Sikkes G. G., et al. Online adaptive MR-guided radiotherapy for rectal cancer; feasibility of the workflow on a 1.5T MR-linac: clinical implementation and initial experience. Radiother Oncol 154 (2021), pp. 172–178. DOI: 10.1016/j.radonc.2020.09.024.

[88] Jezzard P. and Balaban R. S. Correction for geometric distortion in echo planar images from B0 field variations. Magn Reson Med 34 (1) (1995), pp. 65–73. DOI: 10.1002/mrm.1910340111.

[89] Jin J.-Y. Prospect of radiotherapy technology development in the era of immunotherapy. J Natl Cancer Cent 2 (2) (2022), pp. 106–112. DOI: 10.1016/j.jncc.2022.04.001.

[90] Kallehauge J. F., Tanderup K., Haack S., Nielsen T., Muren L. P., Fokdal L., et al. Apparent Diffusion Coefficient (ADC) as a quantitative parameter in diffusion weighted MR imaging in gynecologic cancer: Dependence on b-values used. Acta Oncol 49 (7) (2010), pp. 1017–1022. DOI: 10.3109/0284186X.2010.500305.

[91] Keenan K. E., Carnicka S., Gottlieb S. C., and Stupic K. F. "Assessing changes in MRI measurands incurred in a scanner upgrade: is my study comprised". In: *Proceedings of the ISMRM 25th Annual Meeting, Honolulu*. 2017, pp. 22–24.

[92] Kessler L. G., Barnhart H. X., Buckler A. J., Choudhury K. R., Kondratovich M. V., Toledano A., et al. The emerging science of quantitative imaging biomarkers terminology and definitions for scientific studies and regulatory submissions. Stat Methods Med Res 24 (1) (2015), pp. 9–26. DOI: 10.1177/0962280214537333.

[93] Khalifa F., Soliman A., El-Baz A., Abou El-Ghar M., El-Diasty T., Gimel'farb G., et al. Models and methods for analyzing DCE-MRI: A review. Med Phys 41 (12) (2014), p. 124301. DOI: 10.1118/1.4898202.

[94] Khattab H. M., Montasser M. M., Eid M., Kandil A., and Desouky S. E.-D. Diffusion-weighted magnetic resonance imaging (DWMRI) of head and neck squamous cell carcinoma: could it be an imaging biomarker for prediction of response to chemoradiation therapy. Egypt J Radiol Nucl Med 51 (2020), pp. 1–14. DOI: 10.1186/s43055-020-00323-x.

[95] Kim S., Loevner L., Quon H., Sherman E., Weinstein G., Kilger A., et al. Diffusion weighted MRI for predicting and detecting early response to chemoradiation therapy of squamous cell carcinomas of the head and neck. Clin Cancer Res 15 (3) (2009), p. 986. DOI: 10.1158/1078-0432.CCR-08-1287.

[96] Kim T. H., Baek M. Y., Park J. E., Ryu Y. J., Cheon J. E., Kim I. O., et al. Comparison of DWI Methods in the Pediatric Brain: PROPELLER Turbo Spin-Echo Imaging Versus Readout-Segmented Echo-Planar Imaging Versus Single-Shot Echo-Planar Imaging. AJR Am J Roentgenol 210 (6) (2018), pp. 1352–1358. DOI: 10.2214/ajr.17.18796.

[97] King A. D., Chow K.-K., Yu K.-H., Mo F. K. F., Yeung D. K., Yuan J., et al. Head and neck squamous cell carcinoma: diagnostic performance of diffusion-weighted MR imaging for the prediction of treatment response. Radiology 266 (2) (2013), pp. 531–538. DOI: 10.1148/radiol.12120167.

[98] King A. D. and Thoeny H. C. Functional MRI for the prediction of treatment response in head and neck squamous cell carcinoma: potential and limitations. Cancer Imaging 16 (2016), pp. 1–8. DOI: 10.1186/s40644-016-0080-6.

[99] Klein S., Staring M., Murphy K., Viergever M. A., and Pluim J. P. elastix: a toolbox for intensity-based medical image registration. IEEE Trans Med Imaging 29 (1) (2010), pp. 196–205. DOI: 10.1109/tmi.2009.2035616.

[100] Klüter S. Technical design and concept of a 0.35 T MR-Linac. Clin Transl Radiat Oncol 18 (2019), pp. 98–101. DOI: 10.1016/j.ctro.2019.04.007.

[101] Koh D.-M. and Collins D. J. Diffusion-weighted MRI in the body: applications and challenges in oncology. AJR Am J Roentgenol 188 (6) (2007), pp. 1622–1635. DOI: 10.2214/AJR.06.1403.

[102] Kolff-Gart A. S., Pouwels P. J., Noij D. P., Ljumanovic R., Vandecaveye V., Keyzer F. de, et al. Diffusion-weighted imaging of the head and neck in healthy subjects: reproducibility of ADC values in different MRI systems and repeat sessions. AJNR Am J Neuroradiol 36 (2) (2015), pp. 384–390. DOI: 10.3174/ajnr.A4114.

[103] Koo T. K. and Li M. Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med 15 (2) (2016), pp. 155–163. DOI: 10.1016/j.jcm.2016.02.012.

[104] Kooreman E. S., Houdt P. J. van, Keesman R., Pelt V. W. J. van, Nowee M. E., Pos F., et al. Daily Intravoxel Incoherent Motion (IVIM) In Prostate Cancer Patients During MR-Guided Radiotherapy-A Multicenter Study. Front Oncol 11 (2021), p. 705964. DOI: 10.3389/fonc.2021.705964.

[105] Kooreman E. S., Houdt P. J. van, Keesman R., Pos F. J., Pelt V. W. J. van, Nowee M. E., et al. ADC measurements on the Unity MR-linac - A recommendation on behalf of the Elekta Unity MR-linac consortium. Radiother Oncol 153 (2020), pp. 106–113. DOI: 10.1016/j.radonc.2020.09.046.

[106] Kooreman E. S., Houdt P. J. van, Nowee M. E., Pelt V. W. J. van, Tijssen R. H. N., Paulson E. S., et al. Feasibility and accuracy of quantitative imaging on a 1.5 T MR-linear accelerator. Radiother Oncol 133 (2019), pp. 156–162. DOI: 10.1016/j.radonc.2019.01.011.

[107] Kooreman E. S., Pelt V. van, Nowee M. E., Pos F., Heide U. A. van der, and Houdt P. J. van. Longitudinal Correlations Between Intravoxel Incoherent Motion (IVIM) and Dynamic Contrast-Enhanced (DCE) MRI During Radiotherapy in Prostate Cancer Patients. Front Oncol 12 (2022), p. 897130. DOI: 10.3389/fonc.2022.897130.

[108] Kupelian P. and Sonke J.-J. "Magnetic resonance–guided adaptive radiotherapy: a solution to the future". In: *Semin Radiat Oncol*. Vol. 24. 3. Elsevier. 2014, pp. 227–232. DOI: 10.1016/j.semradonc.2014.02.013.

[109] Lambrecht M., Van Calster B., Vandecaveye V., De Keyzer F., Roebben I., Hermans R., et al. Integrating pretreatment diffusion weighted MRI into a multivariable prognostic model for head and neck squamous cell carcinoma. Radiother Oncol 110 (3) (2014), pp. 429–434. DOI: 10.1016/j.radonc.2014.01.004.

[110] Lambregts D. M., Beets G. L., Maas M., Curvo-Semedo L., Kessels A. G., Thywissen T., et al. Tumour ADC measurements in rectal cancer: effect of ROI methods on ADC values and interobserver variability. Eur Radiol 21 (12) (2011), pp. 2567–2574. DOI: 10.1007/s00330-011-2220-5.

[111] Lassen P., Lacas B., Pignon J.-P., Trotti A., Zackrisson B., Zhang Q., et al. Prognostic impact of HPV-associated p16-expression and smoking status on outcomes following radiotherapy for oropharyngeal cancer: The MARCH-HPV project. Radiother Oncol 126 (1) (2018), pp. 107–115. DOI: 10.1016/j.radonc.2017.10.018.

[112] Lauterbur P. C. Image formation by induced local interactions: examples employing nuclear magnetic resonance. Nature 242 (5394) (1973), pp. 190–191.

[113] Lavigne D., Ng S. P., O'sullivan B., Nguyen-Tan P. F., Filion E., Létourneau-Guillon L., et al. Magnetic resonance-guided radiation therapy for head and neck cancers. Curr Oncol 29 (11) (2022), pp. 8302–8315. DOI: 10.3390/curroncol29110655.

[114] Lawrence L. S. P., Chan R. W., Chen H., Keller B., Stewart J., Ruschin M., et al. Accuracy and precision of apparent diffusion coefficient measurements on a 1.5 T MR-Linac in central nervous system tumour patients. Radiother Oncol 164 (2021), pp. 155–162. DOI: 10.1016/j.radonc.2021.09.020.

[115] Lawrence L. S., Chan R. W., Chen H., Stewart J., Ruschin M., Theriault A., et al. Diffusion-weighted imaging on an MRI-linear accelerator to identify adversely prognostic tumour regions in glioblastoma during chemoradiation. Radiother Oncol 188 (2023), p. 109873. DOI: 10.1016/j.radonc.2023.109873.

[116] Le N. N., Li W., Onishi N., Newitt D. C., Gibbs J. E., Wilmes L. J., et al. Effect of Inter-Reader Variability on Diffusion-Weighted MRI Apparent Diffusion Coefficient Measurements and Prediction of Pathologic Complete Response for Breast Cancer. Tomography 8 (3) (2022), pp. 1208–1220. DOI: 10.3390/tomography8030099.

[117] Le Bihan D., Poupon C., Amadon A., and Lethimonnier F. Artifacts and pitfalls in diffusion MRI. J Magn Reson Imaging 24 (3) (2006), pp. 478–488. DOI: 10.1002/jmri.20683.

[118] Le Bihan D. What can we see with IVIM MRI? Neuroimage 187 (2019), pp. 56–67. DOI: 10.1016/j.neuroimage.2017.12.062.

[119] Le Bihan D., Breton E., Lallemand D., Grenier P., Cabanis E., and Laval-Jeantet M. MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. Radiology 161 (2) (1986), pp. 401–407. DOI: 10.1148/radiology.161.2.3763909.

