Aus der

Universitätsklinik für Radioonkologie mit Poliklinik Tübingen
Sektion Biomedizinische Physik

Development and validation of a robust deformable dose accumulation approach in magnetic resonance-guided radiation therapy

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Humanwissenschaften

der Medizinischen Fakultät
der Eberhard Karls Universität
zu Tübingen

vorgelegt von

Murr, Martina Gisela

2025

Dekan:	Professor Dr. B. Pichler
1. Berichterstatter:	Professorin Dr. D. Thorwarth
2. Berichterstatter:	Professorin Dr. M. Knipper-Breer
Tag der Disputation:	19.03.2025

Contents

List of Figures

List of Tables

Abbreviations

ART	Adaptive radiation therapy
ATP	Adapt to position
ATS	Adapt to shape
CBCT	Cone-beam computed tomography
conv-IGRT	Conventional image-guided radiation therapy
CT	Computed tomography
CTV	Clinical target volume
DDA	Deformable dose accumulation
DIR	Deformable image registration
DMA	Dose mapping/accumulation
DMAL	Dose mapping/accumulation landscape
DVF	Deformation vector field
DVH	Dose-volume histograms
DVP	Dose-volume parameters
ED	Electron densities
EQD2	Equivalent dose to 2 Gy fractions
FF	Fractionation factor
GTV	Gross tumor volume
iDVF	Invers deformation vector field
IGRT	Image-guided radiation therapy
IMRT	Intensity-modulated radiation therapy
LET	Linear energy transfer
LINAC	Linear accelerators
LQM	Linear quadratic model
MR	Magnetic resonance
MRgRT	Magnetic resonance guided radiation therapy
MRI	Magnetic resonance imaging
MTRA	Medical-technical radiology assistants
NTCP	Normal tissue complication probability
OA-MRgRT	Online adaptive magnetic resonance guided radiation therapy
OAR	Organs at risk
PC	Prostate carcinoma
PTV	Planning target volume
RIR	Rigid image registration
RT	Radiation therapy
TCP	Tumor control probability

1 Introduction, aims and objectives of the thesis

1.1 The role and improvement of radiation therapy in cancer treatment

Cancer is one of the most common diseases worldwide and the second cause of death in Europe today [178].Treatment of cancer increases the chances of survival for those affected [178].Various forms of treatment have been established: the three most common treatments are local therapies such as surgery or radiation therapy, which is the focus of this thesis, and systemic therapies e.g. chemotherapy and immunotherapy.

Radiation therapy (RT) focuses on a specific organ or a particular part of the body. Despite this targeted radiation technique using ionizing radiation, side effects cannot be ruled out. Toxicities caused by ionizing radiation have already been observed with the first applications [74].Different techniques and procedures to protect healthy tissue during RT have been continuously developed. These include methods like fractionation and intensity-modulated radiotherapy (IMRT) [228] and technological developments such as image-guided RT (IGRT) [259] using cone beam computed tomography (CBCT) or magnetic resonance-guided RT (MRgRT) [127].

The prediction and prevention of toxicities in RT is an unmet medical need. Toxicities can occur months to years after radiation [111], [213], [253]. Various symptoms of toxicities occur, which are currently assessed and reported by the patient and the treating physician [55]. The reported functional disorders and the clinical-biological effects of radiation, e.g. dose-volume effects of radiation-induced urinary toxicity in prostate RT, are poorly understood [130]. In RT planning, models such as tumor control probability (TCP) and normal tissue complication probability (NTCP) are used to minimize the dose to the corresponding organs, such as the bladder and rectum in prostate RT [130], [146], [208].However, Bentzen et al. reported the inaccuracy of the NTCP model for genitourinary (GU) organs in prostate RT due to problems related to the classification of side effects, selection of appropriate statistical methods, testing of internal and external model validity, and quantification of predictive power and statistical uncertainty [21].

In order to better quantify possible side effects and toxicities of irradiation, it is essential to precisely determine the dose deposition in the target volume and the organs at risk (OAR). However, the frequency and the presence of inter- and intrafraction organ movements increase the complexity of this task. Therefore, developing methods for deformable accumulation of the radiation dose which had been delivered in a fractionated manner over several irradiation sessions is an important prerequisite for precisely describing and determining dose-response models in the OAR. Based on the deformable image registration (DIR) of e.g. daily MR images to account for the organ movements

occurring between treatment sessions, the dose accumulation summarizes the total dose that deformably summarizes the fractional dose contributions calculated from the data sets [45]. Based on this, initial study results showed that the cumulative dose from voxel to voxel could be determined by image registration Based on this, initial study results showed that the accumulated dose from voxel to voxel could be determined by image registration [45], [99], [104].

In the following introduction to this thesis, curative percutaneous radiation therapy is briefly described, followed by a brief overview of radiobiology, covering the cell cycle and the effects of ionizing radiation. Followed by the importance of fractionation of ionizing radiation. To quantify the therapeutic effect of radiation treatments, the following chapter will introduce the concept of tumor control probability (TCP) and normal tissue complication probability (NTCP) models. On this basis, the treatment planning process is outlined, which is followed by a brief description of adaptive radiation therapy, including MRgRT, in a separate chapter. Subsequently, the limitations and possibilities are discussed. Finally, the aim of the work is presented.

1.2 Radiation Therapy

Alongside surgery and chemotherapy, radiation therapy (RT) is one of the three main pillars in the treatment of cancer, in combination or alone. RT is applied in various ways depending on the treatment objective (curative or palliative) and the therapeutic status of the patient (preoperative, postoperative, and palliative). Curative percutaneous RT aims is to deliver a sufficient amount of radiation to the tumor in order to destroy it, assuming that the tumor is sufficiently radiosensitive, while avoiding irradiating healthy tissue to an extent that would lead to severe side effects or morbidity. Its preoperative uses include reduction of tumor size, and better differentiation of the tumor from healthy tissue to avoid local recurrence. Its postoperative uses include the reduction of recurrences such as lymphatic channels. Palliative percutaneous RT aims to achieve freedom from symptoms, quality of life, and prevention of tumor-related symptoms.

Percutaneous RT for cancer usually relies on linear accelerators (LINAC). A LINAC injects electrons into a microwave beam at a specific location and at a specific time. Each electron is impacted by the force applied by the electric field and is carried along by the wave at a higher velocity. The ionizing radiation is emitted by the injector system, radio-frequency power generation system, accelerating waveguide, auxiliary system, beam transport system, beam collimator and monitor system. The kinetic energy of the accelerated electrons ranges from 4 to $25\,\mathrm{MeV}$ [2]. The patient is positioned on the treatment couch in preparation for the irradiation. With the control console the medical-technical radiology assistants (MTRA) operates the irradiation. The tissue is exposed to the distributed beams, which interact with the cells.

1.3 Radiobiology

Cancer is a pathological condition characterized by the uncontrolled growth and metastasis of certain cells in the body. Cancer arises due to certain genetic mutations, which are the fundamental hereditary units responsible for physical characteristics. Chromosomes are composed of tightly packed DNA strands that contain genes. One possible cause of genetic changes that lead to cancer is an error in cell division as part of the cell cycle.

The cell cycle is divided into four functional outputs and one resting phase, and the corresponding phases of the cycle are differently sensitive to radiation. The first and shortest phase of the cycle is called mitosis, which is the actual nuclear and cell division. The mitosis involves the division of the duplicated DNA into two daughter cells. During mitosis the cell is most sensitive to the effects of radiation. The second phase is the growth phase known as G1. This phase lasts for different lengths of time depending on the type of tissue. At the beginning of this phase, the cell is relatively insensitive to radiation. However, during transitioning to the third phase, the synthesis (S-phase), the cell is very sensitive to radiation. In the S-phase, the DNA is replicated, and the cell is again largely resistant to radiation. The S-phase is the longest phase apart from G1. In the fourth and last phase, called G2, the cell continues to grow, giving the cell time to ensure that DNA replication is complete. At the end of this phase, which marks the transitioning back to mitosis, and hence back the beginning of the cycle, the cell becomes radiosensitive again. The transition points from G1 to S-phase and from G2 to mitosis, are partially overruled in tumor cells, allowing cells with damaged DNA to escape apoptosis and senescence, and continue to divide [97]. Between cell cycles, a resting phase called G0 exists in which no cell division takes place. In healthy tissue cells, this represents the active functional phase. With the exception of the damage repair (passing through mitosis, G1, S- and G2 phase), most cells, e.g., muscles or nerve cells, are present in this phase. In tumor cells, this phase is a reversible resting phase, the cell can remain with reduced metabolic activity for days to years.

The aim of RT is to destroy the tumor, which can be achieved by inducing cell death, DNA damage or more precisely DNA double-strand breaks, which are the most difficult to repair. The primary cause of radiation's biological effects is damage to DNA, the cell's most important target. Nevertheless, there are additional aspects within the cell that can result in cell death. On the one hand, damage to the cell can occur directly when the radiation interacts directly with the critical target. The coulomb interaction can ionize or excite the atoms of the target, triggering a series of physical and chemical processes that ultimately lead to biological damage. On the other hand, the radiation interacts indirectly with other molecules and atoms within the cell, especially with water, a cell consists of $80\,\%$ water, which leads to the formation of free radicals that can damage the critical target within the cell by diffusion [2]. The biological effectiveness of different types of radiation is evaluated based on the linear energy transfer (LET). It is described as the ratio of dE/dl in a medium, where dE is the average amount of energy that a charged particle with a specific energy transmits to the medium as it travels a distance dl.

When sterilizing the testicles of a ram with ionizing radiation, it was found that a high single dose did not have the desired effect, in contrast to lower single doses spread over several days. Furthermore, this did not cause any major burns to the testicular skin in comparison [108]. This observation has led to fractionation, which is discussed in the following chapter.

1.4 Fractionation

In the context of RT, fractionation is the process of distributing the radiation dose over time [88], [128], [245]. The total dose is divided into multiple fractions due to the radiobiological characteristics of the tumor and normal tissue [65]. This method aims to optimize the damage to malignant cells while simultaneously reducing the level of damage to healthy tissues. With fractionation, the total radiation dose must be increased compared to single-time irradiation. With the equal effect of the irradiation in terms of cell damage, the fractionation factor (FF) is equal to the dose of fractionated irradiation divided by the dose of single irradiation (FF<1). Various fractionation schemes are employed in RT, including hyper-fractionation with a dose per fraction less than $1.8\,\mathrm{Gy}$ and hypo-fractionation with a dose per fraction greater than $2\,\mathrm{Gy}$. It is assumed that the tolerance of healthy tissue is a limiting factor in RT.

Various effects of the fractionation contribute to effectiveness of the treatment fraction, known as “The Four R’s of Radiotherapy”: repair, repopulation, redistribution, and reoxygenation [248]. Two additional R’s, radiosensitivity [224], and only recently introduced, reactivation [31], have to be considered. These effects influence normal tissue and tumor cells differently. The tolerance of healthy tissue is increased by repairing sublethal damage and repopulation of healthy tissue, while reoxygenation and redistribution of cells result in greater damage to tumor tissue. Normal tissue recovers better from radiation damage than tumor tissue, this is described using models such as tumor control probability (TCP) and normal tissue complication probability (NTCP), which are described in the next chapter.

1.5 Tumor control probability and normal tissue complication probability

The link between the surviving fraction of cells, the fraction of irradiated cells that maintain their ability to reproductive integrity, and the total absorbed radiation dose is expressed by the cell survival curve. In addition to the fractionation, the type of radiation determines the form and steepness of the curve. Mathematically, the cell survival curve is described by the Linear Quadratic Model (LQM), assuming that there are two tissue-type specific characteristic constants α and β. Here, α represents the initial slope of the cell survival curve, and β is a smaller constant that explains the quadratic component of cell death by radiation [2]. $S(D)$ represents the fraction of cells that survive a dose

of D (Equation 1 and Figure 1).

$$S(D) = e^{(-\alpha D - \beta D^2)} \tag{1}$$

The observed biological effect - tumor induction and tissue reaction - against the given

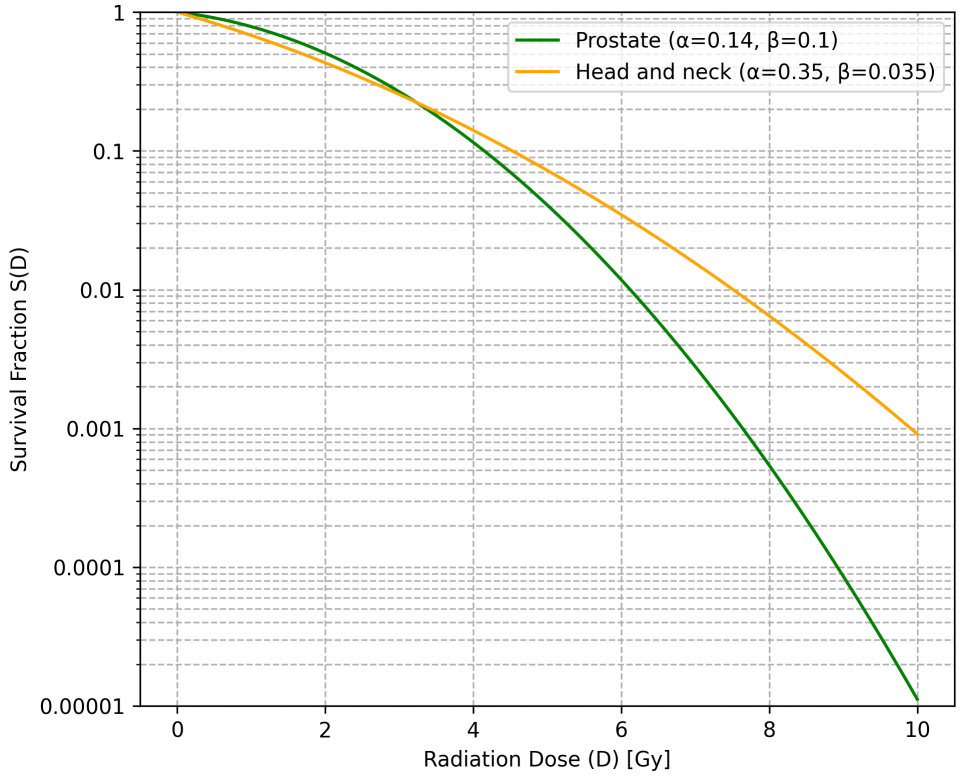

Figure 1: Linear quadratic model (LQM) describing the cell survival curve on the example of the prostate with late and head and neck with early radiation reaction.

dose are described by the dose-response curve. There is a clear distinction between an early radiation reaction, where the α/β ratio is larger for e.g. head and neck with $10\,\mathrm{Gy}$ [134] and late radiation reaction, α/β ratio is smaller with e.g. $1.4\,\mathrm{Gy}$ in prostate [156].

Due to varying fraction schemes in clinical practice, the equivalent dose to $2\,\mathrm{Gy}$ fractions (EQD2) is often used to translate radiobiological concepts into concrete, interpretable values. EQD2 refers to a biologically equivalent effect as would be obtained with a fraction dose of $2\,\mathrm{Gy}$ and also depends on the total dose (D_i) and the volume (V_i) at dose level i. EQD2 obtained with a fraction dose (d), the α/β ratio representing the dose-response relationship, and the tissue-specific parameter n that is associated with the tissue's sensitivity to dose inhomogeneity is calculated as follows:

$$EQD2 = \left(\frac{\sum_i (D_i^n \cdot V_i)}{\sum_i V_i} \right)^{\frac{1}{n}} \cdot \frac{d + \frac{\alpha}{\beta}}{2 + \frac{\alpha}{\beta}} \tag{2}$$

Again, the objective of RT is to deliver an adequate dose of radiation to the tumor in order to destroy it while avoiding a dose that would result in serious damage to the surrounding tissue. This is illustrated by two sigmoidal functions of dose, the tumor control probability (TCP) (Equation 3) [148] and the normal tissue complication probability (NTCP) (Equation 4) using the Lyman-Kutcher-Burman (LKB) model [37], [124], [125], [142], as illustrated in Figure 2.

$$TCP = P_{N=0} = e^{(-N_0 \cdot S(D))} \tag{3}$$

with $S(D)$ (Equation 1), the poison distribution (P_N) to quantify the likelihood of N out of N_0 clonogenic cells surviving a treatment.

$$NTCP = \frac{1}{1 + e\left(-\frac{EQD2 - TD_{50}}{m \cdot TD_{50}}\right)} \tag{4}$$

with EQD2 (Equation 2), TD_{50} dose at which $50\,\%$ of the population is expected to experience a complication and the slope parameter m describing the steepness of the dose-response curve.

The optimal radiation technique for treating a specific tumor is to maximize TCP while minimizing NTCP with the largest possible dose gap between the tumor and surrounding tissue, represented by the therapeutic index [18]. NTCP models are prediction tools

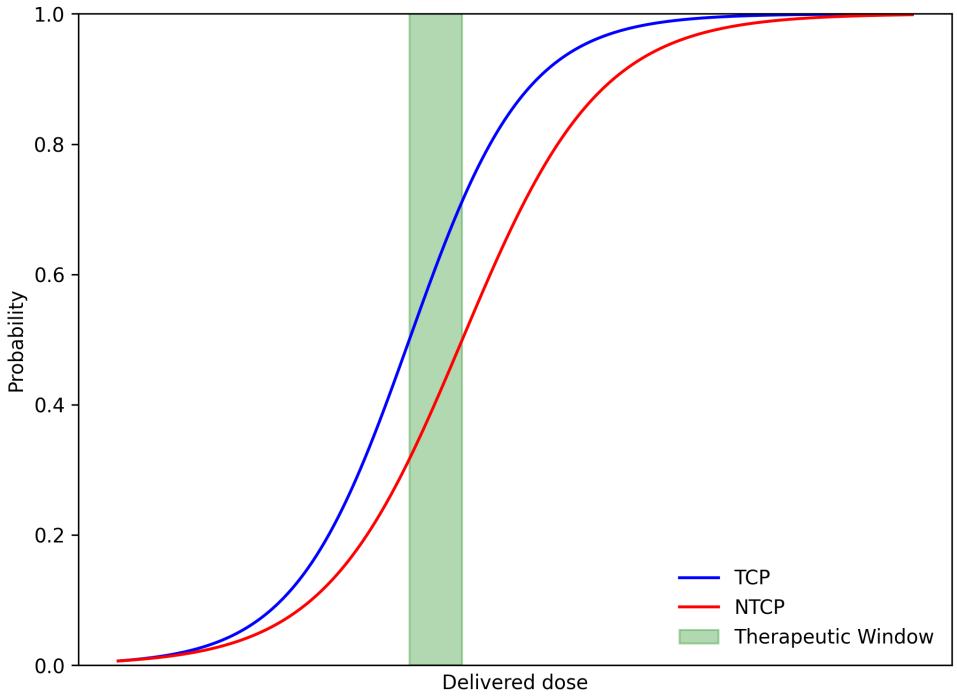

Figure 2: Exemplary illustration of the tumor control probability (TCP) and normal tissue complication probability (NTCP) with the therapeutic window.

used to estimate the likelihood of radiation-induced normal tissue complications. Together with patient and treatment data, these models attempt to transform radiation

dose distributions into a forecast of the likelihood of a problem developing [237]. NTCP models have become more often used in daily clinical practice in recent years. However, there is ongoing discussion about these models due to challenges in development and validation processes [21], [237]. In order to reduce side effects, a precise initial treatment planning is necessary, addressed in the next chapter.

1.6 Radiation Dose Planning

A prerequisite for RT is radiation treatment planning, initiated following the diagnosis and the decision to undergo RT. Based on a computed tomography (CT) scan that displays the patient's unique anatomy as a three-dimensional density distribution, 3D radiation treatment planning has been possible since the late 1980s [89]. The definition of oncological and radiotherapeutic volumes provides the basis for accurate 3D radiation treatment planning, precise dose calculation, and reporting.

The primary tumor, defined as gross tumor volume (GTV), is determined from a combination of imaging modalities, like CT, MR, PET and/or US, diagnostic modalities, and clinical examination. The clinical target volume (CTV) encompasses the immediate surrounding area of the GTV, which may include the smallest pathological conditions and other regions considered at risk and requiring medical intervention. A planning target volume (PTV) is created for the treatment, a geometrical concept that includes an internal target margin and an additional margin for set-up uncertainties, machine tolerances, and intra-treatment variations. The PTV is linked to the frame of reference of the treatment machine and is often described as CTV plus a fixed or variable margin. Furthermore, organs at risk are defined as those that are in close proximity to the PTV and are so sensitive to radiation that the dose received from a treatment may be significant compared to its tolerance, possibly requiring a change in the radiation arrangement or a change in dose [96]. Organs whose radiation tolerance depends on the fractionation scheme should be fully described to avoid bias in evaluating the treatment plan.

A dose indication is a clearly defined prescription and planning constraints such as minimum, maximum or mean dose (Dmin, Dmax, Dmean) to the target and organs at risk (OAR), as well as volume constraints like VxGy, defining a certain relative volume per organ which may not receive more than x Gy , with detailed information on the total dose, fraction dose, and treatment days determined by the oncologist. On this basis, the medical physicist calculates the treatment plan utilizing a complex software solution offered by the treatment planning system (TPS), considering the dosimetric properties of the available radiation types, radiation field shaping, and geometric irradiation options. The irradiation sequence, field directions, field shapes, and respective doses are determined. The dose distribution is calculated on the CT with the organs' differently determined electron densities (ED) and visualized.

Dose-volume histograms (DVH) show cumulatively the percentage of the total volume and the distribution of dose values within a defined volume. These histograms are used to assess the distribution of radiation doses during RT. The defined dose-volume pa-

rameters (DVP), such as Dmin and Dmax, are calculated based on the DVH. Figure 3 a-c illustrates the CT planning image with delineated target volume and relevant OARs of a prostate cancer case, including a 3D dose distribution and beam angles in a representative slice. The corresponding DVH is illustrated in Figure 3 d.

Intensity-modulated RT (IMRT) is an advanced form of three-dimensional conformal RT. It is particularly beneficial for target volumes that are complex in its shape and close to radiosensitive normal tissue. In contrast to conformal RT, it has two different features: first, non-uniform intensity of the beams shaped by a multi-leaf collimator (MLC)[227], [228]; and, second, inverse planning [32], [227], [228] by maintaining tumor dose and normal tissue dose constraints, calculating optimal beam intensity configuration using a TPS. The step and shoot irradiation procedure is used for IMRT, where the gantry stops at defined angles. A special form of this technique is volumetric modulated arc therapy (VMAT), in which the gantry rotates in full or partial arcs around the patient [179]. Both techniques have one or more beams in common which consist of several segments modulated by the MLC.

However, the treatment plan represents the planning situation, which is a point in time. The RT method in which the treatment plan can be modified based on systematic feedback from measurements is called adaptive RT (ART) [261]. ART aims to improve radiotherapy treatment by systematically monitoring treatment variability and using it to re-optimize the treatment plan on a daily basis. This will be discussed in the next chapter.

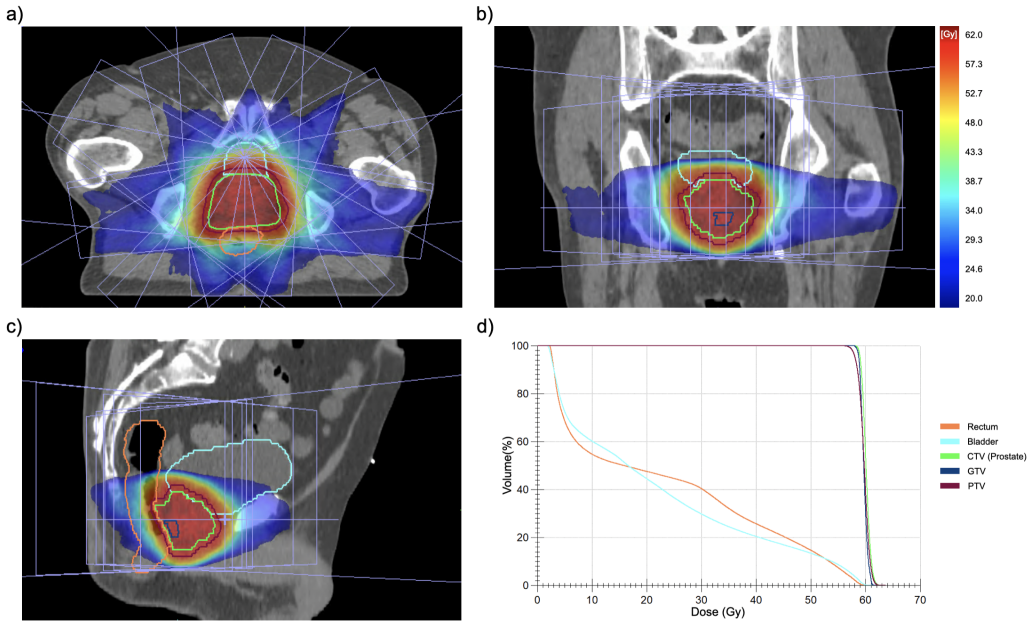

Figure 3: 3D dose planning based on a CT image a) axial, b) coronal c) sagittal view, and d) DVH of the contour's rectum, bladder, CTV, GTV and PTV. The nine beams 25°, 60°, 100°, 135°, 180°, 225°, 260°, 300° *and* 340° *are presented in purple in all three views.*

1.7 Adaptive radiation therapy

The precision of dose delivery using approaches such as IMRT is limited by the imprecision in accurately determining the location of the target during treatment. This uncertainty has two causes. The first cause is originating from the combined effects of inter- and intra-fractional motion of the target relative to anatomical reference points [259] (Figure 4). The second cause of uncertainties is due to setup errors and other inaccuracies. The conventional IMRT approach involves incorporating target volume margins, often resulting in the loss of the majority of the potential advantages offered by more accurate treatment procedures.

Image-guided RT (IGRT), using e.g. on-board cone-beam CT (CBCT), enables visualization of the patient's anatomy, including accurate knowledge of the daily placement of the target volume before delivering a fraction dose. By comparing the patient's position image acquired with the integrated LINAC imaging system to the CT scan for treatment planning and then correcting the patient's position accordingly, a more precise placement of the target volume is achieved. Thus, with IGRT, a reduction in treatment margins, toxicity reduction in OAR, dose escalation, and the avoidance of geographic misses were achieved.

A complete realization of IMRT and IGRT [2] led to the concept of adaptive RT (ART), which makes it possible to adapt to individual patient variations. The parameters of an initial treatment plan - such as the treatment dose and field margin, as well as the field shape, beam intensity, and geometry - are modified if necessary, considering the characterized variation [261]. In detail, three types of ART are distinguished, offline, online, and real-time ART [94], [256], illustrated in Figure 5. In offline ART, the position of the target is recorded at one or more specific points in time. The procedure that initiates the adaptation is performed on a separate imaging system that is not integrated into the LINAC. Changes that are not amendable via positional adjustments, particularly systematic and progressive changes such as weight loss and tumor shape, are addressed by modifying the treatment plan according to the observed changes after the current treatment fraction. The identical procedure is employed as that of the initial planning ART [94], [256].

Online ART seeks to address the temporal and stochastic changes that are identified within a single treatment session while the patient remains in treatment position. Consequently, online ART can only be realized using integrated systems offering online imaging and radiation treatment in the same system and ideally at the same time. Imaging, image registration and comparison, fast replanning, plan evaluation, and quality assurance are necessary to adapt the RT plan on the fly. Hunt et al. [94] categorizes online ART into several stages. The simplest of which is the shift of the plan to the current anatomy, by translation and optional rotation, which does not require optimization or replanning. Another option is to modify the gantry and collimator angle. Subsequent treatment fractions are carried out using the adapted treatment plan [261]. A complete online adjustment includes the following steps [76]. First contouring, which can be done manually, automatically or through the three-dimensional (3D) geometric transformation

matrix which matches planning and fraction image. If the dose delivery prediction does not meet the specifications, replanning is performed until the clinical criteria are met. Subsequently, the plan undergoes a quality assessment. Finally, the treatment is administered. Real-time ART has been implemented to address anatomical differences that occur throughout a treatment session (intrafraction alterations), such as changes in breathing, internal conditions, and peristalsis movements. This method automatically adjusts the treatment plan while delivering the beam, using continuous imaging without requiring any manual interaction from the operator. Gating and tracking techniques are employed to effectively account for these motions. The gating administers the dosage exclusively when the target is located inside a certain geographical position, which is confirmed using imaging techniques such as CBCT, where identifiable markers are used. During tracking, the beam continuously follows the target so that the target always remains in the beam's field of view and the treatment parameters are adjusted accordingly. In contrast to gating, the beam always remains on.

However, despite high-resolution images, the CBCT leaks image quality regarding soft tissue contrast. In contrast to CT images, magnetic resonance (MR) images provide greater soft tissue contrast, which is necessary for some disease diagnoses or the precise localization of tumors, without the need for additional radiation. A recent development of IGRT is MR-guided RT (MRgRT). In the following chapter, MRgRT is outlined in more detail.

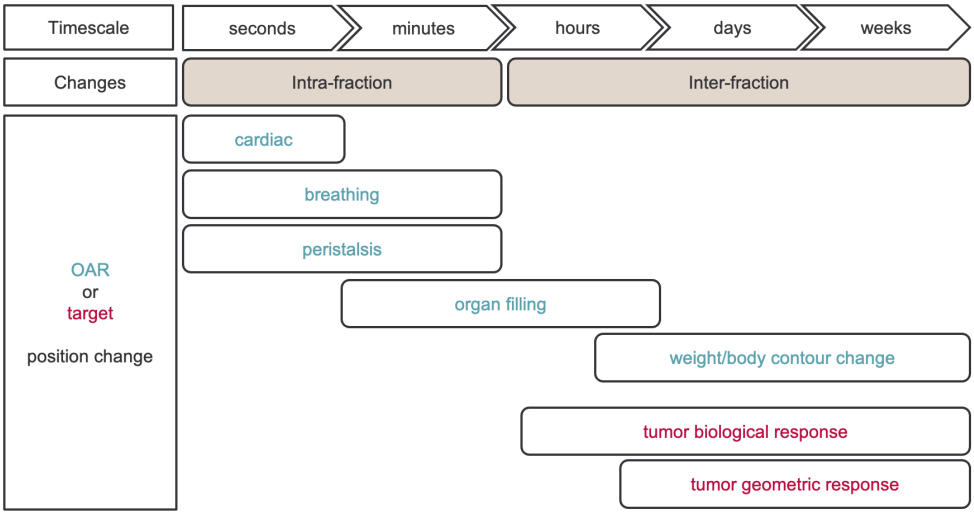

Figure 4: Timescale of OAR and target position changes, based on [94] with permission from Elsevier.

1.8 Magnetic resonance guided radiation therapy

MR imaging (MRI) is based on the excitation of spin magnetization in the body by external magnetic fields. The hydrogen atom is most frequently used for MRI contrast formation, which consists in its nucleus of a single particle, a proton, from which an

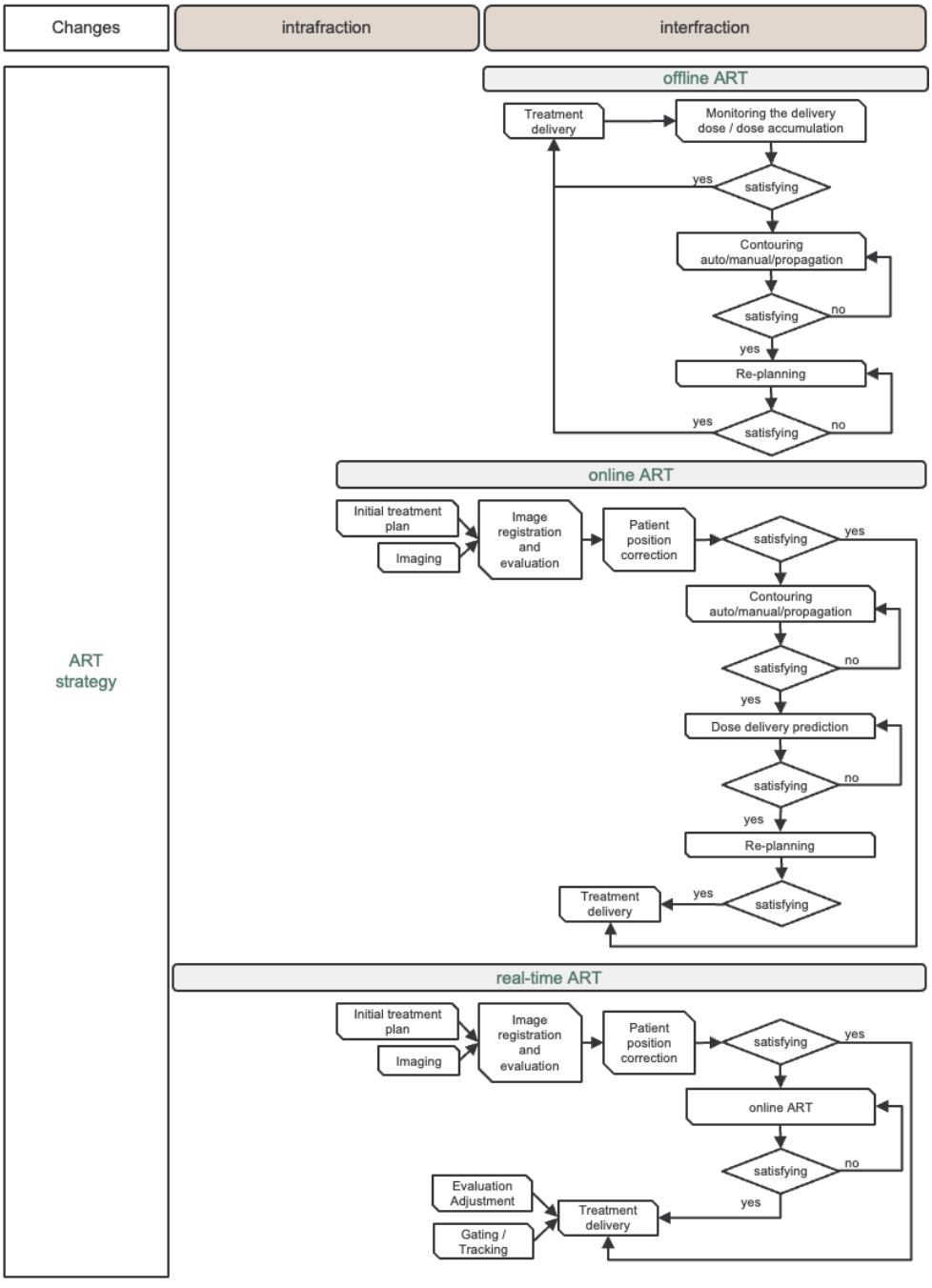

Figure 5: Basic steps in ART strategies, offline, online and real-time ART, based on [76], [94] with permission from Elsevier.

electron orbits. The contrast in the MR image is determined by two relaxation time constants, the longitudinal (T_1) and the transverse (T_2) relaxation time, which are characterized by the tissue.

The MRI acquisition sequence has two basic parameters, which influence the image contrast. The first parameter to do so is the echo time (T_E), which is the time between the midpoint of the radiofrequency (RF) pulse and the midpoint of the echo. The second parameter is the repetition time (T_R), which is the duration between successive points on a recurring sequence of pulses and echoes. The signal intensity (S) of a spin echo sequence is described by

$$S = S_0 \cdot e^{\left(-\frac{T_E}{T_2}\right)} \cdot \left[1 - e^{\left(-\frac{T_R}{T_1}\right)}\right] \tag{5}$$

with S_0 being the density of protons in a volume. MRI also offers several other image contrasts or modalities to determine movement, function, and physiology, used in cancer diagnosis and follow-up assessment, which is why a combined MR-LINAC system was developed [126], [127], [190].

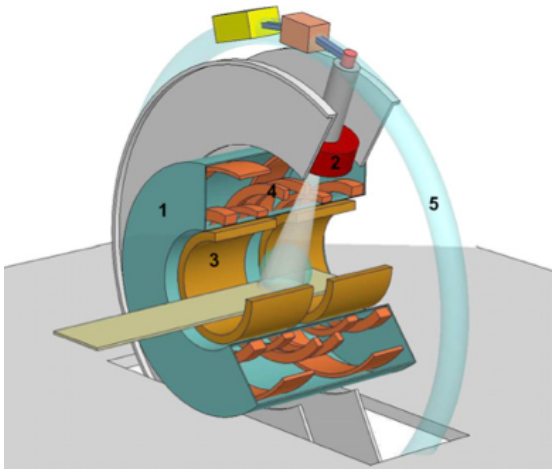

Figure 6: Concept of an MR-LINAC System [190]. 1) The 1.5 T MRI is illustrated in blue, 2) the 6 MV accelerator is situated in a ring surrounding the MRI, 3) split gradient coil presented in yellow, 4) superconducting coils denotes in orange and 5) the toroid in light blue. Reprinted from [190] with permission from Physics in Medicine & Biology.

At the University Hospital Tübingen in Germany, the first patient was treated with the combined $1.5\,\mathrm{T}$ MR-LINAC (Elekta Unity System) in 2018. The MR-system is equipped with a superconducting $1.5\,\mathrm{T}$ magnet that has a diameter of $70\,\mathrm{cm}$ and a closed bore. Without compromising B0 homogeneity, the MRI was modified for radiation beam transmission through positioning the superconducting coils outside the transverse center plane and utilizing a split gradient coil. The MR-LINAC is equipped with a helical gantry that houses all the electrical and technical components of the LINAC. This gantry spins outside the RF cage and is designed to minimize the impact of the static magnetic field by active shielding (see Figure 6). Due to the MRI's space requirements, the photon

source is 143.5 cm from the isocenter, in contrast to typical LINACs which usually have a source-to-isocenter distance of 100 cm. The accelerator system, a standing waveguide, utilizes a single photon energy of 7 MeV.