[120] Le Bihan D., Mangin J.-F., Poupon C., Clark C. A., Pappata S., Molko N., et al. Diffusion tensor imaging: concepts and applications. J Magn Reson Imaging 13 (4) (2001), pp. 534–546. DOI: 10.1002/jmri.1076.

[121] Lecler A., Savatovsky J., Balvay D., Zmuda M., Sadik J. C., Galatoire O., et al. Repeatability of apparent diffusion coefficient and intravoxel incoherent motion parameters at 3.0 Tesla in orbital lesions. Eur Radiol 27 (12) (2017), pp. 5094–5103. DOI: 10.1007/s00330-017-4933-6.

[122] Lee N. Y., Sherman E. J., Schöder H., Wray R., Boyle J. O., Singh B., et al. Hypoxia-directed treatment of human papillomavirus–related oropharyngeal carcinoma. J Clin Oncol 42 (8) (2024), pp. 940–950. DOI: 10.1200/JCO.23.01308.

[123] Leibfarth S., Winter R. M., Lyng H., Zips D., and Thorwarth D. Potentials and challenges of diffusion-weighted magnetic resonance imaging in radiotherapy. Clin Transl Radiat Oncol 13 (2018), pp. 29–37. DOI: 10.1016/j.ctro.2018.09.002.

[124] Lewis B., Guta A., Mackey S., Gach H. M., Mutic S., Green O., et al. Evaluation of diffusion-weighted MRI and geometric distortion on a 0.35T MR-LINAC at multiple gantry angles. J Appl Clin Med Phys 22 (2) (2021), pp. 118–125. DOI: 10.1002/acm2.13135.

[125] Li Q., Tie Y., Alu A., Ma X., and Shi H. Targeted therapy for head and neck cancer: signaling pathways and clinical studies. Signal Transduct Target Ther 8 (1) (2023), p. 31. DOI: 10.1038/s41392-022-01297-0.

[126] Ligtenberg H., Schakel T., Dankbaar J. W., Ruiter L. N., Peltenburg B., Willems S. M., et al. Target Volume Delineation Using Diffusion-weighted Imaging for MR-guided Radiotherapy: A Case Series of Laryngeal Cancer Validated by Pathology. Cureus 10 (4) (2018), e2465. DOI: 10.7759/cureus.2465.

[127] Lillo S., Mirandola A., Vai A., Camarda A. M., Ronchi S., Bonora M., et al. Current Status and Future Directions of Proton Therapy for Head and Neck Carcinoma. Cancers (Basel 16 (11) (2024), p. 2085. DOI: 10.3390/cancers16112085.

[128] Lim I., Tan J., Alam A., Idrees M., Brenan P. A., Coletta R. D., et al. Epigenetics in the diagnosis and prognosis of head and neck cancer: A systematic review. J Oral Pathol Med 53 (2) (2024), pp. 90–106. DOI: 10.1111/jop.13513.

[129] Linge A., Lohaus F., Löck S., Nowak A., Gudziol V., Valentini C., et al. HPV status, cancer stem cell marker expression, hypoxia gene signatures and tumour volume identify good prognosis subgroups in patients with HNSCC after primary radiochemotherapy: A multicentre retrospective study of the German Cancer Consortium Radiation Oncology Group (DKTK-ROG). Radiother Oncol 121 (3) (2016), pp. 364–373. DOI: 10.1016/j.radonc.2016.11.008.

[130] Lis E., Saha A., Peck K. K., Zatcky J., Zelefsky M. J., Yamada Y., et al. Dynamic contrast-enhanced magnetic resonance imaging of osseous spine metastasis before and 1 hour after high-dose image-guided radiation therapy. Neurosurg Focus 42 (1) (2017), E9. DOI: 10.3171/2016.9.Focus16378.

[131] Löck S., Perrin R., Seidlitz A., Bandurska-Luque A., Zschaeck S., Zöphel K., et al. Residual tumour hypoxia in head-and-neck cancer patients undergoing primary radiochemotherapy, final results of a prospective trial on repeat FMISO-PET imaging. Radiother Oncol 124 (3) (2017), pp. 533–540. DOI: 10.1016/j.radonc.2017.08.010.

[132] Loimu V., Seppälä T., Kapanen M., Tuomikoski L., Nurmi H., Mäkitie A., et al. Diffusion-weighted magnetic resonance imaging for evaluation of salivary gland function in head and neck cancer patients treated with intensity-modulated radiotherapy. Radiother Oncol 122 (2) (2017), pp. 178–184. DOI: 10.1016/j.radonc.2016.07.008.

[133] Lu Y., Hatzoglou V., Banerjee S., Stambuk H. E., Gonen M., Shankaranarayanan A., et al. Repeatability Investigation of Reduced Field-of-View Diffusion-Weighted Magnetic Resonance Imaging on Thyroid Glands. J Comput Assist Tomogr 39 (3) (2015), pp. 334–339. DOI: 10.1097/rct.0000000000000227.

[134] Mahmood F., Johannesen H. H., Geertsen P., and Hansen R. H. Repeated diffusion MRI reveals earliest time point for stratification of radiotherapy response in brain metastases. Phys Med Biol 62 (8) (2017), pp. 2990–3002. DOI: 10.1088/1361-6560/aa5249.

[135] Malyarenko D., Amouzandeh G., Pickup S., Zhou R., Manning H. C., Gammon S. T., et al. Evaluation of Apparent Diffusion Coefficient Repeatability and Reproducibility for Preclinical MRIs Using Standardized Procedures and a Diffusion-Weighted Imaging Phantom. Tomography 9 (1) (2023), pp. 375–386. DOI: 10.3390/tomography9010030.

[136] Malyarenko D., Galbán C. J., Londy F. J., Meyer C. R., Johnson T. D., Rehemtulla A., et al. Multi-system repeatability and reproducibility of apparent diffusion coefficient measurement using an ice-water phantom. J Magn Reson Imaging 37 (5) (2013), pp. 1238–1246. DOI: 10.1002/jmri.23825.

[137] Mansfield P. and Grannell P. K. NMR'diffraction'in solids? J Phys C: Solid State Phys 6 (22) (1973), p. L422. DOI: 10.1088/0022-3719/6/22/007.

[138] Martens R. M., Koopman T., Lavini C., Brug T. V., Zwezerijnen G. J. C., Marcus J. T., et al. Early Response Prediction of Multiparametric Functional MRI and (18)F-FDG-PET in Patients with Head and Neck Squamous Cell Carcinoma Treated with (Chemo)Radiation. Cancers (Basel) 14 (1) (2022), p. 216. DOI: 10.3390/cancers14010216.

[139] Martens R. M., Noij D. P., Ali M., Koopman T., Marcus J. T., Vergeer M. R., et al. Functional imaging early during (chemo) radiotherapy for response prediction in head and neck squamous cell carcinoma; a systematic review. Oral Oncol 88 (2019), pp. 75–83. DOI: 10.1016/j.oraloncology.2018.11.005.

[140] Marzi S., Piludu F., Sanguineti G., Marucci L., Farneti A., Terrenato I., et al. The prediction of the treatment response of cervical nodes using intravoxel incoherent motion diffusion-weighted imaging. Eur J Radiol 92 (2017), pp. 93–102. DOI: 10.1016/j.ejrad.2017.05.002.

[141] Matoba M., Tuji H., Shimode Y., Toyoda I., Kuginuki Y., Miwa K., et al. Fractional change in apparent diffusion coefficient as an imaging biomarker for predicting treatment response in head and neck cancer treated with chemoradiotherapy. AJNR Am J Neuroradiol 35 (2) (2014), pp. 379–385. DOI: 10.3174/ajnr.A3706.

[142] McDonald B. A., Vedam S., Yang J., Wang J., Castillo P., Lee B., et al. Initial feasibility and clinical implementation of daily MR-guided adaptive head and neck cancer radiation therapy on a 1.5 T MR-Linac system: prospective R-IDEAL 2a/2b systematic clinical evaluation of technical innovation. Int J Radiat Oncol Biol Phys 109 (5) (2021), pp. 1606–1618. DOI: 10.1016/j.ijrobp.2020.12.015.

[143] McDonald B. A., Salzillo T., Mulder S., Ahmed S., Dresner A., Preston K., et al. Prospective Evaluation of In Vivo and Phantom Repeatability and Reproducibility of Diffusion-Weighted MRI Sequences on 1.5T MRI-Linear Accelerator (MR-Linac) and MR Simulator Devices for Head and Neck Cancers. Radiother Oncol 185 (2023), p. 109717. DOI: 10.1016/j.radonc.2023.109717.

[144] Merhemic Z., Imsirovic B., Bilalovic N., Stojanov D., Boban J., and Thurnher M. M. Apparent diffusion coefficient reproducibility in brain tumors measured on 1.5 and 3 T clinical scanners: A pilot study. Eur J Radiol 108 (2018), pp. 249–253. DOI: 10.1016/j.ejrad.2018.10.010.

[145] Michaelidou A., Adjogatse D., Suh Y., Pike L., Thomas C., Woodley O., et al. 18F-FDG-PET in guided dose-painting with intensity modulated radiotherapy in oropharyngeal tumours: A phase I study (FiGaRO). Radiother Oncol 155 (2021), pp. 261–268. DOI: 10.1016/j.radonc.2020.10.039.

[146] Michoux N. F., Ceranka J. W., Vandemeulebroucke J., Peeters F., Lu P., Absil J., et al. Repeatability and reproducibility of ADC measurements: a prospective multicenter whole-body-MRI study. Eur Radiol 31 (7) (2021), pp. 4514–4527. DOI: 10.1007/s00330-020-07522-0.

[147] Mikayama R., Yabuuchi H., Sonoda S., Kobayashi K., Nagatomo K., Kimura M., et al. Comparison of intravoxel incoherent motion diffusion-weighted imaging between turbo spin-echo and echo-planar imaging of the head and

neck. Eur Radiol 28 (1) (2018), pp. 316–324. DOI: 10.1007/s00330-017-4990-x.

[148] Moffat B. A., Chenevert T. L., Lawrence T. S., Meyer C. R., Johnson T. D., Dong Q., et al. Functional diffusion map: a noninvasive MRI biomarker for early stratification of clinical brain tumor response. Proc Natl Acad Sci USA 102 (15) (2005), pp. 5524–5529. DOI: 10.1073/pnas.050153210.

[149] Mohamed A. S. R., Abusaif A., He R., Wahid K. A., Salama V., Youssef S., et al. Prospective validation of diffusion-weighted MRI as a biomarker of tumor response and oncologic outcomes in head and neck cancer: Results from an observational biomarker pre-qualification study. Radiother Oncol 183 (2023), p. 109641. DOI: 10.1016/j.radonc.2023.109641.

[150] Mohamed A. S., Bahig H., Aristophanous M., Blanchard P., Kamal M., Ding Y., et al. Prospective in silico study of the feasibility and dosimetric advantages of MRI-guided dose adaptation for human papillomavirus positive oropharyngeal cancer patients compared with standard IMRT. Clin Transl Radiat Oncol 11 (2018), pp. 11–18. DOI: 10.1016/j.ctro.2018.04.005.

[151] Nachbar M., Mönnich D., Boeke S., Gani C., Weidner N., Heinrich V., et al. Partial breast irradiation with the 1.5 T MR-Linac: First patient treatment and analysis of electron return and stream effects. Radiother Oncol 145 (2020), pp. 30–35. DOI: 10.1016/j.radonc.2019.11.025.

[152] Nakajo M., Nakajo M., Kajiya Y., Tani A., Kamiyama T., Yonekura R., et al. FDG PET/CT and diffusion-weighted imaging of head and neck squamous cell carcinoma: comparison of prognostic significance between primary tumor standardized uptake value and apparent diffusion coefficient. Clin Nucl Med 37 (5) (2012), pp. 475–480. DOI: 10.1097/RLU.0b013e318248524a.

[153] Newitt D. C., Zhang Z., Gibbs J. E., Partridge S. C., Chenevert T. L., Rosen M. A., et al. Test-retest repeatability and reproducibility of ADC measures by breast DWI: Results from the ACRIN 6698 trial. J Magn Reson Imaging 49 (6) (2019), pp. 1617–1628. DOI: 10.1002/jmri.26539.

[154] Ng S. P., Cardenas C. E., Bahig H., Elgohari B., Wang J., Johnson J. M., et al. Changes in Apparent Diffusion Coefficient (ADC) in Serial Weekly MRI during Radiotherapy in Patients with Head and Neck Cancer: Results from the PREDICT-HN Study. Curr Oncol 29 (9) (2022), pp. 6303–6313. DOI: 10.3390/curroncol29090495.

[155] Noij D. P., Pouwels P. J. W., Ljumanovic R., Knol D. L., Doornaert P., Bree R. de, et al. 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. Eur J Radiol 84 (1) (2015), pp. 108–116. DOI: 10.1016/j.ejrad.2014.10.015.

[156] O'Connor J. P., Aboagye E. O., Adams J. E., Aerts H. J., Barrington S. F., Beer A. J., et al. Imaging biomarker roadmap for cancer studies. Nat Rev Clin Oncol 14 (3) (2017), pp. 169–186. DOI: 10.1038/nrclinonc.2016.162.

[157] O'Connor J. P. B., Robinson S. P., and Waterton J. C. Imaging tumour hypoxia with oxygen-enhanced MRI and BOLD MRI. Br J Radiol 92 (1095) (2019), p. 20180642. DOI: 10.1259/bjr.20180642.

[158] O'Connor J. P., Jackson A., Asselin M.-C., Buckley D. L., Parker G. J., and Jayson G. C. Quantitative imaging biomarkers in the clinical development of targeted therapeutics: current and future perspectives. The Lancet Oncology 9 (8) (2008), pp. 766–776. DOI: 10.1016/S1470-2045(08)70196-7.

[159] Obuchowski N. A. Interpreting Change in Quantitative Imaging Biomarkers. Acad Radiol 25 (3) (2018), pp. 372–379. DOI: 10.1016/j.acra.2017.09.023.

[160] Ogura A., Tamura T., Ozaki M., Doi T., Fujimoto K., Miyati T., et al. Apparent Diffusion Coefficient Value Is Not Dependent on Magnetic Resonance Systems and Field Strength Under Fixed Imaging Parameters in Brain. J Comput Assist Tomogr 39 (5) (2015), pp. 760–765. DOI: 10.1097/rct.0000000000000266.

[161] Ohnishi K., Shioyama Y., Hatakenaka M., Nakamura K., Abe K., Yoshiura T., et al. Prediction of local failures with a combination of pretreatment tumor volume and apparent diffusion coefficient in patients treated with definitive radiotherapy for hypopharyngeal or oropharyngeal squamous cell carcinoma. J Radiat Res 52 (4) (2011), pp. 522–530. DOI: 10.1269/jrr.10178.

[162] Oldenhuis C., Oosting S., Gietema J., and De Vries E. Prognostic versus predictive value of biomarkers in oncology. Eur J Cancer 44 (7) (2008), pp. 946–953. DOI: 10.1016/j.ejca.2008.03.006.

[163] Orlandi M., Botti A., Sghedoni R., Cagni E., Ciammella P., Iotti C., et al. Feasibility of voxel-based dose painting for recurrent glioblastoma guided by ADC values of diffusion-weighted MR imaging. Phys Med 32 (12) (2016), pp. 1651–1658. DOI: 10.1016/j.ejmp.2016.11.106.

[164] Padhani A. R., Liu G., Mu-Koh D., Chenevert T. L., Thoeny H. C., Takahara T., et al. Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations. Neoplasia 11 (2) (2009), pp. 102–125. DOI: 10.1593/neo.81328.

[165] Panyarak W., Chikui T., Yamashita Y., Kamitani T., and Yoshiura K. Image Quality and ADC Assessment in Turbo Spin-Echo and Echo-Planar Diffusion-Weighted MR Imaging of Tumors of the Head and Neck. Acad Radiol 26 (10) (2019), e305–e316. DOI: 10.1016/j.acra.2018.11.016.

[166] Paudyal R., Konar A. S., Obuchowski N. A., Hatzoglou V., Chenevert T. L., Malyarenko D. I., et al. Repeatability of Quantitative Diffusion-Weighted Imaging Metrics in Phantoms, Head-and-Neck and Thyroid Cancers: Preliminary Findings. Tomography 5 (1) (2019), pp. 15–25. DOI: 10.18383/j.tom.2018.00044.

[167] Paudyal R., Oh J. H., Riaz N., Venigalla P., Li J., Hatzoglou V., et al. Intravoxel incoherent motion diffusion-weighted MRI during chemoradiation therapy to characterize and monitor treatment response in human papillomavirus head and neck squamous cell carcinoma. J Magn Reson Imaging 45 (4) (2017), pp. 1013–1023. DOI: 10.1002/jmri.25523.

[168] Peltenburg B., Driessen J. P., Vasmel J. E., Pameijer F. A., Janssen L. M., Terhaard C. H., et al. Pretreatment ADC is not a prognostic factor for local recurrences in head and neck squamous cell carcinoma when clinical T-stage is known. Eur Radiol 30 (2020), pp. 1228–1231. DOI: 10.1007/s00330-019-06426-y.

[169] Peng Y., Jiang Y., Antic T., Sethi I., Schmid-Tannwald C., Eggener S., et al. Apparent diffusion coefficient for prostate cancer imaging: impact of B values. AJR Am J Roentgenol 202 (3) (2014), W247–53. DOI: 10.2214/ajr.13.10917.

[170] Plewes D. B. and Kucharczyk W. Physics of MRI: a primer. J Magn Reson Imaging 35 (5) (2012), pp. 1038–1054. DOI: 10.1002/jmri.23642.

[171] Pöttgen C., Levegrün S., Theegarten D., Marnitz S., Grehl S., Pink R., et al. Value of 18F-fluoro-2-deoxy-D-glucose-positron emission tomography/computed tomography in non–small-cell lung cancer for prediction of pathologic response and times to relapse after neoadjuvant chemoradiotherapy. Clin Cancer Res 12 (1) (2006), pp. 97–106. DOI: 10.1158/1078-0432.CCR-05-0510.

[172] Powers M., Baines J., Crane R., Fisher C., Gibson S., Marsh L., et al. Commissioning measurements on an Elekta Unity MR-linac. Phys Eng Sci Med 45 (2) (2022), pp. 457–473. DOI: 10.1007/s13246-022-01113-7.

[173] Quaresma M., Coleman M. P., and Rachet B. 40-year trends in an index of survival for all cancers combined and survival adjusted for age and sex for each cancer in England and Wales, 1971–2011: a population-based study. Lancet 385 (9974) (2015), pp. 1206–1218. DOI: 10.1016/S0140-6736(14)61396-9.

[174] Raaijmakers A. J., Raaymakers B. W., and Lagendijk J. J. Integrating a MRI scanner with a 6 MV radiotherapy accelerator: dose increase at tissue–air interfaces in a lateral magnetic field due to returning electrons. Phys Med Biol 50 (7) (2005), p. 1363. DOI: 10.1088/0031-9155/50/7/002.

[175] Raaymakers B. W., Jürgenliemk-Schulz I. M., Bol G. H., Glitzner M., Kotte A., Asselen B. van, et al. First patients treated with a 1.5 T MRI-Linac: clinical proof of concept of a high-precision, high-field MRI guided radiotherapy treatment. Phys Med Biol 62 (23) (2017), pp. L41–l50. DOI: 10.1088/1361-6560/aa9517.

[176] Raaymakers B. W., Lagendijk J., Overweg J., Kok J., Raaijmakers A., Kerkhof E., et al. Integrating a 1.5 T MRI scanner with a 6 MV accelerator: proof of concept. Phys Med Biol 54 (12) (2009), N229. DOI: 10.1088/0031-9155/54/12/N01.

[177] Raunig D. L., McShane L. M., Pennello G., Gatsonis C., Carson P. L., Voyvodic J. T., et al. Quantitative imaging biomarkers: a review of statistical methods for technical performance assessment. Stat Methods Med Res 24 (1) (2015), pp. 27–67. DOI: 10.1177/0962280214537344.

[178] Ree A. and Redalen K. Personalized radiotherapy: concepts, biomarkers and trial design. Br J Radiol 88 (1051) (2015), p. 20150009. DOI: 10.1259/bjr.20150009.