MR-guided RT (MRgRT) allows for daily online adaptation of target volume and OARs, consequently adapting or recalculating the dose planning through the adapt-to-position (ATP) or adapt-to-shape (ATS) workflow to account for more precise adaptations of the patient's anatomy of the day (Figure 7) [247]. Both workflows include positioning the patient on the treatment table, which was determined during treatment planning, followed by verifying the patient's current anatomy using MRI. In ATP, the planning or a previously acquired fraction image, CT or MR, respectively, is registered rigid to the day's MR image; the contours are propagated accordingly, followed by rigid dose mapping of the corresponding dose. After the plan is approved by the medical physicist and oncologist, obeying the planning DVPs, the patient is irradiated. If the anatomy of the planning CT or a previous fraction does not represent the current anatomy, the ATS workflow is performed by adapting the contours to the patient's anatomy of the day, followed by re-optimizing and re-calculating the plan and dose based on the treatment planning. Similarly, the plan is applied after the physicist's and oncologist's approval. The MR-only planning approach was introduced recently to reduce the patient's sojourn, based on a population-based ED concept [48]. Automatic contouring [12], [200] and planning [123] were developed to shorten the planning process.

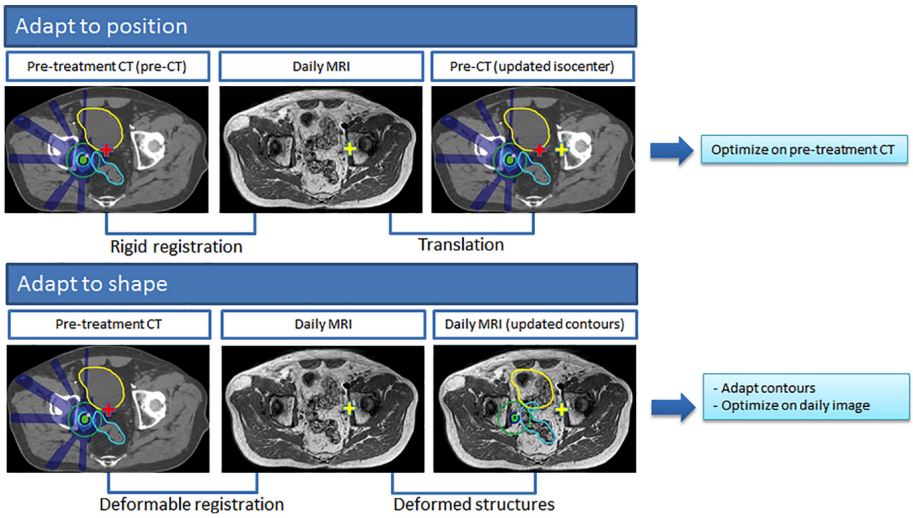

Figure 7: Adapt to position (ATP) and adapt to shape (ATS) illustration. Reprinted from [247] with permission from Elsevier.

1.9 Potentials and limitations

In recent decades, major advances have been made in RT for cancer patients. The treatment has proven successful. Further developments have opened new fields and opportunities to further increase the success rate. However, side effects are still re-

ported by patients. Due to fractionation the total delivered dose during treatment remains unclear.

Methods like rigid image registration (RIR) and deformable image registration (DIR) have been developed to superimpose two different images into one coordinate system, also known as gird. Typically, the images are called source and destination, where the source image is registered to the destination image grid. These images can be, either monomodal, e.g., MRI to MRI or multimodal e.g., CBCT to MRI. The simplest form of image registration is RIR, which involves 3 to 6 degrees of freedom, including translation in the x-, y-, and z-directions and optionally, rotation in spatial directions. This process results in a uniform translation and rotation. A more complex fusion is the DIR, which has several degrees of freedom allowing each voxel displacement to be described by a vector, resulting in a non-uniform transformation. This non-uniform transformation is often presented as a deformation vector field (DVF) and its inverse (iDVF). Over time, numerous algorithms have been developed for DIR.

To calculate the total dose, the transformation is applied to the corresponding doses, assuming the transformation valid. Subsequently, a summation of the transformed doses is performed. Various methods are discussed for dose mapping to achieve a voxel-based dose mapping.

With the introduction of online adaptive MRgRT (OA-MRgRT), detailed daily imaging, precise delineation of tumor and OAR, and dose recalculation became possible. Deformable dose accumulation (DDA), which is defined by DIR with subsequently dose mapping and accumulation, promises the potential to calculate the effectively delivered total treatment dose, an improved TCP and NTCP modelling might be possible, and hence a precise toxicity prediction. However, despite constant further development, DDA has not yet found its way into clinical practice. A robust DDA method and process is still lacking.

1.10 Aims of this thesis

The aim of this thesis is to investigate the use of deformable dose accumulation in adaptive radiation therapy and to develop a robust deformable dose accumulation methodology and workflow for online adaptive MR-guided radiation therapy. To achieve this overarching goal, three dedicated objectives were defined for this thesis.

The first objective is to analyze the applicability and usage of dose mapping and accumulation in radiation therapy, including identification and classification of the challenges that hinder their integration into clinical practice. As part of this objective, current dose mapping and accumulation use cases are categorized, including applicability and usage, to identify the strengths and limitations of these use cases. This phase involves a thorough review of the current literature, which includes clinical studies and software realizations. We analyze these challenges in detail to further explore their causes and impact on the dose mapping and accumulation use cases, including technical, method-

ological, and biological issues. This detailed analysis leads to developing comprehensive guidelines and requirements tailored to researchers, clinical radiotherapy users and also to healthcare providers. These guidelines aim to address the identified challenges by offering practical solutions and best practices for overcoming them. Chapter 2.1 presents the publication [160] that summarizes the results of the first objective.

The second objective is to develop a robust deformable dose accumulation solution to determine the total delivered dose over the course of online adaptive MR-guided radiation therapy. A multi-center study will compare and evaluate the developed solution with several other solutions currently used in research and clinical practice. The deformable dose accumulation procedure consists of three steps: deformable image registration, dose mapping, and dose summation. Deformable dose accumulation solutions are based on different mathematical and physical models, which lead to different results in determining the total dose delivered. Comparison of the different applied solutions is difficult because the ground truth is unknown. Therefore, the multi-center study includes both a gold standard case, which provides the ground truth, and several clinical patient cases from different parts of the body treated with online adaptive MR-guided radiation therapy. The study analyzes the agreement between different algorithms used by different institutions and investigates whether there are differences that depend on the algorithms used or on specific anatomical regions. The project's results in relation to the second objective in chapter 2.2 were published in [159].

The third objective is to develop a deformable dose accumulation workflow that uses the developed deformable dose accumulation solution in the second objective to determine the total dose delivered over the course of treatment. This objective includes the comparison of non-adaptive and adaptive radiation therapy, focusing on classical image-guided and online adaptive MR-guided radiation therapy. The comparison will evaluate the two approaches in terms of tumor dose coverage, normal tissue sparing, and differences in clinical effects. Adaptive radiation therapy, like online-adaptive MR-guided radiation therapy, promises better normal tissue sparing compared to non-adaptive radiation therapy. This is due to the daily acquisition of MR images, which allows daily dose optimization and calculation based on the anatomy of the day. This study uses the deformable dose accumulation solution from the second project in ten prostate cancer patients receiving online adaptive MR-guided radiation therapy. Dosimetrically, we compare the accumulated doses regarding the target volume and normal tissues, such as the rectum, bladder, and urethra, after classical image-guided radiation therapy and online adaptive MR-guided radiation therapy in comparison to the initially planned doses. The dosimetric differences are the analyzed using a normal tissue complication probability model to determine the clinical impact. Chapter 2.3 presents the results of the third objective, summarized in a manuscript [161] which has recently been submitted to Physics and Imaging in Radiation Oncology, and is currently under review.

2 Results and discussion

2.1 Applicability and usage of dose mapping/accumulation in radiotherapy

Autor: **Martina Murr**, Kristy K. Brock, Marco Fusella, Nicholas Hardcastle, Mohammad Hussein, Michael G. Jameson, Isak Wahlstedt, Johnson Yuen, Jamie R. McClelland, Eliana Vasquez Osorio

Published in: Radiotherapy and Oncology 2023, Vol. 182, Pages 109527

DOI: https://doi.org/10.1016/j.radonc.2023.109527

2.1.1 Abstract

Dose mapping/accumulation (DMA) is a topic in radiotherapy (RT) for years, but has not yet found its widespread way into clinical RT routine. During the ESTRO Physics workshop 2021 on "commissioning and quality assurance of deformable image registration (DIR) for current and future RT applications", we built a working group on DMA from which we present the results of our discussions in this article. Our aim in this manuscript is to shed light on the current situation of DMA in RT and to highlight the issues that hinder consciously integrating it into clinical RT routine.

As a first outcome of our discussions, we present a scheme where representative RT use cases are positioned, considering expected anatomical variations and the impact of dose mapping uncertainties on patient safety, which we have named the DMA landscape (DMAL). This tool is useful for future reference when DMA applications get closer to clinical day-to-day use.

Secondly, we discussed current challenges, lightly touching on first-order effects (related to the impact of DIR uncertainties in dose mapping), and focusing in detail on second-order effects often dismissed in the current literature (as resampling and interpolation, quality assurance considerations, and radiobiological issues).

Finally, we developed recommendations, and guidelines for vendors and users. Our main point include: Strive for context-driven DIR (by considering their impact on clinical decisions/judgements) rather than perfect DIR; be conscious of the limitations of the implemented DIR algorithm; and consider when dose mapping (with properly quantified uncertainties) is a better alternative than no mapping.

2.1.2 Introduction

Radiotherapy (RT) aims at treating cancer by delivering radiation to the affected zones (target). In RT planning, a dose distribution is generated on a medical image. Dose distributions are often highly heterogeneous and conformal with steep dose gradients in anatomy close to the target which may move, deform, and/or respond to treatment. It is useful to transfer the dose distribution from one image to another to compare and/or accumulate doses to specific anatomical sub-volumes. Transferring or mapping dose distributions requires aligning the underlying images, assuming any transformation describing the geometric mapping between the two frames of references of the images, can be applied to their corresponding dose distributions [34]. In Fig. 8 a-b, we present a scheme detailing this assumption. Note we refer to the source image as the image associated with the dose to be mapped. Conversely, the destination image defines the grid where we want to map the dose.

Aligning images is done using image registration, which aims at finding the transformation that optimally aligns corresponding anatomy between two images. Rigid transformations, consisting of rotations and translations, are often used to correct for global misalignments (correcting patient positioning). However, rigid transformations are not sufficient to resolve the misalignment due to local changes [45], or large position changes in patient setup [50], [51]. In these cases, deformable image registration (DIR), is preferred [5], [58], [195]. The inputs for DIR include two images, a fixed and a floating image, and the result is a non-rigid transformation. The transformation (T in Fig. 8) is often represented as a deformation vector field (DVF). DVFs are composed by a set of vectors originating from the voxel positions in the fixed image and pointing to the corresponding locations in the floating image (typically not coinciding with the centre of a voxel). Specifically, T maps point coordinates from the fixed grid to the floating frame of reference.

Mapping the source dose distribution to the destination grid can be achieved by either gridded resampling (which pulls or "fishes out" the dose from the floating image), or scattered resampling (which pushes or shoots the dose, as in archery, to the floating image). Depending on the selected strategy, the source/destination image takes the role of the fixed or floating image in the registration, Fig. 8 c. For several applications, the mapped dose is then summed up with another dose, which is known as dose accumulation.

Uncertainties in the registration will introduce dose mapping uncertainties. An important factor affecting registrations is the degree of anatomical variation. Small differences may be accurately registered by most algorithms while complex changes may not. However, a limited number of algorithms attempt to address complex changes [45], [50], such as sliding tissue [51], [58], (dis)appearance of tissue (including different tumour regression modes [5] and surgical interventions [56], [174], [187], [195], [235]). Often these advanced algorithms are only available for research. Other factors can also affect the quality of the registration, such as image acquisition artefacts, lack of contrast, choice of parameters, etc. [34], [210]. These complex changes and other

factors lead to large uncertainties in the registration, and hence uncertainties in the mapped dose, particularly in regions where high dose gradients are present (Fig. 9). Therefore, quantifying dose mapping uncertainties would need to account for both registration uncertainties and dose characteristics.

Dose mapping can be complex and challenging. It is often not clear what the "correct" transformation is, and exactly how this should be used to resample and accumulate the doses. Uncertainties in dose mapping could impact treatment outcomes. An extreme example is if new fractions are re-optimised accounting for estimated delivered dose (mapped with significant errors), which can ultimately result in relapse (if the target was seriously underdosed) or severe side effects (if normal tissues are overdosed). Consequently, determining the proper registration needs to consider at least these two aspects: degree of anatomical variations and the impact of dose mapping uncertainties.

In this manuscript, we will present considerations demonstrating that DMA is more than just applying the results of a registration, as correct dose mapping is not guaranteed even with "perfect" registrations. We also propose a visualisation aid to exemplify the variation in anatomy and impact of uncertainties on the patient's treatment for a range of use cases. We finish the paper with a list of recommendations aimed at users and vendors. Notice that we will not discuss image registration in depth, as several reviews are available in the literature [34], [50], [176], [177], including reviews focussed on DMA [45], [258].

2.1.3 Workshop discussions

Discussions were held during the ESTRO Physics workshop 2021 on "commissioning and quality assurance of DIR for current and future RT applications". This workshop included two focussed online sessions with more than 20 participants. During the second session aspects identified in the first session were discussed in groups. The result of the discussions in the "DMA" group forms the basis of this article.

2.1.4 Dose mapping/accumulation landscape (DMAL)

A direct outcome of these discussions was a scheme representing the current landscape of DMA use cases. The scheme was based on two axes: the degree of anatomical variations expected, and the impact of dose mapping uncertainties on patient safety.

Anatomical variations
The degree of anatomical variations between the images may challenge the registration algorithms [198]. Anatomical variations can be divided into two main categories in RT applications, intra- and inter-patient (Fig. 10 a).

For intra-patient applications, the degree of anatomical variations often increases with the time difference between images: in an extreme example, large anatomical variations are observed between medical scans over lifetime (newborn vs adult). This is particularly relevant in the re-irradiation context, where dramatic changes can be

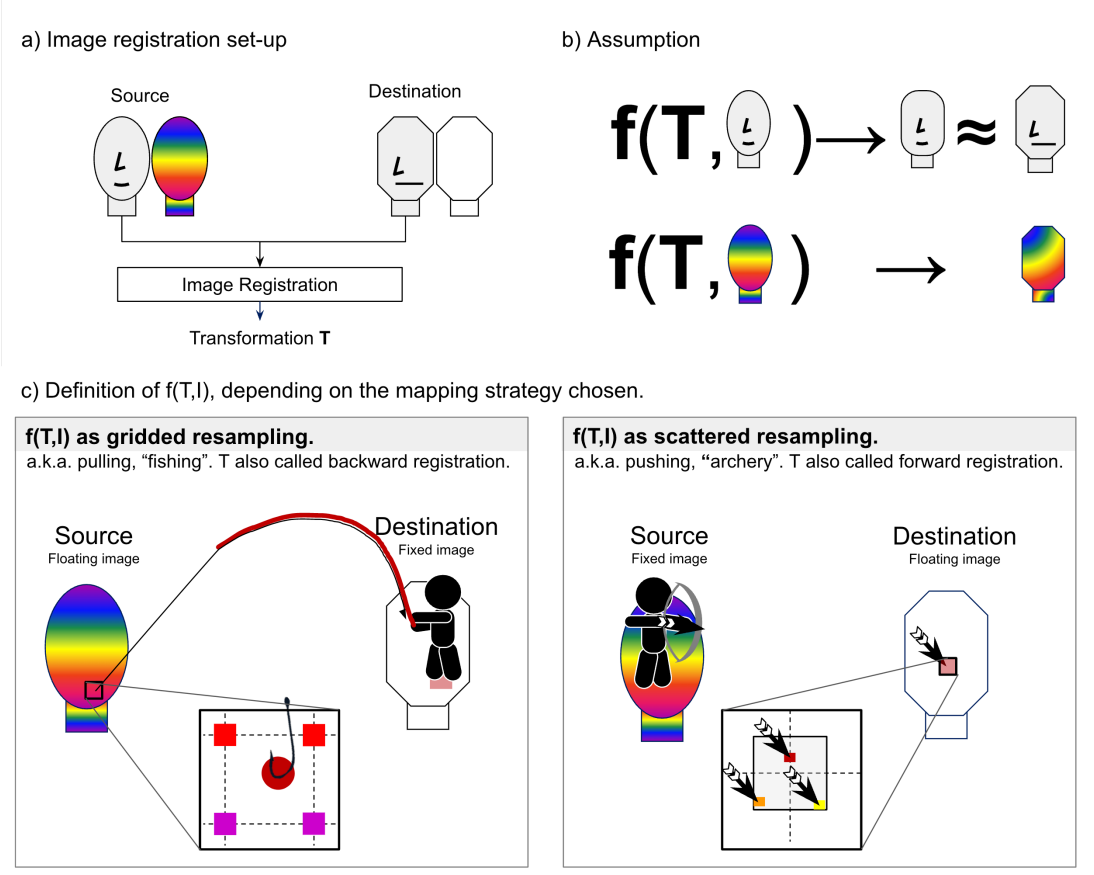

Figure 8: Schematic presenting the use of image registration to map dose distributions. a) Registration set-up. Registration is performed between the source and destination images, here presented in black and white drawings, aiming at mapping the source dose distribution to the destination grid. b) Dose mapping using transformation T, showing the underlying assumption that the registration aligning the pair of images is valid to map any spatially correlated image, such as dose distributions. c) Application of the transformation, depending on the used mapping strategy. Note we distinguish between source/destination and fixed/floating images. Source image is the image that is associated with the dose to be mapped. Destination image is where we want to map the dose. Fixed and floating images are the roles these images take in the registration process.

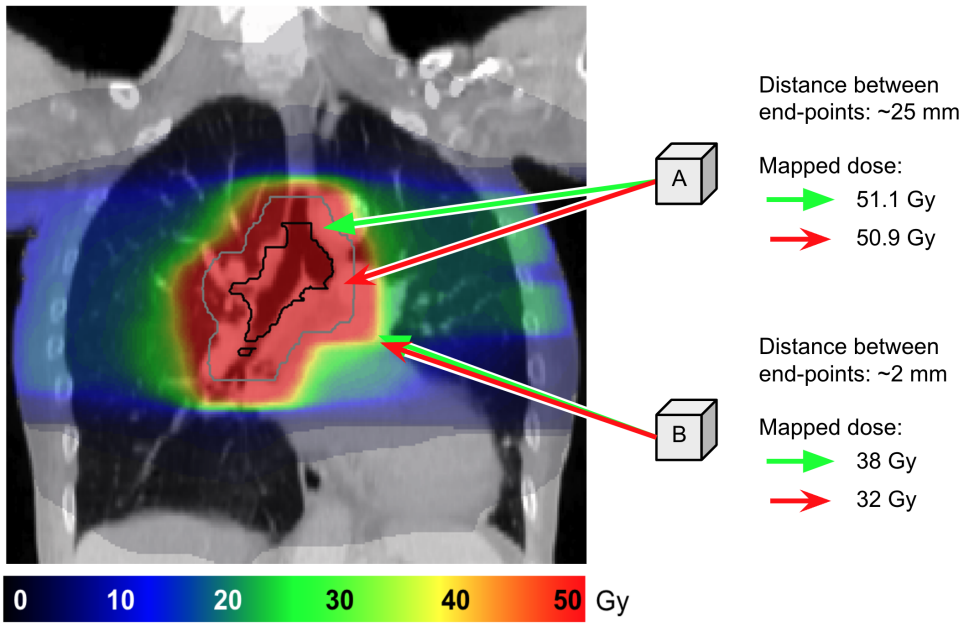

Figure 9: An illustrative case demonstrating the effect of registration uncertainties on dose mapping, and its interplay with dose gradients. For each voxel, A and B, two arrows are shown: 1) a red arrow representing an "erroneous" vector resulting after image registration, 2) a green arrow representing the "correct" vector. Even though there is a large distance between the end-points of two arrows for voxel A, its mapped dose differs slightly. On the other hand, the distance between the end-points for voxel B is small (below "accepted thresholds"), but the mapped dose differs considerably.

caused by treatment and time. Alternatively, smaller changes can be expected when registering images taken minutes apart, such as intra-fraction images [194].

For variations during treatment, we distinguish between inter- [93] and intra-fraction changes [91], [133]. These include disappearance and appearance of tissue and content (e.g. bladder/rectum filling [57], [196]), and tissue changes (e.g. tumour regression/growth, weight loss) [81], [221], [222]. The changes vary from small (millimetres) to large (centimetres) [118], [164], [221] (Fig. 10 a). The range varies depending on several factors like tumour site, and healthy surrounding tissues [36], [154]. Furthermore, anatomical variations occur in all directions often leading to complex deformations [24], [193], [268]. These complex changes inherently challenge the assumption of 1-to-1 anatomical mapping made by most DIR algorithms.

For inter-patient dose mapping applications, anatomical variations have a different meaning, representing anatomical differences between individuals rather than changes over time. Inter-patient dose mapping is used to explore local association of doses and outcomes [238], or assess biological effects [219]. The magnitude of inter-patient differences tends to be larger than most intra-patient changes. Moreover, "corresponding" anatomy is not well defined, as there is not a strict 1-to-1 anatomical mapping due to the natural variability between patients.

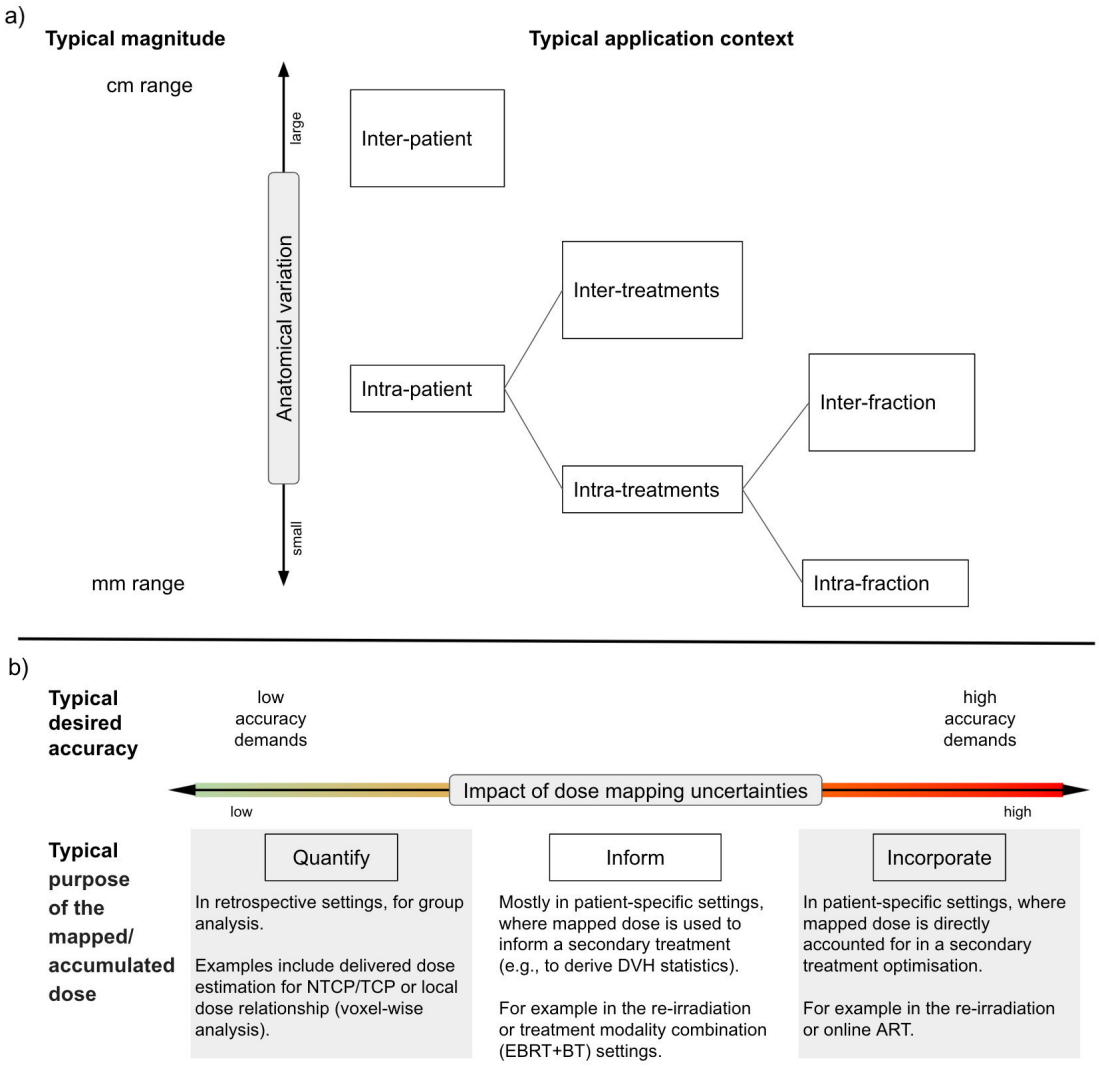

Figure 10: Illustration of the DMAL axes. a) Expected anatomical variations. b) Impact of dose mapping uncertainties on patient safety.

Impact of uncertainties

The impact of uncertainties depends on the intended use of the mapped dose, with increasing impact to applications close to a particular patient's pathway, Fig. 10 b. For this axis, we propose three main levels with no strict borders. With low impact for patient's safety, we considered the use cases where mapped doses are used to quantify, like delivered doses in retrospective studies [25], or dose variations across large groups [183], [216]. With an intermediate impact, we considered applications where the mapped/accumulated dose is used to inform by estimating global statistics. For example, when dose from an initial treatment is mapped to a subsequent image for re-treatment, and only DVH statistics are used to inform dose constraints in a re-irradiation plan [8]. Last, with the highest impact for patient's safety, we consider cases when the (voxel-by-voxel) mapped doses are incorporated to change a patient's treatment. For instance, online adaptive RT (ART), when a fraction is adapted on the spot using an estimation of the delivered dose. Uncertainties in the mapped dose due to inelastic tumour regression (visible tumour regression with the healthy surrounding tissue staying in place [222]), may result in underdosing invisible disease which could increase the risk of local recurrence. These are some of the reasons for cautious clinical integration of this tool.

Current landscape

The dose mapping/accumulation landscape (DMAL), Fig. 11, is populated with relevant example use cases that highlight how these dimensions vary for different use cases. A selection of three cases are discussed in Supplement 9.1.1.

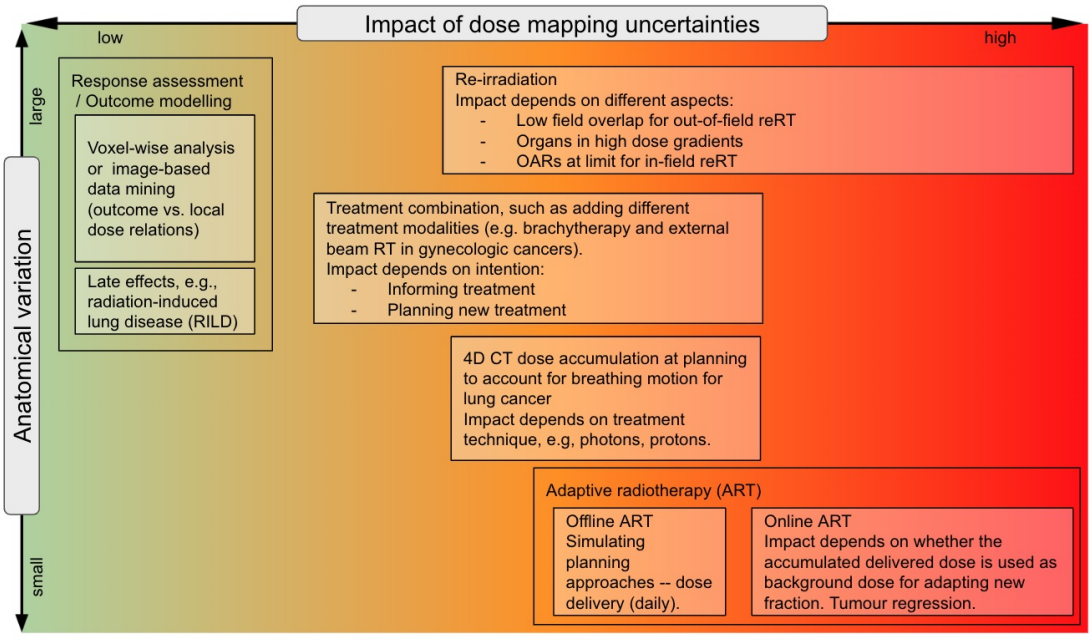

Figure 11: DMAL, presenting the current landscape of use cases. The span of each box represents the typical ranges in the anatomical variation and expected impact of dose mapping uncertainties for a given use case.

2.1.5 Identification of current challenges

Nine participants from the original workshop met five times between January and April 2022. During these meetings, we discussed articles highlighting critical issues and current challenges in implementation and clinical adoption of dose mapping, Supplement 9.1.2. Due to space constraints, we present our considerations for four issues and challenges in detail.

We hypothesise that there are first- and second-order effects on the uncertainties of DMA. The first-order effects relate to the impact of registration uncertainties on dose mapping. Mentioned earlier, it is often impossible for the registration result to "accurately" map the anatomy between two images. We argue that considering what is an appropriate registration and the impact of uncertainties for a particular application is more relevant than discussing its absolute "accuracy", as we lack a "ground truth" mapping with which to validate the accuracy. Second-order effects, including issues such as energy/mass transfer vs dose mapping, resampling/interpolation, etc., also contribute to dose mapping uncertainties, but with a smaller impact. For dose accumulation, biological uncertainties such as using a given α/β value for EQD2 or BED, or the validity of the LQ model, are relevant.

Moreover, uncertainties could impact the applications in a systematic or random manner. For example, when a prior treatment's dose distribution is transferred onto a subsequent re-irradiation scan any mapping uncertainty will have a systematic impact on the second treatment. Conversely, uncertainties associated with accumulating the dose of a fractionated treatment can typically be considered random.

First-order effects
Evaluating the correctness of a registration is extremely challenging as ground truth is commonly unavailable. Phantoms or biomechanical models are proposed to quantify registration errors at commissioning [34], [272]. For patient-specific applications, tools still need to be developed and used in daily practice [181]. Alternatively, methods to quantify dose mapping uncertainty as maps [158], [204] are available, section Quality Assurance. However, dose uncertainty estimations only assess consistency. Determining whether the registration is correct goes beyond quantifying consistency. Attempts to study this aspect with deformable phantoms have been reported [265]. Auxiliary structures/landmarks can also be used to determine the degree of accuracy, as suggested in TG-132 [34] or dense landmark clouds, as suggested by Paganelli [181]. However, the validity of these metrics is limited to the region where these contours/points are defined. Therefore, this is still an open issue which requires careful consideration when applying DIR in daily clinical applications.

As shown in Fig. 9, the registration uncertainties interplay with the dose distribution characteristics, particularly dose gradients [203], [238]. The theoretical impact of registration uncertainties to map a dose distribution can be quantified using the distance to dose difference (DTD) [203]. DTD indicates the local admissible registration uncertainty that would keep dose mapping uncertainty below a given tolerance. They found

that registration uncertainties of maximum $1\,\mathrm{mm}$ would be allowed in regions of steep dose gradients, while uncertainties of $> 20\,\mathrm{mm}$ are acceptable in regions of low dose gradients, for an IMRT lung dose distribution using a $5\,\%$ of the prescribed dose as tolerance.

Generally, DIR is performed for mono- (such as CT-CT) or multi-modal (CT-CBCT or CT-MR, etc.) images. In multimodal DIR, the differences in the images pose an additional challenge. For example, due to CBCT noise, artefacts, reduced image quality and limited field of view/length, impairs CT-CBCT DIR [15], [205]. Challenges in MR-CT DIR occur e.g. by MRI of the lungs which show less structure than CT [263] or by CT images of the prostate which do not show clear boundaries of the organ [273].

Mentioned earlier, there is often no true 1-to-1 mapping between images, especially for (dis)appearance of tissue or anatomies from different individuals. However, the vast majority of registration algorithms attempt to obtain a 1-to-1 mapping between the images (especially diffeomorphic registration algorithms) [10], [11], [239]. Even though this may not be a good representation of the true mapping between the images, it is recommended to use transformations that represent a 1-to-1 mapping when resampling dose (positive Jacobian determinant).

Finally, RT accounts for estimates of accuracy at each stage, including dose distributions calculated on grids of $1 - 3\,\mathrm{mm}$ and machines that are accurate to within $2\,\%$ or $2\,\mathrm{mm}$. A first goal should be to aim for DIR uncertainties that align with the estimates of the other steps, so that DIR-based DMA is not an outlier in the overall process [169].

Second-order effects

Resampling and interpolation

There has been some discussion in the literature on how dose distributions (represented as images) should be resampled when deforming them with a transformation. Two distinct approaches have been proposed: "direct dose mapping" (DDM) and "energy/mass transfer" (EMT) [45], [136]. DDM directly resamples the dose, whereas EMT first converts the dose into energy and mass, resamples the energy and mass separately, and then calculates the dose from the resampled energy and mass. Some publications claimed or implied that the EMT method is based on more sound physical principles than the DDM method, and therefore gives more correct results [45], [136]. We argue that this is based on a misconception that voxels in an image represent discrete anatomical units rather than a discrete sampling of an underlying continuous function. Better understanding and rigorous application of sampling theory would help demonstrate that if the dose, energy, and mass images are resampled correctly, then the same results should be achieved regardless of whether the dose is first converted into energy and mass or not. Therefore, the differences between DDM and EMT comes down to how the resampling is implemented.

Differences between results of the DDM and EMT methods presented in the litera-

ture are due to differences in the way the dose images and energy/mass images are interpolated, with neither of the methods following what is commonly considered as best practice for resampling images. We will now highlight some of the main sources of confusion and issues that should be considered when resampling a dose/energy/mass image using a 1D example, Fig. 12. Then, we will provide recommendations on how to best resample dose distributions. For a more detailed discussion of these issues and others related to image resampling see [39], [54], [230], [231].

One essential difference between the DDM and EMT methods, as presented in the literature, is the use of "pull" vs "push" resampling or interpolation, Fig. 8 c. DDM uses pull interpolation whereas EMT uses push interpolation. In general, pull interpolation is preferred when resampling images as the interpolation step occurs on a regular grid, whereas push interpolation requires interpolating scattered data points. Depending on the method used, scattered data interpolation can lead to undulating artefacts and even "holes" in the resampled image (shown in the EMT method [136] exemplified in Fig. 12 c) or relies on methods that are considerably more computationally demanding. This is the reason that all registration algorithms use pull interpolation in their internal optimisation (to the best of our knowledge).

Push interpolation has been used for EMT to ensure that the overall energy/mass was conserved during resampling. Notice that mass can also be preserved when using pull interpolation by multiplying the resampled image by the local volume change (Jacobian determinant), as proposed for mass-preserving registration algorithms [232], [266]. Note, such mass-preserving registration algorithms should not be used for most dose resampling applications, as the assumption of mass preservation between the images is not valid (for example due to (dis)appearing tissue or different patients). Our recommendation is to always use pull interpolation when resampling either dose, energy or mass while accounting for local volume changes.

Another cause of the differences between DDM and EMT results in the literature is that the mappings are required in the opposite directions, so separate registrations are performed for each method (swapping the fixed and floating images, Fig. 8 c). Inconsistencies between the registration results used for each method (unless an inverse consistent registration is used, which was not the case in [136]), further contribute to the differences in the results observed between the two methods.

On a more technical side, the choice of interpolation method or kernel is key when resampling an image. Linear interpolation is the most well-known and widely used interpolation method. It is very fast but it introduces a small amount of blurring to the resampled image. Other interpolation methods, such as cubic convolution, windowed-sinc (with Lanczos, Blackman or Welch kernels) or spline-based methods, introduce less blurring but can lead to ringing, where the maximum/minimum values in the resampled images are larger/smaller than the maximum/minimum values in the original image [152], [230], [231], potentially introducing negative dose/energy/mass values. The blurring introduced from a single application of linear interpolation is small and unlikely to have a large impact, however, if the image is resampled multiple times the

blurring will accumulate and cause noticeable degradation of the resampled images. This effect will have impacted the evaluation of the DDM method presented in [136] where the dose was resampled 3 times (first resampled onto the source CT image grid, then onto the target CT image grid, and finally resampled onto the target dose image grid) introducing unnecessary blurring to the result. Therefore, we strongly recommend avoiding resampling an image multiple times, instead composing the transformations and only resampling with the final composed transformation.

The final issue we want to highlight is aliasing. This occurs when the original image contains higher frequency information than can be represented in the resampled image, either because the resampled image has a lower resolution or because the transformation causes parts of the image to be compressed. When aliasing occurs the high frequency information incorrectly appears as lower frequencies, which in practice can lead to structures arbitrarily appearing brighter or darker than they should in the resampled image [230], [231]. Therefore, we would recommend using a spatially varying interpolation kernel, as described in [39] when the transformation contains regions with large compressions (by a factor of 2 or more) within regions of a high dose gradient.