[179] Sabbatini P., Larson S., Kremer A., Zhang Z.-F., Sun M., Yeung H., et al. Prognostic significance of extent of disease in bone in patients with androgen-independent prostate cancer. J Clin Oncol 17 (3) (1999), pp. 948–948. DOI: 10.1200/JCO.1999.17.3.948.

[180] Sakamoto J., Yoshino N., Okochi K., Imaizumi A., Tetsumura A., Kurohara K., et al. Tissue characterization of head and neck lesions using diffusion-weighted MR imaging with SPLICE. Eur J Radiol 69 (2) (2009), pp. 260–268. DOI: 10.1016/j.ejrad.2007.10.008.

[181] Sasaki M., Yamada K., Watanabe Y., Matsui M., Ida M., Fujiwara S., et al. Variability in absolute apparent diffusion coefficient values across different platforms may be substantial: a multivendor, multi-institutional comparison study. Radiology 249 (2) (2008), pp. 624–630. DOI: 10.1148/radiol.2492071681.

[182] Sattler B., Lee J. A., Lonsdale M., and Coche E. PET/CT (and CT) instrumentation, image reconstruction and data transfer for radiotherapy planning. Radiother Oncol 96 (3) (2010), pp. 288–297. DOI: 10.1016/j.radonc.2010.07.009.

[183] Schakel T., Hoogduin J. M., Terhaard C. H. J., and Philippens M. E. P. Technical Note: Diffusion-weighted MRI with minimal distortion in head-and-neck radiotherapy using a turbo spin echo acquisition method. Med Phys 44 (8) (2017), pp. 4188–4193. DOI: 10.1002/mp.12363.

[184] Schick F. SPLICE: sub-second diffusion-sensitive MR imaging using a modified fast spin-echo acquisition mode. Magn Reson Med 38 (4) (1997), pp. 638–644. DOI: 10.1002/mrm.1910380418.

[185] Shukla-Dave A., Obuchowski N. A., Chenevert T. L., Jambawalikar S., Schwartz L. H., Malyarenko D., et al. Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials. J Magn Reson Imaging 49 (7) (2019), e101–e121. DOI: 10.1002/jmri.26518.

[186] Siegel R. L., Miller K. D., Wagle N. S., and Jemal A. Cancer statistics, 2023. CA Cancer J Clin 73 (1) (2023), pp. 17–48. DOI: 10.3322/caac.21763.

[187] Sijtsema N. D., Lauwers I., Verduijn G. M., Hoogeman M. S., Poot D. H., Hernandez-Tamames J. A., et al. Relating pre-treatment non-Gaussian intravoxel incoherent motion diffusion-weighted imaging to human papillomavirus status and response in oropharyngeal carcinoma. Phys Imaging Radiat Oncol 30 (2024), p. 100574. DOI: 10.1016/j.phro.2024.100574.

[188] Smits H. J., Vink S. J., Ridder M. de, Philippens M. E., and Dankbaar J. W. Prognostic value of pretreatment radiological MRI variables and dynamic contrast-enhanced MRI on radiotherapy treatment outcome in laryngeal and hypopharyngeal tumors. Clin Transl Radiat Oncol 49 (2024), p. 100857. DOI: 10.1016/j.ctro.2024.100857.

[189] Spick C., Bickel H., Pinker K., Bernathova M., Kapetas P., Woitek R., et al. Diffusion-weighted MRI of breast lesions: a prospective clinical investigation of the quantitative imaging biomarker characteristics of reproducibility, repeatability, and diagnostic accuracy. NMR Biomed 29 (10) (2016), pp. 1445–1453. DOI: 10.1002/nbm.3596.

[190] Stehling M. K., Turner R., and Mansfield P. Echo-planar imaging: magnetic resonance imaging in a fraction of a second. Science 254 (5028) (1991), pp. 43–50. DOI: 10.1126/science.1925560.

[191] Stejskal E. O. Use of Spin Echoes in a Pulsed Magnetic-Field Gradient to Study Anisotropic, Restricted Diffusion and Flow. J Chem Phys 43 (10) (1965), pp. 3597–3603. DOI: 10.1063/1.1696526.

[192] Subesinghe M., Scarsbrook A. F., Sourbron S., Wilson D. J., McDermott G., Speight R., et al. Alterations in anatomic and functional imaging parameters with repeated FDG PET-CT and MRI during radiotherapy for head and neck cancer: a pilot study. BMC Cancer 15 (2015), pp. 1–11. DOI: 10.1186/s12885-015-1154-8.

[193] Sugahara T., Korogi Y., Kochi M., Ikushima I., Shigematu Y., Hirai T., et al. Usefulness of diffusion-weighted MRI with echo-planar technique in the evaluation of cellularity in gliomas. J Magn Reson Imaging 9 (1) (1999), pp. 53–60. DOI: 10.1002/(sici)1522-2586(199901)9:1<53::aid-jmri7>3.0.co;2-2.

[194] Sumikawa T., Yabuuchi H., Sumikawa C., Nakashima Y., and Miura G. Influence of blade width and magnetic field strength on the ADC on PROPELLER DWI in head and neck. J Neuroradiol 33 (1) (2020), pp. 39–47. DOI: 10.1177/1971400919870178.

[195] Sung H., Ferlay J., Siegel R. L., Laversanne M., Soerjomataram I., Jemal A., et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 71 (3) (2021), pp. 209–249. DOI: 10.3322/caac.21660.

[196] Takatsu Y., Nakamura M., Suzuki Y., and Miyati T. Dependence of apparent diffusion coefficient on slice position in magnetic resonance diffusion imaging. Magn Reson Imaging 99 (2023), pp. 41–47. DOI: 10.1016/j.mri.2023.01.009.

[197] Taylor F. G., Quirke P., Heald R. J., Moran B. J., Blomqvist L., Swift I. R., et al. Preoperative magnetic resonance imaging assessment of circumferential resection margin predicts disease-free survival and local recurrence: 5-year follow-up results of the MERCURY study. J Clin Oncol 32 (1) (2014), pp. 34–43. DOI: 10.1200/JCO.2012.45.3258.

[198] Thoeny H. C., De Keyzer F., and King A. D. Diffusion-weighted MR imaging in the head and neck. Radiology 263 (1) (2012), pp. 19–32. DOI: 10.1148/radiol.11101821.

[199] Thoeny H. C., Keyzer F. D., Claus F. G., Sunaert S., and Hermans R. Gustatory Stimulation Changes the Apparent Diffusion Coefficient of Salivary Glands: Initial Experience. Radiology 235 (2) (2005), pp. 629–634. DOI: 10.1148/radiol.2352040127.

[200] Thorwarth D., Ege M., Nachbar M., Mönnich D., Gani C., Zips D., et al. Quantitative magnetic resonance imaging on hybrid magnetic resonance linear accelerators: Perspective on technical and clinical validation. Phys Imaging Radiat Oncol 16 (2020), pp. 69–73. DOI: 10.1016/j.phro.2020.09.007.

[201] Thorwarth D., Welz S., Mönnich D., Pfannenberg C., Nikolaou K., Reimold M., et al. Prospective evaluation of a tumor control probability model based on dynamic 18F-FMISO PET for head and neck cancer radiotherapy. J Nucl Med 60 (12) (2019), pp. 1698–1704. DOI: 10.2967/jnumed.119.227744.

[202] Tijssen R. H., Philippens M. E., Paulson E. S., Glitzner M., Chugh B., Wetscherek A., et al. MRI commissioning of 1.5 T MR-linac systems–a multi-institutional study. Radiother Oncol 132 (2019), pp. 114–120. DOI: 10.1016/j.radonc.2018.12.011.

[203] Trada Y., Keall P., Jameson M., Moses D., Lin P., Chlap P., et al. Changes in serial multiparametric MRI and FDG-PET/CT functional imaging during radiation therapy can predict treatment response in patients with head and neck cancer. Eur Radiol 33 (12) (2023), pp. 8788–8799. DOI: 10.1007/s00330-023-09843-2.

[204] Treiber J. M., White N. S., Steed T. C., Bartsch H., Holland D., Farid N., et al. Characterization and correction of geometric distortions in 814 diffusion weighted images. PLoS One 11 (3) (2016), e0152472. DOI: 10.1371/journal.pone.0152472.

[205] Tsao J. Ultrafast imaging: principles, pitfalls, solutions, and applications. J Magn Reson Imaging 32 (2) (2010), pp. 252–266. DOI: 10.1002/jmri.22239.

[206] Vandecaveye V., Dirix P., De Keyzer F., Op de Beeck K., Vander Poorten V., Roebben I., et al. Predictive value of diffusion-weighted magnetic resonance imaging during chemoradiotherapy for head and neck squamous cell carcinoma. Eur Radiol 20 (7) (2010), pp. 1703–1714. DOI: 10.1007/s00330-010-1734-6.

[207] Varga R., Fueger B. J., Ferrara F., Kapetas P., Pötsch N., Helbich T. H., et al. Evaluation of apparent diffusion coefficient (ADC) with regards to reproducibility and diagnostic accuracy as well as possible significance of pre-and post-contrast acquisition and employment of different b values. Eur J Radiol 181 (2024), p. 111730. DOI: 10.1016/j.ejrad.2024.111730.

[208] Varoquaux A., Rager O., Lovblad K. O., Masterson K., Dulguerov P., Ratib O., et al. Functional imaging of head and neck squamous cell carcinoma with diffusion-weighted MRI and FDG PET/CT: quantitative analysis of ADC and SUV. Eur J Nucl Med Mol Imaging 40 (6) (2013), pp. 842–852. DOI: 10.1007/s00259-013-2351-9.

[209] Vera P., Thureau S., Chaumet-Riffaud P., Modzelewski R., Bohn P., Vermandel M., et al. Phase II study of a radiotherapy total dose increase in hypoxic lesions identified by 18F-misonidazole PET/CT in patients with non–small cell lung carcinoma (RTEP5 study). J Nucl Med 58 (7) (2017), pp. 1045–1053. DOI: 10.2967/jnumed.116.188367.

[210] Vokes E. E., Agrawal N., and Seiwert T. Y. HPV-associated head and neck cancer. J Natl Cancer Inst 107 (12) (2015), djv344. DOI: 10.1093/jnci/djv344.

[211] Wahid K. A., Ahmed S., He R., Dijk L. V. van, Teuwen J., McDonald B. A., et al. Evaluation of deep learning-based multiparametric MRI oropharyngeal primary tumor auto-segmentation and investigation of input channel effects: Results from a prospective imaging registry. Clin Transl Radiat Oncol 32 (2022), pp. 6–14. DOI: 10.1016/j.ctro.2021.10.003.

[212] Weiger M. and Pruessmann K. P. Short-T2 MRI: Principles and recent advances. Prog Nucl Magn Reson Spectrosc 114 (2019), pp. 237–270. DOI: 10.1016/j.pnmrs.2019.07.001.

[213] Weller A., Papoutsaki M. V., Waterton J. C., Chiti A., Stroobants S., Kuijer J., et al. Diffusion-weighted (DW) MRI in lung cancers: ADC test-retest repeatability. Eur Radiol 27 (11) (2017), pp. 4552–4562. DOI: 10.1007/s00330-017-4828-6.

[214] Welz S., Paulsen F., Pfannenberg C., Reimold M., Reischl G., Nikolaou K., et al. Dose escalation to hypoxic subvolumes in head and neck cancer: A randomized phase II study using dynamic [18F] FMISO PET/CT. Radiother Oncol 171 (2022), pp. 30–36. DOI: 10.1016/j.radonc.2022.03.021.

[215] Widmann G., Henninger B., Kremser C., and Jaschke W. MRI sequences in head & neck radiology–state of the art. Rofo 189 (05) (2017), pp. 413–422. DOI: 10.1055/s-0043-103280.

[216] Winkel D., Bol G. H., Kroon P. S., Asselen B. van, Hackett S. S., Werensteijn-Honingh A. M., et al. Adaptive radiotherapy: The Elekta Unity MR-linac concept. Clin Transl Radiat Oncol 18 (2019), pp. 54–59. DOI: 10.1016/j.ctro.2019.04.001.

[217] Winter R. M., Leibfarth S., Schmidt H., Zwirner K., Mönnich D., Welz S., et al. Assessment of image quality of a radiotherapy-specific hardware solution for PET/MRI in head and neck cancer patients. Radiother Oncol 128 (3) (2018), pp. 485–491. DOI: 10.1016/j.radonc.2018.04.018.

[218] Winter R. M., Schmidt H., Leibfarth S., Zwirner K., Welz S., Schwenzer N. F., et al. Distortion correction of diffusion-weighted magnetic resonance imaging of the head and neck in radiotherapy position. Acta Oncol 56 (11) (2017), pp. 1659–1663. DOI: 10.1080/0284186x.2017.1377347.

[219] Winter R. M., Boeke S., Leibfarth S., Habrich J., Clasen K., Nikolaou K., et al. Clinical validation of a prognostic preclinical magnetic resonance imaging biomarker for radiotherapy outcome in head-and-neck cancer. Radiother Oncol 204 (2025), p. 110702. DOI: 10.1016/j.radonc.2024.110702.

[220] Wong O. L., Yuan J., Poon D. M., Chiu S. T., Yang B., Chiu G., et al. Prostate diffusion-weighted imaging (DWI) in MR-guided radiotherapy: Reproducibility assessment on 1.5 T MR-Linac and 1.5 T MR-simulator. Magn Reson Imaging 111 (2024), pp. 47–56. DOI: 10.1016/j.mri.2024.03.020.

[221] Wright P., Arnesen M. R., Lønne P.-I., Suilamo S., Silvoniemi A., Dale E., et al. Repeatability of hypoxia dose painting by numbers based on EF5-PET in head and neck cancer. Acta Oncol 60 (11) (2021), pp. 1386–1391. DOI: 10.1080/0284186X.2021.1944663.

[222] Xu G., Liu H., Ling D., Li Y., Lu N., Li X., et al. Acquisition and reconstruction with motion suppression DWI enhance image quality in nasopharyngeal carcinoma patients: Non-echo-planar DWI comparison with single-shot echo-planar DWI. Eur J Radiol (2024), p. 111752. DOI: 10.1016/j.ejrad.2024.111752.

[223] Yang Y., Cao M., Sheng K., Gao Y., Chen A., Kamrava M., et al. Longitudinal diffusion MRI for treatment response assessment: Preliminary experience using an MRI-guided tri-cobalt 60 radiotherapy system. Med Phys 43 (3) (2016), pp. 1369–1373. DOI: 10.1118/1.4942381.

[224] Yaromina A., Thames H., Zhou X., Hering S., Eicheler W., Dörfler A., et al. Radiobiological hypoxia, histological parameters of tumour microenvironment and local tumour control after fractionated irradiation. Radiother Oncol 96 (1) (2010), pp. 116–122. DOI: 10.1016/j.radonc.2010.04.020.

[225] Yıldırım H. C., Şahin M., Ergen Ş. A., Karaçam S. Ç., and Öksüz D. Ç. Recurrence patterns of pancreatic cancer treated with adjuvant intensity modulated radiotherapy. Rep Pract Oncol Radiother 27 (3) (2022), pp. 440–448. DOI: 10.5603/RPOR.a2022.0045.

[226] Yoshizako T., Yoshida R., Asou H., Nakamura M., and Kitagaki H. Comparison between turbo spin-echo and echo planar diffusion-weighted imaging of the female pelvis with 3T MRI. Acta Radiol Open 10 (2) (2021), p. 2058460121994737. DOI: 10.1177/2058460121994737.

[227] Yung J. P., Ding Y., Hwang K.-P., Cardenas C. E., Ai H., Fuller C. D., et al. Quantitative Evaluation of apparent diffusion coefficient in a large multi-unit institution using the QIBA diffusion phantom. medRxiv (2020). DOI: 10.1101/2020.09.09.20191403.

[228] Zahra M. A., Hollingsworth K. G., Sala E., Lomas D. J., and Tan L. T. Dynamic contrast-enhanced MRI as a predictor of tumour response to radiotherapy. Lancet Oncol 8 (1) (2007), pp. 63–74. DOI: 10.1016/s1470-2045(06)71012-9.

[229] Zhang S., Zeng N., Yang J., He J., Zhu F., Liao W., et al. Advancements of radiotherapy for recurrent head and neck cancer in modern era. Radiation Oncology 18 (1) (2023), p. 166. DOI: 10.1186/s13014-023-02342-0.

[230] Zijlema S. E., Tijssen R. H., Malkov V. N., Van Dijk L., Hackett S. L., Kok J. G., et al. Design and feasibility of a flexible, on-body, high impedance coil receive array for a 1.5 T MR-linac. Phys Med Biol 64 (18) (2019), p. 185004. DOI: 10.1088/1361-6560/ab37a8.

[231] Zschaeck S., Löck S., Hofheinz F., Zips D., Mortensen L. S., Zöphel K., et al. Individual patient data meta-analysis of FMISO and FAZA hypoxia PET scans from head and neck cancer patients undergoing definitive radiochemotherapy. Radiother Oncol 149 (2020), pp. 189–196. DOI: 10.1016/j.radonc.2020.05.022.

[232] Zukauskaite R., Rumley C. N., Hansen C. R., Jameson M. G., Trada Y., Johansen J., et al. Delineation uncertainties of tumour volumes on MRI of head and neck cancer patients. Clin Transl Radiat Oncol 36 (2022), pp. 121–126. DOI: 10.1016/j.ctro.2022.08.005.

8 Statement on own contributions

Statement of own contribution of the dissertation

The work for this thesis was conducted in the section for Biomedical Physics at the Department of Radiation Oncology in the University Hospital Tübingen under supervision of Prof. Dr. rer. nat. Daniela Thorwarth.

The study was designed in collaboration with Prof. Dr. rer. nat. Daniela Thorwarth and Prof. Dr. med. Daniel Zips.

Parts of this thesis are published as peer-reviewed papers and correlating statement of contributions for each of the collective work are listed as follows.

I assure, that I have written the manuscript myself and not to have used any other sources than those given by me.

Tübingen, the 07.01.2025

Repeatability of diffusion-weighted magnetic resonance imaging in head and neck cancer at the 1.5T MR-Linac

Statement of own contribution: **Jonas Habrich**

Title: Repeatability of diffusion-weighted magnetic resonance imaging in head and neck cancer at the 1.5T MR-Linac

Authors: J. Habrich, S. Boeke, M. Nachbar, K. Nikolaou, F. Schick, C. Gani, D. Zips and D. Thorwarth

Published in: Radiotherapy and Oncology 2022 Vol. 174 Pages 141-148

The original research article was conducted in the Section for Biomedical Physics, Department of Radiation Oncology in the University Hospital Tübingen under supervision of Prof. Dr. rer. nat. Daniela Thorwarth.

The study was planned in collaboration with Prof. Dr. rer. nat. Daniela Thorwarth, Prof. Dr. rer. nat. Dr. med. Fritz Schick, Prof. Dr. med. Konstantin Nikolaou and Prof. Dr. med. Daniel Zips.

All data postprocessing including image registration was done by me. Development of analysis methodology, image analysis and statistical evaluation was conducted by myself as well as writing of the manuscript.

Dr. med Simon Boeke and Prof. Dr. med. Cihan Gani recruited and treated all patients. Dr. med. Simon Boeke delineated all patient images while Dr. Marcel Nachbar supported in patient treatment planning.

All co-authors provided critical proofreading of the manuscript.

Reproducibility of diffusion-weighted magnetic resonance imaging in head and neck cancer assessed on a 1.5 T MR-Linac and comparison to parallel measurements on a 3 T diagnostic scanner

Statement of own contribution: **Jonas Habrich**

Title: Reproducibility of diffusion-weighted magnetic resonance imaging in head and neck cancer assessed on a 1.5T MR-Linac and comparison to parallel measurements on a 3T diagnostic scanner

Authors: J. Habrich, S. Boeke, V. Fritz, E. Koerner, K. Nikolaou, F. Schick, C. Gani, D. Zips, D. Thorwarth

Published in: Radiotherapy and Oncology 2024 Vol. 191 Pages 110046

The original research article was conducted in the Section for Biomedical Physics, Department of Radiation Oncology in the University Hospital Tübingen under supervision of Prof. Dr. rer. nat. Daniela Thorwarth.