Quality assurance considerations

Quality assurance (QA), defined as the procedures and processes followed to ensure that the quality of each dose mapping is maintained, is essential in clinical practice. However, this is difficult to implement due to unknown ground truth and uncertainties that arise from patient-specific characteristics (for example, cervix-uterus interfractional changes [144]). In addition, dose accumulation QA consists of several essential steps: ensuring the appropriate mapping/accumulation workflow was followed, QA of DIR, QA of dose mapping, and finally a review of all steps [34]. Determining the best workflow for DMA is not trivial, and to the best of the authors' knowledge, there are no resources available for this specific matter. We define mapping/accumulation workflow as the selection of a given algorithm, its parameters, directionality, resampling strategy, and the definition of the minimum set of metrics and/or procedures for QA. Prior to defining a workflow, we propose to clearly define 1) what is the intended use of the DMA, and 2) which is the region of interest (ROI), which will then inform the expected anatomical variation. With these aspects, it is easier to identify the potential impact of dose uncertainty, such as illustrated in our DMAL, Fig. 11. Particularly, consider defining how to handle extreme anatomical variations (such as missing organs in the re-irradiation settings, or completely different anatomy as in inter-patient registrations).

Next, it is needed to assess whether DIR has fulfilled its task. Ideally, DIR would have aligned all the corresponding anatomy between the images, properly accounting for missing tissue (or organs). This could be translated to level 0 in the registration uncertainty assessment levels proposed in the TG-132 [34]. However, current registration algorithms often performed less optimal than this, with a recent publication comparing commercial systems reporting mean target registration error (TRE) ranging between 2.8 and 6.8 mm [198]. For DMA, locally aligned DIR, focused on the ROI (and/or the regions containing dose) would suffice (level 1, TG-132 [34]). However, these levels rely

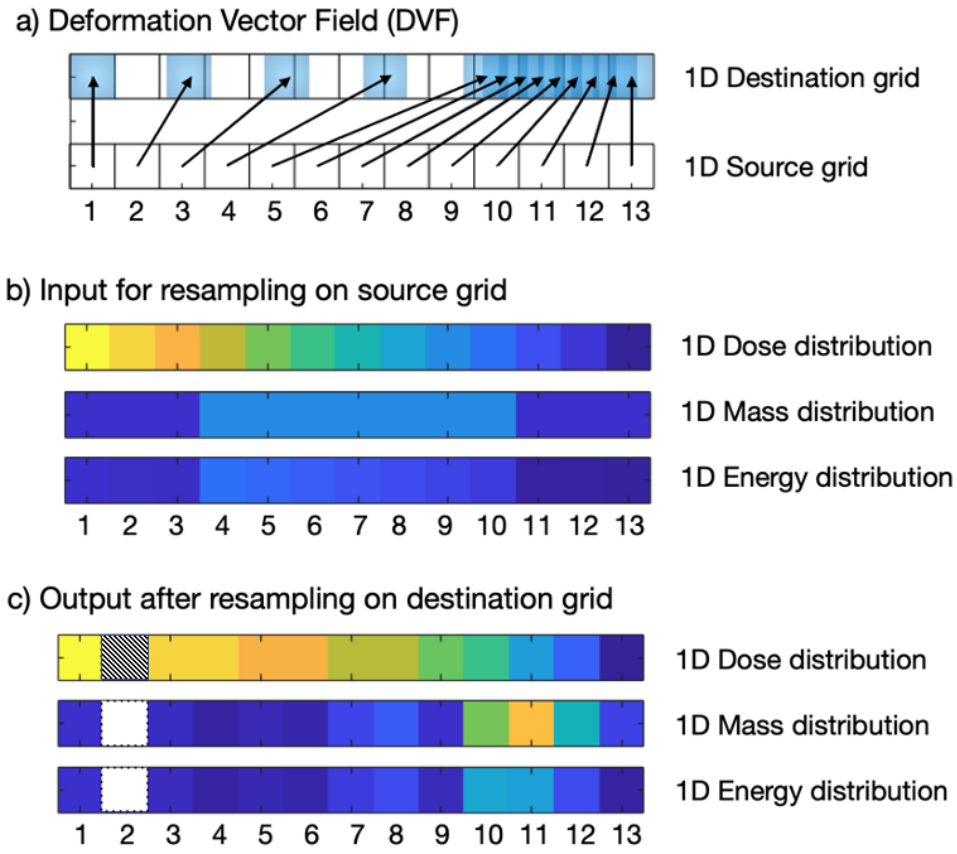

Figure 12: 1D scheme highlighting the main problems of using "energy/mass transfer" resulting from naive use of push interpolation. a) The DVF used to push the data representing a constant expansion over 1/3 of the image and a constant compression over the other 2/3, with the 2nd pixel having no mapped pixels. b) The input data. c) The resampled data using the EMT method described in [136]. For each voxel in the dose distributions shown in b) and c), the value corresponds to energy divided by mass. Note the 2nd pixel of the resampled energy and mass distributions corresponds to a "hole" (value of 0), which results in an undefined mapped dose value (hatched pixel). Additionally, the energy and mass distributions contain undulations that would not be expected from the DVF (as it represents constant expansion/compression) and the dose has an undesirable "step-like" appearance (pixel 3 vs 4 having the same value, same for 5 vs 6 and 7 vs 8).

on quantifying the alignment of the anatomy, assumed to be defined on both images, using metrics such as TRE, mean distance to agreement, dice similarity coefficient (DSC). This presents a challenge as these anatomical landmarks/contours are often not available for one or both images. The validity of these geometrical metrics is limited to their direct vicinity, and cannot account for distortions far from their location [112]. DVF-based metrics (Jacobian determinant, inverse consistency error (ICE) and transitivity error (TE)) provide information in all locations where the DVF is defined. However, they don't assess registration correctness but local volume changes (Jacobian) and consistency (ICE and TE), therefore they are insufficient on their own for QA [181]. We strongly recommend visual inspection of the deformed image and the transformation (deformed grid/DVF visualisation), focused on the ROI, as a strict minimum to provide a global assessment of DIR performance.

Testing the selected algorithm and parameters to phantoms (either physical or digital) also provides insight, as done at commissioning in clinical applications. Several phantoms have been developed and made available for the community. These include known ground truth features, such as landmarks, DVF and dose distribution [34]. For digital phantoms, a current limitation is that they are often provided in file formats other than DICOM, and treatment planning systems are relatively closed to import data beyond DICOM. Converting between file formats requires tools and custom scripts that may not be available on clinical systems, increases workload and potential for making mistakes, and hinders clinical adoption.

Next comes QA dose mapping itself. Where corresponding anatomical landmarks are visible in each image, the dose to the landmark can be sampled from both the original and mapped dose distributions. With accurate registrations, these doses should be equal. Assessing the discrepancies between their TRE and dose deviations may give an indication of the cause of the uncertainty, whether due to spatial registration inaccuracy or a second-order effect. The confidence in the landmark identification is a limitation. Furthermore, this assessment is valid in the landmarks' proximity.

Other strategies were proposed to assess the impact of inverse (in)consistent registrations, for instance the DVH overlap method [106]. Using the terminology introduced in Fig. 8, two transformations are used: source-to-destination and its inverse, destination-to-source. Structures defined on the destination grid are propagated onto the source dose distribution, and the source dose distribution is mapped to the destination grid. The impact of inconsistencies in the registration within the structures is assessed by comparing the derived dose volume histograms (DVH) of dose/contour sets. If the volume of the structure is conserved, the mapped DVH should be equal to the original DVH and imply minimal inverse consistency error (ICE). For this method, structures must be present in the destination image and the quantification would only be valid for the region delineated. Moreover, variations between original and mapped DVHs will depend on the structure volume; with small mapping errors exacerbated in small volume structures, while larger mapping errors may be masked in large volume structures. Another strategy used the inconsistencies of registration defined as the net displacement of every voxel in different structures after successive application of the forward

and backward transformations, and summarised differences using DVH bands [141]. However, these methods in isolation do not provide enough information for QA.

Methods were proposed to quantify dose mapping uncertainty in a local scale, by generating uncertainty maps [92], [158], [204]. In general, these maps contain the standard deviations of different mapped doses, calculated on each voxel. Salguero and colleagues [204] proposed a general framework to estimate voxel-wise dose mapping uncertainty: First, a cluster of points is obtained for each voxel from an iterative DIR, where each iteration included an "artificial" perturbation in the registration. Next, the dispersion of these points is used to compute spatial uncertainty. Last, the spatial uncertainty estimates are used in combination with the mapped dose distribution to compute the point-by-point dose standard deviation. Another strategy relies on a deformation model created using principal component analysis to sample spatially-correlated uncertainties and quantify their impact in daily dose mapping [158]. Hub and colleagues [92] estimated dose mapping uncertainties based on varying b-spline coefficients. Dose mapping uncertainties maps allow to keep spatial components of uncertain regions, which can help decide whether the mapped dose is of use. Therefore, we recommend quantifying the impact of registration uncertainties on dose mapping after a global assessment of DIR performance using one of these strategies [92], [158], [204]. In this context, a clear challenge is the lack of software tools facilitating the quantification and visualisation of these uncertainty maps in clinical practice.

Finally, QA carried out with developer methods and tools cannot replace the assessment of the user. Therefore, verification of DIR and DMA results by the user is essential.

Radiobiological issues

An important point raised during workshop discussions is that dose is at the end of the day a surrogate for what we really want to quantify, that is the radiobiological effect of the treatment. As outlined by Jaffray and colleagues [98] in their QUANTEC vision paper, the delivered dose has been poorly understood. Recent advances in DIR are allowing the delivered dose to be more accurately defined. Now that the delivered physical dose can be investigated, how do we incorporate this information into our understanding of radiobiology? A good example of a study linking radiobiological endpoints to the delivered - not planned dose - was performed by Bohoudi and colleagues [25]. They sought to identify delivered dose parameters linked with bladder toxicity. Dose was accumulated for 101 prostate SBRT patients treated with ART. The accumulated bladder V20-32 Gy showed better correlation than the planned V20-32 Gy with an increase in International Prostate Symptom Score. However, there were large bladder volume variations, potentially impacting the accuracy of the DIR.

Estimating biological accumulated dose is not straight forward. The Linear quadratic model (LQM) assumes fractional doses of equal magnitude used in a power-to-n law. Given the dose, especially to OARs, can vary each treatment, it is not valid to accumulate linearly using DIR. The concept of total biological dose, $bEQD_d$, has been introduced by Niebuhr et al. [171] to address this issue. Briefly, the $bEQD_d$, represents the total treatment dose that yields a given biological effect but takes each dose

per fraction into account rather than an average. Niebuhr et al. report that the $bEQD_d$ was systematically higher than conventionally accumulated dose with differences in hot spots of 3.3 - 4.9 Gy for conventional and 8.4 Gy for hypofractionated prostate cancer treatment plans. Determining the impact of these differences for outcome modelling and adaptive strategies is still unclear.

2.1.6 Discussion

In this manuscript, we presented the current landscape of dose mapping/accumulation in RT, visually via the DMAL, Fig. 11. DMAL connects the expected anatomical variations and the impact of dose mapping uncertainties for patient safety and can be used in the future for analysis and safety considerations of new use cases. Unlike the magnitude of anatomical variations, metrics to quantify the impact of uncertainties on patient safety are lacking. Open discussions in a multi-disciplinary team are essential to define the position of any new use case. In our approach, we identified three main levels in the impact of uncertainties on patient safety, which are related to how the mapped/accumulated dose is used in the use cases. For use cases in the right side of the DMAL, a full quantification of dose mapping uncertainties (including both correctness and consistency) is required. However, tools to quantify these are not available in commercial packages. Therefore, we recommend extreme caution, with a very limited use in clinical practice. Vendors are encouraged to develop and implement tools to streamline this quantification soon, and ideally incorporate these uncertainties in treatment plan optimisation.

We also presented considerations on the current challenges in DMA, going beyond DIR uncertainties, which is the main focus of published literature [45], [258]. We argue that the discussion on energy/mass transfer vs direct dose mapping is irrelevant, since, if done correctly, the results should not differ. This is at the end of the day, an implementation choice. We recommend using DDM in DMA implementations. However, we acknowledge that this decision is not in the hands of most users. Thus, it is further recommended transparent communication from the vendors on the selected strategy implemented in their clinical software to raise awareness of the limitation and possible impact of these for the users.

To get a comprehensive understanding of a DMA system, we recommend following four QA steps: ensuring the appropriate mapping/accumulation workflow was followed, QA of the DIR result, QA of the DMA result, and finally a review of the impact the DMA uncertainties will have on the clinical application. Implementing a DMA workflow involves many inter-connected tasks and decisions, and requires a well-coordinated team including dosimetrists, physicists, and radiation oncologists each playing different roles to ensure patient safety. We discussed the limited use and availability of digital phantoms for clinical systems commissioning, especially as non-standard file formats hinder their use clinically. Moreover, specific phantoms to enable commissioning, and automatic identification of corresponding landmarks would be of use for dose mapping QA.

The last major issue discussed in our manuscript was the impact of radiobiological uncertainties in dose accumulation. Dose, as a surrogate of the radiobiological effect of the treatment, is poorly understood. However, with DMA tools, we can start gaining insights on the true correlation between delivered dose and treatment outcomes. Furthermore, converting to radiobiologically-corrected doses (EQD2 or BED) relies on the use of the LQM, which can be an oversimplification of the real biological effect of dose in different tissues. Even with their shortcomings, we recommend using biologically corrected doses when accumulating doses.

Advanced registration techniques including diffeomorphic, symmetric, and inverse consistent algorithms, have been available in research software for years [11], and continue to be further developed [132], but few commercial packages include these features, limiting their use in clinical practice. There are also still several open research questions, such as how to best account for (dis)appearing tissue and sliding motion, process longitudinal data, and utilise recent advances in learning-based approaches [132]. The challenge of producing better registrations for DMA requires understanding exactly how the registration will be used to map/accumulate dose, and how the results will be interpreted and utilised in the clinical workflow. Multi-disciplinary work, where physicists, radiobiologist, clinicians, and computer scientists/engineers collaborate closely together is key to make real progress with dose mapping addressing the issues raised in this paper and beyond.

Several aspects cannot be improved by algorithms. As mentioned before, dose mapping relies on the registration of the underlying anatomy captured in medical images. Therefore, dose mapping applications will be hindered by any limitations in image acquisition and reconstruction. Additionally, complex anatomical changes such as inelastic tumour regression happening during treatment [5], [82] will completely mislead any intensity-based registration and may result in highly uncertain dose mapping. Even though methods were proposed to identify these regression modes [5], these have not yet been actively incorporated in DMA. Mitigation strategies are required in the meantime; again, quantification of dose mapping uncertainties is a must.

Our last take-home message is to strive for a context-driven DIR rather than a perfect DIR. This may mean sacrificing global registration accuracy to favour locally accurate registrations or have multiple registrations depending on the organ/application. Here, it is important to be aware of the limitations of the implemented DIR algorithm, which is essential to understanding the impact of DIR and dose mapping uncertainties in the context of clinical decisions and judgements. For some use cases, keep in mind that dose mapping (with properly quantified uncertainties) is a better alternative than no mapping.

Recommendations for vendors

- Develop and implement tools and visualisation means, which can be run quickly and easily after every registration in the system, aiming at:
  - Analysing the resulting transformations. Whenever contour/landmarks are

available (or created via automatic segmentation), we recommend including distance metrics, as recommended by TG-132, but warning the user on their limited validity. We also recommend visualising the DVF, as well as generating the map of the determinant of the Jacobian (to highlight contractions, expansions and registration folds). Additionally, visualisation of the inverse-consistency metric, bending energy, harmonic energy, curl would be desirable.

  - Identifying dose gradients to highlight the regions of high dose mapping uncertainty. We recommend overlaying this on the images as well as on the DVF/Jacobian maps.
  - Applying multiple algorithms (or parameters for the same algorithm) to provide a range of plausible registrations and estimate dose mapping uncertainty from these. Visualization can be done as confidence bands around individual DVHs or uncertainty maps (such as proposed in [92], [158], [204]). This is key to enabling QA.

- Integrate tools for performing and evaluating dose mapping with other clinical software such as auto-contouring and treatment planning systems to provide a seamless and automated clinical workflow for adaptive radiotherapy.
- Allow the user to export registration results (including initial rigid/affine and subsequent DVFs) for external evaluation and/or comparisons. This would ideally be in a standardised, well-documented, and easy to read format. For the DVF, we recommend using standard file formats, such as the DICOM extension for Deformable Registration in Radiation Oncology (DRRO) proposed by the IHE Radiation Oncology Technical Framework [47].
- Generate clear descriptions of the workflow applied, to allow traceability for mapped doses.
- Include in all training/tutorials resources clear indications of the limitations of the registration algorithm (e.g., inability of registering disappearing tissues) and dose mapping strategy implemented (e.g., direction of registration and resampling strategy) and their impact in a variety of cases, including "simple" and challenging cases, such that users can easily identify possible shortcomings during normal operation.
- Allow the user to specify regions of interest where the registration should be as accurate as possible.
- Implement built-in workflows for resampling dose following the advice in this paper.
- Benchmark tools and algorithms on public data sets and publish results.
- Implement state-of-the-art DIR algorithms into commercial products, especially if their aim is dose mapping/accumulation. Examples include diffeomorphic, inverse consistent, symmetric and sliding motion.

Recommendations for users

- Clearly identify the purpose of the DMA, and relate it to the DMAL, to determine the potential impact of the dose mapping uncertainties. This should be done both, at commissioning and at QA time.
- Consider local dose mapping uncertainties, ideally quantified with tools provided by vendors, and the impact this uncertainty will have on the specific application. Getting the appropriate registration for the application/organ/region should be the first priority, which may mean that you sacrifice global registration accuracy, or have multiple registrations. Always, consider which alternative is better for the patient: uncertain dose mapping or not mapping the dose at all (e.g., using DVH statistics) in the context of extra workload.
- Be aware of the software limitations and how to assess uncertainties, both for DIR and for dose mapping. For this we recommend training to develop a clear and comprehensive understanding on:
  - The DIR software used in their practice, including limitations, and implementation decisions,
  - The evaluation of DIR, both qualitatively and quantitatively, [34]
  - The dose mapping workflow used in their clinical software, including highlighting the regions of high impact due to uncertainties such as regions of high dose gradient
  - The evaluation of DMA uncertainties, ideally with tools provided by vendors,
  - The proper use of tools provided by vendors for QA for DIR and DMA, and their importance in daily practice.
- If feasible, apply multiple algorithms (or multiple feasible parameters) for the same task to provide a range of plausible registrations, from which a dose mapping uncertainty measure can be derived for every registration run which can have direct patient impact.
- Document clearly the registration workflow followed to map/accumulate dose distributions for each patient and the QA results.
- Develop departmental procedures and policies to perform DMA consistently, ideally following international recommendations/guidelines. Important aspects to account for include achievable and required accuracies for use cases in different anatomical sites (keeping patient safety in mind), alternative approaches when the registration is not successful or the dose mapping is highly uncertain, and how to handle tissue not present in both images.

2.1.7 Supplementary Material

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

2.2 A multi-institutional comparison of retrospective deformable dose accumulation for online adaptive MR-guided radiotherapy

Autor: **Martina Murr**, Uffe Bernchou, Edyta Bubula-Rehm, Mark Ruschin, Parisa Sadeghi, Peter Voet, Jeff D. Winter, Jinzhong Yang, Eyesha Younus, Cornel Zachiu, Yao Zhao, Hualiang Zhong, Daniela Thorwarth

Published in: Physics and Imaging in Radiation Oncology 2024, Vol. 30, Pages 100588

DOI: https://doi.org/10.1016/j.phro.2024.100588

2.2.1 Abstract

Background and Purpose: Application of different deformable dose accumulation (DDA) solutions makes institutional comparisons after online-adaptive magnetic resonance-guided radiotherapy (OA-MRgRT) challenging. The aim of this multi-institutional study was to analyze accuracy and agreement of DDA-implementations in OA-MRgRT.

Material and Methods: One gold standard (GS) case deformed with a biomechanical-model and five clinical cases consisting of prostate (2x), cervix, liver, and lymph node cancer, treated with OA-MRgRT, were analyzed. Six centers conducted DDA using institutional implementations. Deformable image registration (DIR) and DDA results were compared using the contour metrics Dice Similarity Coefficient (DSC), surface-DSC, Hausdorff-distance (HD95%), and accumulated dose-volume histograms (DVHs) analyzed via intraclass correlation coefficient (ICC) and clinical dosimetric criteria (CDC).

Results: For the GS, median DDA errors ranged from 0.0 to $2.8\,\mathrm{Gy}$ across contours and implementations. DIR of clinical cases resulted in DSC > 0.8 for up to $81.3\,\%$ of contours and a variability of surface-DSC values depending on the implementation. Maximum HD95% = $73.3\,\mathrm{mm}$ was found for duodenum in the liver case. Although DVH ICC > 0.90 was found after DDA for all but two contours, relevant absolute CDC differences were observed in clinical cases: Prostate I/II showed maximum differences in bladder V28Gy ($10.2/7.6\,\%$), while for cervix, liver, and lymph node the highest differences were found for rectum D2cm^3 ($2.8\,\mathrm{Gy}$), duodenum Dmax ($7.1\,\mathrm{Gy}$), and rectum D0.5cm^3 ($4.6\,\mathrm{Gy}$).

Conclusion: Overall, high agreement was found between the different DIR and DDA implementations. Case- and algorithm-dependent differences were observed, leading to potentially clinically relevant results. Larger studies are needed to define future DDA-guidelines.

2.2.2 Introduction

Online magnetic resonance-guided radiotherapy (MRgRT) enables optimal plan adaptation concerning the patient's daily anatomy, allowing improved target coverage and organ at risk (OAR) sparing [247]. Deformable dose accumulation (DDA) is a method to evaluate doses from several treatment fractions deformed according to daily images to account for anatomical changes during treatment [45]. According to several recent studies, DDA promises precise adaption of delivered organ-specific doses to reduce OAR toxicity [4], [25], [42], more precisely determined dose tolerances [191], and/or additional target dose escalation [4]. However, DDA still requires investigation of dose mapping uncertainties and anatomical or contouring variations [160].

Deformable image registration (DIR) is performed as first step of DDA. Several DIR algorithms, categorized by matching criteria and deformation model, are available for clinical and research use, like intensity or hybrid methods and free-form deformations or diffusion/deformation models [223]. As recently published, underlying algorithms itself typically works differently in various anatomical regions [68], [105], [168], [188]. Additionally, different algorithms applied to one anatomical region can lead to different results [105], [157], and algorithm settings such as different configuration values or optimization methods may change registration results [275].

The deformation vector field (DVF) generated by DIR is applied to the corresponding 3D dose map in the second step of DDA. In literature, several resampling and interpolation approaches are discussed, especially direct dose mapping (DDM) and energy mass transfer (EMT) [45], [136]. DDA algorithms are currently being applied using various implementations across institutions, which may provide different results for the same problem, as several studies show [27], [241]. For MR-guided liver stereotactic body radiation (SBRT), Wahlstedt et al. [241] observed that DDA algorithms are highly patient- and fraction-dependent and recommend using numerous algorithms in tandem. Bosma et al. [27] found that in MRgRT of prostate cancer, EMT led to lower dose errors than DDM.

This multi-institutional investigation aimed to assess several different implementations for DDA utilizing the same datasets including a gold standard (GS) case deformed using a biomechanical model and five clinical cases. The goal was to identify differences in results based on the algorithms used.

2.2.3 Material and Methods

The multicenter DDA study was conducted by a working group of the Elekta MR-Linac consortium involving six institutions with clinical expertise in MRgRT using the 1.5 T MR-Linac (Unity, Elekta AB, Sweden).

Data characteristics

Initially, DDA was conducted on a prostate case where deformations were created by biomechanical simulations. This process included loading clinical contours into finite

element modeling software (FEBioStudio v1.7.1) [143], followed by assigning the physical attributes to each anatomical tissue [27]. Simulations of five clinically representative bladder and rectal fillings were used to actuate surrounding tissue and apply displacements and deformations to a T2-weighted MR image (MRI) (Supplementary Figure S2). For each simulated anatomical situation, RT plans were created by recalculating the dose on the new anatomy, resulting in five total dose distributions. The known underlying deformations enabled the calculation of a GS accumulated dose (GS-DDA) using DDM.

Subsequently, five clinical datasets of patients with different tumor entities were analyzed, including (I, II) two prostate, one (III) cervix, (IV) liver, and (V) lymph node cases, as detailed in Supplementary Table S2 and Fig. S3. Patients were treated at one participating institute with SBRT on the 1.5 T MR-Linac using the "adapt-to-shape" (ATS) workflow [247]. Experienced radiation oncologists re-contoured the MRI for each fraction offline to address minor issues eventually caused by the tight timeline for online contouring. Subsequently, wall contours for rectum, colon, and duodenum, likewise for bladder and stomach, were created using negative margins of 3 and 4 mm to investigate DIR and DDA results in the hollow organs with homogeneous fillings. These fraction data sets, including MRI, re-delineated contours, and dose files, were shared using ProKnow D (Elekta AB, Sweden, V1.33.0). Clinical dosimetric criteria (CDC) used for plan approval were provided by the institution for each specific case (Supplementary Table S3). Ethical approval and data sharing agreement existed for all clinical cases.

Deformable image registration, contour propagation, and dose accumulation
Each institute performed DIR using the software available at their institute. Two institutes (A, C) had a hybrid intensity/structure-based algorithm (Monaco ADMIRE Research, Elekta, Sweden) [83] or ANACONDA (RaySearch Laboratories AB, Sweden) [244]. Mixed/hybrid and contour-guided DIR algorithms [186] (MIM Software Inc., USA, versions 6.8.5/7.0.6) were used by institutes B and D, respectively. Institute E used an in-house intensity-based DIR algorithm [71], [242]. RTTracker (UMC, Utrecht, The Netherlands) [211], [268] was used by institute F. Supplementary 9.2.1, including Table S4, describes each algorithm in more detail. Five institutions used contour-guided DIR, while one institution (E) did not (cf. Table 2). All MRIs were deformably registered to the first fraction. Each institution propagated all offline re-delineated contours of fractions 2–5 to fraction 1 using the resulting DVF.

Dose mapping was conducted by applying the DVF to the dose file of each fraction. For dose resampling, two distinct methods were used: DDM [86] (A-E) and EMT [217] (F). The software solutions Slicer3D (V4.11) [63], Raystation (8B), MIM, an in-house developed solution, and RTTracker (V4.0) were used, respectively. No additional DIR/DDA requirements were made for the institutes.

Data analysis
Each institute's accumulated dose maps for the GS case were compared with the known GS-DDA, calculating voxel-wise absolute differences within the target volume

Table 2: Overview of contours used for DIR guidance per institute (A-F) for the cases gold standard, (I) prostate 1, (II) prostate 2, (III) cervix, (IV) liver, (V) lymph node. Abbreviations; CTV: clinical target volume, GTV: gross tumor volume, CTV_{HR}: high risk CTV, CTV_{IR}: intermediate risk CTV, CTV_R: right lymph node CTV, CTV_L: left lymph node CTV.

Institute	A	B	C	D	E	F
Gold Standard (I) Prostate 1 (II) Prostate 2	Bladder, femurs, rectum	CTV, bladder, rectum	CTV, bladder, rectum	CTV, bladder, rectum, sphincter	None	CTV, GTV, bladder, rectum
(III) Cervix	Bladder, rectum, sigmoid	CTV_{HR}, bladder	Bladder, rectum, sigmoid	CTV_{HR}, CTV_{IR}, GTV, bladder, rectum, sigmoid	None	CTV_{HR}, bladder, rectum
(IV) Liver	Colon, duodenum, kidney, small bowel, spinal cord, stomach	GTV, liver	GTV, duodenum, liver, pancreas, stomach	GTV, colon, duodenum, kidney, pancreas, spinal cord, small bowel	None	GTV, duodenum, stomach
(V) Lymph node	Bladder, rectum, sigmoid	rectum	CTV_L, CTV_R, bladder, rectum	Bladder, femurs, pelvis, rectum, sacrum	None	CTV_L, CTV_R, bladder, rectum

and OARs. Median (interquartile range, IQR) were compared, as a normal distribution could not be assumed. Contour propagation of all cases was evaluated by the Dice Similarity Coefficient (DSC) [53], surface DSC (SDSC) with $2\,\mathrm{mm}$ threshold [173] and $95\,\%$ Hausdorff-distance (HD95%) [95], [198] using MATLAB (R2020b). Corresponding dose-volume-histograms (DVH) were generated to evaluate the accumulated doses. All relevant CDC derived from these DVHs were calculated and compared for all cases and institutions using ProKnow. Furthermore, the DVHs and CDCs for the wall structures were calculated.

The Kruskal-Wallis test was used to determine whether there were significantly different in the dose errors reported by the six institutes for the GS. The institutes were also compared pairwise via a Mann-Whitney test [145]. A repeated measures ANOVA was conducted to ascertain variances among the means of distinct algorithms, while a post-hoc test was employed to discern specific group disparities. These analyses were undertaken to evaluate significant differences across institutes concerning DSC and CDC metrics. The tests were conducted at a significance level of $5\,\%$.

Intraclass correlation coefficients (ICC) were computed to test the similarity of the different accumulated DVH curves, indicating the level of correlation and agreement between the DVHs determined using the various algorithms [119]. A high ICC indicates a high level of agreement between the DVHs, while a low ICC indicates a low level of agreement. ICC estimates and their $95\,\%$ confidence intervals (CI) were computed using SPSS statistical package version 23 (SPSS Inc, Chicago, IL), using a single-rating, consistency, 2-way mixed-effects model [119].

2.2.4 Results

For the GS, very good agreement was achieved by 5/6 institutes. Institute E reported higher median (IQR) dose accumulation errors for clinical target volume (CTV) and rectum, 0.2 (0.7) and 2.8 (3.7) Gy, respectively, compared to the other institutes. In contrast, the highest median errors were found for bladder by institutes C and F with 0.2 (0.8) and 0.2 (0.7) Gy, respectively, whereas institute E presented the highest IQR of $1.1\,\mathrm{Gy}$. Fig. 13 presents the full statistical distribution of the voxel-based dose differences, with further details available in Supplementary Table S5. The Kruskal-Wallis test rejected the null hypothesis for all three contours, indicating statistically significant institution differences, only institutes C and F demonstrated statistically insignificant differences within the CTV. Similar observations were obtained within the rectum for institutes A and D.

Comparing DSC results per contour and institute for the GS and clinical cases, all DSC values were above 0.8 for GS, whereas a DSC > 0.8 was met in $81.3\,\%$, $75.7\,\%$, $34.6\,\%$, $64.6\,\%$, and $68.8\,\%$ of contours by institutes A, C, D, E, and F for the clinical cases. Fig. 14 illustrates the DSCs for each patient. The ANOVA for repeated measures resulted in statistically non-significant DSC differences for GTV in case (I), and for sphincter and GTV in case (II). Similarly, in case (III), non-significant differences were found for $\mathrm{CTV_{HR}}$, $\mathrm{CTV_{IR}}$, and GTV, as well as for spinal cord in case (IV). Detailed DSC results for each case are summarized in Supplementary Tables S6 - S11. The ANOVA and

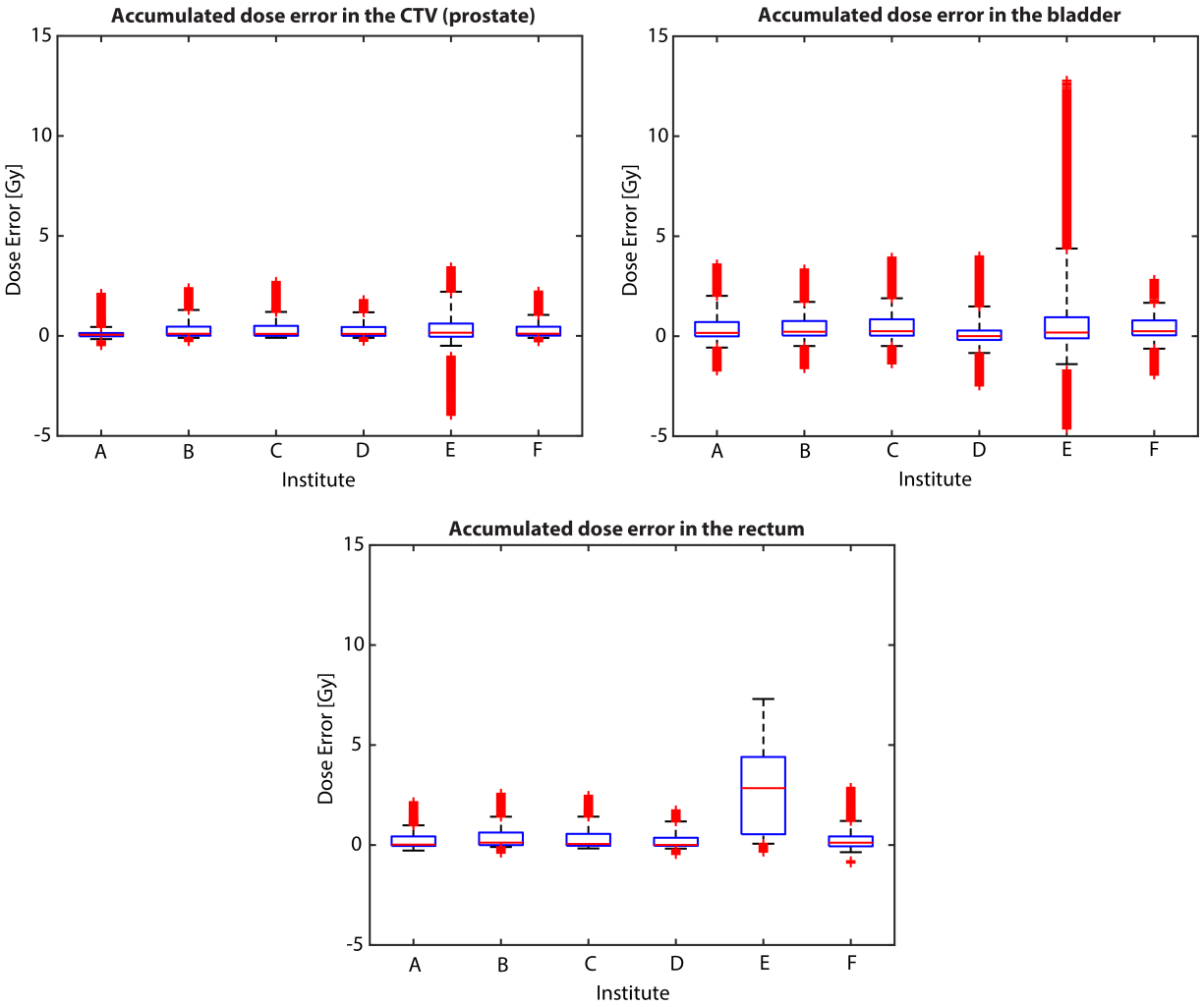

Figure 13: Statistical distribution of the accumulated dose errors for each participating institute. The illustrations are made individually for the bladder, prostate, and rectum. For each of the boxplots, the box limits correspond to the 25th and the 75th percentiles of the set, while the whiskers are the 5th and the 95th percentiles. The red line inside each box indicates the median of the set, with the red markers beyond each whisker showcasing the outliers.

pairwise post-hoc test results are given in Supplementary Table 9.2.2.The SDSC values for the GS case showed slight differences in the algorithm performances, with F consistently achieving higher SDSC values in most contours. Conversely, in the clinical cases, more pronounced disparities were noted in the SDSC values among different algorithms, notably with A, C, and E reaching higher values. For the GS, a maximum HD95% of 7.2 mm was found for femur$_L$. In contrast, maximum HD95% of 73.3 mm was observed for duodenum in clinical case (IV), for details see Supplementary Tables S6 - S11. The most challenging area seemed to be the left lymph node in case (V), where the target consisted of two independently moving lymph nodes. An in-depth motion analysis of the two lymph nodes is presented in Supplementary Fig. S4.

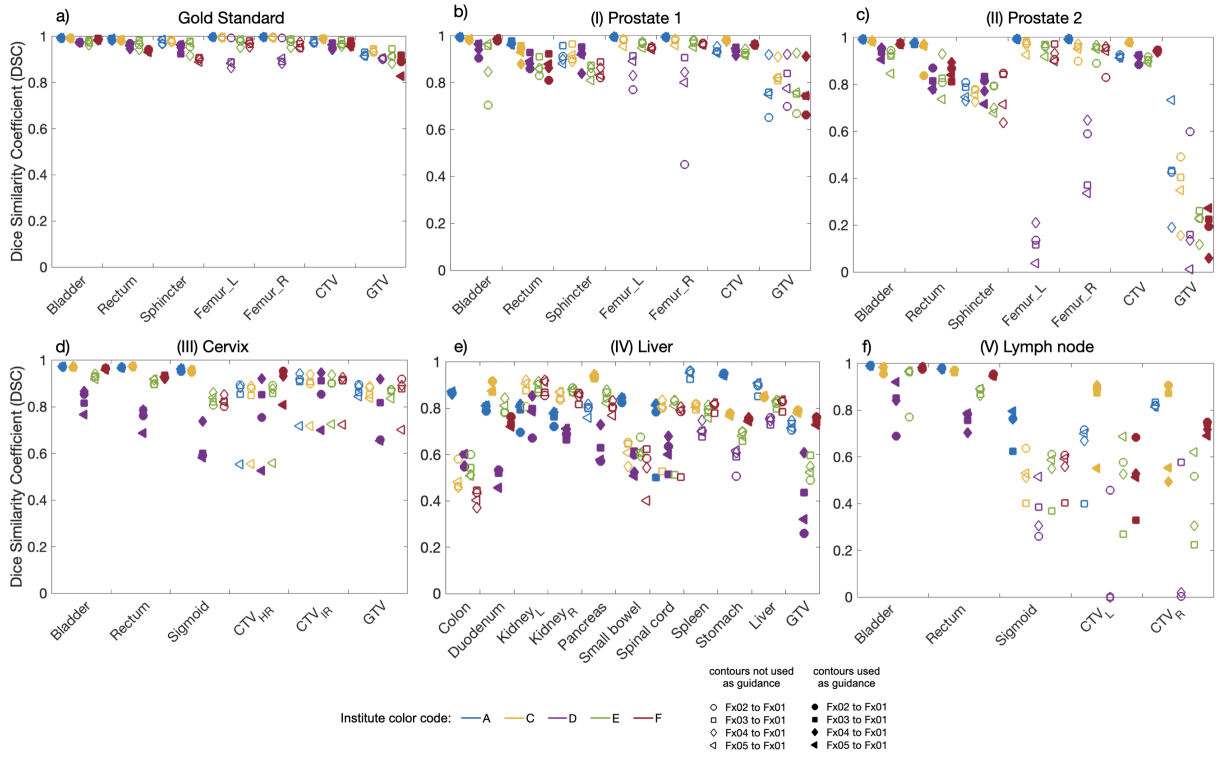

Figure 14: Illustration of the dice similarity coefficient (DSC) results for the cases gold standard (a), (I) prostate 1 (b), (II) prostate 2 (c), (III) cervix (d), (IV) liver (e), and (V) lymph node (f). Case-related deformed contours are listed on the x-axis. The results of the institutes (A, C, D, E, F) are shown in different colors, institute B did not provide deformed contours. The results of the different fractions are shown in different marker types. Abbreviations; CTV: clinical target volume, GTV: gross tumor volume, CTV$_{HR}$: High risk clinical target volume, CTV$_{IR}$: Intermediate risk clinical target volume, CTV$_R$: right lymph node, CTV$_L$: left lymph node.