The study was planned in collaboration with Prof. Dr. rer. nat. Daniela Thorwarth, Prof. Dr. rer. nat. Dr. med. Fritz Schick, Prof. Dr. med. Konstantin Nikolaou and Prof. Dr. med. Daniel Zips.

Data postprocessing was conducted by me with support of Victor Fritz. Elisa Koerner helped with phantom measurements and analysis.

All image registration was done by me. Development of analysis methodology, image analysis and statistical evaluation was conducted by myself as well as writing of the manuscript.

Dr. med Simon Boeke and Prof. Dr. med. Cihan Gani recruited and treated all patients. Dr. med. Simon Boeke delineated all patient images.

All co-authors provided critical proofreading of the manuscript.

Longitudinal assessment of diffusion-weighted imaging during magnetic resonance-guided radiotherapy in head and neck cancer

Statement of own contribution: **Jonas Habrich**

Title: Longitudinal assessment of diffusion-weighted imaging during magnetic resonance-guided radiotherapy in head and neck cancer

Authors: S. Boeke[#], J. Habrich[#], S. Kübler, J. Boldt, F. Schick, K. Nikolaou, J. Kübler, C. Gani, M. Niyazi, D. Zips, D. Thorwarth

[#] contributed equally

Published in: Radiation Oncology 2025 Vol. 20(1) Pages 15

The original research article was conducted in the Section for Biomedical Physics, Department of Radiation Oncology in the University Hospital Tübingen under supervision of Prof. Dr. rer. nat. Daniela Thorwarth.

The study was planned in collaboration with Prof. Dr. rer. nat. Daniela Thorwarth, Prof. Dr. rer. nat. Dr. med. Fritz Schick, Prof. Dr. med. Konstantin Nikolaou and Prof. Dr. med. Daniel Zips.

All data postprocessing including image registration was done by me while Dr. med Jens Kübler helped with data acquisition. Development of analysis methodology, image analysis and complete statistical evaluation was conducted by myself.

Writing of the manuscript was done together with Dr. med. Simon Boeke.

Dr. med Simon Boeke and Prof. Dr. med. Cihan Gani recruited and treated all patients. Sarah Kübler and Jessica Boldt supported in clinical patient treatment. Dr. med. Simon Boeke delineated all patient images.

Prof. Dr. med. Maximilian Niyazi, Dr. med. Simon Boeke, Prof. Dr. rer. nat. Dr. med. Fritz Schick, Prof. Dr. med. Konstantin Nikolaou, Prof. Dr. med. Daniel Zips, Prof. Dr. rer. nat. Daniela Thorwarth provided critical discussion of study methodology.

All co-authors provided critical proofreading of the manuscript.

9 Supplementary Materials

9.1 Repeatability of diffusion-weighted magnetic resonance imaging in head and neck cancer at the 1.5T MR-Linac

Table S1: Imaging timepoints of all patients. Retest images were acquired either during daily irradiation (white background) or before daily irradiation (grey background) within the same treatment fraction as the corresponding test image.

Fx	Pat1	Pat2	Pat3	Pat4	Pat5	Pat6	Pat7	Pat8	Pat9	Pat10	Pat11
1											
2										X	
3	X	X	X			X	X	X	X		X
4				X						X	
5											
6											
7			X								
8		X		X		X	X		X	X	
9					X			X			
10	X										
11											
12			X								

13						X		X				X		X		X
14			X			X						X				
15	X															
16														X		
17				X			X									
18																X
19		X										X		X		
20			X				X				X					
21										X				X		
22				X												X
23		X				X								X		
24												X				X
25			X				X			X				X		
26											X					
27				X												
28										X						X
29						X						X		X		
30																X
31	X													X		
32			X		X			X			X					
33						X										X
34		X										X				
35						X				X						

Table S2: Mean ADC value [$\cdot 10^{-6}$mm^2/s] of test and retest and the difference between test and retest [$\cdot 10^{-6}$mm^2/s] of all patients and every VOI. Missing values are due to VOIs outside the field-of-view. The left submandibular gland of patient 4 could not be delineated because of an adjecant lymph node with a suspected infiltration of the left submandibular gland.

patient	fx	left submand.		right submand.		left parotid		right parotid		GTV		LN	
		mean	difference	mean	difference	mean	difference	mean	difference	mean	difference	mean	difference
1	3	1052	-94	1360	-51	749	13	782	79	1326	230	1386	-60
	10	1383	-132	1480	-136	941	25	877	-17	1563	-214	1504	-145
	15	1588	-270	1638	-21	1045	-64	1111	18	2043	-138	1483	-6
	19	1544	21	1497	143	1028	-65	1129	13	1970	33	1263	7
	23	1589	-73	1487	-31	—	—	1235	65	1975	-78	1287	-240
	31	1733	-240	1675	-190	—	—	1376	-177	2231	-111	1446	-180
	34	1826	109	1659	-256	1294	-140	1536	-179	2561	35	1424	8
2	3	1046	-146	1229	-181	823	8	1218	-30	1134	-429	1293	167
	8	1249	21	1204	-18	871	-32	1170	-12	1509	-227	1120	64
	14	1332	-62	1433	-47	785	-49	1176	72	1782	-161	1321	-373
	20	1506	7	1584	90	940	36	1219	-35	1759	19	1291	-69
	25	1438	161	1591	84	860	-86	1187	13	1685	518	1374	411
	32	1404	-184	1428	-315	777	-43	1017	-24	1518	-787	1544	73
3	3	996	8	1094	30	768	0	837	-18	1180	-10	929	105
	7	1091	-65	1027	-140	824	-24	864	-19	1393	-110	1221	-110
	12	1259	2	1285	51	907	-34	942	52	1685	-265	1230	140
	17	1481	-121	1522	-92	960	-36	1161	-38	1796	-493	1234	69

	22	1318	-257	1493	0	907	-3	1039	-25	1699	-558	1203	-85
	27	1473	-186	1511	-140	842	-50	905	6	1843	-101	1056	-13
	32	1518	-325	1509	-140	792	-111	818	-11	1738	-416	1178	-185
												897	-87
	4	—	—	1097	234	434	-22	516	-53	1002	207	1117	229
												1106	-65
												1099	-19
	8	—	—	969	371	653	-44	655	23	1218	-24	1163	254
												1217	166
												1307	-7
	14	—	—	1109	-35	404	-32	564	-12	1433	20	1261	-60
												1321	-19
												1235	19
4	23	—	—	1132	117	492	-14	568	-10	1519	87	1316	-225
												1248	-49
												1236	-208
	29	—	—	1303	-238	465	-31	537	-115	1433	-67	1300	-46
												1330	-24
												1333	-24
	32	—	—	1210	369	490	-72	491	17	1693	450	1348	143
												1079	-112
												1604	-96
	35	—	—	1134	-37	418	-38	483	-56	1886	-229	1350	0

											1136	-63
5	9	970	-214	1046	-149	679	-48	773	-114	965	-589	—	—
	13	1229	-15	1262	21	778	48	890	-35	1399	21	—	—
	17	1255	-248	1136	-138	704	-45	819	37	1478	16	—	—
	20	1265	-180	1476	28	712	-18	951	-63	1707	87	—	—
	25	1255	108	1291	26	688	-30	877	-47	1692	-121	—	—
	32	1418	-34	1436	-15	863	-24	957	-94	1555	-149	—	—
6	3	1130	-271	1034	-358	765	-55	765	-54	1202	-50	898	-5
												1174	-353
	8	1064	-205	929	-237	867	-44	893	-113	1270	22	929	-275
												1263	-250
	13	1447	-210	1274	8	1063	139	957	130	1525	-338	1097	89
												1721	257
	21	1523	-180	1449	-149	1030	-52	1039	-60	1454	-653	1134	-72
												1531	-11
	25	1233	-3	1083	52	1053	25	1029	-117	1648	-444	1079	-93
												1437	-374
	28	1410	-142	1231	30	1153	19	961	-97	1679	-20	1275	129
												1602	124
	35	1605	-110	1522	-384	1443	-232	1118	-124	1761	-393	1425	-410
												1705	-49
7	3	1187	-1	1275	-132	825	-38	894	51	1253	-13	1471	-166
												961	298
	8	1199	109	1299	-13	870	-41	919	90	1319	3	1381	27

											1116	6
	20	1681	-296	1603	-135	961	-130	1011	-265	1652	-202	1775	-291
												1797	-429
	26	1626	-129	1806	-5	972	-56	932	-116	1712	1	1809	-232
												1448	348
	32	1725	-103	1845	-81	927	-92	961	-8	1749	-33	1922	-54
												1457	-59
	3	1244	-37	1353	128	763	-57	912	-21	1031	-3	765	-35
												1545	-38
	9	1371	-39	1609	114	815	13	957	44	1523	-15	1187	-13
												1685	-25
	14	1410	-176	1552	76	725	-32	782	-61	1466	-348	1318	11
												1675	-78
8	19	1500	-255	1752	-124	844	-20	975	-70	1614	-31	1659	-6
												1744	-45
	24	1435	13	1903	-235	959	47	1065	-49	1665	-60	1778	-69
												1785	-168
	29	1798	-212	2104	-262	1014	-132	1090	-80	1921	23	1877	-112
												1887	-125
	34	1595	-135	1981	-1	909	-7	1008	20	1697	11	1600	-81
												1865	60
	3	1115	3	1241	29	816	-21	909	-39	1230	-2	885	31
	8	1370	90	1523	43	1043	96	1034	79	1200	-20	1010	7
	13	1394	156	1687	-95	1197	-86	1149	36	1286	214	1019	99

	19	1451	-229	1403	-47	1231	-75	1177	-65	1391	-206	1148	21
	22	1525	-206	1585	-92	1151	-72	1147	27	1551	-390	1255	-261
	27	1449	-134	1398	-85	1221	31	1231	29	1623	-365	1207	-99
	30	1278	22	1323	-37	1257	-16	1231	-43	1559	-88	1102	-10
	33	1311	24	1372	-43	1302	-23	1287	-46	1660	10	1251	-197
	35	1352	-82	1417	-251	1346	88	1311	2	1664	-43	1705	71
10	2	1369	80	1542	262	1004	-18	1013	15	955	54	957	-13
	4	1195	3	1207	33	912	-59	945	0	1020	-17	893	-161
	8	1426	29	1329	87	1118	-53	1065	37	1163	18	979	-13
	13	1441	40	1713	59	1035	29	1172	-11	1516	100	1122	136
	16	1553	52	1830	-59	1186	16	1223	14	1627	-65	1336	-41
	21	1570	-60	1772	-38	1317	-159	1297	0	1648	-91	1368	104
	23	1555	-32	1737	-166	1417	-117	1351	102	1638	-20	1204	-12
	25	1533	1	1802	23	1328	-102	1302	198	1605	49	1404	-278
	27	1686	-45	1893	-30	1285	-59	1404	50	1794	31	1386	148
	29	1860	-210	2036	63	1319	-48	1461	-124	1931	35	1386	-330
	31	1649	-31	1850	-91	1278	-8	1359	-121	1755	10	1346	-128
	35	1852	-35	2206	-108	1245	-44	1452	-4	1995	-16	1480	-28
11	3	1335	-79	1153	6	963	-59	956	-78	1396	201	1245	55
	13	1400	-10	1233	5	1212	5	1127	-81	1983	-160	1429	-148
	18	1435	-108	1253	-85	1285	-137	1241	-146	1952	-51	1509	-103
	22	1540	-87	1277	-115	1355	-92	1179	-12	1994	-7	1545	-107
	24	1400	58	1282	-7	1293	-31	1159	-6	2020	129	1556	122
	28	1467	-107	1469	-142	1206	-124	1130	-116	1902	-86	1550	-124

30	1556	-11	1408	-71	1342	-121	1329	-47	1900	109	1551	74
32	1555	15	1204	22	1288	-96	1217	-156	1922	-5	1372	-48

9.2 Reproducibility of diffusion-weighted magnetic resonance imaging in head and neck cancer assessed on a 1.5 T MR-Linac and comparison to parallel measurements on a 3 T diagnostic scanner

Signal-to-noise evaluation

Mean signal-to-noise ratio (SNR) was evaluated for all acquired sequences of the 1.5 T MR-Linac (EPI4b and EPI3b) and the 3 T diagnostic scanner (Siemens Vida) over all 15 patients. SNR was calculated as:

$$SNR = \frac{mean\ signal\ value\ GTV}{standard\ deviation\ in\ noise\ ROI}$$

For the diagnostic scanner, the noise ROI was delineated outside the patient in the image background. For the MR-Linac sequences, noise ROIs were placed in a signal-free region inside the patient due to automatic vendor-induced image filtering outside the patient. Noise ROIs were delineated inside the lung for EPI4b images while for EPI3b, due to smaller field-of-view, no signal-free regions could be identified, and noise ROIs were delineated in a homogenous muscle tissue as noise substitute.

Table S3: Mean signal-to-noise ratios for all diffusion-weighted sequences from the diagnostic scanner (Vida) and the MR-Linac (EPI4b, EPI3b) and analyzed b-values over all 15 patients.

	b150	b200	b500	b1000
Vida	—	238.1	147.3	100.9
EPI4b	—	91.8	67.4	—
EPI3b	63.9	—	60.4	—

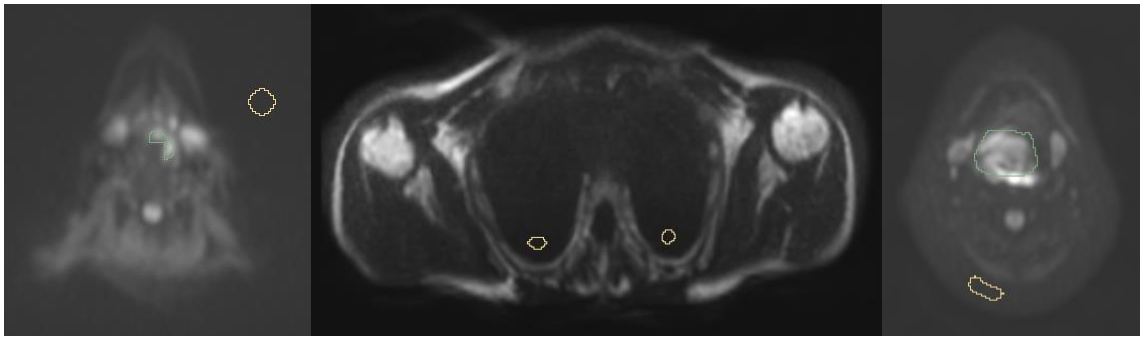

Figure S1: Exemplary images of the GTVs (green) and noise ROIs (brown) used for the SNR calculation in the image background of Vida image from the diagnostic scanner (left), inside the lung of the EPI4b (middle) and in muscle tissue of the EPI3b (right) of the MR-Linac.

Comparability EPI3b and EPI4b

To test comparability of both MR-Linac sequences (EPI3b and EPI4b), mean ADC values of ten head and neck cancer patients were analyzed. For all ten patients both sequences were acquired before the start of radiotherapy on the same day and in the same imaging session. GTVs were delineated on the EPI3b images and EPI4b images were rigidly registered to the EPI3b images. ADC maps were calculated using b-values of 150 and 500 s/mm^2 for the EPI3b and 200 and 500 s/mm^2 for the EPI4b. Mean GTV ADC values were analyzed.

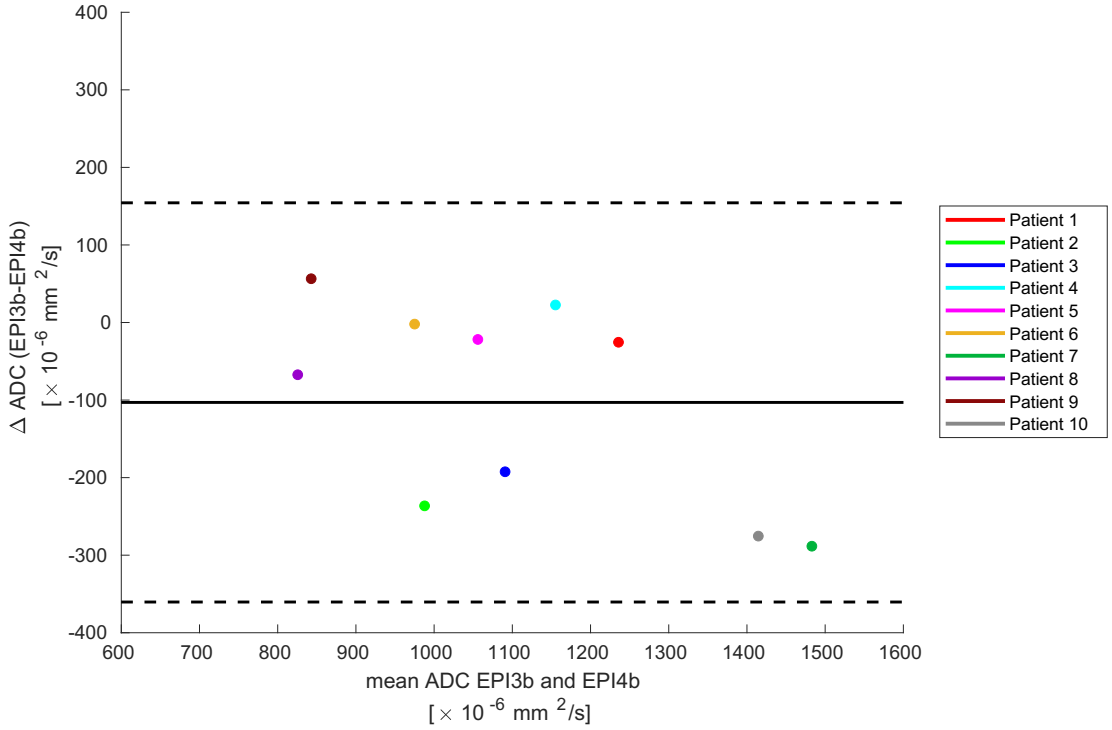

Figure S2: Bland-Altman plot for mean GTV ADC values calculated with EPI3b and EPI4b, acquired on the MR-Linac, for ten head and neck cancer patients. The mean ADC difference is visualized with a solid line while the dotted lines represent the limits of agreement defined as the mean ADC difference ± 1.96 * standard deviation.

Phantom measurements

To eliminate many possible sources for uncertainties and variations in ADC evaluation, measurements with the QIBA DWI phantom were performed. Therefore, all three sequences, evaluated in the patient analysis, were acquired at two timepoints on the MR-Linac and the Diagnostic scanner with two acquisition per sequence and timepoint. The QIBA DWI phantom was prepared according to the QIBA handbook (https://qibawiki.rsna.org/images/a/a5/QIBA_DWI_Profile_Conformance _Testing_Supplement_1_20210401a.pdf). ROIs of all 13 vials were delineated in 3D Slicer and mean ADC values for all vials were calculated. ADC maps were calculated using b-values of 150 and 500 s/mm^2 for the EPI3b of the MR-Linac, 200 and 500 s/mm^2 for the EPI4b of the MR-Linac and 200, 500 and 1000 s/mm^2 for the Diagnostic scanner sequence (Vida). ADC values were analyzed in comparison to the given reference ADC value and all sequences were compared by Bland-Altman analysis.