In Fig. 15, DDA DVHs for every case and contour are displayed. ICC > 0.90 indicated excellent correlation for all DVHs, except for CTV$_L$ and CTV$_R$ in case (V), with ICCs (95 % CI) of 0.84 (0.84 - 0.85) and 0.72 (0.71 - 0.73), respectively. However, significant differences (ANOVA, $p < 0.05$) were found for all contour DVHs, except for CTV case (I), GTV and spinal cord case (IV), and bladder case (V). ICC, ANOVA and pairwise post-hoc test results are summarized in Supplementary Table 9.2.2. Fig. 16 visualizes the

different deformed accumulated dose distributions exemplarily for case (IV). Resulting DHVs of the deformed fraction doses are shown in Supplementary Fig. S5. The DVHs of the wall structures are shown in Supplementary Fig. S6.

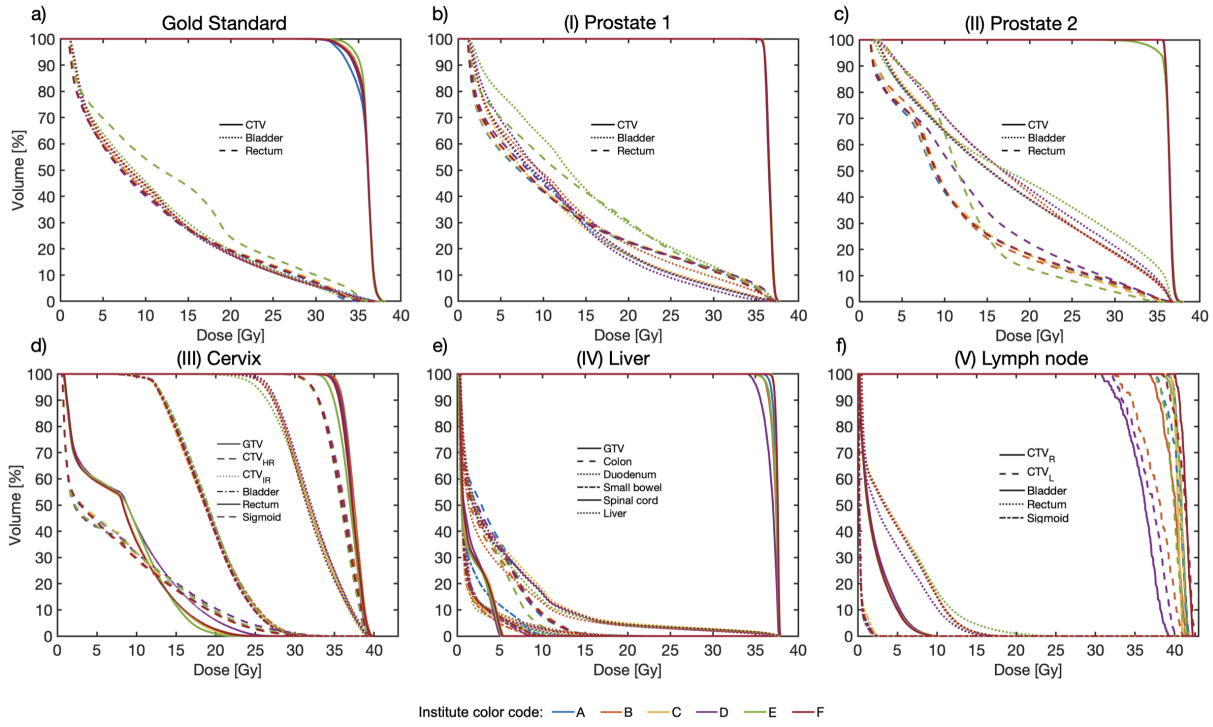

Figure 15: Results of the accumulated DVH per structure and institutes; for the cases gold standard (a), (I) prostate 1 (b), (II) prostate 2 (c), (III) cervix (d), (IV) liver (e) and (V) lymph node (f). The institutes (A-F) are presented in different colors, the case-related contours are presented in different line styles. Contours of fraction one were used for the respective DVH calculation.

The largest range of OAR CDC values in the GS was observed for rectum V32Gy with 4.1 % [3.5 % - 7.6 %]. In clinical cases (I) and (II), largest OAR CDC range was found for bladder V28Gy with 10.2 % [5.9 % - 16.1 %] and 7.6 % [22.4 % - 30.0 %], respectively. Similarly, the range of rectum D2cm^3 in case (III) was 2.8 [16.5 - 19.3] Gy. For the duodenum, case (IV), differences in Dmax of up to 7.1 [16.6 - 23.7] Gy were observed. For case (V), the largest range of CDC values was found for rectum D0.5cm^3 with 4.6 [14.1 - 18.7] Gy. CTV CDC showed largest variation for GS V34.3 Gy and case (V) D98% (CTV_R and CTV_L) with 9.5 % [86.3 % - 95.8 %], 8.1 [31.9 - 40.0] Gy, and 6.5 [33.1 - 39.6] Gy, respectively. Details on CDC results are given in Supplementary Table S6, results of wall structures are summarized in Supplementary Table S14.

2.2.5 Discussion

In this study, DDA was performed on a GS with known deformations based on simulations using a biomechanical modelling, in addition to five distinct clinical cases. A total of six different DDA approaches were investigated, using various contour guiding

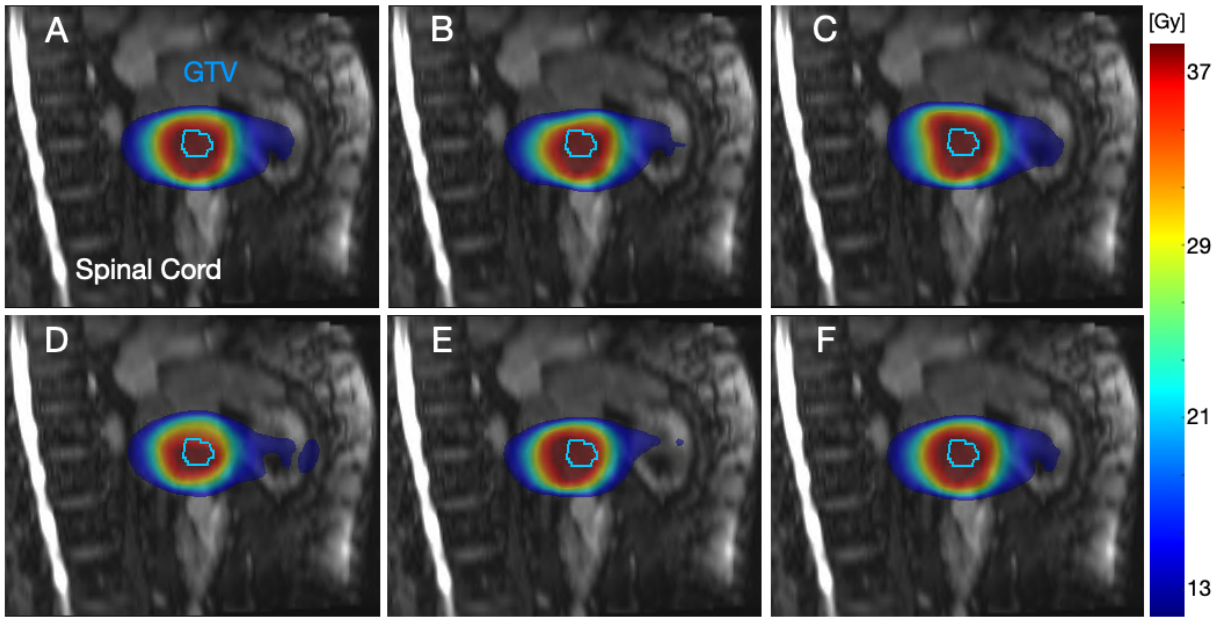

Figure 16: Sagittal plane view of institutional deformed dose accumulation (DDA) distribution results of the liver case (subfigures A-F are presenting the different institutes A-F). The gross tumor volume (GTV) contour of fraction one is presented in all images.

methods and two different dose resampling techniques, DDM and EMT. In general, a high degree of agreement was found between the different DIR and DDA assessments. However, absolute differences of potentially clinically relevant magnitude were found depending on the clinical cases and algorithms.

The finite element simulation enabled the calculation of GS-DVF of both forward and backward transformations, which are nearly inverse to each other within the size of a voxel. GS-DVF enabled consistent mapping of image and dose from the first fraction to the others and vice versa, allowing accuracy evaluation of the different DDA solutions compared to the GS-DDA. It is crucial to note that while the simulated deformations show a high level of anatomical accuracy [27], they do not completely replicate the intricacy of the true deformations in the pelvis. Nevertheless, the DIR solution without contour guidance (E) resulted for CTV and rectum in notably higher DDA errors and for bladder in the highest IQR but showed similar contour propagation performance to other solutions. This might be due to algorithm regularization and poor anatomical contrast inside the contour boundaries, causing large uncertainty in the estimated deformations and in turn leading to dose wrapping/accumulation errors. This highlights the need for DIR and DDA QA metrics that detect errors not just in high-contrast image areas like organ boundaries but also in near-isointense areas often seen in soft tissue boundaries [215], [264]. Statistical testing has largely shown significant differences in dose accumulation accuracy among the institutions. On the other hand, it is worth noting that such differences are often between 1 and $2\,\%$ of the maximum prescribed dose and thus their clinical relevance depends on the specific requirements.

By nature, for the clinical cases no DDA ground truth was available. A DVF and dose distribution analysis would provide a precise estimation of deformation results [160]. Unfortunately, this analysis was not possible due to export and reading issues and requires additional investigation. Alternatively, the percentages of contours yielding DSC > 0.8 were reported per algorithm, in addition to SDSC and HD95% were calculated. Nevertheless, it is worth noting that DSC may have restricted sensitivity to local registration errors, particularly for volumes of varying sizes. The clinical cases presented more variations in SDSC values among different algorithms as in the GS. Algorithms A, C, and E presented higher SDSC values compared to others in the clinical cases, indicating their relatively better performance in clinical scenarios.

The largest HD95% of $73.3\,\mathrm{mm}$ was found for duodenum in case (IV) which was not clinically concerning as this contour region was far from the high- or near-high-dose area. Nevertheless, this indicates remaining registration uncertainties underlining the need for appropriate registration strategies depending on clinical priorities. In contrast, for the CTVs of case (V), a maximum HD95% of $4.3\,\mathrm{mm}$ was reported, showing more robust results in high-dose regions.

In our study, offline re-delineations were used for DIR and DDA analysis. Instead of using propagated contours used for DIR guidance, re-delineation of deformed images and subsequent comparison to reference contours may be an alternative approach for investigating DIR quality. Unfortunately, such contours were not available for our analysis. However, it should be noted that the contours from the first fraction were used as guidance. Consequently, observed DSC reduction may be caused by registration errors or delineation variation. We acknowledge that manual contouring is subject to inter- and intra-observer variability [23], [28]. However, this may also be observed in clinical practice and thus requires thorough retrospective analysis. In contrast, re-delineation in online adaptive RT must be fast and accurate to satisfy the precision standards for DDA analysis. Automated deep-learning algorithms are currently developed for fast and robust auto-contouring [16], [73], [252]. These tools may soon be available for online adaptive RT, enabling contour-guidance for online DIR.

Analysis of DSC, SDSC and HD95% showed that the DIR algorithm relaying on contours only performed poorest overall. This algorithm generally works well in regions near contours but deteriorates significantly in more distant regions, such that the distance between the assessed OAR and the guiding contours can lead to significantly decreased registration results. Furthermore, the accuracy of this algorithm seems to decrease if more contours are used for guidance, most likely due to overfitting and increasing model complexity. However, the study of Wahlstadt et al. [241] revealed best performance for contour-only and hybrid algorithms in terms of DSC and HD95%, while the hybrid and intensity-based algorithms performed best for image similarity metrics in liver. In our study, consistent DIR results were found for hybrid intensity/structure-based algorithms, suggesting this might be well suited for clinical DDA. Similar results were found by Bosma et al. [28], who showed that contour-guidance significantly increased registration accuracy. However, due to the low number of investigated approaches no conclusion about the optimal registration strategy can be derived from this study.

An alternative comparison of DDA strategies was performed by comparing resulting DVHs of the different strategies using ICC and ANOVA. Even though ICCs indicated high correlation of DVHs, absolute differences in CDC were observed. Although the CDC presents a single DVH point, it may be of clinical relevance and observed differences require further clinical evaluation as they might impact toxicity and target coverage after online adaptive RT DDA. Low ICCs were observed for left and right CTV in case (V), due to inter-fraction motion of the lymph nodes [23], their very small volumes [207], and image resolution [263]. As an alternative approach, separate rigid registration of each lymph node with dose summation might be considered. Nevertheless, to estimate potential toxicity in the genitourinary system, DDA should be preferred.

The evaluation of the largest differences in the considered OAR CDC showed that the algorithm without contour-guidance (E) yielded minimum or maximum values, even though demonstrating better DSC, SDSC and HD95% results compared to the contour-only algorithm (D). This might be related to registration uncertainties near the high-dose gradient, i.e. a larger distance between points in high-dose areas results in small dose differences as opposed to points in gradient areas near the high-dose [160].

Regarding DDM and EMT, no valid statement can be made, as there was not enough variation in dose mapping approaches in this study. Furthermore, the decision about EMT or DDM can become rather complex, depending on the deformation characteristics over the course of the treatment [27], [160]. For the GS simulation, DDM- and EMT-based accumulated doses did not show any significant differences. Thus, for the institutional comparison DDM was employed, as it is currently the more commonly used approach.

The study is subject to limitations; only one of the institutions used EMT for dose warping and one used DIR without contour guidance. This makes the comparison of algorithms challenging but also reflects the present clinical situation without clear recommendations for DDA usage. Furthermore, some of the employed software solutions lack regulatory approval, thereby restricting their clinical use.

In conclusion, the compared algorithms for DIR and DDA yielded a generally high level of agreement. Nevertheless, absolute differences of potentially clinically relevant magnitude were observed depending on the clinical cases and algorithms. Before using DDA for online adaptive MRgRT in clinical practice, further studies are needed to provide recommendations and guidelines. In addition, robust methods for uncertainty quantification should be further investigated to determine areas of variation in the future.

2.2.6 Supplementary Material

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

2.3 Comparison of online adaptive and non-adaptive magnetic resonance image-guided radiation therapy in prostate cancer using dose accumulation

Autor: **Martina Murr**, Daniel Wegener, Simon Böke, Cihan Gani, David Mönnich, Maximilian Niyazi, Moritz Schneider, Daniel Zips, Arndt-Christian Müller and Daniela Thorwarth

Submitted to: Physics and Imaging in Radiation Oncology 2024

Manuscript Number: PHIRO-D-24-00218

Current Status: Review

2.3.1 Abstract

Background and Purpose Conventional image-guided radiotherapy (conv-IGRT) is standard in prostate cancer (PC) but does not account for inter-fraction anatomical changes. Online-adaptive magnetic resonance-guided RT (OA-MRgRT) may improve organ-at-risk (OARs) sparing and clinical target volume (CTV) coverage. The aim of this study was to analyze accumulated OAR and target doses in PC after OA-MRgRT and conv-IGRT in comparison to pre-treatment reference planning (refPlan).

Material and Methods Ten patients with PC, previously treated with OA-MRgRT at the 1.5 T MR-Linac (20 x 3 Gy), were included. Accumulated OA-MRgRT doses were determined by deformably registering all fraction's MR-images. Conv-IGRT was simulated through rigid registration of the planning computed tomography with each fraction's MR-image for dose mapping/accumulation. Dose-volume parameters (DVPs), including CTV D50% and D98%, rectum, bladder, urethra, Dmax and V56Gy for OA-MRgRT, conv-IGRT and refPlan were compared using the Wilcoxon signed-rank test. Clinical relevance of accumulated dose differences was analyzed using a normal-tissue complication-probability model.

Results CTV-DVPs were comparable, whereas OA-MRgRT yielded decreased median OAR-DVPs compared to conv-IGRT, except for bladder V56Gy. OA-MRgRT demonstrated significantly lower median rectum Dmax over conv-IGRT (59.1/59.9 Gy, p=0.006) and refPlan (60.1 Gy, p=0.012). Similarly, OA-MRgRT yielded reduced median bladder Dmax compared to conv-IGRT (60.0/60.4 Gy, p=0.006), and refPlan (61.2 Gy, p=0.002). Overall, accumulated dose differences were small and did not translate into clinically relevant effects.

Conclusion Deformably accumulated OA-MRgRT using 20 x 3 Gy in PC showed significant but small dosimetric differences comparted to conv-IGRT. Feasibility of a dose

accumulation methodology was demonstrated, which may be relevant for evaluating future hypo-fractionated OA-MRgRT approaches.

2.3.2 Introduction

Conventional image-guided radiation therapy (conv-IGRT) is the standard of care in precise irradiation of the target volume and sparing of healthy tissue in prostate cancer (PC) treatment [135].

Currently, radiation therapy (RT) is based on planning computed tomography (pCT) images and sometimes additional magnetic resonance imaging (MRI) scans. The target volume and organs at risk (OARs) are identified and contoured on pCT and/or MRI, and the treatment dose is optimized according to clinical dosimetric criteria. Prior to irradiation, the patient's position is validated and corrected, if necessary, based on cone-beam CT (CBCT) images acquired on the treatment couch. Correction vectors are determined via rigid registration of the CBCT to the reference anatomy defined by the pCT. Despite the precision of this targeted radiation technique, PC patients can experience urogenital side effects such as strictures, obstruction, hematuria, dysuria, and incontinence [111], [253].

With the introduction of conv-IGRT target coverage was increased, whereas the actual dose delivered to normal tissues remained highly uncertain because of volume as well as positional changes of the OARs with respect to the target. This is due to the fact that conv-IGRT lacks the ability to account for deformations and anatomical changes between fractions and changes in volume, such as bladder and rectum filling, which are critical for healthy tissue sparing [45]. Online-adaptive RT (OA-RT) techniques, such as magnetic resonance-guided radiotherapy (MRgRT), promise better sparing of OARs while ensuring high tumor dose coverage [4], as MRI scans are acquired daily, allowing for dose re-optimization and -calculation based on the daily anatomy [247].

Fractionated OA-RT necessitates deformable dose accumulation (DDA) to determine the total dose deposited in both the tumor and surrounding tissue [182]. DDA involves several steps: deformable image registration (DIR) to accommodate anatomical variations, dose mapping to determine the delivered dose to normal tissues and the target volume, and summation of deformed fractional doses to calculate the total delivered dose [160]. However, integrating DDA into the clinical workflow for OA-RT faces challenges, primarily due to uncertainties associated with DIR [168] and its technical implementation [160]. Despite these challenges, studies have demonstrated the potential of DDA in MRgRT, including predicting toxicity and facilitating treatment adaptation [4], [25], [42], [260].

The aim of this study was to investigate total deformably accumulated doses in PC after OA-MRgRT and conv-IGRT, in comparison to doses from the pre-treatment reference plan (refPlan), using relevant planning dose-volume parameters (DVPs) for clinical target volume (CTV) and OARs. The clinical significance of the differences was assessed using a normal tissue complication probability (NTCP) model.

2.3.3 Material and Methods

Patient characteristics
This prospective study included data from ten PC patients who were treated with OA-MRgRT (20 x 3 Gy) at the 1.5 T MR-Linac (Unity, Elekta AB, Stockholm, Sweden) between March 2019 and March 2020 at the University Hospital Tübingen. The cohort's median (range) age at treatment was 76.4 (58.8 - 83.1) years. All patients were part of a clinical study approved by the institutional review board (NCT02724670) and gave written informed consent. Patient inclusion criteria were histologically confirmed PC with indication for curative RT, cT1b-cT3a cN0 cM0, ECOG PS 0-2, and IPSS$\leq$12 (possibly only achieved after neoadjuvant hormone therapy or tamsulosin administration).

Reference Planning
Before treatment initiation, a refPlan was created for each patient based on the pCT (voxel size 1.17 x 1.17 x 3.00 mm^3) (Figure 17 a), matched to a T2-weighted (T2w) simulation MRI for precise MR contouring. The radiation oncologist delineated the CTV defined by the prostate contour. The planning target volume (PTV60) was created using a 5 mm margin around the CTV (except 3 mm dorsally towards the rectum). The CTV plus up to 20 mm of the seminal vesicle base determent the CTV_{SV}, as well as a respective planning target volume prescribed with 57.6 Gy (PTV57.6) with a 6 mm margin around CTV_{SV} (except 5 mm dorsally towards the rectum). Bladder, rectum, and urethra were delineated as OARs. In addition to dose prescriptions of 60.0/57.6 Gy, maximum dose constraints for rectum, bladder, and urethra were set to Dmax $\leq$ 61.0 Gy. Further OARs constraints were V56Gy < 13.5 % for the rectum and V56Gy < 18.0 % for the bladder. Dose planning and calculation was performed on a voxel grid size of 3 x 3 x 3 mm^3, employing nine IMRT beams with the MONACO treatment planning system (Version 5.51.11, Elekta AB, Stockholm, Sweden).

Conventional-IGRT workflow simulation
Before each daily radiation fraction using MRgRT, a T2w-MRI scan (voxel size 0.83 x 0.83 x 2.00 mm^3) was acquired. To simulate a conv-IGRT workflow, the pCT was rigidly registered to the daily T2w-MRI of each fraction (cf. Figure 17 b) using a bounding box around the CTV taking into account translations only. After rigid image registration, the refPlan dose was mapped accordingly to each fraction's T2w-MRI.

Online adaptive MRgRT
OA-MRgRT was applied based on the daily acquired T2w-MRI. For each fraction, the refPlan was adapted to the patient's anatomy using the adapt-to-shape (ATS) or, in some cases, the adapt-to-position (ATP) workflow (cf. Figure 17 c and Table 3). ATP was only used after visual inspection by a radiation oncologist and no relevant anatomical and position changes after initial image registration.
In the ATS workflow, the pCT and contours were deformably registered to the current T2w-MRI. Contours were propagated, corrected, or re-delineated, if necessary, by the radiation oncologist. Each adaptive plan was re-optimized based on the respective T2w-MRI and the new fraction contours using a synthetic CT with assigned bulk electron density [48]. For further details regarding the workflows, please refer to [247]. In

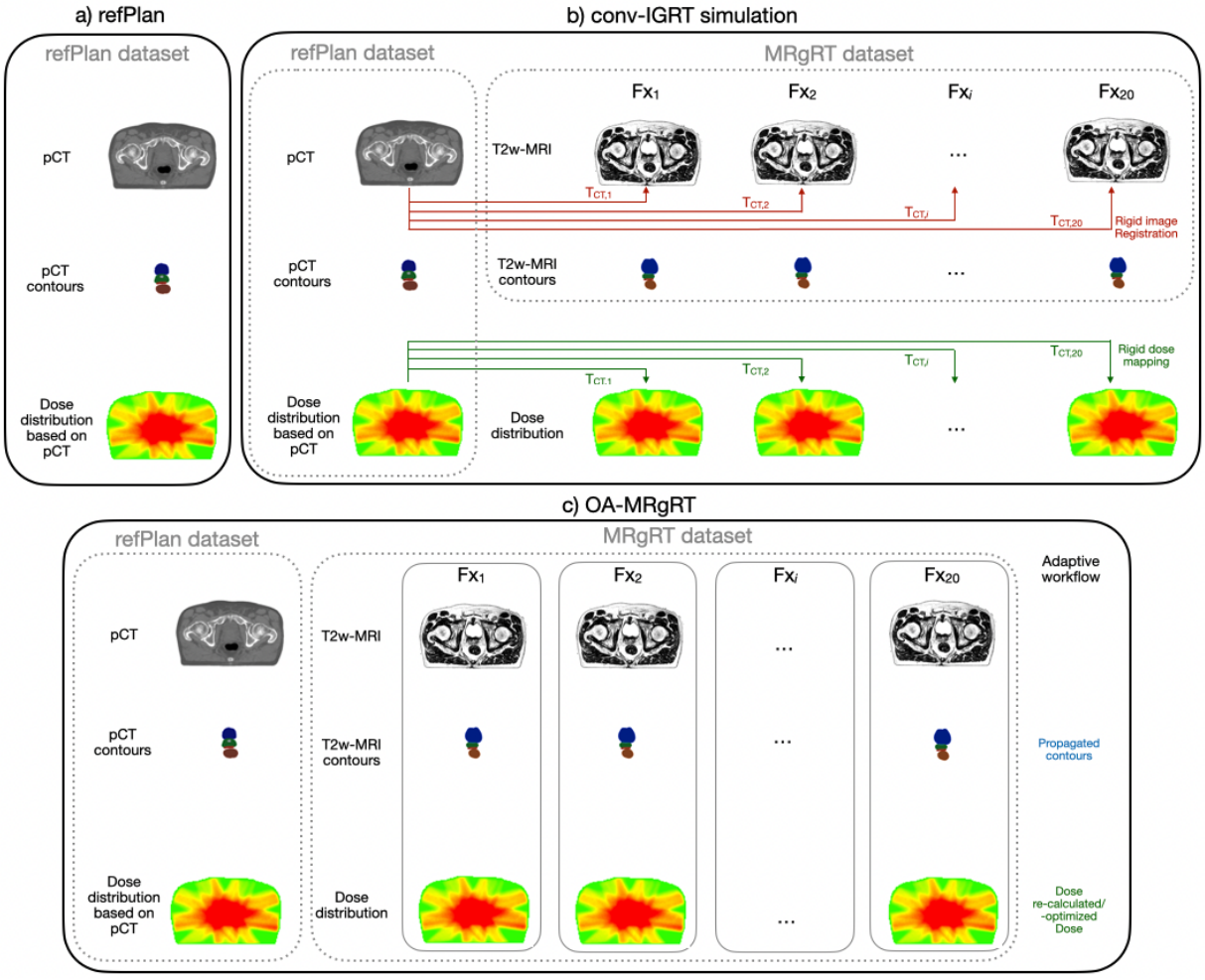

Figure 17: Schematic illustration of (a) the refPlan including pCT, contours and dose distribution, (b) conv-IGRT simulation with rigid registration of the pCT to each T2w-MRI and the resulting translation matrix T, and (c) OA-MRgRT with new online adaptive dose planning for each fraction.

Table 3: Summary of the OA-MRgRT workflows used per patient. Fraction numbers using ATS or ATP based on the pCT or previous T2w-MRI are given.

Patient	ATS	ATP	
		to the pCT scan	to a T2w-MRI of a previous treatment
P01	17	3	0
P02	9	11	0
P03	19	1	0
P04	20	0	0
P05	12	7	1
P06	14	6	0
P07	8	12	0
P08	16	3	1
P09	17	3	0
P10	13	0	7
Total	**145**	**46**	**9**

contrast, if the ATP workflow had been used, the daily T2w-MRI was rigidly aligned with either the pCT or an individually selected T2w-MRI from a prior treatment fraction. The location of the isocenter was updated accordingly, and the planned dose of the respective fraction was then recalculated for this position.

Dose mapping and accumulation

For the DDA procedure and analysis, the contours of the CTV created during clinical OA-MRgRT treatment on the daily T2w-MRI were used after cross-checking by an experienced radiation oncologist, who also delineated the urethra on each T2w-MRI manually. Bladder and rectum contours were created on the T2w-MRIs using automatic contouring (ADMIRE auto contouring, V1.0, Elekta AB, Stockholm, Sweden) to ensure high anatomical accuracy of the DDA analysis.

As a first step of DDA for both, conv-IGRT and OA-MRgRT, the daily T2w-MRIs of the OA-MRgRT were deformably registered to the first fraction, cf. Figure 18 a, using the auto-contours of the prostate, bladder, and rectum of the first fraction as registration guidance. Subsequently, the fraction doses were deformably mapped according to the deformation vector field (DVF) resulting from DIR (Figure 18 b). ADMIRE research (Version 3.48, Elekta AB, Stockholm, Sweden) was employed to perform DIR and dose mapping, using a hybrid intensity/structure-based algorithm [159]. The accumulation was carried out using an in-house-developed Python script (Python Version 3).

Analysis

The comparison between the refPlan and the two adaptive treatment strategies, namely the accumulated conv-IGRT simulation and OA-MRgRT doses, was conducted through a dosimetric analysis. This involved evaluating DVPs derived from the planning objectives used during routine plan approval in our department for the OARs rectum, bladder

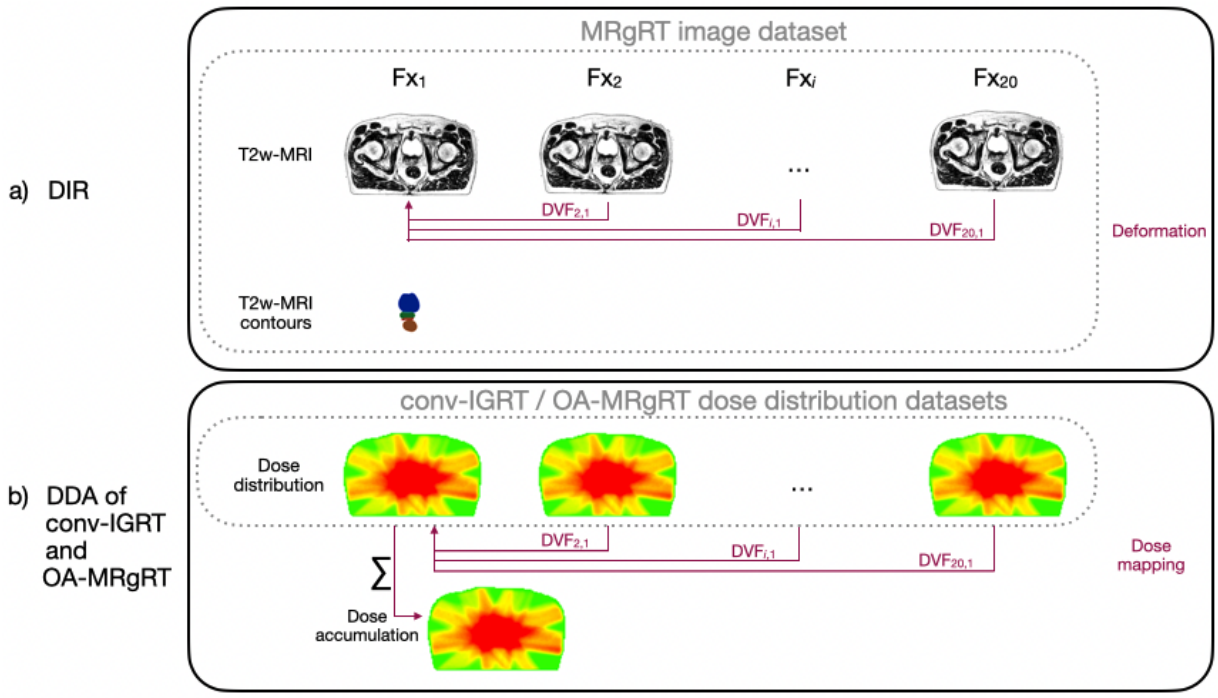

Figure 18: Schematic overview of a) the DIR and b) DDA procedure for conv-IGRT and OA-MRgRT. The DIR resulted in a DVF used for dose mapping for both accumulation approaches rigid registered refPlan doses of conv-IGRT and the daily adaptive doses of OA-MRgRT.

and urethra as well as for the prostate CTV (D98% $\geq$ 57 Gy, D50% $\geq$ 60 Gy). CTV_{SV} was excluded from this analysis due to the non-anatomical definition of this volume. All relevant DVPs, CTV D50% and D98%, rectum, bladder, and urethra Dmax, and rectum and bladder V56Gy, were extracted from corresponding dose-volume histograms (DVHs), computed with a bin width of 0.01 Gy.

Whether dosimetric differences between the refPlan and the two adaptive treatment strategies lead to clinically meaningful differences was calculated on the basis of the NTCP using the Lyman-Kutcher-Burman (LKB) model

$$NTCP = \frac{1}{1 + e\left(-\frac{EQD2 - TD_{50}}{m \cdot TD_{50}}\right)} \tag{6}$$

with EQD2 equivalent dose in 2 Gy fractions (cf. Supplementary 9.3.1), TD_{50} dose at which 50 % of the population is expected to experience a complication and m slope parameter describing the steepness of the dose-response curve [37], [124], [125], [142]. The NTCP was calculated for bladder incontinence G2+ ($\alpha/\beta = 1.5$ Gy, $TD_{50} = 108.9$ Gy, $m = 0.24$, $n = 0.02$) [33], and for rectum late toxicity G2+ ($\alpha/\beta = 3.0$ Gy [147], $TD_{50} = 80.8$ Gy [59], $m = 0.15$ [37], $n = 0.12$ [37]).

Statistical analysis of the differences in DVPs and NTCP obtained for the different treatment approaches was carried out using the non-parametric Wilcoxon signed-rank test,

with α = 0.05 and a Bonferroni correction accounting for the three different treatment approaches. Additionally, Cohen's d was calculated for the DVPs (cf. Supplementary 9.3.1).

2.3.4 Results

Variability of bladder and rectum volumes per treatment fraction are presented in Supplementary Fig. S7. The two investigated DDA approaches, OA-MRgRT and conv-IGRT, resulted in comparable median (range) accumulated doses for CTV D50% with 60.0 (59.6 - 60.6) Gy and 59.9 (59.6 - 60.6) Gy, similar to the refPlan with 60.1 (59.7 - 60.6) Gy. Additionally, the approaches achieved the CTV D98% planning objective in 100 % of the patients, OA-MRgRT with a median (range) of 58.6 (57.6 - 59.5) Gy, conv-IGRT with 58.3 (56.9 - 59.4) Gy and refPlan with 58.6 (57.8 - 59.2) Gy. Comparing OA-MRgRT to conv-IGRT resulted in statistically significant differences for CTV D98% (p = 0.029), whereas for CTV D50% was different between conv-IGRT and refPlan (p = 0.012).

In compliance with the planning objectives defined for OARs, the rectum Dmax was attained for 10/10 patients for both the OA-MRgRT and the refPlan with a median (range) of 59.1 (58.4 - 60.9) Gy and 60.1 (59.6 - 60.8) Gy respectively, while this objective would have been reached in 90 % of the cases using conv-IGRT with 59.9 (59.2 - 61.6) Gy. Statistically significant differences were found between OA-MRgRT and conv-IGRT (p = 0.006) or refPlan (p = 0.012). Conversely, adherence to the rectum planning objective V56Gy was achieved with OA-MRgRT for all patients with a median (range) of 3.5 % (1.4 % - 7.9 %), whereas 90 % of all patients achieved the planning constraint for conv-IGRT with 4.2 % (2.3 % - 14.9 %) and refPlan with 6.0 % (2.4 % - 15.4 %). Comparing OA-MRgRT and refPlan, significant statistical differences were observed (p = 0.012). The bladder Dmax planning constraint was fulfilled in 100 % of the cases by OA-MRgRT, 90 % by conv-IGRT, and 30 % by refPlan, with a median (range) of 60.0 (59.5 - 60.9) Gy, 60.4 (59.4 - 61.4) Gy and 61.2 (60.9 - 62.4) Gy, respectively. Additionally, bladder Dmax demonstrated significant differences comparing OA-MRgRT and conv-IGRT to refPlan (both p = 0.006) and OA-MRgRT to conv-IGRT (p = 0.018). However, all approaches achieved 100 % for the V56Gy planning objective of the bladder with a median (range) 3.6 % (1.6 % - 9.7 %) for OA-MRgRT, 3.5 % (1.7 % - 9.0 %) for conv-IGRT, and 5.6 % (4.0 % - 8.9 %) for refPlan. For the urethra Dmax, attainment levels were 90 %, 80 %, and 50 % for OA-MRgRT with a median (range) of 60.3 (59.7 - 61.2) Gy, conv-IGRT with 60.6 (60.1 - 61.1) Gy, and the refPlan with 61.0 (60.4 - 62.5) Gy, respectively. For urethra Dmax, a significantly lower value was observed for conv-IGRT compared to the refPlan (p = 0.029). The detailed analysis and distribution of planning constraints and DVPs for all investigated approaches are presented in Figure 19, including Cohen's d results. Additionally, population DVHs for CTV and all OARs are visualized in Figure 20. Figure 21 demonstrates the resulting accumulated dose distributions of the three different approaches for a representative patient case, illustrating better rectal sparing in the high-dose area with OA-MRgRT. No significant differences were found between OA-MRgRT and conv-IGRT using DDA, with regards to NTCP (cf. Supplementary Fig. S8).