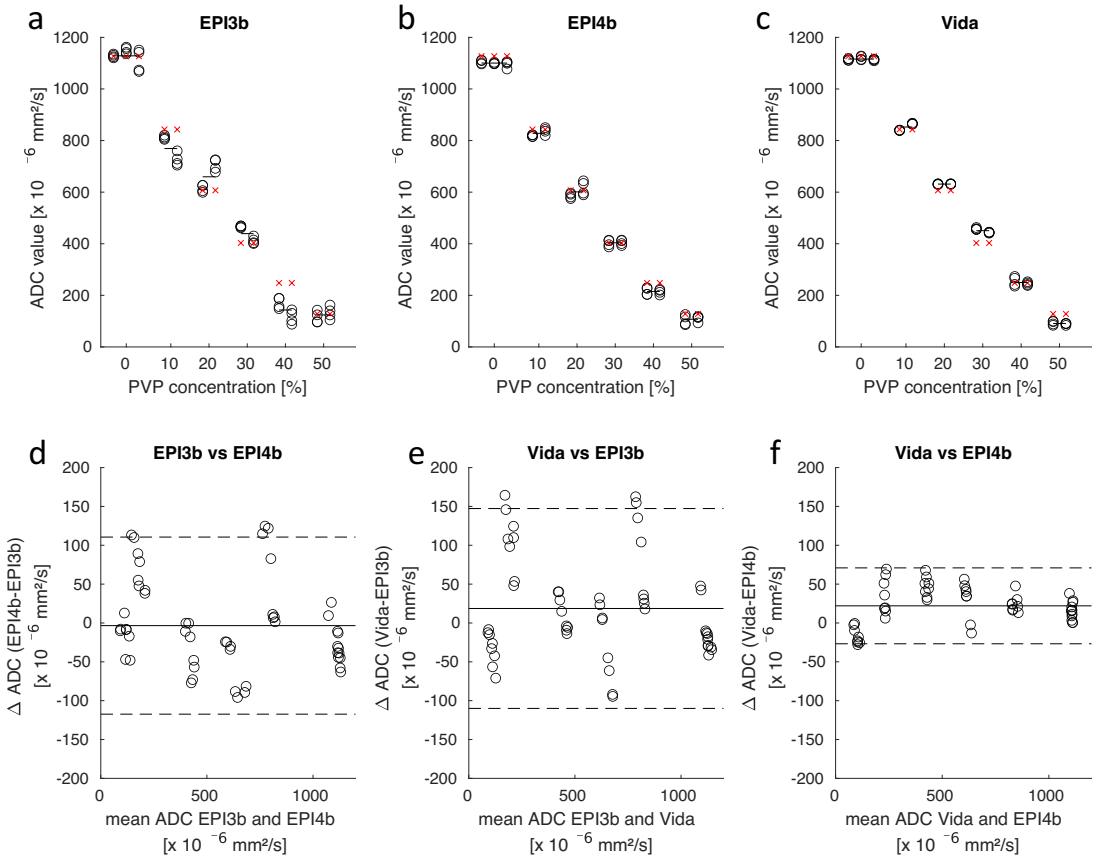

Figure S3: Phantom measurements with the QIBA DWI phantom. The top row displays mean ADC values of both MR-Linac sequences (a, b) and the Diagnostic scanner (c) against the reference ADC values, indicated as red crosses. Solid lines in the top row visualize mean ADC values over all measurements per polyvinylpyrrolidone (PVP) concentration while data points varied from left to right per PVP concentration by being further away from the isocenter. The bottom row represents Bland-Altman analysis comparing mean ADC values from both MR-Linac sequences (d) and the Diagnostic scanner sequence against each MR-Linac sequence (e, f). Here, the solid line visualizes the mean ADC difference and the dotted lines the limits of agreement defined as the mean ADC difference ± 1.96 * standard deviation.

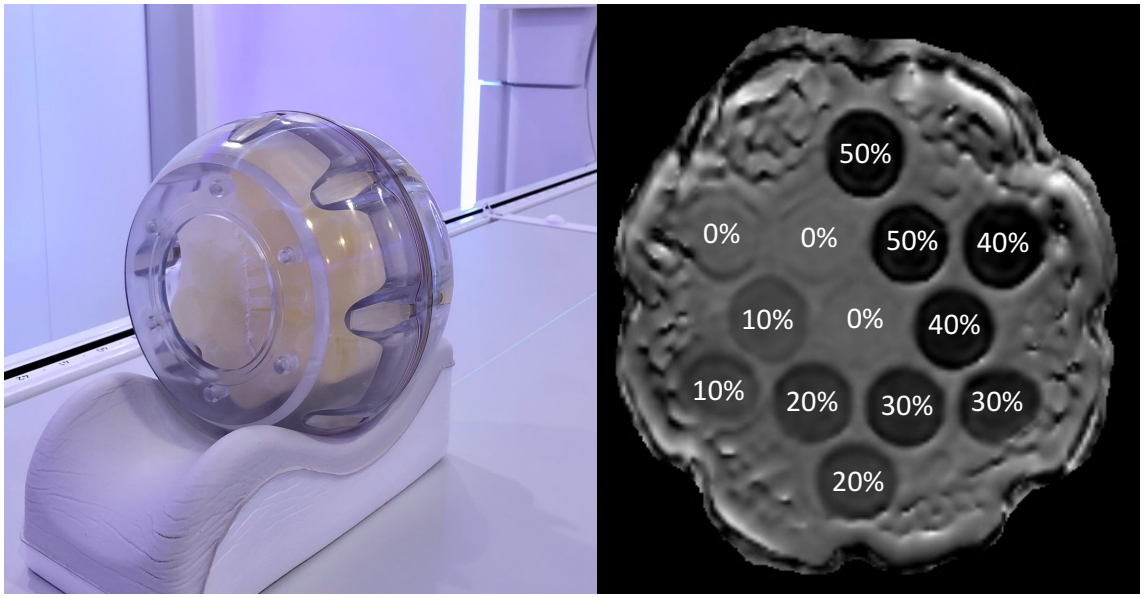

Figure S4: Photograph of the QIBA diffusion phantom (left) and ADC map calculated from an EPI3b acquisition with all vials labeled with their according polyvinylpyrrolidone (PVP) concentration (right).

Table S4: Organ-at-risk coverage over all diffusion-weighted acquisitions from the MR-Linac (MRL) and the Diagnostic scanner (DS) at all timepoints. The order of coverages is given as MRL pretreatment/ DS pretreatment/ MRL week 2/ DS week 2. Volumes used for the subsequent analysis are indicated in bold numbers.

patients	left submandibular gland		right submandibular gland		left parotid gland		right parotid gland	
	CT volume [cm^3]	coverage [%]	CT volume [cm^3]	coverage [%]	CT volume [cm^3]	coverage [%]	CT volume [cm^3]	coverage [%]
Pat01	11.3	**100**/100/—/—	7.3	**100**/100/—/—	37.9	**67.8**/74.7/—/—	36.8	**45.1**/52.2/—/—
Pat02	9.6	**100**/100/100/100	9.9	**100**/100/100/100	24.8	**100**/100/100/100	29.4	**100**/100/100/100
Pat03	3.5	**100**/100/100/100	3.7	**100**/100/100/100	19.8	**100**/100/100/100	17.5	**100**/100/100/100
Pat04	8.8	**100**/100/—/—	7.8	**100**/100/—/—	22.4	**91.5**/100/—/—	23.0	**100**/100/—/—
Pat05	—	—/—/—/—	8.0	**100**/100/—/—	18.7	**100**/100/—/—	22.3	**100**/100/—/—
Pat06	3.2	**100**/100/—/—	2.2	**100**/100/—/—	18.9	**100**/100/—/—	18.9	**100**/100/—/—
Pat07	8.3	—/—/**100**/100	6.2	—/—/**100**/100	16.4	—/—/70.7/**59.1**	22.9	—/—/94.3/**88.2**
Pat08	9.0	**100**/100/—/—	9.1	**100**/100/—/—	27.9	**100**/100/—/—	26.6	**100**/100/—/—
Pat09	11.7	**100**/100/100/100	7.9	**100**/100/100/100	28.1	**92.9**/100/98.9/100	25.1	**99.6**/100/100/100
Pat10	12.4	**100**/100/100/100	13.5	**100**/100/100/100	23.4	**65.8**/98.7/65.8/100	26.8	**60.8**/97.4/60.8/100
Pat11	5.4	**100**/100/100/100	6.3	**100**/100/100/100	17.2	100/100/**99.4**/100	11.5	100/100/**99.1**/100
Pat12	8.5	**95.3**/100/—/—	8.5	**98.8**/100/—/—	30.3	**100**/100/—/—	26.0	**100**/100/—/—
Pat13	10.8	**100**/100/100/100	9.7	**100**/100/100/100	19.1	97.9/100/**41.4**/100	18.8	100/100/**51.5**/100
Pat14	8.2	**100**/100/100/100	10.5	**100**/100/100/100	19.8	**27.9**/52.1/52.1/81.8	21.0	**37.6**/64.3/64.3/91.9
Pat15	13.6	**100**/100/100/100	11.0	**100**/100/100/100	41.0	67.8/99.0/**41.0**/100	39.6	72.7/100/**46.0**/100

10 Acknowledgments

During the last four years of working in the Section for Biomedical Physics and writing this thesis, I had the pleasure of being majorly supported by a great group of people, both in my daily work as well as in my private life.
First of all, I would like to thank my supervisor Professor Daniela Thorwarth for giving me the chance to work in the field of scientific research and providing me with continuous support, new ideas and detailed feedback. It really helped me to broaden my view and grow as a researcher and person.
Furthermore, my sincere thanks go to Professor Daniel Zips, Professor Maximilian Niyazi, Professor Konstantin Nikolaou and Professor Fritz Schick for making this work possible and for the constant critical scrutiny of all results and projects, showing me another view on my work.
While I was working in the Department of Radiation Oncology, I could always rely on a great and productive exchange within the whole department, but from the MR-Linac group in particular. The collaboration with this great group of people from different professions was always productive but no less fun. While all of these people have contributed their part to this work, my special thanks go to Dr. Simon Böke for his continued extra effort put into my research in the framework of doing the delineations for all of my data as well as support with his medical expertise and Dr. David Mönnich for taking on the role as my mentor and always being there with help and advice.
By being part of the biomedical physics research group, I had the pleasure of enjoying great collaborations with colleagues, but almost more important spending the week working with an amazing group of friends such as Moritz, Ivan, Martina, Benni, Karen, Dominik, Cora and Marcel. Traveling to conferences together, the shared stress before an abstract deadline, celebrating each publication with a cake or the occasional drink after work, made going to work much more fun every day.
Although coming to Tübingen over six years ago without knowing anyone, the decision of living in a shared apartment has truly made Tübingen into a second home for me. By sharing stories of the day in the evening, cooking together or binge watching a new netflix series, I have not only found another home, but also many

lifelong friends with Kathi, Vincent, Nici, JK, Lara and Nele.
Lastly, but most important I am deeply grateful for the steadily support I have received by my girlfriend Celine, my sister Eva and my parents Judith and Wolfgang, who always had an open ear for my complains when times were stressful and motivated me to keep going. Without you, this thesis would not been possible.