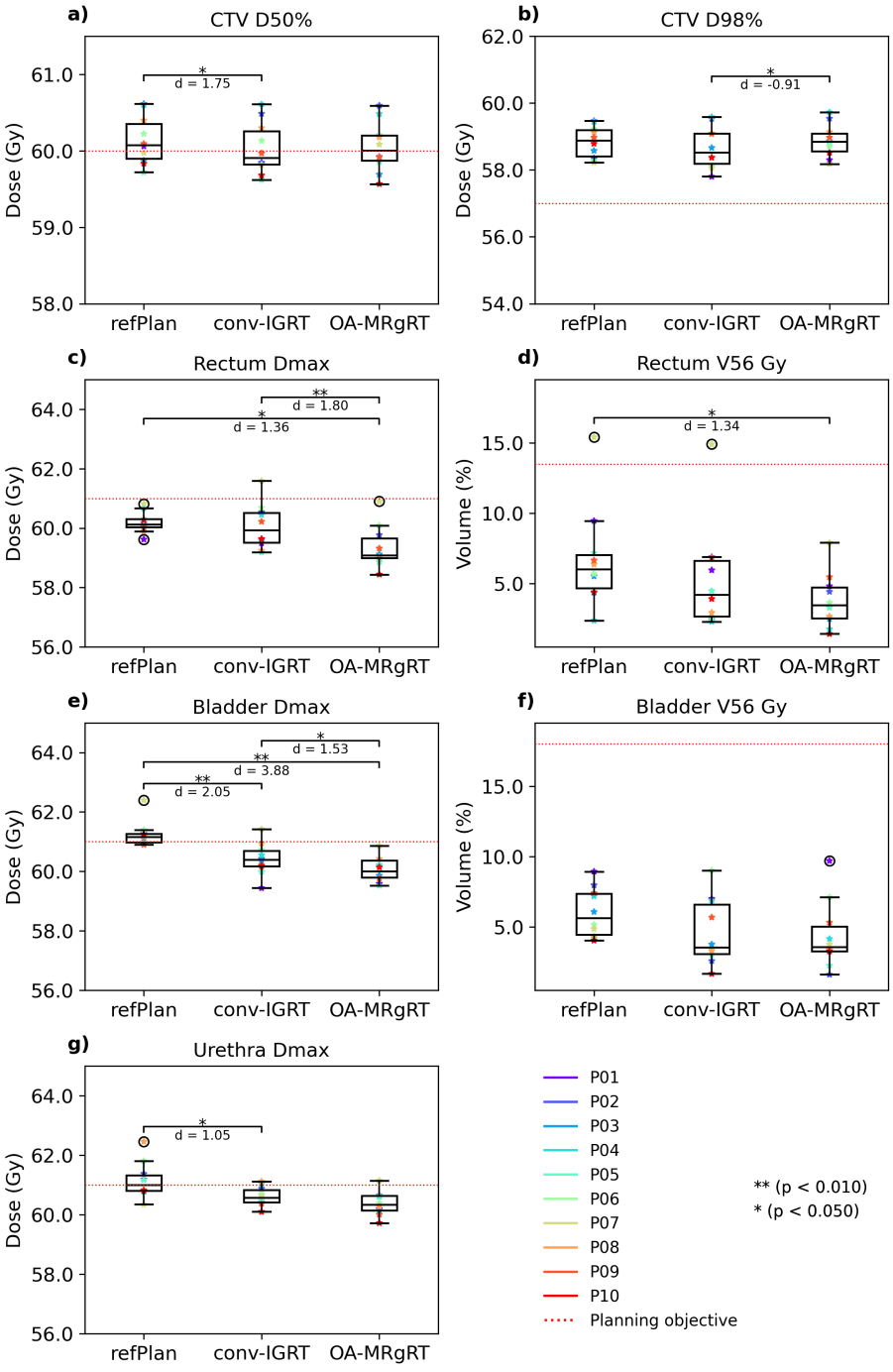

Figure 19: Boxplots showing the differences in the DVPs for CTV and OARs between refPlan, conv-IGRT, and OA-MRgRT. The boxes represent interquartile range (IQR), which is the range between the first quartile (Q1) and the third quartile (Q3). The black line inside the box is the median of the dataset. The length of the whiskers is set to 1.5 times the IQR. Individual data points beyond the whiskers are considered potential outliers and plotted as black circles. The color-coded points are the patient-individual results. The red dotted lines represent the planning objectives: CTV D50%≥60 Gy (a), D98%≥57 Gy (b), rectum, bladder, and urethra Dmax<61 Gy (c, e, g), rectum V56Gy<13.5% (d), and bladder V56Gy<18% (f). Statistical analysis was performed using the non-parametric Wilcoxon signed-rank test, with α = 0.05 and a Bonferroni correction accounting for the three different treatment approaches, in addition to Cohen's d.

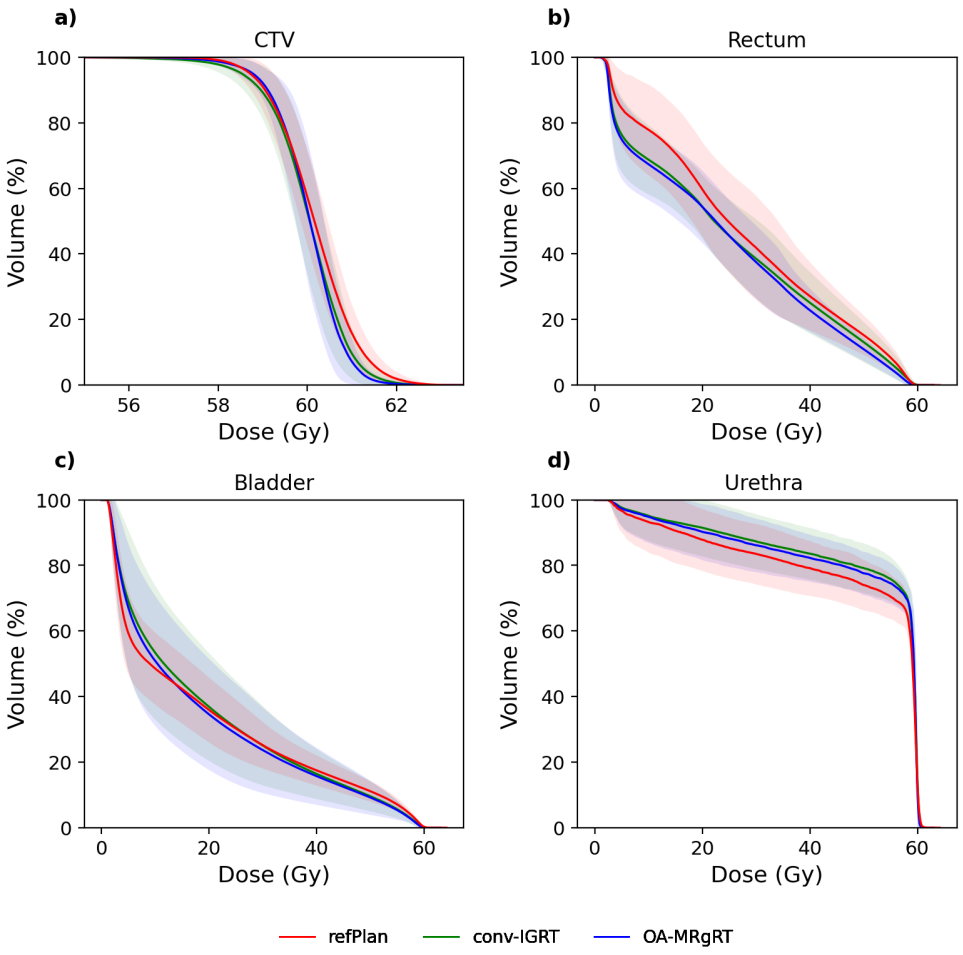

Figure 20: Illustration of the mean DVHs results across all patients for refPlan (red), conv-IGRT (green) and OA-MRgRT (blue) in the CTV (a), rectum (b), bladder (c), and urethra (d). The transparent area for each approach represents the standard deviation across all patients.

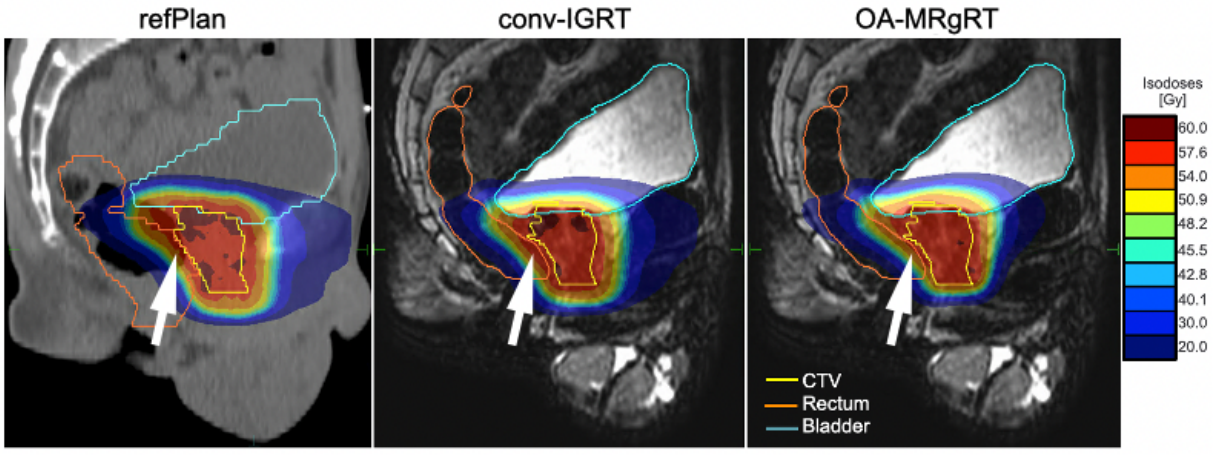

Figure 21: Sagittal view of the refPlan and the different accumulated dose distribution results, conv-IGRT and OA-MRgRT, for patient 10. The contours of CTV, rectum, and bladder are shown.

2.3.5 Discussion

The aim of this study was to compare DDA for two adaptive treatment strategies, conv-IGRT and OA-MRgRT, with refPlan. While the refPlan represents the intended dose for targets and OARs, a conv-IGRT approach was simulated illustrating the total accumulated dose after e.g. CBCT adaptation. In contrast, the total accumulated dose delivered to patients during OA-MRgRT including daily replanning was calculated. Our DDA study showed that OA-MRgRT allows increased OAR sparing while maintaining similar target coverage compared to the refPlan, indicating its potential clinical benefits compared to conv-IGRT. However, for the treatment schedule of $20 \times 3\,\mathrm{Gy}$ the overall effects and thus clinical relevance were small. Nevertheless, the established workflow will allow investigation of dosimetric differences for future hypo-fractionated treatments.

The process to generate the refPlan consists of two main phases: acquisition of the pCT including patient positioning, which typically is done a week prior to the start of the treatment, followed by contouring of targets and OARs in addition to dose planning, which require approximately one working day for the physicist and the radiation oncologist. The duration of each treatment fraction varies depending on the chosen approach. In the conv-IGRT workflow, which includes patient setup, position verification using on board imaging such as CBCT and irradiation, each treatment session takes approximately 5 minutes [72]. In contrast, total prostate OA-MRgRT treatment times typically range from 20 to 45 minutes per session, with the duration influenced by the necessity for recontouring and the speed of adaptive as well as the replanning process [243]. Throughout this time, it is crucial to reduce intra-fraction motion. Efforts to expedite the online-adaptive planning process have led to the development of automated contouring [102], [109], [167], planning [123] and dose calculation [165] tools. However, the management of intra-fraction motion during OA-RT remains challenging. To overcome this, technologies such as real time motion management were recently proposed to accommodate intra-fraction motion for patients undergoing OA-MRgRT [62], [77]. These advancements might contribute to a reduction of the overall OA-MRgRT treatment time in the future.

In this study, the decision to analyze CTV planning constraints instead of PTV after DDA was based on the clinical rationale to ensure CTV dose coverage by defining an artificially constructed PTV contour [155]. Consequently, analyzing planning objectives, particularly CTV D50% and D98%, offers valuable insights into the dosimetric performance of treatment planning and delivery after DDA. Our findings demonstrated that both approaches, OA-MRgRT and conv-IGRT achieved similar median CTV doses compared to the refPlan, indicating consistent target coverage across the different approaches. This consistency suggests that, irrespective of the specific treatment approach employed in our study, maintaining a comparable level of target coverage was prioritized to ensure similar tumor control.

In the assessment of OAR-DVPs, differences were found between refPlan, conv-IGRT, and OA-MRgRT, particularly in rectum, bladder, and urethra. Although the OAR planning constraints were not fully achieved for the refPlan across patients, OA-MRgRT

effectively met these constraints in all but one OAR-DVPs (Dmax urethra). The comparison between OA-MRgRT and conv-IGRT showed a decrease in median high dose constraints using OA-MRgRT. This indicates that OA-MRgRT may be more effective at sparing OARs due to its ability to adapt the treatment plan for each fraction based on daily high resolution anatomical imaging feedback, in contrast to conv-IGRT, where the plan is based on a single pre-treatment image. We hypothesize that DDA considering multiple fractions results in smoothing high dose tails of DVHs and thus leads to reduced maximum OAR-DVPs. Our results are consistent with the findings of recent studies [46], [103]. However, the study by Xiong et al. [260] found that OA-MRgRT additionally enhances target coverage. These varying results highlight differing priorities in adaptive planning, specifically, the optimization of planning for target volume coverage versus minimizing OAR Dmax.

In our study, twenty fraction doses were accumulated. With a smaller number of fractions, e.g. five, the contribution of each fraction dose map will increase. Consequently, averaging of maximum DVPs will be less pronounced [241]. As a result, any discrepancies between the planned dose and the delivered dose in individual fractions may be more apparent in the accumulated dose distribution. Conversely, with a larger number of fractions, there are more opportunities for the accumulated dose to integrate and average out variations between fractions. This may result in a more uniform accumulated dose distribution, as the cumulative effect of multiple fractions tends to mitigate individual discrepancies. However, the number of fractions is not the only factor influencing the accuracy of DDA [160]. Other factors, such as the magnitude and nature of anatomical changes [220], the accuracy of DIR [168], and the robustness of the treatment planning process [87], [267], also play critical roles. On the other hand, the PACE-B trial showed that smaller fraction numbers are not inferior to the fractionation used in this study in terms of biochemical and clinical outcome [9]. Additionally, it reduces patient attendances and shortens treatment time. For DDA, this implies a requirement for robust and precise implementation and performance.

In this study, a hybrid intensity/structure-based algorithm was employed for DIR. The literature proposes two methods for dose resampling: direct dose mapping (DDM) and energy/mass transfer (EMT) [160]. We utilized the DDM method, as a recent study from our group [159] demonstrated that both methods yield comparable results. Additionally, their study indicated that the algorithm used for this study provides robust quality assurance outcomes.

The NTCP model revealed no significant differences between the two RT approaches, OA-MRgRT and conv-IGRT, and thus no clinical impact despite significant differences in the DVP evaluation. However, in this study, patients were treated with medium to long course RT applying twenty fractions of $3\,\mathrm{Gy}$. We hypothesize that the difference will be higher when applying fewer RT fractions. However, further studies are needed to investigate this. Nevertheless, our study demonstrated feasibility and transferability of the established DDA-method and -workflow to various RT scenarios.

There are limitations to the study. In the OA-MRgRT workflow, anatomical changes

were compensated by the online-adaptive workflow, in contrast larger anatomical deviations such as bladder and/or rectum filling which were accepted for certain treatment fractions may not have been accepted during clinical conv-IGRT using a CBCT workflow. Ultimately this may imply limitations regarding our conv-IGRT simulations. Additionally, the simulated conv-IGRT workflow did not account for any rotation during rigid registration. However, in practice, this could be compensated with a hexapod table. Moreover, the calculation of accumulated doses currently lacks error and confidence interval estimation. A further limitation is the small cohort size, which may have influenced the statistical analysis.

In conclusion, the retrospective comparison of OA-MRgRT and conv-IGRT in PC patients demonstrated feasibility and transferability of the established DDA-method and -workflow. Significant dosimetric differences were found, but due to the small effect size when applying long course treatments, these did not translate into clinical relevance with respect to NTCP. However, clinical effects might be more pronounced for hypo-fractionated RT using only few fractions.

2.3.6 Supplementary Material

The supplementary material of the publication is depicted in the supplementary material of the dissertation at section 9.3.

3 Discussion

In this present work, deformable dose accumulation in adaptive radiation therapy was investigated to develop a robust deformable dose accumulation methodology and workflow for online adaptive MR-guided radiation therapy (OA-MRgRT).

In a first step, a current landscape was effectively developed by identifying and classifying the current applications of dose mapping/accumulation (DMA). Consequently, the obstacles to DMA that hinder clinical implementation and adoption were identified and analyzed. This led to the development of guidelines and requirements for providers and users.

The second goal was to develop a robust deformable dose accumulation (DDA) solutions. In a multi-center study, we compared and evaluated the solution with various other solutions currently used in research and clinical practice. For RT, there are various deformable image registration (DIR) and DDA-solutions available. The solutions use different settings, resampling methods, and registration guidance. Moreover, these solutions find application in diverse anatomical regions. The multi-center study investigated potential differences between the DIR and DDA solutions used. Six institutes from the MR-LINAC consortium participated in this study. The study included five clinical cases from the participating institutions: two prostate cases, one cervix, liver, and lymph node case that were treated with online adaptive MR-guided radiation therapy, and one gold standard (GS) case based on a biomechanical model since there is no known ground truth for DDA quality assurance (QA). We analyzed DIR using established QA-metrics and DDA using the dose-volume histogram (DVH) and dose-volume parameters (DVP).

The third objective included the elaboration of a DDA-workflow. Following the successful development and positive test results of the robust DDA solutions, the work concludes with applying the developed DDA solution in a retrospective study, comparing non-adaptive and adaptive RT of prostate cancer patients who received OA-MRgRT. Adaptive RT aims to spare normal tissue more effectively than non-adaptive. We dosimetrically compared the accumulated doses of the target volume and normal tissues after classical image-guided and OA-MRgRT to the initial treatment plan doses. A normal tissue complication probability model was used to determine clinical impact.

3.1 Current state and challenges of dose mapping and accumulation in radiation therapy

The first objective of this thesis was to analyze the current status and the challenges of dose mapping/accumulation (DMA) in radiation therapy (RT) [160]. By identifying the challenges that hinder its integration into clinical practice, the aim was to provide clarity to drive its integration forward.

To analyze the current status, the most recent DMA applications in RT were presented in a landscape called DMAL. It combines the expected anatomical variations and the impact of dose mapping uncertainties on patient safety of the use cases. The DMAL presents use cases that provide an assessment on which the user can make decisions. In this context, the DMAL can also serve as a tool for categorizing new use cases or revising the ones already presented. The anatomical variations on the y-axis, such as inter- and intra-fractional, are well known and have been the subject of previous and ongoing research [30], [91], [196]. We identified three main categories for the impact of uncertainties in dose mapping: "quality" which refers to retrospective group analysis; "inform" which provides information for secondary treatment; and "include" which implies direct consideration in secondary treatment. For use cases categorized as "incorporate" the DMAL requires full quantification of uncertainties in dose mapping, including correctness and consistency. Nevertheless, tools to quantify these uncertainties are not yet widely available in commercial solutions, and there is a lack of metrics to quantify the impact of uncertainties on patient safety. As a result, we recommend extreme caution and highly targeted use in clinical practice.

In order to shed light on the impact of uncertainties on patient safety in detail, the first-order effects, i.e. the DIR uncertainties, were first identified, which are difficult to quantify because no "ground truth" is known for verification. To begin with, DIR uncertainties have various sources, which can be categorized into image-derived, including anatomical changes, image quality and multimodal registration, and algorithm-derived, which includes parameter settings [168]. Furthermore, because of their nature, proposed metrics have limited meaning [168]. The selection of suitable metrics, among other things, depends on the subsequent DIR application. For DIR quality assurance for contour propagation, visual inspection of the registered image and contour is at least recommended [29], [160]. As part of the DIR QA for deformation of quantitative information such as dose, in addition to the visual inspection of the registered image, it is recommended to assess a plausibility check of the Jacobi determinant, the bending energy, and the dose (gradient) [29], [160], [168]. However, this necessitates the tools' availability and applicability in commercial software solutions, as well as an assessment of the interpretation of the results. In the near future, providers should develop and implement tools to rationalize this quantification, ideally incorporating these uncertainties.

In addition to the quantification of DIR uncertainties as first-order effects, second-order effects associated with dose mapping and accumulation were also identified as a challenge, which includes application methods of dose mapping. Two resampling and inter-

polation methods are currently under discussion, DDM and EMT. We argue that correct application of both methods yields the same result. However, the study by Bosma et al. [27] found differences between the two methods and concluded that EMT performs better in terms of decreasing dose errors for prostate treatment. DDM is hypothesized to have limitations in deforming surfaces with significant fluctuations in tissue density, as well as in tissue areas with internal density variations, resulting in a Jacobian determinant that deviates from one, indicating compression or expansion. The intention of EMT is to compensate for the changes in density caused by realignment. The results of this thesis's second objective [159], in contrast, support the assertion that there were no significant differences between the methods. However, an important point is that not only the method of resampling and interpolation is subject to uncertainties, but that it is a sum of uncertainties (first- and second-order effects).

We further named and investigated radiobiological issues as a challenge posed by second-order effects. The biological response to radiation may change over time and is influenced by factors such as the fractionation regimen, the total treatment duration and the individual patient recovery process. There is a complex interplay of various biological factors, such as the microenvironment, genetic predisposition and cellular repair mechanisms. NTCP models are based on certain mathematical assumptions and parameters described by the LQM, which describes fractionated therapy using a power-to-n rule. These assumptions may not always be correct, leading to possible inaccuracies in predicting events. In dose accumulation, the power-to-n rule is only applicable if the dose of each fraction is identical. Another reason may be an oversimplification of the actual biological effect of the dose in the different tissues [134]. Due to the nonlinear nature of the cell survival curve, linear accumulation of dose is not useful for predicting the biological effect of dose variations in the fractions [237]. Nieburg et al. introduced the biological dose tool, $\mathrm{bEQD_d}$, which takes into account the dose of each fraction and the radiobiology of the tissue, α/β values [171]. Uncertainties in the NTCP model, as well as uncertainties in the actual applied dose distribution, have a significant impact on accuracy achievements and challenges [60].

We formulated and developed specific recommendations and guidelines for providers and users in light of the findings. The primary objectives are to prioritize context-driven DIR and highlight its impact on clinical decisions and assessments, rather than striving for unachievable perfection in DIR accuracy. It is important to acknowledge the limitations inherent in the implemented DIR algorithms. Furthermore, the analysis suggests that dose mapping, even without precise quantification of uncertainties, could be a better alternative than omitting dose mapping. These considerations underline the importance of open discussion within a multidisciplinary team to thoroughly assess and resolve the uncertainties associated with each specific use case. Such collaborative assessments are critical to optimize clinical outcomes and ensure the reliability of DIR applications in practice.

3.2 Development and evaluation of a robust DDA solution by comparing different solutions

The second objective of this thesis was to develop a robust in-house solution for DDA and investigate different available DDA algorithms [159]. We specifically investigated the DDA solutions by comparing the differences between various available DDA algorithms in different anatomical areas.

The study observed differences in the results when comparing the DSC results between a gold standard (GS) case with known ground truth and all clinical cases across all contours. For the GS, all institute's results presented with DSC > 0.8, while for the clinical cases, one institute obtained a maximum of $81.3\,\%$, and one institute yielded a minimum DSC of $34.6\,\%$ of all contours in question with DSC > 0.8. The clinical cases' DSC analysis revealed that using too many contours for registration guidance resulted in lower DSC values. Carefully selecting the regions of interest can yield a robust and context-driven registration result. However, the community highly debates the use of DSC as a primary evaluation metric, arguing that it is not a suitable metric for evaluating DIR or contour propagation, particularly in the case of small volumes. Bosma et al. [27] recently presented classifications of metrics for the DIR analysis. Therefore, the study also analyzed the SDSC and HD95% metrics. However, the study noted that there is still room for improvement in both the selection of metrics for analyzing the DIR and DDA results, as well as the interpretation of these results. DIR and DDA algorithms are very complex, and an objective comparison of the results in a clinical context is challenging. Nenoff et al. [168] discussed in detail the impact of DIR uncertainty on RT use cases and how to deal with these uncertainties for patient-related applications. The first objective of this dissertation [160] emphasized the need for analysis and comparison of the DVF. Unfortunately, in this study, it was impossible to export the DVF for some solutions, and some DVFs could not be imported for analysis due to the file format. This demonstrates the need for continued implementation of the software solutions' requirements. Additionally, the analysis still requires the DICOM tags for the DVF to be sufficiently defined; other parts of the DIR have already sufficiently defined DICOM tags [47].

The ICC analysis of the DVH results showed a good correlation across the solutions. However, the OAR DVP analysis resulted in ranges of $10.2/7.6\,\%$ for bladder V28Gy (prostate I/II), $2.8\,\mathrm{Gy}$ for rectum D2cm^3 (cervix), $7.1\,\mathrm{Gy}$ for duodenum Dmax (liver) and up to $4.6\,\mathrm{Gy}$ for rectum D0.5cm^3 (lymph nodes). Several studies [160], [168], [241]recommend using different algorithms to estimate the accuracy of the dose accumulation application. Nevertheless, what conclusions can we draw from such differences as those mentioned above? The ground truth is unknown, and it is not known which of the algorithm results comes closest to the unknown truth. The algorithm results can yield a measure of the uncertainty in dose mapping. Focus should be on areas that directly and critically affect patients. Again, for the subsequent clinical evaluation, a multidisciplinary team is required, and more research on dose differences and their effects is required.

3.3 Application of the robust DDA solution in a clinical ART scenario

In relation to the third objective of this thesis [161], we developed a deformable dose accumulation workflow that allows retrospective comparison of treatment strategies. We demonstrated the workflow's feasibility and transferability by comparing conventional image-guided RT (conv-IGRT) and online adaptive MR-guided radiation therapy (OA-MRgRT) with the initial treatment plan (refPlan).

The analysis of the DVPs for the target and the OAR revealed that the two investigated DDA approaches, OA-MRgRT and conv-IGRT, resulted in comparable accumulated doses for the target volume, like the refPlan. Compared with refPlan, OA-MRgRT and conv-IGRT reduced median OAR DVPs; with the exception of bladder V56Gy, OA-MRgRT also resulted in lower median OAR DVPs than conv-IGRT. The comparison of different RT methods, such as MRgRT and IGRT, reveals a dosimetric sparing of the organs at risk with MRgRT, which confirms earlier study results [46], [103]. However, the selection of adaptive planning priorities is crucial. Specifically, it is about optimizing planning in terms of target volume coverage and minimizing the maximum dose to the OAR. The study by Xiong et al. [260], which demonstrated that OA-MRgRT also enhanced target coverage, reflects this. However, these studies all confirm the potential of OA-MRgRT.

Although there were significant differences in DVP assessment, the NTCP model showed no significant differences between the two RT approaches, OA-MRgRT and conv-IGRT, supporting the hypothesis of negligible clinical impact. The overall complications analyzed in the NTCP model were minimal. The NTCP model used the DVH as input to predict the probability of toxicity based on dose thresholds and tissue-specific parameters [37], [124], [125], [142]. These analytical models are mathematically efficient but may simplify the complexity of radiation-induced side effects [21], [237]. This includes the fact that the DVH does not capture the spatial dose distribution. Furthermore, different patients may respond differently due to patient-specific factors, such as genetics, even if they receive the same DVH curve. In addition, the method used did not take into account patient-reported toxicity outcomes, which may be important for a more comprehensive assessment of side effects. In this study, we treated the patients with twenty fractions of $3\,\mathrm{Gy}$ each. We assume that the differences will be greater when applying fewer RT fractions. In any case, further investigations are required to clarify these questions.

Two techniques for mapping and accumulating doses in the context of ART are currently under discussion, the "star" and the "cascade" techniques. The "star" technique maps all fractions to one reference fraction, such as the first fraction, while the "cascade" technique maps the previous fraction to the current fraction. The impact of registration uncertainty varies depending on the technique employed. The cascade approach propagates errors in previous registrations throughout the entire process. On the other hand, with the star technique, it is more difficult to register images taken towards the

end of the treatment, as these may show greater anatomical variations compared to the reference image. Furthermore, the sequence in which the doses are mapped can alter the accumulated dose outcome unless the procedure is both inverse-consistent and transitive [20]. A preliminary study [162] analyzed the differences between the two DDA mapping techniques, "start" and "cascade" techniques. While individual patients showed large variations, particularly in the maximum OAR dose levels, the mean differences between these approaches were only minor. This thesis utilized DDA's "star" techniques.

OA-MRgRT is one option of IGRT. A recently introduced integrated CT-LINAC for online CBCT-guided ART uses artificial intelligence and machine learning to create and adjust the treatment plan for highly targeted external beam RT [7], [131]. An iterative CBCT reconstruction process reduces noise and improves the image quality of the CBCT images, resulting in improved soft tissue presentation and dose calculation [131]. The developed DDA solution and workflow are feasible and transferable to this treatment system. However, it's important to note that, given the differences in contrast between CT and MR images, the DIR algorithm must also ensure a suitable DIR quality for CT image registration.

In this study, we performed the DDA retrospectively after the patient's irradiation was complete, using the "star" DDA approach. There was no time pressure in offline DDA, allowing for the testing and adjustment of various DIR parameter settings as needed to achieve a satisfactory registration. To perform the DIR and dose mapping steps, we had to export and load the data sets, which included images, structure sets, plans, and dose distributions, into our research system. Subsequently, we exported and analyzed the datasets. It was always possible to check the results of the individual steps for plausibility and quality using external tools. In this process, time did not play a decisive role. It is also conceivable to implement an online DDA workflow with the following scenario: From the second day of treatment, the previous fractional doses are accumulated offline; on the day of irradiation, the plan and dose are adapted to the day's anatomy, and then the accumulated dose is calculated with the current dose of the day and revised if necessary if criteria are not met. If DDA is to be integrated into the online workflow, several requirements must be met, as a short fraction of treatment time is crucial for the patient. Firstly, the software must integrate the DDA solution to prevent import-export. Moreover, the system must incorporate QA tools [160] to promptly determine and evaluate the registration outcome in a reasonable time. The system must also be straightforward to use.

3.4 Further Application of DDA

The second and third objective of this thesis were to investigate and develop a solution and workflow for deformable dose accumulation (DDA) in adaptive radiation therapy (RT). However, the DMAL from the first project shows additional use cases for dose mapping and accumulation. These other use cases face different challenges.

Treatment planning for 4D dose accumulation in lung cancer should explicitly account for intra-fraction motion, such as respiratory movement. However, the time required for 4D dose accumulation presents a challenge, as it necessitates recalculating the planned dose in several respiratory phases of a 4D CT, up to approximately 10 [236]. This approach benefits from accounting for the differences in tissue density between the first reference phase and the subsequent 4D CT phases. Studies have demonstrated that phase-based 4D dose planning is more accurate than three-dimensional (3D) dose planning based on average-intensity images due to the inclusion of respiratory movement [1], [67], [75], [78], [153], [236]. Recently, He et al. [85] demonstrated that 4DCT-based longitudinal dose accumulation in lung cancer patients provides a more accurate estimate of the delivered dose and this approach may be especially advantageous for patients who have a significant longitudinal response. However, it was found that most studies did not investigate DIR-based dosimetric uncertainties, and those that did report used different metrics across the studies to quantify these uncertainties [168].

When re-irradiating a patient due to recurrence, challenges arise with respect to anatomical changes over time, including surgical removal of an organ, e.g., prostatectomy. For instance, when radiating the affected lymph nodes in the pelvis, the dose from the previous RT for the prostate carcinoma must be considered. Calculating the total dose for the risk organs already exposed to radiation from the previous irradiation and the planned irradiation is the goal of dose accumulation during re-irradiation, along with ensuring compliance with the re-irradiation's dose limits. Studies found that re-irradiation requires individual assessment, including fractionation correction, and registration can provide relevant information on significant anatomical changes, offering the potential for a better understanding of accumulated OAR tolerances in the future [175], [184]. The recently published ESTRO-EORTC consensus statement on re-irradiation [6], [246] also confirms this. In 2022, an ESTRO physics workshop on "Reirradiation: Improving dose summation for plan optimization, assessment, and outcome analysis" took place. The workshop resulted in a comprehensive evaluation of the variability in estimating accumulated doses in scenarios involving re-irradiation. The findings revealed substantial variations in reporting accumulated doses, indicating a level of uncertainty. These variations have important implications for outcome analysis and the interpretation of published doses. Using a uniform technique that includes spatially mapped doses has the potential to improve accuracy in estimating accumulated dose in re-irradiation scenarios [84].

The goal of DMA in response assessment and outcome modeling is to link the dose received by an organ to a clinical outcome. Currently, two approaches are used: inter-patient and intra-patient modeling. For intra-patient, voxel-based analysis is one of the

few approaches. This is a challenge because the individual patient's anatomy, such as height and weight, varies much more than the variability between patients. An NTCP model is then created from the accumulated dose data and the reported side effects [151], [182]. Beasley et al. [19] presented an alternative technique for correlating dose-response relationships, image-based data mining.

Another area involves the combination or summation of various radiation treatments, such as external beam RT and brachytherapy (BT), commonly used in gynecologic cancer treatment. In this context, DIR faces large and complex deformations, including tumor shrinkage, bladder and rectum filling, BT application, and packing material. According to Swamidas et al. [226], DIR algorithms lack the necessary robustness to effectively manage the deformation complexity. Additionally, they noticed a broad range of uncertainty when utilizing DDA. A recently published study [271] concluded that utilizing DIR dose accumulation might serve as an alternative approach to demonstrate and assess the delivered doses received by the OARs in the context of combination treatment. Nevertheless, it is important to use DIR with caution in order to overcome or at least respect its limits [271].

3.5 Weaknesses and risks

This thesis identified weaknesses and risks of the DDA. While there are several commercial and open source DDA software solutions available, none, to the best of our knowledge, enable easy use and complete DDA implementation, let alone provide QA solutions. Furthermore, DIR and DDA solutions struggle with accuracy and uncertainty, failing to accurately capture complex deformations. These, in turn, have a direct effect on the accumulated dose, which can lead to an incorrect assessment of the delivered dose. Uncertainties in the accuracy of accumulated dose distributions may affect clinical decision-making regarding treatment adjustment (cf. DMAL 2.1.4 category "include" on impact of uncertainties in dose mapping). This, in turn, can have a negative impact on treatment and patient outcomes, such as underdosing the target volume or overdosing the OAR. Furthermore, the QA suffers from a lack of ground truth [160]. Although tools and metrics are available [29], the effective use of DIR requires expertise in the selection of suitable algorithms, the setting of parameters, and the interpretation of results. This also applies to evaluating subsequent dose accumulation results.

3.6 Outlook

Currently, DDA predominantly uses classic DIR algorithms that are based on mathematical models. These are based on techniques such as optimization, features and physical models. Optimizing these algorithms requires knowledge of how to construct features and adjust parameters. However, caution is required in this process. The complexity of the chosen mathematical models limits the algorithms. We can usually interpret the internal workings of classical algorithms as they are based on explicit mathematical formulations and principles.

Similar to other areas of RT such as automatic planning [3], [43], [44], [61], [107], [110], [214], [225] and contour generation [12], [100], [102], [167], [197], [206], deep learning (DL) models have also found their way into DIR. These models learn representations and patterns directly from the data. They use artificial neural networks with multiple layers to automatically extract features and learn complex mappings from inputs to outputs. In DIR, DL approaches use e.g. bayesian frameworks [52], [113], [114] implicit neural representations [251] or convolutional neural networks (CNNs) [13], [38], [64], [90], [101], [115], [122], [137], [140], [251], [257] such as VoxelMorph [13], [14] and U-NET [137] to directly learn from DVFs, resulting in a DL algorithm that directly determines the DVF. A second approach is to learn similarity metrics trained on image pairs [218], [254], [270]. In DL training, there is typically a distinction between supervised learning with available ground truth and unsupervised learning without additional data. DL models are flexible and can adapt to different data types and tasks. They can learn complex mappings and representations from large amounts of data without explicit feature engineering. However, they require data for training and can be computationally intensive for training and inference. DL models are often considered black boxes due to their complex architectures and the distributed nature of the learning representations. Therefore, understanding why a DL model makes a particular decision or prediction can be challenging.

For the second objective of this thesis [159], we used a gold standard case based on a biomechanical model to provide a ground truth [143]. For this purpose, we simulated the moving images by applying representative bladder and rectal fillings [27], resulting in known DVFs. This approach can serve as an alternative to the classic DIR algorithms by reversing the method. Biomechanical modeling aims to model the physical processes and characteristics that determine tissue deformation to achieve a more precise alignment that closely matches the actual reality. It offers a physics-based method for aligning images by simulating the mechanical behavior of tissues. The biomechanical model uses the segmented data, mesh, material properties, and boundary conditions for the simulation. It also uses a numerical method like the finite element method (FEM) [14], [143] to solve the biomechanical equations for tissue deformation. An alternative approach is to use kinematic body models in conjunction with a mass-spring model [170] or a chainmail algorithm [17], [229], [240] to deform the soft tissue. The results of the biomechanical simulation derive a DVF that describes how each point in the moving image should be transformed to match the fixed image. Using physical principles for registration leads to more realistic and anatomically faithful alignments. However, this approach is more computationally intensive and requires detailed anatomical and material data.

3.7 Conclusion

Dose accumulation is and will be a challenge in RT. Aiming for context-driven DIR rather than an idealized, perfect one is important. This approach might necessitate favoring local registration accuracy over global accuracy or using multiple registrations tailored to specific organs or applications. Before using the DIR algorithm, it is important to know what its limitations are. This is necessary to fully grasp how DIR and its related DMA uncertainties affect clinical decisions and assessments [160]. For certain applications, it is crucial to consider that DMA, provided the uncertainties are properly quantified, may be a more advantageous alternative than not using DMA at all.

The analysis of various DIR and DDA implementations revealed a significant level of agreement among them [159]. Nevertheless, there were noticeable variations in the absolute differences, which might have substantial therapeutic implications, depending on the specific clinical scenarios and algorithms employed.

The retrospective comparison of OA-MRgRT and conv-IGRT in PC patients demonstrated the feasibility and transferability of the established DDA-method and -workflow [161]. The study revealed significant dosimetric differences. Nevertheless, the small effect size in applying long-term treatments did not result in clinical relevance for NTCP. However, therapeutic effects may be more pronounced in hypo-fractionated RT.

Despite the challenges of DMA, algorithms from various institutions demonstrated high consistency, although with case- and algorithm-dependent variations that could potentially have significant clinical impact. In comparison to conv-IGRT, the deformably accumulated OA-MRgRT dose revealed significant but small dosimetric differences. We demonstrated the feasibility and transferability of a dose accumulation methodology, which could be relevant for evaluating future hypo-fractionated OA-MRgRT approaches.

4 Summary

Cancer is a common disease worldwide and the second cause of death in Europe. In addition to chemotherapy and surgery, radiation therapy (RT) is used to kill cancer cells by delivering a high dose of radiation. Today, innovative adaptive RT techniques, such as image-guided RT (IGRT), enable precise treatment. However, IGRT, such as cone beam computed tomography, often does not account for deformations between fractions and volume variations, leading to uncertainties in dose delivery. In recent years, developments in RT, including the combined magnetic resonance linear accelerator (MR-LINAC) system, have revolutionized adaptive RT. Online adaptive magnetic resonance-guided RT (OA-MRgRT) allows daily adaptation of the treatment plan based on the anatomy of the respective day. Despite these technological innovations, such as daily anatomical imaging and plan adaptation, the accurate calculation of the total dose delivered to the patient during treatment remains a challenge due to the deformations between each fraction; this underlines the urgent need for robust deformable dose accumulation (DDA) solutions to accurately calculate the total dose delivered to the target volume and surrounding tissue.

The first objective of this thesis was to investigate the applicability and use of dose mapping and accumulation (DMA) in RT. This first phase involved the identification and classification of current DMA applications in RT to develop a comprehensive landscape, including the description of strengths and limitations. We then identified and analyzed the significant barriers to clinical implementation and broad use for dose accumulation. This phase included a thorough review of existing literature, case studies, and implementations to identify the current use of DMA. We closely examined the identified barriers, which included technical, methodological, and biological issues, to understand their causes and their impact on the applicability of DMA. This detailed analysis led to the development of comprehensive guidelines and requirements tailored to researchers, healthcare providers and manufacturers. These guidelines address the identified barriers by providing practical solutions and best practices to overcome them.

Deformable dose accumulation includes deformable image registration, corresponding dose mapping, and dose summation. Different mathematical models are available for the first two applications. The second study of this dissertation involved the development and investigation of a robust DDA solution, along with a comparison the DDA implementations from other research groups. This was achieved by conducting a multicenter study evaluating several solutions in different anatomical areas. The study evaluated a gold standard case that served as known ground truth and analyzed five clinical cases treated with online adaptive MRgRT. The study showed significant agreement between the different implementations, but also revealed differences that depend on specific cases and algorithms and may have significant consequences in the therapeutic context. The study also showed that our proposed DDA solution worked consistently and reliably, producing results equivalent to those of the other algorithms.

Adaptive RT, such as IGRT and OA-MRgRT, aims to account for inter-fraction anatomic

variations that differ from situation during treatment planning. Fractionated irradiation requires the use of DDA to compare different adaptive approaches in terms of the total dose delivered to the tumor and surrounding tissue. In the third objective of this thesis, we used the deformed dose accumulation solution from the second objective of this thesis to investigate the accumulated doses in a clinical study treating patient with prostate cancer with adaptive, moderately hypo-fractionated RT. The total deformably accumulated doses of two adaptive treatment strategies - conventional IGRT (conv-IGRT) and OA-MRgRT - and the treatment plan generated during simulation were compared dosimetrically to investigate the hypothesis that OA-MRgRT allows more effective sparing of organs-at-risk compared to reference planning and conv-IGRT. All techniques, including reference planning, conv-IGRT, and OA-MRgRT, showed similar mean doses in the target volume. Although all accumulated approaches resulted in lower maximum organ-at-risk doses compared to the treatment plan, OA-MRgRT in particular showed encouraging results compared to conv-IGRT. However, the dosimetric differences did not translate into clinical relevance according to normal tissue complication probability models. Overall, the study demonstrated feasibility and transferability of the established DDA method and workflow.

This thesis provided a comprehensive methodological basis for the development of a dedicated DDA approach and applied it for the first time to a clinical study dataset using DDA in fractionated prostate cancer RT. Clinical evaluation and further research focusing on dose differences and their effects are currently being developed to initiate the next phase of OA-MRgRT, which may include online or real-time dose accumulation in order to further personalize cancer treatment with RT.

Location, Date, Signature Prof. Dr. rer. nat. Daniela Thorwarth

5 Zusammenfassung

Krebs ist eine weltweit verbreitete Krankheit und die zweithäufigste Todesursache in Europa. Neben der Chemotherapie und der Chirurgie wird die Strahlentherapie (RT) eingesetzt, um Krebszellen durch Abgabe einer hohen Strahlendosis abzutöten. Heute ermöglichen innovative adaptive RT-Techniken wie die bildgesteuerte RT (IGRT) eine präzise Behandlung. Allerdings werden bei der IGRT, z. B. bei der Kegelstrahl-Computertomographie, Verformungen zwischen den Fraktionen und Volumenschwankungen oft nicht berücksichtigt, was zu Unsicherheiten bei der Dosisabgabe führt. In den letzten Jahren haben Entwicklungen in der RT, einschließlich des kombinierten Magnetresonanz-Linearbeschleuniger Systems (MR-LINAC), die adaptive RT revolutioniert. Die online adaptiven Magnetresonanztomographie-geführte RT (OA-MRgRT) ermöglicht eine tägliche Anpassung des Behandlungsplans auf der Grundlage der Anatomie des jeweiligen Tages. Trotz dieser technologischen Innovationen, wie der täglichen anatomischen Bildgebung und der Plananpassung, bleibt die genaue Berechnung der Gesamtdosis, die dem Patienten während der Behandlung verabreicht wird, aufgrund der Verformungen zwischen den einzelnen Fraktionen eine Herausforderung; dies unterstreicht den dringenden Bedarf an robusten Lösungen für die deformierbare Dosisakkumulation (DDA) zur genauen Berechnung der Gesamtdosis, die dem Zielvolumen und dem umgebenden Gewebe verabreicht wird.

Das erste Ziel dieser Arbeit bestand darin, die Anwendbarkeit und den Einsatz von Dosiskartierung und -akkumulation (DMA) in der RT zu untersuchen. In dieser ersten Phase wurden die aktuellen DMA-Anwendungen in der RT identifiziert und klassifiziert, um eine umfassende Landschaft zu entwickeln, einschließlich der Beschreibung von Stärken und Grenzen. Anschließend wurden die wesentlichen Hindernisse für die klinische Umsetzung und die breite Anwendung der Dosisakkumulation ermittelt und analysiert. Diese Phase umfasste eine gründliche Durchsicht der vorhandenen Literatur, Fallstudien und Implementierungen, um den derzeitigen Einsatz von DMA zu ermitteln. Wir untersuchten die identifizierten Hindernisse, zu denen technische, methodische und biologische Fragen gehörten, genau, um ihre Ursachen und ihre Auswirkungen auf die Anwendbarkeit der DMA zu verstehen. Diese detaillierte Analyse führte zur Entwicklung umfassender Leitlinien und Anforderungen, die auf Forscher, Gesundheitsdienstleister und Hersteller zugeschnitten sind. Diese Leitlinien befassen sich mit den ermittelten Hindernissen, indem sie praktische Lösungen und bewährte Verfahren zu ihrer Überwindung anbieten.

Die deformierbare Dosisakkumulation umfasst die deformierbare Bildregistrierung (DIR), die entsprechende Dosiskartierung und die Dosissummierung. Für die ersten beiden Anwendungen gibt es unterschiedliche mathematische Modelle. Die zweite Studie im Rahmen dieser Dissertation befasste sich mit der Entwicklung und Untersuchung einer robusten DDA-Lösung sowie mit einem Vergleich der DDA-Implementierungen anderer Forschungsgruppen. Zu diesem Zweck wurde eine multizentrische Studie durchgeführt, in der mehrere Lösungen in verschiedenen anatomischen Bereichen evaluiert wurden. In der Studie wurde ein Goldstandard-Fall ausgewertet, der als bekannte

Grundwahrheit diente, und es wurden fünf klinische Fälle analysiert, die mit der online adaptiven MRgRT (OA-MRgRT) behandelt wurden. Die Studie zeigte eine signifikante Übereinstimmung zwischen den verschiedenen Implementierungen, aber auch Unterschiede, die von spezifischen Fällen und Algorithmen abhängen und im therapeutischen Kontext erhebliche Konsequenzen haben können. Die Studie zeigte auch, dass die von uns vorgeschlagene DDA-Lösung konsistent und zuverlässig funktioniert und gleichwertige Ergebnisse wie die anderen Algorithmen liefert.

Adaptive RT, wie IGRT und OA-MRgRT, zielt darauf ab, anatomische Variationen zwischen den Fraktionen zu berücksichtigen Variationen zu berücksichtigen, die von der Situation während der Behandlungsplanung abweichen. Die fraktionierte Bestrahlung erfordert den Einsatz von DDA, um verschiedene adaptive Ansätze in Bezug auf die Gesamtdosis für den Tumor und das umliegende Gewebe zu vergleichen. In der dritten Zielsetzung dieser Arbeit haben wir die deformierte Dosisakkumulationslösung aus der zweiten Zielsetzung dieser Arbeit verwendet, um die akkumulierten Dosen in einer klinischen Studie zu untersuchen, in der Patienten mit Prostatakrebs mit adaptiver, moderat hypo-fraktionierter RT behandelt wurden. Die deformierbar akkumulierten Gesamtdosen von zwei adaptiven Behandlungsstrategien - konventionelle IGRT (conv-IGRT) und OA-MRgRT - und der während der Simulation generierte Behandlungsplan wurden dosimetrisch verglichen, um die Hypothese zu untersuchen, dass OA-MRgRT im Vergleich zur Referenzplanung und conv-IGRT eine effektivere Schonung der Risikoorgane ermöglicht. Alle Techniken, einschließlich Referenzplanung, conv-IGRT und OA-MRgRT, zeigten ähnliche mittlere Dosen im Zielvolumen. Obwohl alle akkumulierten Ansätze im Vergleich zum Behandlungsplan zu niedrigeren maximalen Organ-Risiko-Dosen führten, zeigte insbesondere die OA-MRgRT im Vergleich zur conv-IGRT ermutigende Ergebnisse. Die dosimetrischen Unterschiede führten jedoch nicht zu einer klinischen Relevanz gemäß den Modellen für die Wahrscheinlichkeit von Komplikationen im Normalgewebe. Insgesamt zeigte die Studie die Durchführbarkeit und Übertragbarkeit der etablierten DDA-Methode und des Arbeitsablaufs.

Diese Arbeit lieferte eine umfassende methodische Grundlage für die Entwicklung eines speziellen DDA-Ansatzes und wendete diesen erstmals auf einen klinischen Studiendatensatz an, der DDA bei der fraktionierten Prostatakrebs-RT verwendet. Die klinische Bewertung und weitere Forschung mit Schwerpunkt auf Dosisunterschieden und deren Auswirkungen werden derzeit entwickelt, um die nächste Phase der OA-MRgRT einzuleiten, die eine Online- oder Echtzeit-Dosisakkumulation umfassen könnte, um die Krebsbehandlung mit RT weiter zu personalisieren.

6 Publications related to the dissertation

1. Applicability and usage of dose mapping/accumulation in radiotherapy

   Martina Murr, Kristy K. Brock, Marco Fusella, Nicholas Hardcastle, Mohammad Hussein, Michael G. Jameson, Isak Wahlstedt, Johnson Yuen, Jamie R. McClelland, Eliana Vasquez Osorio

   Published in: Radiotherapy and Oncology 2023 Vol. 182, Pages 109527

2. Review and recommendations on deformable image registration uncertainties for radiotherapy applications

   Lena Nenoff, Florian Amstutz, **Martina Murr**, Ben Archibald-Heeren , Marco Fusella , Mohammad Hussein, Wolfgang Lechner , Ye Zhang , Greg Sharp, Eliana Vasquez Osorio

   Published in: Physics in Medicine & Biology 2023, Vol. 68, Pages 24TR01

3. A multi-institutional comparison of retrospective deformable dose accumulation for online adaptive MR-guided radiotherapy

   Martina Murr, Uffe Bernchou, Edyta Bubula-Rehm, Mark Ruschin, Parisa Sadeghi, Peter Voet, Jeff D. Winter, Jinzhong Yang, Eyesha Younus, Cornel Zachiu, Yao Zhao, Hualiang Zhong, Daniela Thorwarth

   Published in: Physics and Imaging in Radiation Oncology 2024 Vol. 30, Pages 100588

4. Comparison of online adaptive and non-adaptive magnetic resonance image-guided radiation therapy in prostate cancer using dose accumulation

   Martina Murr, Daniel Wegener, Simon Böke, Cihan Gani, David Mönnich, Maximilian Niyazi, Moritz Schneider, Daniel Zips, Arndt-Christian Müller, Daniela Thorwarth

   Submitted to: Physics and Imaging in Radiation Oncology 2024, manuscript number: PHIRO-D-24-00218, current status: Revise

7 References

[1] M. A. Admiraal, D. Schuring, and C. W. Hurkmans, Dose calculations accounting for breathing motion in stereotactic lung radiotherapy based on 4D-CT and the internal target volume, *Radiat Oncol*, vol. 86, no. 1, pp. 55–60, 2008. DOI: 10.1016/j.radonc.2007.11.022.

[2] I. A. E. Agency, Radiation Oncology Physics. International Atomic Energy Agency, 2005.

[3] S. H. Ahn, E. Kim, C. Kim, W. Cheon, M. Kim, S. B. Lee, *et al.*, Deep learning method for prediction of patient-specific dose distribution in breast cancer, *Radiat Oncol*, vol. 16, no. 1, p. 154, 2021. DOI: 10.1186/s13014-021-01864-9.

[4] S. Alam, H. Veeraraghavan, K. Tringale, E. Amoateng, E. Subashi, A. J. Wu, *et al.*, Inter- and intrafraction motion assessment and accumulated dose quantification of upper gastrointestinal organs during magnetic resonance-guided ablative radiation therapy of pancreas patients, *Phys Imaging Radiat Oncol*, vol. 21, pp. 54–61, 2022. DOI: 10.1016/j.phro.2022.02.007.

[5] L. M. Amugongo, A. Green, D. Cobben, M. van Herk, A. McWilliam, and E. V. Osorio, Identification of modes of tumor regression in non-small cell lung cancer patients during radiotherapy, *Med Phys*, vol. 49, no. 1, pp. 370–381, 2022. DOI: 10.1002/mp.15320.

[6] N. Andratschke, J. Willmann, A. L. Appelt, N. Alyamani, P. Balermpas, B. G. Baumert, *et al.*, European Society for Radiotherapy and Oncology and European Organisation for Research and Treatment of Cancer consensus on re-irradiation: definition, reporting, and clinical decision making, *Lancet Oncol*, vol. 23, no. 10, e469–e478, 2022. DOI: 10.1016/S1470-2045(22)00447-8.

[7] Y. Archambault, C. Boylan, D. Bullock, T. Morgas, J. Peltola, E. Ruokokoski, *et al.*, Making on-line adaptive radiotherapy possible using artificial intelligence and machine learning for efficient daily re-planning, *Med Phys Int J*, vol. 8, 2020.

[8] S. Armstrong and P. Hoskin, Complex Clinical Decision-Making Process of Re-Irradiation, *Clin Oncol*, vol. 32, no. 11, pp. 688–703, 2020. DOI: 10.1016/j.clon.2020.07.023.

[9] N. v. As, A. Tree, J. Patel, P. Ostler, H. V. D. Voet, D. A. Loblaw, *et al.*, 5-Year Outcomes from PACE B: An International Phase III Randomized Controlled Trial Comparing Stereotactic Body Radiotherapy (SBRT) vs. Conventionally Fractionated or Moderately Hypo Fractionated External Beam Radiotherapy for Localized Prostate Cancer, *Int J Radiat Oncol Biol Phys*, vol. 117, no. 4, e2–e3, 2023. DOI: 10.1016/j.ijrobp.2023.08.027.

[10] J. Ashburner, A fast diffeomorphic image registration algorithm, *NeuroImage*, vol. 38, no. 1, pp. 95–113, 2007. DOI: 10.1016/j.neuroimage.2007.07.007.

[11] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee, Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain, *Med Image Anal*, vol. 12, no. 1, pp. 26–41, 2008. DOI: 10.1016/j.media.2007.06.004.

[12] A. Balagopal, M. Dohopolski, Y. S. Kwon, S. Montalvo, H. Morgan, T. Bai, *et al.*, Deep learning based automatic segmentation of the Internal Pudendal Artery in definitive radiotherapy treatment planning of localized prostate cancer, *Phys Imaging Radiat Oncol*, vol. 0, no. 0, 2024. DOI: 10.1016/j.phro.2024.100577.

[13] G. Balakrishnan, A. Zhao, M. R. Sabuncu, A. V. Dalca, and J. Guttag, An Unsupervised Learning Model for Deformable Medical Image Registration, in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 9252–9260. DOI: 10.1109/CVPR.2018.00964.

[14] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca, VoxelMorph: A Learning Framework for Deformable Medical Image Registration, *IEEE Trans. Med. Imaging*, vol. 38, no. 8, pp. 1788–1800, 2019. DOI: 10.1109/TMI.2019.2897538.

[15] J. Barber, J. Yuen, M. Jameson, L. Schmidt, J. Sykes, A. Gray, *et al.*, Deforming to Best Practice: Key considerations for deformable image registration in radiotherapy, *J Med Radiat Sci*, vol. 67, no. 4, pp. 318–332, 2020. DOI: 10.1002/jmrs.417.

[16] H. Baroudi, K. K. Brock, W. Cao, X. Chen, C. Chung, L. E. Court, *et al.*, Automated Contouring and Planning in Radiation Therapy: What Is 'Clinically Acceptable'?, *Diagnostics*, vol. 13, no. 4, p. 667, 2023. DOI: 10.3390/diagnostics13040667.

[17] C. J. Bauer, H. Teske, A. Walter, P. Hoegen, S. Adeberg, J. Debus, *et al.*, Biofidelic image registration for head and neck region utilizing an in-silico articulated skeleton as a transformation model, *Phys Med Biol*, vol. 68, no. 9, p. 095 006, 2023. DOI: 10.1088/1361-6560/acc7f1.

[18] M. Baumann, C. Petersen, and M. Krause, TCP and NTCP in preclinical and clinical research in Europe, *Rays*, vol. 30, pp. 121–6, 2005.

[19] W. Beasley, M. Thor, A. McWilliam, A. Green, R. Mackay, N. Slevin, *et al.*, Image-based Data Mining to Probe Dosimetric Correlates of Radiation-induced Trismus, *Int J Radiat Oncol Biol Phys*, vol. 102, no. 4, pp. 1330–1338, 2018. DOI: 10.1016/j.ijrobp.2018.05.054.

[20] E. T. Bender, N. Hardcastle, and W. A. Tomé, On the dosimetric effect and reduction of inverse consistency and transitivity errors in deformable image registration for dose accumulation, *Med Phys*, vol. 39, no. 1, pp. 272–280, 2012. DOI: 10.1118/1.3666948.

[21] S. M. Bentzen, L. S. Constine, J. O. Deasy, A. Eisbruch, A. Jackson, L. B. Marks, *et al.*, Quantitative Analyses of Normal Tissue Effects in the Clinic (QUANTEC): An Introduction to the Scientific Issues, *Int J Radiat Oncol Biol Phys*, vol. 76, no. 3, S3–S9, 2010. DOI: 10.1016/j.ijrobp.2009.09.040.

[22] J. Bertholet, G. Anastasi, D. Noble, A. Bel, R. van Leeuwen, T. Roggen, *et al.*, Patterns of practice for adaptive and real-time radiation therapy (POP-ART RT) part II: Offline and online plan adaption for interfractional changes, *Radiother Oncol*, vol. 153, pp. 88–96, 2020. DOI: 10.1016/j.radonc.2020.06.017.

[23] U. Björeland, J. Jonsson, M. Alm, L. Beckman, T. Nyholm, and C. Thellenberg-Karlsson, Inter-fraction movements of the prostate and pelvic lymph nodes during IGRT, *J Radiat Oncol*, vol. 7, no. 4, pp. 357–366, 2018. DOI: 10.1007/s13566-018-0366-3.

[24] M. Bleeker, M. C. C. M. Hulshof, A. Bel, J.-J. Sonke, and A. v. d. Horst, Gastric deformation models for adaptive radiotherapy: Personalized vs population-based strategy, *Radiother Oncol*, vol. 166, pp. 126–132, 2022. DOI: 10.1016/j.radonc.2021.11.028.

[25] O. Bohoudi, A. M. E. Bruynzeel, S. Tetar, B. J. Slotman, M. A. Palacios, and F. J. Lagerwaard, Dose accumulation for personalized stereotactic MR-guided adaptive radiation therapy in prostate cancer, *Radiother Oncol*, vol. 157, pp. 197–202, 2021. DOI: 10.1016/j.radonc.2021.01.022.

[26] E. Boman, M. Kapanen, L. Pickup, and S.-L. Lahtela, Importance of deformable image registration and biological dose summation in planning of radiotherapy retreatments, *Med Dosim*, vol. 42, no. 4, pp. 296–303, 2017. DOI: 10.1016/j.meddos.2017.06.006.

[27] L. S. Bosma, C. Zachiu, M. Ries, B. D. d. Senneville, and B. W. Raaymakers, Quantitative investigation of dose accumulation errors from intra-fraction motion in MRgRT for prostate cancer, *Phys Med Biol*, vol. 66, no. 6, p. 065 002, 2021. DOI: 10.1088/1361-6560/abe02a.

[28] L. S. Bosma, M. Ries, B. Denis de Senneville, B. W. Raaymakers, and C. Zachiu, Integration of operator-validated contours in deformable image registration for dose accumulation in radiotherapy, *Phys Imaging Radiat Oncol*, vol. 27, p. 100 483, 2023. DOI: 10.1016/j.phro.2023.100483.

[29] L. S. Bosma, M. Hussein, M. G. Jameson, S. Asghar, K. K. Brock, J. R. McClelland, *et al.*, Tools and recommendations for commissioning and quality assurance of deformable image registration in radiotherapy, *Phys Imaging Radiat Oncol*, vol. 32, 2024. DOI: 10.1016/j.phro.2024.100647.

[30] T. Bostel, I. Sachpazidis, M. Splinter, N. Bougatf, T. Fechter, C. Zamboglou, *et al.*, Dosimetric Impact of Interfractional Variations in Prostate Cancer Radiotherapy—Implications for Imaging Frequency and Treatment Adaptation, *Front Oncol*, vol. 9, 2019.

[31] J. Boustani, M. Grapin, P.-A. Laurent, L. Apetoh, and C. Mirjolet, The 6th R of Radiobiology: Reactivation of Anti-Tumor Immune Response, *Cancers*, vol. 11, no. 6, p. 860, 2019. DOI: 10.3390/cancers11060860.

[32] A. Brahme, Optimization of stationary and moving beam radiation therapy techniques, *Radiother Oncol*, vol. 12, no. 2, pp. 129–140, 1988. DOI: 10.1016/0167-8140(88)90167-3.

[33] D. H. Brand, S. C. Brüningk, A. Wilkins, O. Naismith, A. Gao, I. Syndikus, *et al.*, The Fraction Size Sensitivity of Late Genitourinary Toxicity: Analysis of Alpha/Beta (α/β) Ratios in the CHHiP Trial, *Int J Radiat Oncol Biol Phys*, vol. 115, no. 2, pp. 327–336, 2023. DOI: 10.1016/j.ijrobp.2022.08.030.

[34] K. K. Brock, S. Mutic, T. R. McNutt, H. Li, and M. L. Kessler, Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the AAPM Radiation Therapy Committee Task Group No. 132, *Med Phys*, vol. 44, no. 7, e43–e76, 2017. DOI: 10.1002/mp.12256.

[35] E. D. Brooks, X. Wang, B. De, V. Verma, T. D. Williamson, R. Hunter, *et al.*, An algorithm for thoracic re-irradiation using biologically effective dose: a common language on how to treat in a "no-treat zone", *Radiat Oncol*, vol. 17, no. 1, p. 4, 2022. DOI: 10.1186/s13014-021-01977-1.

[36] C. L. Brouwer, R. J. H. M. Steenbakkers, J. A. Langendijk, and N. M. Sijtsema, Identifying patients who may benefit from adaptive radiotherapy: Does the literature on anatomic and dosimetric changes in head and neck organs at risk during radiotherapy provide information to help?, *Radiother Oncol*, vol. 115, no. 3, pp. 285–294, 2015. DOI: 10.1016/j.radonc.2015.05.018.

[37] C. Burman, G. J. Kutcher, B. Emami, and M. Goitein, Fitting of normal tissue tolerance data to an analytic function, *Int J Radiat Oncol Biol Phys*, vol. 21, no. 1, pp. 123–135, 1991. DOI: 10.1016/0360-3016(91)90172-Z.

[38] X. Cao, J. Yang, J. Zhang, Q. Wang, P.-T. Yap, and D. Shen, Deformable Image Registration Using a Cue-Aware Deep Regression Network, *IEEE Trans Biomed Eng*, vol. 65, no. 9, pp. 1900–1911, 2018. DOI: 10.1109/TBME.2018.2822826.

[39] M. J. Cardoso, M. Modat, T. Vercauteren, and S. Ourselin, Scale Factor Point Spread Function Matching: Beyond Aliasing in Image Resampling, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. Frangi, Eds., Cham: Springer International Publishing, 2015, pp. 675–683. DOI: 10.1007/978-3-319-24571-3_81.

[40] L. Cella, S. Monti, T. Xu, R. Liuzzi, A. Stanzione, M. Durante, *et al.*, Probing thoracic dose patterns associated to pericardial effusion and mortality in patients treated with photons and protons for locally advanced non-small-cell lung cancer, *Radiother Oncol*, vol. 160, pp. 148–158, 2021. DOI: 10.1016/j.radonc.2021.04.025.

[41] M. Chao, J. Penagaricano, Y. Yan, E. G. Moros, P. Corry, and V. Ratanatharathorn, Voxel-Based Dose Reconstruction for Total Body Irradiation With Helical TomoTherapy, *Int J Radiat Oncol Biol Phys*, vol. 82, no. 5, pp. 1575–1583, 2012. DOI: 10.1016/j.ijrobp.2011.01.021.

[42] J. Chen, J.-P. Bissonnette, T. Craig, P. Munoz-Schuffenegger, T. Tadic, L. A. Dawson, *et al.*, Liver SBRT dose accumulation to assess the impact of anatomic variations on normal tissue doses and toxicity in patients treated with concurrent sorafenib, *Radiother Oncol*, vol. 182, p. 109 588, 2023. DOI: 10.1016/j.radonc.2023.109588.

[43] X. Chen, K. Men, Y. Li, J. Yi, and J. Dai, A feasibility study on an automated method to generate patient-specific dose distributions for radiotherapy using deep learning, *Med Phys*, vol. 46, no. 1, pp. 56–64, 2019. DOI: 10.1002/mp.13262.

[44] X. Chen, K. Men, J. Zhu, B. Yang, M. Li, Z. Liu, *et al.*, DVHnet: A deep learning-based prediction of patient-specific dose volume histograms for radiotherapy planning, *Med Phys*, vol. 48, no. 6, pp. 2705–2713, 2021. DOI: 10.1002/mp.14758.

[45] I. J. Chetty and M. Rosu-Bubulac, Deformable Registration for Dose Accumulation, *Semin Radiat Oncol*, vol. 29, no. 3, pp. 198–208, 2019. DOI: 10.1016/j.semradonc.2019.02.002.

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

[47] I. R. O. T. Committee, *Technical Frameworks, IHE Radiation Oncology, Deformable Registration in Radiation Oncology (DRRO)*, 2022.

[48] I. Coric, K. Shreshtha, T. Roque, N. Paragios, C. Gani, D. Zips, *et al.*, Dosimetric Evaluation of Dose Calculation Uncertainties for MR-Only Approaches in Prostate MR-Guided Radiotherapy, *Front Phys*, vol. 10, p. 897 710, 2022. DOI: 10.3389/fphy.2022.897710.

[49] C. H. Crane, Balancing Fractionation and Advanced Technology in Consideration of Reirradiation, *Semin Radiat Oncol*, vol. 30, no. 3, pp. 201–203, 2020. DOI: 10.1016/j.semradonc.2020.02.009.

[50] W. R. Crum, T. Hartkens, and D. L. G. Hill, Non-rigid image registration: theory and practice, *Brit J Radiol*, vol. 77, no. suppl_2, S140–S153, 2004. DOI: 10.1259/bjr/25329214.

[51] V. Delmon, S. Rit, R. Pinho, and D. Sarrut, Registration of sliding objects using direction dependent B-splines decomposition, *Phys Med Biol*, vol. 58, no. 5, p. 1303, 2013. DOI: 10.1088/0031-9155/58/5/1303.

[52] V. S. Deshpande and J. S. Bhatt, Bayesian Deep Learning for Deformable Medical Image Registration, in *Pattern Recognition and Machine Intelligence*, B. Deka, P. Maji, S. Mitra, D. K. Bhattacharyya, P. K. Bora, and S. K. Pal, Eds., Cham: Springer International Publishing, 2019, pp. 41–49. DOI: 10.1007/978-3-030-34872-4_5.

[53] L. R. Dice, Measures of the Amount of Ecologic Association Between Species, *Ecol*, vol. 26, no. 3, pp. 297–302, 1945. DOI: 10.2307/1932409.

[54] N. A. Dodgson, Image resampling, University of Cambridge, Computer Laboratory, UCAM-CL-TR-261, 1992. DOI: 10.48456/tr-261.

[55] J. L. Donovan, F. C. Hamdy, J. A. Lane, M. Mason, C. Metcalfe, E. Walsh, *et al.*, Patient-Reported Outcomes after Monitoring, Surgery, or Radiotherapy for Prostate Cancer, *N Engl J Med*, vol. 375, no. 15, pp. 1425–1437, 2016. DOI: 10.1056/NEJMoa1606221.

[56] D. Drobny, H. Carolus, S. Kabus, and J. Modersitzki, Handling Non-Corresponding Regions in Image Registration, in *Bildverarbeitung für die Medizin 2015*, H. Handels, T. M. Deserno, H.-P. Meinzer, and T. Tolxdorff, Eds., Berlin, Heidelberg: Springer, 2015, pp. 107–112. DOI: 10.1007/978-3-662-46224-9_20.

[57] S. Dutta, A. Dewan, S. Mitra, M. K. Sharma, S. Aggarwal, S. Barik, *et al.*, Dosimetric impact of variable bladder filling on IMRT planning for locally advanced carcinoma cervix, *J Egypt Natl Canc Inst*, vol. 32, no. 1, p. 31, 2020. DOI: 10.1186/s43046-020-00033-5.

[58] B. Eiben, J. Bertholet, M. J. Menten, S. Nill, U. Oelfke, and J. R. McClelland, Consistent and invertible deformation vector fields for a breathing anthropomorphic phantom: a post-processing framework for the XCAT phantom, *Phys Med Biol*, vol. 65, no. 16, p. 165 005, 2020. DOI: 10.1088/1361-6560/ab8533.

[59] B. Emami, J. Lyman, A. Brown, L. Cola, M. Goitein, J. E. Munzenrider, *et al.*, Tolerance of normal tissue to therapeutic irradiation, *Int J Radiat Oncol Biol Phys*, vol. 21, no. 1, pp. 109–122, 1991. DOI: 10.1016/0360-3016(91)90171-Y.

[60] G. M. Engeseth, C. Stokkevåg, and L. P. Muren, Achievements and challenges in normal tissue response modelling for proton therapy, *Phys Imaging Radiat Oncol*, vol. 24, pp. 118–120, 2022. DOI: 10.1016/j.phro.2022.11.004.

[61] J. Fan, J. Wang, Z. Chen, C. Hu, Z. Zhang, and W. Hu, Automatic treatment planning based on three-dimensional dose distribution predicted from deep learning technique, *Med Phys*, vol. 46, no. 1, pp. 370–381, 2019. DOI: 10.1002/mp.13271.

[62] M. F. Fast, M. Cao, P. Parikh, and J.-J. Sonke, Intrafraction Motion Management With MR-Guided Radiation Therapy, *Semin Radiat Oncol*, vol. 34, no. 1, pp. 92–106, 2024. DOI: 10.1016/j.semradonc.2023.10.008.

[63] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J.-C. Fillion-Robin, S. Pujol, *et al.*, 3D Slicer as an image computing platform for the Quantitative Imaging Network, *Magn Reson Imaging*, vol. 30, no. 9, pp. 1323–1341, 2012. DOI: 10.1016/j.mri.2012.05.001.

[64] E. Ferrante, O. Oktay, B. Glocker, and D. H. Milone, On the Adaptability of Unsupervised CNN-Based Deformable Image Registration to Unseen Image Domains, in *Machine Learning in Medical Imaging*, Y. Shi, H.-I. Suk, and M. Liu, Eds., Cham: Springer International Publishing, 2018, pp. 294–302. DOI: 10.1007/978-3-030-00919-9_34.

[65] J. F. Fowler, Biological Factors Influencing Optimum Fractionation in Radiation Therapy, *Acta Oncol*, vol. 40, no. 6, pp. 712–717, 2001. DOI: 10.1080/02841860152619124.

[66] A. Frederick, S. Quirk, P. Grendarova, L. v. Dyke, T. Meyer, S. Weppler, *et al.*, An updated approach for deriving PTV margins using image guidance and deformable dose accumulation, *Phys Med Biol*, vol. 67, no. 7, p. 075 004, 2022. DOI: 10.1088/1361-6560/ac5ce5.

[67] P. Freislederer, A. von Münchow, F. Kamp, C. Heinz, S. Gerum, S. Corradini, *et al.*, Comparison of planned dose on different CT image sets to four-dimensional Monte Carlo dose recalculation using the patient's actual breathing trace for lung stereotactic body radiation therapy, *Med Phys*, vol. 46, no. 7, pp. 3268–3277, 2019. DOI: 10.1002/mp.13579.

[68] N. Fukumitsu, K. Nitta, T. Terunuma, T. Okumura, H. Numajiri, Y. Oshiro, *et al.*, Registration error of the liver CT using deformable image registration of MIM Maestro and Velocity AI, *BMC Med Imaging*, vol. 17, no. 1, p. 30, 2017. DOI: 10.1186/s12880-017-0202-z.

[69] R. García-Mollá, N. d. Marco-Blancas, J. Bonaque, L. Vidueira, J. López-Tarjuelo, and J. Perez-Calatayud, Validation of a deformable image registration produced by a commercial treatment planning system in head and neck, *Phys Med*, vol. 31, no. 3, pp. 219–223, 2015. DOI: 10.1016/j.ejmp.2015.01.007.

[70] B. J. Gebhardt, J. A. Vargo, D. Ling, B. Jones, M. Mohney, D. A. Clump, *et al.*, Carotid Dosimetry and the Risk of Carotid Blowout Syndrome After Reirradiation With Head and Neck Stereotactic Body Radiation Therapy, *Int J Radiat Oncol Biol Phys*, vol. 101, no. 1, pp. 195–200, 2018. DOI: 10.1016/j.ijrobp.2017.11.045.

[71] R. B. Ger, J. Yang, Y. Ding, M. C. Jacobsen, C. D. Fuller, R. M. Howell, *et al.*, Accuracy of deformable image registration on magnetic resonance images in digital and physical phantoms, *Med Phys*, vol. 44, no. 10, pp. 5153–5161, 2017. DOI: 10.1002/mp.12406.

[72] A. I. Ghanem, A. A. Elsaid, M. A. Elshaikh, and G. A. Khedr, Volumetric-Modulated Arc Radiotherapy with Daily Image-Guidance Carries Better Toxicity Profile for Higher Risk Prostate Cancer, *Asian Pac J Cancer Prev*, vol. 22, no. 1, pp. 61–68, 2021. DOI: 10.31557/APJCP.2021.22.1.61.

[73] E. Gibbons, M. Hoffmann, J. Westhuyzen, A. Hodgson, B. Chick, and A. Last, Clinical evaluation of deep learning and atlas-based auto-segmentation for critical organs at risk in radiation therapy, *J Med Radiat Sci*, vol. 70, no. S2, pp. 15–25, 2023. DOI: 10.1002/jmrs.618.

[74] O. Glasser, Wilhelm Conrad Röntgen und die Geschichte der Röntgenstrahlen. Berlin, Heidelberg: Springer, 1995. DOI: 10.1007/978-3-642-79312-7.

[75] C. K. Glide-Hurst, G. D. Hugo, J. Liang, and D. Yan, A simplified method of four-dimensional dose accumulation using the mean patient density representation, *Med Phys*, vol. 35, no. 12, pp. 5269–5277, 2008. DOI: 10.1118/1.3002304.

[76] C. K. Glide-Hurst, P. Lee, A. D. Yock, J. R. Olsen, M. Cao, F. Siddiqui, *et al.*, Adaptive Radiation Therapy (ART) Strategies and Technical Considerations: A State of the ART Review From NRG Oncology, *Int J Radiat Oncol Biol Phys*, vol. 109, no. 4, pp. 1054–1075, 2021. DOI: 10.1016/j.ijrobp.2020.10.021.

[77] G. Grimbergen, S. L. Hackett, F. van Ommen, A. L. H. M. W. van Lier, P. T. S. Borman, L. T. C. Meijers, *et al.*, Gating and intrafraction drift correction on a 1.5 T MR-Linac: Clinical dosimetric benefits for upper abdominal tumors, *Radiother Oncol*, vol. 189, p. 109 932, 2023. DOI: 10.1016/j.radonc.2023.109932.

[78] M. Guckenberger, J. Wilbert, T. Krieger, A. Richter, K. Baier, J. Meyer, *et al.*, Four-Dimensional Treatment Planning for Stereotactic Body Radiotherapy, *Int J Radiat Oncol Biol Phys*, vol. 69, no. 1, pp. 276–285, 2007. DOI: 10.1016/j.ijrobp.2007.04.074.

[79] M. G. Guren, C. Undseth, B. L. Rekstad, M. Brændengen, S. Dueland, K.-L. G. Spindler, *et al.*, Reirradiation of locally recurrent rectal cancer: A systematic review, *Radiother Oncol*, vol. 113, no. 2, pp. 151–157, 2014. DOI: 10.1016/j.radonc.2014.11.021.

[80] E. Haber and J. Modersitzki, Intensity Gradient Based Registration and Fusion of Multi-modal Images, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006*, R. Larsen, M. Nielsen, and J. Sporring, Eds., Berlin, Heidelberg: Springer, 2006, pp. 726–733. DOI: 10.1007/11866763_89.

[81] O. Hamming-Vrieze, S. R. van Kranen, W. D. Heemsbergen, C. A. H. Lange, M. W. M. van den Brekel, M. Verheij, *et al.*, Analysis of GTV reduction during radiotherapy for oropharyngeal cancer: Implications for adaptive radiotherapy, *Radiother Oncol*, vol. 122, no. 2, pp. 224–228, 2017. DOI: 10.1016/j.radonc.2016.10.012.

[82] O. Hamming-Vrieze, S. R. v. Kranen, S. v. Beek, W. Heemsbergen, M. v. Herk, M. W. M. v. d. Brekel, *et al.*, Evaluation of Tumor Shape Variability in Head-and-Neck Cancer Patients Over the Course of Radiation Therapy Using Implanted Gold Markers, *Int J Radiat Oncol Biol Phys*, vol. 84, no. 2, e201–e207, 2012. DOI: 10.1016/j.ijrobp.2012.03.014.

[83] X. Han, L. S. Hibbard, and V. Willcut, An Efficient Inverse-Consistent Diffeomorphic Image Registration Method for Prostate Adaptive Radiotherapy, in *Prostate Cancer Imaging. Computer-Aided Diagnosis, Prognosis, and Intervention*, A. Madabhushi, J. Dowling, P. Yan, A. Fenster, P. Abolmaesumi, and N. Hata, Eds., Berlin, Heidelberg: Springer, 2010, pp. 34–41. DOI: 10.1007/978-3-642-15989-3_5.

[84] N. Hardcastle, E. V. Osorio, A. Jackson, C. Mayo, A. E. Aarberg, M. Ayadi, *et al.*, Multi-centre evaluation of variation in cumulative dose assessment in reirradiation scenarios, *Radiother Oncol*, vol. 194, 2024. DOI: 10.1016/j.radonc.2024.110184.

[85] Y. He, G. Cazoulat, C. Wu, S. Svensson, L. Almodovar-Abreu, B. Rigaud, *et al.*, Quantifying the Effect of 4-Dimensional Computed Tomography–Based Deformable Dose Accumulation on Representing Radiation Damage for Patients with Locally Advanced Non-Small Cell Lung Cancer Treated with Standard-Fractionated Intensity-Modulated Radiation Therapy, *Int J Radiat Oncol Biol Phys*, vol. 118, no. 1, pp. 231–241, 2024. DOI: 10.1016/j.ijrobp.2023.07.016.

[86] E. Heath and J. Seuntjens, A direct voxel tracking method for four-dimensional Monte Carlo dose calculations in deforming anatomy, *Med Phys*, vol. 33, no. 2, pp. 434–445, 2006. DOI: 10.1118/1.2163252.

[87] V. Hernandez, C. R. Hansen, L. Widesott, A. Bäck, R. Canters, M. Fusella, *et al.*, What is plan quality in radiotherapy? The importance of evaluating dose metrics, complexity, and robustness of treatment plans, *Radiother Oncol*, vol. 153, pp. 26–33, 2020. DOI: 10.1016/j.radonc.2020.09.038.

[88] J.-C. Horiot, P. Bontemps, W. v. d. Bogaert, R. L. Fur, D. v. d. Weijngaert, M. Bolla, *et al.*, Accelerated fractionation (AF) compared to conventional fractionation (CF) improves loco-regional control in the radiotherapy of advanced head and neck cancers: results of the EORTC 22851 randomized trial, *Radiother Oncol*, vol. 44, no. 2, pp. 111–121, 1997. DOI: 10.1016/S0167-8140(97)00079-0.

[89] G. N. Hounsfield, Computerized transverse axial scanning (tomography): Part 1. Description of system, *Br J Radiol*, vol. 46, no. 552, pp. 1016–1022, 1973. DOI: 10.1259/0007-1285-46-552-1016.

[90] Y. Hu, M. Modat, E. Gibson, W. Li, N. Ghavami, E. Bonmati, *et al.*, Weakly-supervised convolutional neural networks for multimodal image registration, *Med Image Anal*, vol. 49, pp. 1–13, 2018. DOI: 10.1016/j.media.2018.07.002.

[91] E. Huang, L. Dong, A. Chandra, D. A. Kuban, I. I. Rosen, A. Evans, *et al.*, Intrafraction prostate motion during IMRT for prostate cancer, *Int J Radiat Oncol Biol Phys*, vol. 53, no. 2, pp. 261–268, 2002. DOI: 10.1016/S0360-3016(02)02738-4.

[92] M. Hub and C. P. Karger, Estimation of the uncertainty of elastic image registration with the demons algorithm, *Phys Med Biol*, vol. 58, no. 9, pp. 3023–3036, 2013. DOI: 10.1088/0031-9155/58/9/3023.

[93] S. C. Huijskens, I. W. E. M. van Dijk, R. de Jong, J. Visser, R. D. Fajardo, C. M. Ronckers, *et al.*, Quantification of renal and diaphragmatic interfractional motion in pediatric image-guided radiation therapy: A multicenter study, *Radiother Oncol*, vol. 117, no. 3, pp. 425–431, 2015. DOI: 10.1016/j.radonc.2015.09.020.

[94] A. Hunt, V. N. Hansen, U. Oelfke, S. Nill, and S. Hafeez, Adaptive Radiotherapy Enabled by MRI Guidance, *Clin Oncol*, vol. 30, no. 11, pp. 711–719, 2018. DOI: 10.1016/j.clon.2018.08.001.

[95] D. Huttenlocher, G. Klanderman, and W. Rucklidge, Comparing images using the Hausdorff distance, *IEEE Trans Pattern Anal Mach Intell*, vol. 15, no. 9, pp. 850–863, 1993. DOI: 10.1109/34.232073.

[96] *ICRU Report 62, Prescribing, Recording and Reporting Photon Beam Therapy (Supplement to ICRU 50) – ICRU*.

[97] E. Israels and L. Israels, The Cell Cycle, *Oncologist*, vol. 5, no. 6, pp. 510–513, 2000. DOI: 10.1634/theoncologist.5-6-510.

[98] D. A. Jaffray, P. E. Lindsay, K. K. Brock, J. O. Deasy, and W. A. Tomé, Accurate Accumulation of Dose for Improved Understanding of Radiation Effects in Normal Tissue, *Int J Radiat Oncol Biol Phys*, vol. 76, no. 3, Supplement, S135–S139, 2010. DOI: 10.1016/j.ijrobp.2009.06.093.

[99] G. Janssens, J. Orban de Xivry, S. Fekkes, A. Dekker, B. Macq, P. Lambin, *et al.*, Evaluation of nonrigid registration models for interfraction dose accumulation in radiotherapy, *Med Phys*, vol. 36, no. 9Part1, pp. 4268–4276, 2009. DOI: 10.1118/1.3194750.

[100] S. Jeong, W. Cheon, S. Kim, W. Park, and Y. Han, Deep-learning-based segmentation using individual patient data on prostate cancer radiation therapy, *PLOS ONE*, vol. 19, no. 7, e0308181, 2024. DOI: 10.1371/journal.pone.0308181.

[101] B. Jian, M. F. Azampour, F. De Benetti, J. Oberreuter, C. Bukas, A. S. Gersing, *et al.*, *Weakly-supervised Biomechanically-constrained CT/MRI Registration of the Spine*, 2022. DOI: 10.48550/arXiv.2205.07568.

[102] J. Jiang, C. Min Seo Choi, J. O. Deasy, A. Rimner, M. Thor, and H. Veeraraghavan, Artificial intelligence-based automated segmentation and radiotherapy dose mapping for thoracic normal tissues, *Phys Imaging Radiat Oncol*, vol. 29, p. 100 542, 2024. DOI: 10.1016/j.phro.2024.100542.

[103] V. Jóhannesson, A. Gunnlaugsson, P. Nilsson, P. Brynolfsson, E. Kjellén, and E. Wieslander, Dose-volume relationships of planned versus estimated delivered radiation doses to pelvic organs at risk and side effects in patients treated with salvage radiotherapy for recurrent prostate cancer, *Tech Innov Patient Support Radiat Oncol*, vol. 29, 2024. DOI: 10.1016/j.tipsro.2023.100231.

[104] N. Kadoya, Y. Miyasaka, T. Yamamoto, Y. Kuroda, K. Ito, M. Chiba, *et al.*, Evaluation of rectum and bladder dose accumulation from external beam radiotherapy and brachytherapy for cervical cancer using two different deformable image registration techniques, *J Radiat Res*, vol. 58, no. 5, pp. 720–728, 2017. DOI: 10.1093/jrr/rrx028.

[105] N. Kadoya, Y. Nakajima, M. Saito, Y. Miyabe, M. Kurooka, S. Kito, *et al.*, Multi-institutional Validation Study of Commercially Available Deformable Image Registration Software for Thoracic Images, *Int J Radiat Oncol Biol Phys*, vol. 96, no. 2, pp. 422–431, 2016. DOI: 10.1016/j.ijrobp.2016.05.012.

[106] K. Kainz, J. Garcia Alvarez, H. Zhong, S. Lim, E. Ahunbay, A. Tai, *et al.*, Use of a DVH overlay technique for quality assurance of deformable image registration-based dose accumulation, *Med Phys*, vol. 49, no. 1, pp. 611–623, 2022. DOI: 10.1002/mp.15375.

[107] R. N. Kandalan, D. Nguyen, N. H. Rezaeian, A. M. Barragán-Montero, S. Breedveld, K. Namuduri, *et al.*, Dose prediction with deep learning for prostate cancer radiation therapy: Model adaptation to different treatment planning practices, *Radiother Oncol*, vol. 153, pp. 228–235, 2020. DOI: 10.1016/j.radonc.2020.10.027.

[108] D. Kardamakis, S. Baatout, M. Bourguignon, N. Foray, and Y. Socol, History of Radiation Biology, in *Radiobiology Textbook*, S. Baatout, Ed., Cham: Springer International Publishing, 2023, pp. 1–24. DOI: 10.1007/978-3-031-18810-7_1.

[109] M. Kawula, I. Hadi, L. Nierer, M. Vagni, D. Cusumano, L. Boldrini, *et al.*, Patient-specific transfer learning for auto-segmentation in adaptive 0.35 T MRgRT of prostate cancer: a bi-centric evaluation, *Med Phys*, vol. 50, no. 3, pp. 1573–1585, 2023. DOI: 10.1002/mp.16056.

[110] V. Kearney, J. W. Chan, S. Haaf, M. Descovich, and T. D. Solberg, DoseNet: a volumetric dose prediction algorithm using 3D fully-convolutional neural networks, *Phys Med Biol*, vol. 63, no. 23, p. 235 022, 2018. DOI: 10.1088/1361-6560/aaef74.

[111] S. L. Kerns, L. Fachal, L. Dorling, G. C. Barnett, A. Baran, D. R. Peterson, *et al.*, Radiogenomics Consortium Genome-Wide Association Study Meta-Analysis of Late Toxicity After Prostate Cancer Radiotherapy, *J. Natl. Cancer Inst.*, vol. 112, no. 2, pp. 179–190, 2020. DOI: 10.1093/jnci/djz075.

[112] M. L. Kessler, Image registration and data fusion in radiation therapy, *Brit J Radiol*, vol. 79, no. special_issue_1, S99–S108, 2006. DOI: 10.1259/bjr/70617164.

[113] S. Khawaled and M. Freiman, Unsupervised Deep-Learning Based Deformable Image Registration: A Bayesian Framework, *arXiv: Computer Vision and Pattern Recognition*, 2020.

[114] S. Khawaled and M. Freiman, NPBDREG: Uncertainty assessment in diffeomorphic brain MRI registration using a non-parametric Bayesian deep-learning based approach, *Computerized Medical Imaging and Graphics*, vol. 99, p. 102 087, 2022. DOI: 10.1016/j.compmedimag.2022.102087.

[115] B. Kim, J. Kim, J.-G. Lee, D. H. Kim, S. H. Park, and J. C. Ye, Unsupervised Deformable Image Registration Using Cycle-Consistent CNN, in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, D. Shen, T. Liu, T. M. Peters, L. H. Staib, C. Essert, S. Zhou, *et al.*, Eds., Cham: Springer International Publishing, 2019, pp. 166–174. DOI: 10.1007/978-3-030-32226-7_19.

[116] H. Kim, Y. C. Lee, S. H. Benedict, B. Dyer, M. Price, Y. Rong, *et al.*, Dose Summation Strategies for External Beam Radiation Therapy and Brachytherapy in Gynecologic Malignancy: A Review from the NRG Oncology and NCTN Medical Physics Subcommittees, *Int J Radiat Oncol Biol Phys*, vol. 111, no. 4, pp. 999–1010, 2021. DOI: 10.1016/j.ijrobp.2021.06.019.

[117] Y. S. Kim, Reirradiation of head and neck cancer in the era of intensity-modulated radiotherapy: patient selection, practical aspects, and current evidence, *Radiat Oncol J*, vol. 35, no. 1, pp. 1–15, 2017. DOI: 10.3857/roj.2017.00122.

[118] C. Kontaxis, D. M. de Muinck Keizer, L. G. W. Kerkmeijer, T. Willigenburg, M. D. den Hartogh, J. R. N. van der Voort van Zyp, *et al.*, Delivered dose quantification in prostate radiotherapy using online 3D cine imaging and treatment log files on a combined 1.5T magnetic resonance imaging and linear accelerator system, *Phys Imaging Radiat Oncol*, vol. 15, pp. 23–29, 2020. DOI: 10.1016/j.phro.2020.06.005.

[119] T. K. Koo and M. Y. Li, A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research, *J Chiropr Med*, vol. 15, no. 2, pp. 155–163, 2016. DOI: 10.1016/j.jcm.2016.02.012.

[120] A. J. Koong, D. A. S. Toesca, R. von Eyben, E. L. Pollom, and D. T. Chang, Reirradiation with stereotactic body radiation therapy after prior conventional fractionation radiation for locally recurrent pancreatic adenocarcinoma, *Adv Radiat Oncol*, vol. 2, no. 1, pp. 27–36, 2017. DOI: 10.1016/j.adro.2017.01.003.

[121] T. J. Kruser, B. P. McCabe, M. P. Mehta, D. Khuntia, T. C. Campbell, H. M. Geye, *et al.*, Reirradiation for Locoregionally Recurrent Lung Cancer: Outcomes in Small Cell and Non–Small Cell Lung Carcinoma, *Am J Clin Oncol*, vol. 37, no. 1, pp. 70–76, 2014. DOI: 10.1097/COC.0b013e31826b9950.

[122] D. Kuang and T. Schmah, FAIM – A ConvNet Method for Unsupervised 3D Medical Image Registration, in *Machine Learning in Medical Imaging*, H.-I. Suk, M. Liu, P. Yan, and C. Lian, Eds., Cham: Springer International Publishing, 2019, pp. 646–654. DOI: 10.1007/978-3-030-32692-0_74.

[123] L. A. Künzel, S. Leibfarth, O. S. Dohm, A.-C. Müller, D. Zips, and D. Thorwarth, Automatic VMAT planning for post-operative prostate cancer cases using particle swarm optimization: A proof of concept study, *Phys Med*, vol. 69, pp. 101–109, 2020. DOI: 10.1016/j.ejmp.2019.12.007.

[124] G. J. Kutcher, C. Burman, L. Brewster, M. Goitein, and R. Mohan, Histogram reduction method for calculating complication probabilities for three-dimensional treatment planning evaluations, *Int J Radiat Oncol Biol Phys*, vol. 21, no. 1, pp. 137–146, 1991. DOI: 10.1016/0360-3016(91)90173-2.

[125] G. J. Kutcher and C. Burman, Calculation of complication probability factors for non-uniform normal tissue irradiation: The effective volume method gerald, *Int J Radiat Oncol Biol Phys*, vol. 16, no. 6, pp. 1623–1630, 1989. DOI: 10.1016/0360-3016(89)90972-3.

[126] J. J. W. Lagendijk, B. W. Raaymakers, A. J. E. Raaijmakers, J. Overweg, K. J. Brown, E. M. Kerkhof, *et al.*, MRI/linac integration, *Radiother Oncol*, vol. 86, no. 1, pp. 25–29, 2008. DOI: 10.1016/j.radonc.2007.10.034.

[127] J. J. W. Lagendijk, B. W. Raaymakers, and M. van Vulpen, The Magnetic Resonance Imaging–Linac System, *Semin Radiat Oncol*, vol. 24, no. 3, pp. 207–209, 2014. DOI: 10.1016/j.semradonc.2014.02.009.

[128] L. G. Lajtha, R. Oliver, and F. Ellis, Rationalisation of Fractionation in Radiotherapy, *Br J Radiol*, vol. 33, no. 394, pp. 634–635, 1960. DOI: 10.1259/0007-1285-33-394-634.

[129] D. S. Lakomy, J. Yang, S. Vedam, J. Wang, B. Lee, A. Sobremonte, *et al.*, Clinical Implementation and Initial Experience With a 1.5 Tesla MR-Linac for MR-Guided Radiation Therapy for Gynecologic Cancer: An R-IDEAL Stage 1 and 2a First in Humans Feasibility Study of New Technology Implementation, *Pract Radiat Oncol*, vol. 12, no. 4, e296–e305, 2022. DOI: 10.1016/j.prro.2022.03.002.

[130] V. Landoni, C. Fiorino, C. Cozzarini, G. Sanguineti, R. Valdagni, and T. Rancati, Predicting toxicity in radiotherapy for prostate cancer, *Phys Medica*, vol. 32, no. 3, pp. 521–532, 2016. DOI: 10.1016/j.ejmp.2016.03.003.

[131] E. Lavrova, M. D. Garrett, Y.-F. Wang, C. Chin, C. Elliston, M. Savacool, *et al.*, Adaptive Radiation Therapy: A Review of CT-based Techniques, *Radiol Imaging Cancer*, vol. 5, no. 4, e230011, 2023. DOI: 10.1148/rycan.230011.

[132] D. Lee, S. Alam, J. Jiang, L. Cervino, Y.-C. Hu, and P. Zhang, Seq2Morph: A deep learning deformable image registration algorithm for longitudinal imaging studies and adaptive radiotherapy, *Med Phys*, vol. n/a, no. n/a, DOI: 10.1002/mp.16026.

[133] J. Lee, S.-H. Liu, J.-B. Lin, M.-H. Wu, C.-J. Wu, H.-C. Tai, *et al.*, Image-guided study of inter-fraction and intra-fraction set-up variability and margins in reverse semi-decubitus breast radiotherapy, *Radiat Oncol*, vol. 13, no. 1, p. 254, 2018. DOI: 10.1186/s13014-018-1200-1.

[134] C. M. van Leeuwen, A. L. Oei, J. Crezee, A. Bel, N. A. P. Franken, L. J. A. Stalpers, *et al.*, The alfa and beta of tumours: a review of parameters of the linear-quadratic model, derived from clinical radiotherapy studies, *Radiat Oncol*, vol. 13, no. 1, p. 96, 2018. DOI: 10.1186/s13014-018-1040-z.

[135] G. Li, Y. Li, J. Wang, X. Gao, Q. Zhong, L. He, *et al.*, Guidelines for radiotherapy of prostate cancer (2020 edition), *Pract Radiat Oncol*, vol. 5, no. 3, pp. 160–182, 2021. DOI: 10.1002/pro6.1129.

[136] H. S. Li, H. Zhong, J. Kim, C. Glide-Hurst, M. Gulam, T. S. Nurushev, *et al.*, Direct dose mapping versus energy/mass transfer mapping for 4D dose accumulation: fundamental differences and dosimetric consequences, *Phys Med Biol*, vol. 59, no. 1, pp. 173–188, 2014. DOI: 10.1088/0031-9155/59/1/173.

[137] X. Liang, H. Morgan, T. Bai, M. Dohopolski, D. Nguyen, and S. Jiang, Deep learning based direct segmentation assisted by deformable image registration for cone-beam CT based auto-segmentation for adaptive radiotherapy, *Phys Med Biol*, vol. 68, no. 4, p. 045 012, 2023. DOI: 10.1088/1361-6560/acb4d7.

[138] S. Y. Lim, A. Tran, A. N. K. Tran, A. Sobremonte, C. D. Fuller, L. Simmons, *et al.*, Dose accumulation of daily adaptive plans to decide optimal plan adaptation strategy for head-and-neck patients treated with MR-Linac, *Med Dosim*, vol. 47, no. 1, pp. 103–109, 2022. DOI: 10.1016/j.meddos.2021.08.005.

[139] S. Lim-Reinders, B. M. Keller, S. Al-Ward, A. Sahgal, and A. Kim, Online Adaptive Radiation Therapy, *Int J Radiat Oncol Biol Phys*, vol. 99, no. 4, pp. 994–1003, 2017. DOI: 10.1016/j.ijrobp.2017.04.023.

[140] R. Liu, Z. Li, X. Fan, C. Zhao, H. Huang, and Z. Luo, Learning Deformable Image Registration From Optimization: Perspective, Modules, Bilevel Training and Beyond, *IEEE Trans Pattern Anal Mach Intell*, vol. 44, no. 11, pp. 7688–7704, 2022. DOI: 10.1109/TPAMI.2021.3115825.

[141] N. J. Lowther, S. H. Marsh, and R. J. W. Louwe, Quantifying the dose accumulation uncertainty after deformable image registration in head-and-neck radiotherapy, *Radiother Oncol*, vol. 143, pp. 117–125, 2020. DOI: 10.1016/j.radonc.2019.12.009.

[142] J. T. Lyman, Complication Probability as Assessed from Dose-Volume Histograms, *Radiat Res*, vol. 104, no. 2s, S13–S19, 1985. DOI: 10.2307/3576626.

[143] S. A. Maas, B. J. Ellis, G. A. Ateshian, and J. A. Weiss, FEBio: Finite Elements for Biomechanics, *J Biomech Eng*, vol. 134, no. 011005, 2012. DOI: 10.1115/1.4005694.

[144] U. Mahantshetty, P. Naga, A. Nachankar, Y. Ghadi, A. Dheera, L. Scaria, *et al.*, Set-Up Errors, Organ Motion, Tumour Regression and its Implications on Internal Target Volume–Planning Target Volume During Cervical Cancer Radiotherapy: Results From a Prospective Study, *Clin Oncol*, vol. 34, no. 3, pp. 189–197, 2022. DOI: 10.1016/j.clon.2021.10.010.

[145] H. B. Mann and D. R. Whitney, On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other, *Ann Math Stat*, vol. 18, no. 1, pp. 50–60, 1947.

[146] L. B. Marks, E. D. Yorke, A. Jackson, R. K. T. Haken, L. S. Constine, A. Eisbruch, *et al.*, Use of Normal Tissue Complication Probability Models in the Clinic, *Int J Radiat Oncol Biol Phys*, vol. 76, no. 3, S10–S19, 2010. DOI: 10.1016/j.ijrobp.2009.07.1754.

[147] S. Marzi, B. Saracino, M. G. Petrongari, S. Arcangeli, S. Gomellini, G. Arcangeli, *et al.*, Modeling of α/β for late rectal toxicity from a randomized phase II study: conventional versus hypofractionated scheme for localized prostate cancer, *J Exp Clin Cancer Res*, vol. 28, no. 1, p. 117, 2009. DOI: 10.1186/1756-9966-28-117.

[148] P. Mayles, A. Nahum, and J. C. Rosenwald, Eds., Handbook of Radiotherapy Physics: Theory and Practice. Boca Raton: CRC Press, 2007. DOI: 10.1201/9781420012026.

[149] M. W. McDonald, M. G. Moore, and P. A. S. Johnstone, Risk of Carotid Blowout After Reirradiation of the Head and Neck: A Systematic Review, *Int J Radiat Oncol Biol Phys*, vol. 82, no. 3, pp. 1083–1089, 2012. DOI: 10.1016/j.ijrobp.2010.08.029.

[150] A. McWilliam, J. Kennedy, C. Hodgson, E. Vasquez Osorio, C. Faivre-Finn, and M. van Herk, Radiation dose to heart base linked with poorer survival in lung cancer patients, *Eur J Cancer*, vol. 85, pp. 106–113, 2017. DOI: 10.1016/j.ejca.2017.07.053.

[151] A. McWilliam, G. Palma, A. Abravan, O. Acosta, A. Appelt, M. Aznar, *et al.*, Voxel-based analysis: Roadmap for clinical translation, *Radiother Oncol*, vol. 188, 2023. DOI: 10.1016/j.radonc.2023.109868.

[152] E. H. W. Meijering, W. J. Niessen, J. P. W. Pluim, and M. A. Viergever, Quantitative Comparison of Sinc-Approximating Kernels for Medical Image Interpolation, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI'99*, C. Taylor and A. Colchester, Eds., Berlin, Heidelberg: Springer, 1999, pp. 210–217. DOI: 10.1007/10704282_23.

[153] A. Meijers, A.-C. Knopf, A. P. G. Crijns, J. F. Ubbels, A. G. H. Niezink, J. A. Langendijk, *et al.*, Evaluation of interplay and organ motion effects by means of 4D dose reconstruction and accumulation, *Radiother Oncol*, vol. 150, pp. 268–274, 2020. DOI: 10.1016/j.radonc.2020.07.055.

[154] A. Mencarelli, S. R. van Kranen, O. Hamming-Vrieze, S. van Beek, C. R. Nico Rasch, M. van Herk, *et al.*, Deformable Image Registration for Adaptive Radiation Therapy of Head and Neck Cancer: Accuracy and Precision in the Presence of Tumor Changes, *Int J Radiat Oncol Biol Phys*, vol. 90, no. 3, pp. 680–687, 2014. DOI: 10.1016/j.ijrobp.2014.06.045.

[155] J. Michalski, T. M. Pisansky, C. A. Lawton, L. Potters, and D. A. Kuban, Chapter 51 - Prostate Cancer, in *Clinical Radiation Oncology (Third Edition)*, L. L. Gunderson and J. E. Tepper, Eds., Philadelphia: W.B. Saunders, 2012, pp. 1037–1097. DOI: 10.1016/B978-1-4377-1637-5.00051-1.

[156] R. Miralbell, S. A. Roberts, E. Zubizarreta, and J. H. Hendry, Dose-Fractionation Sensitivity of Prostate Cancer Deduced From Radiotherapy Outcomes of 5,969 Patients in Seven International Institutional Datasets: α/β = 1.4 (0.9–2.2) Gy, *Int J Radiat Oncol Biol Phys*, vol. 82, no. 1, e17–e24, 2012. DOI: 10.1016/j.ijrobp.2010.10.075.

[157] H. Miura, S. Ozawa, M. Nakao, K. Furukawa, Y. Doi, H. Kawabata, *et al.*, Impact of deformable image registration accuracy on thoracic images with different regularization weight parameter settings, *Phys Med*, vol. 42, pp. 108–111, 2017. DOI: 10.1016/j.ejmp.2017.09.122.

[158] M. J. Murphy, F. J. Salguero, J. V. Siebers, D. Staub, and C. Vaman, A method to estimate the effect of deformable image registration uncertainties on daily dose mapping, *Med Phys*, vol. 39, no. 2, pp. 573–580, 2012. DOI: 10.1118/1.3673772.

[159] M. Murr, U. Bernchou, E. Bubula-Rehm, M. Ruschin, P. Sadeghi, P. Voet, *et al.*, A multi-institutional comparison of retrospective deformable dose accumulation for online adaptive magnetic resonance-guided radiotherapy, *Phys Imaging Radiat Oncol*, vol. 30, p. 100 588, 2024. DOI: 10.1016/j.phro.2024.100588.

[160] M. Murr, K. K. Brock, M. Fusella, N. Hardcastle, M. Hussein, M. G. Jameson, *et al.*, Applicability and usage of dose mapping/accumulation in radiotherapy, *Radiother Oncol*, vol. 182, p. 109 527, 2023. DOI: 10.1016/j.radonc.2023.109527.

[161] M. Murr, D. Wegener, S. Böke, C. Gani, D. Mönnich, M. Niyazi, *et al.*, Comparison of online adaptive and non-adaptive magnetic resonance image-guided radiation therapy in prostate cancer using dose accumulation, *Phys Imaging Radiat Oncol*, vol. in revision, 2024.

[162] M. Murr, D. Wegener, M. Nachbar, A.-C. Mueller, D. Zips, and D. Thorwarth, Comparison of different strategies for deformable dose accumulation in prostate cancer radiotherapy, ESTRO 2022, 2022.

[163] L. J. Murray, J. Lilley, M. A. Hawkins, A. M. Henry, P. Dickinson, and D. Sebag-Montefiore, Pelvic re-irradiation using stereotactic ablative radiotherapy (SABR): A systematic review, *Radiother Oncol*, vol. 125, no. 2, pp. 213–222, 2017. DOI: 10.1016/j.radonc.2017.09.030.

[164] T. F. Mutanga, H. C. J. d. Boer, V. Rajan, M. L. P. Dirkx, L. Incrocci, and B. J. M. Heijmen, Day-to-Day Reproducibility of Prostate Intrafraction Motion Assessed by Multiple kV and MV Imaging of Implanted Markers During Treatment, *Int J Radiat Oncol Biol Phys*, vol. 83, no. 1, pp. 400–407, 2012. DOI: 10.1016/j.ijrobp.2011.05.049.

[165] C. G. Muurholm, T. Ravkilde, S. Skouboe, E. Worm, R. Hansen, M. Høyer, *et al.*, Real-time dose-guidance in radiotherapy: Proof of principle, *Radiother Oncol*, vol. 164, pp. 175–182, 2021. DOI: 10.1016/j.radonc.2021.09.024.

[166] S. Myrehaug, H. Soliman, C. Tseng, C. Heyn, and A. Sahgal, Re-irradiation of Vertebral Body Metastases: Treatment in the Radiosurgery Era, *Clin Oncol*, vol. 30, no. 2, pp. 85–92, 2018. DOI: 10.1016/j.clon.2017.11.005.

[167] M. Nachbar, M. Io Russo, C. Gani, S. Boeke, D. Wegener, F. Paulsen, *et al.*, Automatic AI-based contouring of prostate MRI for online adaptive radiotherapy, *Z Med Phys*, vol. 34, no. 2, pp. 197–207, 2024. DOI: 10.1016/j.zemedi.2023.05.001.

[168] L. Nenoff, F. Amstutz, M. Murr, B. Archibald-Heeren, M. Fusella, M. Hussein, *et al.*, Review and recommendations on deformable image registration uncertainties for radiotherapy applications, *Phys Med Biol*, vol. 68, no. 24, 24TR01, 2023. DOI: 10.1088/1361-6560/ad0d8a.

[169] L. Nenoff, C. O. Ribeiro, M. Matter, L. Hafner, M. Josipovic, J. A. Langendijk, *et al.*, Deformable image registration uncertainty for inter-fractional dose accumulation of lung cancer proton therapy, *Radiother Oncol*, vol. 147, pp. 178–185, 2020. DOI: 10.1016/j.radonc.2020.04.046.

[170] J. Neylon, X. Qi, K. Sheng, R. Staton, J. Pukala, R. Manon, *et al.*, A GPU based high-resolution multilevel biomechanical head and neck model for validating deformable image registration, *Med Phys*, vol. 42, no. 1, pp. 232–243, 2015. DOI: 10.1118/1.4903504.

[171] N. I. Niebuhr, M. Splinter, T. Bostel, J. Seco, C. M. Hentschke, R. O. Floca, *et al.*, Biologically consistent dose accumulation using daily patient imaging, *Radiat Oncol*, vol. 16, no. 1, p. 65, 2021. DOI: 10.1186/s13014-021-01789-3.

[172] C. Nieder, N. H. Andratschke, and A. L. Grosu, Re-irradiation for Recurrent Primary Brain Tumors, *Anticancer Res*, vol. 36, no. 10, pp. 4985–4995, 2016.

[173] S. Nikolov, S. Blackwell, A. Zverovitch, R. Mendes, M. Livne, J. De Fauw, *et al.*, *Deep learning to achieve clinically applicable segmentation of head and neck anatomy for radiotherapy*, 2021.

[174] S. Nithiananthan, S. Schafer, D. J. Mirota, J. W. Stayman, W. Zbijewski, D. D. Reh, *et al.*, Extra-dimensional Demons: A method for incorporating missing tissue in deformable image registration, *Med Phys*, vol. 39, no. 9, pp. 5718–5731, 2012. DOI: 10.1118/1.4747270.

[175] M. Nix, S. Gregory, M. Aldred, L. Aspin, J. Lilley, B. Al-Qaisieh, *et al.*, Dose summation and image registration strategies for radiobiologically and anatomically corrected dose accumulation in pelvic re-irradiation, *Acta Oncol*, vol. 61, no. 1, pp. 64–72, 2022. DOI: 10.1080/0284186X.2021.1982145.

[176] S. Oh and S. Kim, Deformable image registration in radiation therapy, *Radiat Oncol J*, vol. 35, no. 2, pp. 101–111, 2017. DOI: 10.3857/roj.2017.00325.

[177] F. P. Oliveira and J. M. R. Tavares, Medical image registration: a review, *Comput Methods Biomech Biomed Engin*, vol. 17, no. 2, pp. 73–93, 2014. DOI: 10.1080/10255842.2012.670855.

[178] W. H. Organization, *Face sheets, Cancer*, 2022.

[179] K. Otto, Volumetric modulated arc therapy: IMRT in a single gantry arc, *Med Phys*, vol. 35, no. 1, pp. 310–317, 2008. DOI: 10.1118/1.2818738.

[180] D. Owen, J. Lukovic, A. Hosni, C. H. Crane, T. S. Hong, L. A. Dawson, *et al.*, Challenges in Reirradiation of Intrahepatic Tumors, *Semin Radiat Oncol*, vol. 30, no. 3, pp. 242–252, 2020. DOI: 10.1016/j.semradonc.2020.02.004.

[181] C. Paganelli, G. Meschini, S. Molinelli, M. Riboldi, and G. Baroni, Patient-specific validation of deformable image registration in radiation therapy: Overview and caveats, *Med Phys*, vol. 45, no. 10, e908–e922, 2018. DOI: 10.1002/mp.13162.

[182] G. Palma, S. Monti, and L. Cella, Voxel-based analysis in radiation oncology: A methodological cookbook, *Phys Med*, vol. 69, pp. 192–204, 2020. DOI: 10.1016/j.ejmp.2019.12.013.

[183] G. Palma, S. Monti, V. D'Avino, M. Conson, R. Liuzzi, M. C. Pressello, *et al.*, A Voxel-Based Approach to Explore Local Dose Differences Associated With Radiation-Induced Lung Damage, *Int J Radiat Oncol Biol Phys*, vol. 96, no. 1, pp. 127–133, 2016. DOI: 10.1016/j.ijrobp.2016.04.033.

[184] K. C. Paradis, C. Mayo, D. Owen, D. E. Spratt, J. Hearn, B. Rosen, *et al.*, The Special Medical Physics Consult Process for Reirradiation Patients, *Adv Radiat Oncol*, vol. 4, no. 4, 2019. DOI: 10.1016/j.adro.2019.05.007.

[185] E. S. Paulson, E. Ahunbay, X. Chen, N. J. Mickevicius, G.-P. Chen, C. Schultz, *et al.*, 4D-MRI driven MR-guided online adaptive radiotherapy for abdominal stereotactic body radiation therapy on a high field MR-Linac: Implementation and initial clinical experience, *ctRO*, vol. 23, pp. 72–79, 2020. DOI: 10.1016/j.ctro.2020.05.002.

[186] J. W. Piper, J. H. Richmond, and A. S. Nelson, VoxAlign Deformation Engine®, 2018.

[187] K. Pirpinia, P. A. N. Bosman, J.-J. Sonke, M. v. Herk, and T. Alderliesten, Evolutionary multi-objective meta-optimization of deformation and tissue removal parameters improves the performance of deformable image registration of pre- and post-surgery images, in *Medical Imaging 2019: Image Processing*, vol. 10949, SPIE, 2019, pp. 838–848. DOI: 10.1117/12.2512760.

[188] J. Pukala, P. B. Johnson, A. P. Shah, K. M. Langen, F. J. Bova, R. J. Staton, *et al.*, Benchmarking of five commercial deformable image registration algorithms for head and neck patients, *J Appl Clin Med Phys*, vol. 17, no. 3, pp. 25–40, 2016. DOI: 10.1120/jacmp.v17i3.5735.

[189] A. Qin, J. Liang, X. Han, N. O'Connell, and D. Yan, Technical Note: The impact of deformable image registration methods on dose warping, *Med Phys*, vol. 45, no. 3, pp. 1287–1294, 2018. DOI: 10.1002/mp.12741.

[190] B. W. Raaymakers, J. J. W. Lagendijk, J. Overweg, J. G. M. Kok, A. J. E. Raaijmakers, E. M. Kerkhof, *et al.*, Integrating a 1.5 T MRI scanner with a 6 MV accelerator: proof of concept, *Phys Med Biol*, vol. 54, no. 12, N229, 2009. DOI: 10.1088/0031-9155/54/12/N01.

[191] M. Rabe, M. A. Palacios, J. R. van Sörnsen de Koste, C. Eze, M. Hillbrand, C. Belka, *et al.*, Comparison of MR-guided radiotherapy accumulated doses for central lung tumors with non-adaptive and online adaptive proton therapy, *Med Phys*, vol. 50, no. 5, pp. 2625–2636, 2023. DOI: 10.1002/mp.16319.

[192] N. Riaz, J. C. Hong, E. J. Sherman, L. Morris, M. Fury, I. Ganly, *et al.*, A nomogram to predict loco-regional control after re-irradiation for head and neck cancer, *Radiother Oncol*, vol. 111, no. 3, pp. 382–387, 2014. DOI: 10.1016/j.radonc.2014.06.003.

[193] B. Rigaud, G. Cazoulat, S. Vedam, A. M. Venkatesan, C. B. Peterson, N. Taku, *et al.*, Modeling Complex Deformations of the Sigmoid Colon Between External Beam Radiation Therapy and Brachytherapy Images of Cervical Cancer, *Int J Radiat Oncol Biol Phys*, vol. 106, no. 5, pp. 1084–1094, 2020. DOI: 10.1016/j.ijrobp.2019.12.028.

[194] B. Rigaud, A. Simon, J. Castelli, C. Lafond, O. Acosta, P. Haigron, *et al.*, Deformable image registration for radiation therapy: principle, methods, applications and evaluation, *Acta Oncol*, vol. 58, no. 9, pp. 1225–1237, 2019. DOI: 10.1080/0284186X.2019.1620331.

[195] P. Risholm, E. Samset, I.-F. Talos, and W. Wells, A Non-rigid Registration Framework That Accommodates Resection and Retraction, in *Information Processing in Medical Imaging*, J. L. Prince, D. L. Pham, and K. J. Myers, Eds., Berlin, Heidelberg: Springer, 2009, pp. 447–458. DOI: 10.1007/978-3-642-02498-6_37.

[196] M. Roch, A. Zapatero, P. Castro, D. Büchser, L. Pérez, D. Hernández, *et al.*, Impact of rectum and bladder anatomy in intrafractional prostate motion during hypofractionated radiation therapy, *Clin Transl Oncol*, vol. 21, no. 5, pp. 607–614, 2019. DOI: 10.1007/s12094-018-1960-y.

[197] Y. Rong, Q. Chen, Y. Fu, X. Yang, H. A. Al-Hallaq, Q. J. Wu, *et al.*, NRG Oncology Assessment of Artificial Intelligence Deep Learning–Based Auto-segmentation for Radiation Therapy: Current Developments, Clinical Considerations, and Future Directions, *Int J Radiat Oncol Biol Phys*, vol. 119, no. 1, pp. 261–280, 2024. DOI: 10.1016/j.ijrobp.2023.10.033.

[198] Y. Rong, M. Rosu-Bubulac, S. H. Benedict, Y. Cui, R. Ruo, T. Connell, *et al.*, Rigid and Deformable Image Registration for Radiation Therapy: A Self-Study Evaluation Guide for NRG Oncology Clinical Trial Participation, *Pract Radiat Oncol*, vol. 11, no. 4, pp. 282–298, 2021. DOI: 10.1016/j.prro.2021.02.007.

[199] M. Rosu, I. J. Chetty, J. M. Balter, M. L. Kessler, D. L. McShan, and R. K. Ten Haken, Dose reconstruction in deforming lung anatomy: Dose grid size effects and clinical implications, *Med Phys*, vol. 32, no. 8, pp. 2487–2495, 2005. DOI: 10.1118/1.1949749.

[200] R. Rouhi, S. Niyoteka, A. Carré, S. Achkar, P.-A. Laurent, M. B. Ba, *et al.*, Automatic gross tumor volume segmentation with failure detection for safe implementation in locally advanced cervical cancer, *Phys Imaging Radiat Oncol*, vol. 0, no. 0, 2024. DOI: 10.1016/j.phro.2024.100578.

[201] Y. G. Roussakis, H. Dehghani, S. Green, and G. J. Webster, Validation of a dose warping algorithm using clinically realistic scenarios, *Br J Radiol*, vol. 88, no. 1049, p. 20 140 691, 2015. DOI: 10.1259/bjr.20140691.

[202] R. Rulach, G. G. Hanna, K. Franks, J. McAleese, and S. Harrow, Re-irradiation for Locally Recurrent Lung Cancer: Evidence, Risks and Benefits, *Clin Oncol*, vol. 30, no. 2, pp. 101–109, 2018. DOI: 10.1016/j.clon.2017.11.003.

[203] N. K. Saleh-Sayah, E. Weiss, F. J. Salguero, and J. V. Siebers, A distance to dose difference tool for estimating the required spatial accuracy of a displacement vector field, *Med Phys*, vol. 38, no. 5, pp. 2318–2323, 2011. DOI: 10.1118/1.3572228.

[204] F. J. Salguero, N. K. Saleh-Sayah, C. Yan, and J. V. Siebers, Estimation of three-dimensional intrinsic dosimetric uncertainties resulting from using deformable image registration for dose mapping, *Med Phys*, vol. 38, no. 1, pp. 343–353, 2011. DOI: 10.1118/1.3528201.

[205] D. Sarrut, T. Baudier, M. Ayadi, R. Tanguy, and S. Rit, Deformable image registration applied to lung SBRT: Usefulness and limitations, *Phys Med*, vol. 44, pp. 108–112, 2017. DOI: 10.1016/j.ejmp.2017.09.121.

[206] M. H. F. Savenije, M. Maspero, G. G. Sikkes, J. R. N. van der Voort van Zyp, A. N. T. J. Kotte, G. H. Bol, *et al.*, Clinical implementation of MRI-based organs-at-risk auto-segmentation with convolutional networks for prostate radiotherapy, *Radiat Oncol*, vol. 15, no. 1, p. 104, 2020. DOI: 10.1186/s13014-020-01528-0.

[207] A. Scaggion, C. Fiandra, G. Loi, C. Vecchi, and M. Fusella, Free-to-use DIR solutions in radiotherapy: Benchmark against commercial platforms through a contour-propagation study, *Phys Med*, vol. 74, pp. 110–117, 2020. DOI: 10.1016/j.ejmp.2020.05.011.

[208] W. Schaake, A. v. d. Schaaf, L. V. v. Dijk, A. H. H. Bongaerts, A. C. M. v. d. Bergh, and J. A. Langendijk, Normal tissue complication probability (NTCP) models for late rectal bleeding, stool frequency and fecal incontinence after radiotherapy in prostate cancer patients, *Radiother Oncol*, vol. 119, no. 3, pp. 381–387, 2016. DOI: 10.1016/j.radonc.2016.04.005.

[209] C. Schröder, I. Stiefel, S. Tanadini-Lang, I. Pytko, E. Vu, M. Guckenberger, *et al.*, Re-irradiation in the thorax – An analysis of efficacy and safety based on accumulated EQD2 doses, *Radiother Oncol*, vol. 152, pp. 56–62, 2020. DOI: 10.1016/j.radonc.2020.07.033.

[210] T. E. Schultheiss, W. A. Tome, and C. G. Orton, Point/counterpoint: it is not appropriate to "deform" dose along with deformable image registration in adaptive radiotherapy, *Med Phys*, vol. 39, no. 11, pp. 6531–6533, 2012. DOI: 10.1118/1.4722968.

[211] B. D. d. Senneville, C. Zachiu, M. Ries, and C. Moonen, EVolution: an edge-based variational method for non-rigid multi-modal image registration, *Phys Med Biol*, vol. 61, no. 20, p. 7377, 2016. DOI: 10.1088/0031-9155/61/20/7377.

[212] M. Shanker, B. Chua, C. Bettington, M. C. Foote, and M. B. Pinkham, Re-irradiation for recurrent high-grade gliomas: a systematic review and analysis of treatment technique with respect to survival and risk of radionecrosis, *Neuro-Oncology Practice*, vol. 6, no. 2, pp. 144–155, 2019. DOI: 10.1093/nop/npy019.

[213] L. E. A. Shelley, M. P. F. Sutcliffe, S. J. Thomas, D. J. Noble, M. Romanchikova, K. Harrison, *et al.*, Associations between voxel-level accumulated dose and rectal toxicity in prostate radiotherapy, *Phys Imaging Radiat Oncol*, vol. 14, pp. 87–94, 2020. DOI: 10.1016/j.phro.2020.05.006.

[214] Y. Shen, X. Tang, S. Lin, X. Jin, J. Ding, and M. Shao, Automatic dose prediction using deep learning and plan optimization with finite-element control for intensity modulated radiation therapy, *Med Phys*, vol. 51, no. 1, pp. 545–555, 2024. DOI: 10.1002/mp.16743.

[215] L. Shi, Q. Chen, S. Barley, Y. Cui, L. Shang, J. Qiu, *et al.*, Benchmarking of Deformable Image Registration for Multiple Anatomic Sites Using Digital Data Sets With Ground-Truth Deformation Vector Fields, *Pract Radiat Oncol*, vol. 11, no. 5, pp. 404–414, 2021. DOI: 10.1016/j.prro.2021.02.012.

[216] J. Shortall, G. Palma, H. Mistry, E. V. Osorio, A. McWilliam, A. Choudhury, *et al.*, Flogging a Dead Salmon? Reduced Dose Posterior to Prostate Correlates With Increased PSA Progression in Voxel-Based Analysis of 3 Randomized Phase 3 Trials, *Int J Radiat Oncol Biol Phys*, vol. 110, no. 3, pp. 696–699, 2021. DOI: 10.1016/j.ijrobp.2021.01.017.

[217] J. V. Siebers and H. Zhong, An energy transfer method for 4D Monte Carlo dose calculation, *Med Phys*, vol. 35, no. 9, pp. 4096–4105, 2008. DOI: 10.1118/1.2968215.

[218] M. Simonovsky, B. Gutiérrez-Becker, D. Mateus, N. Navab, and N. Komodakis, A Deep Metric for Multimodal Registration, in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2016*, S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, and W. Wells, Eds., Cham: Springer International Publishing, 2016, pp. 10–18. DOI: 10.1007/978-3-319-46726-9_2.

[219] M. Skaarup, M. J. Lundemann, S. Darkner, M. Jørgensen, L. Marner, D. Mirkovic, *et al.*, A framework for voxel-based assessment of biological effect after proton radiotherapy in pediatric brain cancer patients using multi-modal imaging, *Med Phys*, vol. 48, no. 7, pp. 4110–4121, 2021. DOI: 10.1002/mp.14989.

[220] J.-J. Sonke, M. Aznar, and C. Rasch, Adaptive Radiotherapy for Anatomical Changes, *Semin Radiat Oncol*, vol. 29, no. 3, pp. 245–257, 2019. DOI: 10.1016/j.semradonc.2019.02.007.

[221] J.-J. Sonke, M. Aznar, and C. Rasch, Non-rigid image registration: theory and practice, *Semin Radiat Oncol*, vol. 29, no. 3, pp. 245–257, 2019. DOI: 10.1016/j.semradonc.2019.02.007.

[222] J.-J. Sonke and J. Belderbos, Adaptive Radiotherapy for Lung Cancer, *Semin Radiat Oncol*, vol. 20, no. 2, pp. 94–106, 2010. DOI: 10.1016/j.semradonc.2009.11.003.

[223] A. Sotiras, C. Davatzikos, and N. Paragios, Deformable Medical Image Registration: A Survey, *IEEE Trans. Med. Imaging*, vol. 32, no. 7, pp. 1153–1190, 2013. DOI: 10.1109/TMI.2013.2265603.

[224] G. G. Steel, T. McMillan, and J. Peacock, The 5Rs of Radiobiology, *Int J Radiat Biol*, vol. 56, no. 6, pp. 1045–1048, 1989. DOI: 10.1080/09553008914552491.

[225] Z. Sun, X. Xia, J. Fan, J. Zhao, K. Zhang, J. Wang, *et al.*, A hybrid optimization strategy for deliverable intensity-modulated radiotherapy plan generation using deep learning-based dose prediction, *Med Phys*, vol. 49, no. 3, pp. 1344–1356, 2022. DOI: 10.1002/mp.15462.

[226] J. Swamidas, C. Kirisits, M. D. Brabandere, T. P. Hellebust, F.-A. Siebert, and K. Tanderup, Image registration, contour propagation and dose accumulation of external beam and brachytherapy in gynecological radiotherapy, *Radiother Oncol*, vol. 143, pp. 1–11, 2020. DOI: 10.1016/j.radonc.2019.08.023.

[227] A. Taylor and M. E. B. Powell, Intensity-modulated radiotherapy—what is it?, *Cancer Imaging*, vol. 4, no. 2, pp. 68–73, 2004. DOI: 10.1102/1470-7330.2004.0003.

[228] B. S. Teh, S. Y. Woo, and E. B. Butler, Intensity Modulated Radiation Therapy (IMRT): A New Promising Technology in Radiation Oncology, *Oncologist*, vol. 4, no. 6, pp. 433–442, 1999. DOI: 10.1634/theoncologist.4-6-433.

[229] H. Teske, K. Bartelheimer, J. Meis, R. Bendl, E. M. Stoiber, and K. Giske, Construction of a biomechanical head and neck motion model as a guide to evaluation of deformable image registration, *Phys Med Biol*, vol. 62, no. 12, N271, 2017. DOI: 10.1088/1361-6560/aa69b6.

[230] P. Thevenaz, T. Blu, and M. Unser, Interpolation revisited [medical images application], *IEEE Trans. Med. Imaging*, vol. 19, no. 7, pp. 739–758, 2000. DOI: 10.1109/42.875199.

[231] Image Interpolation and Resampling, *Handbook of Medical Imaging, Processing and Analysis*, P. Thévenaz, T. Blu, and M. Unser, Eds., 2000.

[232] K. Thielemans, E. Asma, and R. M. Manjeshwar, Mass-preserving image registration using free-form deformation fields, in *2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*, 2009, pp. 2490–2495. DOI: 10.1109/NSSMIC.2009.5402070.

[233] J.-P. Thirion, Image matching as a diffusion process: an analogy with Maxwell's demons, *Med Image Anal*, vol. 2, no. 3, pp. 243–260, 1998. DOI: 10.1016/S1361-8415(98)80022-4.

[234] M. Trovo, E. Minatel, E. Durofil, J. Polesel, M. Avanzo, T. Baresic, *et al.*, Stereotactic Body Radiation Therapy for Re-irradiation of Persistent or Recurrent Non-Small Cell Lung Cancer, *Int J Radiat Oncol Biol Phys*, vol. 88, no. 5, pp. 1114–1119, 2014. DOI: 10.1016/j.ijrobp.2014.01.012.

[235] D. Tward, T. Brown, Y. Kageyama, J. Patel, Z. Hou, S. Mori, *et al.*, Diffeomorphic Registration With Intensity Transformation and Missing Data: Application to 3D Digital Pathology of Alzheimer's Disease, *Front Neurosci*, vol. 14, 2020.

[236] G. Valdes, C. Lee, S. Tenn, P. Lee, C. Robinson, K. Iwamoto, *et al.*, The relative accuracy of 4D dose accumulation for lung radiotherapy using rigid dose projection versus dose recalculation on every breathing phase, *Med Phys*, vol. 44, no. 3, pp. 1120–1127, 2017. DOI: 10.1002/mp.12069.

[237] L. Van den Bosch, E. Schuit, H. P. van der Laan, J. B. Reitsma, K. G. M. Moons, R. J. H. M. Steenbakkers, *et al.*, Key challenges in normal tissue complication probability model development and validation: towards a comprehensive strategy, *Radiother Oncol*, vol. 148, pp. 151–156, 2020. DOI: 10.1016/j.radonc.2020.04.012.

[238] C. Veiga, A. M. Lourenço, S. Mouinuddin, M. van Herk, M. Modat, S. Ourselin, *et al.*, Toward adaptive radiotherapy for head and neck patients: Uncertainties in dose warping due to the choice of deformable registration algorithm, *Med Phys*, vol. 42, no. 2, pp. 760–769, 2015. DOI: 10.1118/1.4905050.

[239] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, Diffeomorphic Demons Using ITK's Finite Difference Solver Hierarchy, *Insight J*, 2008. DOI: 10.54294/ux2obj.

[240] F. P. Vidal and P.-F. Villard, Development and validation of real-time simulation of X-ray imaging with respiratory motion, *Computerized Medical Imaging and Graphics*, vol. 49, pp. 1–15, 2016. DOI: 10.1016/j.compmedimag.2015.12.002.

[241] I. Wahlstedt, A. George Smith, C. E. Andersen, C. P. Behrens, S. Nørring Bekke, K. Boye, *et al.*, Interfractional dose accumulation for MR-guided liver SBRT: Variation among algorithms is highly patient- and fraction-dependent, *Radiother Oncol*, p. 109 448, 2022. DOI: 10.1016/j.radonc.2022.109448.

[242] H. Wang, L. Dong, M. F. Lii, A. L. Lee, R. de Crevoisier, R. Mohan, *et al.*, Implementation and validation of a three-dimensional deformable registration algorithm for targeted prostate cancer radiotherapy, *Int J Radiat Oncol Biol Phys*, vol. 61, no. 3, pp. 725–735, 2005. DOI: 10.1016/j.ijrobp.2004.07.677.

[243] D. Wegener, A. Thome, F. Paulsen, C. Gani, J. Boldt, S. Butzer, *et al.*, First Experience and Prospective Evaluation on Feasibility and Acute Toxicity of Online Adaptive Radiotherapy of the Prostate Bed as Salvage Treatment in Patients with Biochemically Recurrent Prostate Cancer on a 1.5T MR-Linac, *J Clin Med*, vol. 11, no. 16, p. 4651, 2022. DOI: 10.3390/jcm11164651.

[244] O. Weistrand and S. Svensson, The ANACONDA algorithm for deformable image registration in radiotherapy, *Med Phys*, vol. 42, no. 1, pp. 40–53, 2015. DOI: 10.1118/1.4894702.

[245] G. Wiernik, Fractionation in radiotherapy (review), *Anticancer Res*, vol. 3, no. 4, pp. 283–297, 1983.

[246] J. Willmann, A. L. Appelt, P. Balermpas, B. G. Baumert, D. de Ruysscher, M. Hoyer, *et al.*, Re-irradiation in clinical practice: Results of an international patterns of care survey within the framework of the ESTRO-EORTC E2-RADIatE platform, *Radiother Oncol*, vol. 189, p. 109 947, 2023. DOI: 10.1016/j.radonc.2023.109947.

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

[248] H. R. Withers, The Four R's of Radiotherapy, in *Advances in Radiation Biology*, J. T. Lett and H. Adler, Eds., vol. 5, Elsevier, 1975, pp. 241–271. DOI: 10.1016/B978-0-12-035405-4.50012-8.

[249] M. Witte, F. Pos, L. Incrocci, and W. Heemsbergen, Association between incidental dose outside the prostate and tumor control after modern image-guided radiotherapy, *Phys Imaging Radiat Oncol*, vol. 17, pp. 25–31, 2021. DOI: 10.1016/j.phro.2020.12.003.

[250] M. G. Witte, W. D. Heemsbergen, R. Bohoslavsky, F. J. Pos, A. Al-Mamgani, J. V. Lebesque, *et al.*, Relating Dose Outside the Prostate With Freedom From Failure in the Dutch Trial 68 Gy vs. 78 Gy, *Int J Radiat Oncol Biol Phys*, vol. 77, no. 1, pp. 131–138, 2010. DOI: 10.1016/j.ijrobp.2009.04.040.

[251] J. M. Wolterink, J. C. Zwienenberg, and C. Brune, Implicit Neural Representations for Deformable Image Registration, in *Proceedings of The 5th International Conference on Medical Imaging with Deep Learning*, PMLR, 2022, pp. 1349–1359.

[252] J. Wong, V. Huang, D. Wells, J. Giambattista, J. Giambattista, C. Kolbeck, *et al.*, Implementation of deep learning-based auto-segmentation for radiotherapy planning structures: a workflow study at two cancer centers, *Radiat Oncol*, vol. 16, no. 1, p. 101, 2021. DOI: 10.1186/s13014-021-01831-4.

[253] R. C. Wortel, L. Incrocci, F. J. Pos, J. V. Lebesque, M. G. Witte, U. A. van der Heide, *et al.*, Acute Toxicity After Image-Guided Intensity Modulated Radiation Therapy Compared to 3D Conformal Radiation Therapy in Prostate Cancer Patients, *Int J Radiat Oncol Biol Phys*, vol. 91, no. 4, pp. 737–744, 2015. DOI: 10.1016/j.ijrobp.2014.12.017.

[254] G. Wu, M. Kim, Q. Wang, Y. Gao, S. Liao, and D. Shen, Unsupervised Deep Feature Learning for Deformable Registration of MR Brain Images, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*, K. Mori, I. Sakuma, Y. Sato, C. Barillot, and N. Navab, Eds., Berlin, Heidelberg: Springer, 2013, pp. 649–656. DOI: 10.1007/978-3-642-40763-5_80.

[255] K.-L. Wu, G.-L. Jiang, H. Qian, L.-J. Wang, H.-J. Yang, X.-L. Fu, *et al.*, Three-dimensional conformal radiotherapy for locoregionally recurrent lung carcinoma after external beam irradiation: A prospective phase I–II clinical trial, *Int J Radiat Oncol Biol Phys*, vol. 57, no. 5, pp. 1345–1350, 2003. DOI: 10.1016/S0360-3016(03)00768-5.

[256] Q. J. Wu, T. Li, Q. Wu, and F.-F. Yin, Adaptive Radiation Therapy: Technical Components and Clinical Applications, *Cancer J*, vol. 17, no. 3, p. 182, 2011. DOI: 10.1097/PPO.0b013e31821da9d8.

[257] J. Xi, L. Luo, J. Li, Y. Zhang, Y. Wang, and D. Jiang, An attention-based residual neural network for deformable image registration, in *Thirteenth International Conference on Information Optics and Photonics (CIOP 2022)*, vol. 12478, SPIE, 2022, pp. 373–383. DOI: 10.1117/12.2653313.

[258] H. Xiao, G. Ren, and J. Cai, A review on 3D deformable image registration and its application in dose warping, *Radiat Med Prot*, vol. 1, no. 4, pp. 171–178, 2020. DOI: 10.1016/j.radmp.2020.11.002.

[259] L. Xing, B. Thorndyke, E. Schreibmann, Y. Yang, T.-F. Li, G.-Y. Kim, *et al.*, Overview of image-guided radiation therapy, *Med Dosim*, vol. 31, no. 2, pp. 91–112, 2006. DOI: 10.1016/j.meddos.2005.12.004.

[260] Y. Xiong, M. Rabe, C. Rippke, M. Kawula, L. Nierer, S. Klüter, *et al.*, Impact of daily plan adaptation on accumulated doses in ultra-hypofractionated magnetic resonance-guided radiation therapy of prostate cancer, *Phys Imaging Radiat Oncol*, vol. 29, p. 100 562, 2024. DOI: 10.1016/j.phro.2024.100562.

[261] D. Yan, F. Vicini, J. Wong, and A. Martinez, Adaptive radiation therapy, *Phys Med Biol*, vol. 42, no. 1, p. 123, 1997. DOI: 10.1088/0031-9155/42/1/008.

[262] J. Yang, S. Vedam, B. Lee, P. Castillo, A. Sobremonte, N. Hughes, *et al.*, Online adaptive planning for prostate stereotactic body radiotherapy using a 1.5 Tesla magnetic resonance imaging-guided linear accelerator, *Phys Imaging Radiat Oncol*, vol. 17, pp. 20–24, 2021. DOI: 10.1016/j.phro.2020.12.001.

[263] Y. X. Yang, S.-K. Teo, E. Van Reeth, C. H. Tan, I. W. K. Tham, and C. L. Poh, A hybrid approach for fusing 4D-MRI temporal information with 3D-CT for the study of lung and lung tumor motion, *Med Phys*, vol. 42, no. 8, pp. 4484–4496, 2015. DOI: 10.1118/1.4923167.

[264] U. J. Yeo, J. R. Supple, M. L. Taylor, R. Smith, T. Kron, and R. D. Franich, Performance of 12 DIR algorithms in low-contrast regions for mass and density conserving deformation, *Med Phys*, vol. 40, no. 10, p. 101 701, 2013. DOI: 10.1118/1.4819945.

[265] U. J. Yeo, M. L. Taylor, J. R. Supple, R. L. Smith, L. Dunn, T. Kron, *et al.*, Is it sensible to "deform" dose? 3D experimental validation of dose-warping, *Med Phys*, vol. 39, no. 8, pp. 5065–5072, 2012. DOI: 10.1118/1.4736534.

[266] Y. Yin, E. A. Hoffman, and C.-L. Lin, Mass preserving nonrigid registration of CT lung images using cubic B-spline, *Med Phys*, vol. 36, no. 9Part1, pp. 4213–4222, 2009. DOI: 10.1118/1.3193526.

[267] A. D. Yock, R. Mohan, S. Flampouri, W. Bosch, P. A. Taylor, D. Gladstone, *et al.*, Robustness Analysis for External Beam Radiation Therapy Treatment Plans: Describing Uncertainty Scenarios and Reporting Their Dosimetric Consequences, *Pract Radiat Oncol*, vol. 9, no. 4, pp. 200–207, 2019. DOI: 10.1016/j.prro.2018.12.002.

[268] C. Zachiu, B. D. d. Senneville, T. Willigenburg, J. R. N. V. v. Zyp, J. C. J. d. Boer, B. W. Raaymakers, *et al.*, Anatomically-adaptive multi-modal image registration for image-guided external-beam radiotherapy, *Phys Med Biol*, vol. 65, no. 21, p. 215 028, 2020. DOI: 10.1088/1361-6560/abad7d.

[269] Y. Zhang, L. Zhang, L. E. Court, P. Balter, L. Dong, and J. Yang, Tissue-specific deformable image registration using a spatial-contextual filter, *Computerized Medical Imaging and Graphics*, vol. 88, p. 101 849, 2021. DOI: 10.1016/j.compmedimag.2020.101849.

[270] L. Zhao and K. Jia, Deep Adaptive Log-Demons: Diffeomorphic Image Registration with Very Large Deformations, *Comput Math Methods Med*, vol. 2015, no. 1, p. 836 202, 2015. DOI: 10.1155/2015/836202.

[271] T. Zhao, Y. Chen, B. Qiu, J. Zhang, H. Liu, X. Zhang, *et al.*, Evaluating the accumulated dose distribution of organs at risk in combined radiotherapy for cervical carcinoma based on deformable image registration, *Brachytherapy*, vol. 22, no. 2, pp. 174–180, 2023. DOI: 10.1016/j.brachy.2022.09.001.

[272] H. Zhong, J. Kim, H. Li, T. Nurushev, B. Movsas, and I. J. Chetty, A finite element method to correct deformable image registration errors in low-contrast regions, *Phys Med Biol*, vol. 57, no. 11, pp. 3499–3515, 2012. DOI: 10.1088/0031-9155/57/11/3499.

[273] H. Zhong, N. Wen, J. J. Gordon, M. A. Elshaikh, B. Movsas, and I. J. Chetty, An adaptive MR-CT registration method for MRI-guided prostate cancer radiotherapy, *Phys Med Biol*, vol. 60, no. 7, p. 2837, 2015. DOI: 10.1088/0031-9155/60/7/2837.

[274] P. Ziegenhein, C. P. Kamerling, M. F. Fast, and U. Oelfke, Real-time energy/mass transfer mapping for online 4D dose reconstruction, *Sci Rep*, vol. 8, no. 1, p. 3662, 2018. DOI: 10.1038/s41598-018-21966-x.

[275] M. Ziegler, M. Nakamura, H. Hirashima, R. Ashida, M. Yoshimura, C. Bert, *et al.*, Accumulation of the delivered treatment dose in volumetric modulated arc therapy with breath-hold for pancreatic cancer patients based on daily cone beam computed tomography images with limited field-of-view, *Med Phys*, vol. 46, no. 7, pp. 2969–2977, 2019. DOI: 10.1002/mp.13566.

8 Statement on own contribution

Statement of own contribution of the dissertation

This thesis was conducted in the Section for Biomedical Physics at the University Hospital for Radiation Oncology Tübingen under the 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. Arndt-Christian Müller.

Parts of this thesis were published as peer-reviewed papers. Statements of contribution for each of the collective works are listed below.

I affirm that I have written the thesis manuscript myself and have used no other sources than the ones indicated.

Tübingen, the 01.10.2024

Applicability and usage of dose mapping/accumulation in radiotherapy

Statement of own contribution: **Martina Murr**

Title: Applicability and usage of dose mapping/accumulation in radiotherapy

Autor: Martina Murr, Kristy K. Brock, Marco Fusella, Nicholas Hardcastle, Mohammad Hussein, Michael G. Jameson, Isak Wahlstedt, Johnson Yuen, Jamie R. McClelland, Eliana Vasquez Osorio

Published in: Radiotherapy and Oncology 2023, Vol. 182, Pages 109527

The research leading to this article was conducted by an international group of researchers that participated at the ESTRO Physics workshop 2021 on "commissioning and quality assurance of deformable image registration (DIR) for current and future RT applications" working group on dose mapping and accumulation.

The working group was led by myself with the supervision of Eliana Vasquez Osorio and Jamie R. McClelland.

Myself, Kristy K. Brock, Marco Fusella, Nicholas Hardcastle, Mohammad Hussein, Michael G. Jameson, Isak Wahlstedt, Johnson Yuen, Jamie R. McClelland, and Eliana Vasquez Osorio collaborated to develop the research concept.

I was supported in the selection of methods by Kristy K. Brock, Marco Fusella, Nicholas Hardcastle, Mohammad Hussein, Michael G. Jameson, Isak Wahlstedt, Johnson Yuen, Jamie R. McClelland and Eliana Vasquez Osorio.

Data analysis and interpretation of results were performed by myself in collaboration with the group.

I conceptualized the manuscript with the support of Eliana Vasquez Osorio, with essential input from all co-authors.

All co-authors contributed to the essential proofreading of the manuscript.

A multi-institutional comparison of retrospective deformable dose accumulation for online adaptive MR-guided radiotherapy

Statement of own contribution: **Martina Murr**

Title: A multi-institutional comparison of retrospective deformable dose accumulation for online adaptive MR-guided radiotherapy

Autor: Martina Murr, Uffe Bernchou, Edyta Bubula-Rehm, Mark Ruschin, Parisa Sadeghi, Peter Voet, Jeff D. Winter, Jinzhong Yang, Eyesha Younus, Cornel Zachiu, Yao Zhao, Hualiang Zhong, Daniela Thorwarth

Published in: Physics and Imaging in Radiation Oncology 2024, Vol. 30, Pages 100588

The original research article was conducted by the dose accumulation working group of the Elekta MR-Linac consortium under the supervision of Prof. Dr. rer. nat. Daniela Thorwarth.

The study was designed in collaboration with Uffe Bernchou, Edyta Bubula-Rehm, Mark Ruschin, Peter Voet, Jeff D. Winter, Jinzhong Yang, Cornel Zachiu, Hualiang Zhong, Daniela Thorwarth and myself.

Edyta Bubula-Rehm, Mark Ruschin, Peter Voet, Jeff D. Winter, Jinzhong Yang, Cornel Zachiu, Hualiang Zhong and Daniela Thorwarth supported me in selecting the methods.

I received assistance in patient recruitment from Uffe Bernchou, Cornel Zachiu, Hualiang Zhong, and Daniela Thorwarth.

Mark Ruschin, Parisa Sadeghi, Jeff D. Winter, Jinzhong Yang, Eyesha Younus, Cornel Zachiu, Yao Zhao and Hualiang Zhong assisted me with the data acquisition.

I analysed the data in collaboration with Cornel Zachiu and with the support of Daniela Thorwarth.

The interpretation of the results was done by me with support of Mark Ruschin, Jeff D. Winter, Hualiang Zhong, Jinzhong Yang, Cornel Zachiu, and Daniela Thorwarth.

The manuscript was written by me with input from Uffe Bernchou, Edyta Bubula- Rehm, Mark Ruschin, Parisa Sadeghi, Peter Voet, Jeff D. Winter, Jinzhong Yang, Eyesha Younus, Cornel Zachiu, Yao Zhao, Hualiang Zhong and Daniela Thorwarth.

All co-authors provided critical proofreading of the manuscript.

Comparison of online adaptive and non-adaptive magnetic resonance image-guided radiation therapy in prostate cancer using dose accumulation

Statement of own contribution: **Martina Murr**

Title: Comparison of online adaptive and non-adaptive magnetic resonance image-guided radiation therapy in prostate cancer using dose accumulation

Autor: Martina Murr, Daniel Wegener, Simon Böke, Cihan Gani, David Mönnich, Maximilian Niyazi, Moritz Schneider, Daniel Zips, Arndt-Christian Müller, Daniela Thorwarth

Submitted to: Physics and Imaging in Radiation Oncology 2024, manuscript number: PHIRO-D-24-00218, current status: In Revision.

The research concept was developed by Arndt-Christian Müller, Daniel Zips and Daniela Thorwarth.

In collaboration with Arndt-Christian Müller and Daniela Thorwarth, I selected the methods used in this study.

Patient recruitment was handled by Daniel Wegener, Simon Böke, Cihan Gani, Daniel Zips, and Arndt-Christian Müller.

I had the support of Daniel Wegener, Simon Böke, Cihan Gani, David Mönnich, Maximilian Niyazi, Moritz Schneider, and Arndt-Christian Müller for the data collection.

The data analysis and interpretation of the results were performed by myself with support from Arndt-Christian Müller and Daniela Thorwarth.

I executed the conception of the manuscript with the assistance of Daniela Thorwarth.

All co-authors contributed to the essential proofreading of the manuscript.

9 Supplementary

9.1 Applicability and usage of dose mapping/accumulation in radiotherapy

9.1.1 Supplementary A: Selected use cases

Three use cases from the DMAL were selected to be discussed in detail, highlighting critical issues, current limitations and safety concerns when using dose mapping/accumulation. Notice that these cases are in the extremes of DMAL, Figure 11 in the main manuscript.

Adaptive radiotherapy (ART)
ART enables adapting a treatment to variations on the patient anatomy, by using treatment images such as CT, CBCT, and MR [139]. The adaptation can be achieved online (same day) or offline (scheduled or ad hoc) [22].

Ideally, dose accumulation in ART would allow to optimise treatment fractions based on dose delivered in previous fractions, while properly accounting for tumour regression (that is, elastic and inelastic regression modes). This use can be translated to "incorporate" in our DMAL uncertainty axis. However, due to several limitations in DIR algorithms, including the proper management of tumour regression, dose accumulation is mostly used to quantify the delivered dose to the tumour and critical organs when each treatment fraction is a different plan and inform replanning. Therefore, dose accumulation is currently not standard for offline/online ART.

Different approaches can be taken when dose is mapped and accumulated in the ART scenario. For example, one can map the dose in a "cascade style", where the previous fraction is mapped to the current fraction [25] (Figure S1, top), or one can map all fractions to a single reference image (1-n registrations, a.k.a. back-propagation, etc), e.g., planning CT or the first fraction image (Figure S1, bottom). Depending on the used approach, registration uncertainties will impact differently: in the cascade approach, errors in early registrations are propagated, while in the 1-n approach progressive changes will result in more challenging registration for images acquired reaching the end of treatment. Moreover, the order in which the doses are mapped can change the accumulated dose result unless the algorithm is inverse-consistent and transitive [20].

In the DMAL, ART is positioned in the anatomical variation axis from small to intermediate changes Figure 11. We note that for some sites sudden large changes can happen such as lung deflation; however, the treatment may be interrupted for a complete re-plan. Particular challenges under this context include (dis)appearing of tissue (e.g. tumour), large volume changes (e.g. bladder or intestines) and deformations as well as small volume for registration (e.g. penile bulb) and image field of view cropped

images and dose warping (either via DDM or EMT). At the impact of dose mapping uncertainties axis, offline ART is positioned at the beginning of high as it is used informative (retrospective, offline) in a patient-specific setting which typically does not affect the patient/treatment directly. However, notice that if a new plan is created mid-way in treatment due to the patient's anatomy changing dramatically, for example lung atelectasis, any uncertainty when estimating the previously given dose would systematically impact the new treatment, increasing the risk of impact of dose mapping on patient safety. Online ART on the other hand is positioned at high, whenever mapped/accumulated dose is directly used for treatment optimisation (such as in "incorporate"), as it may directly affect treatment outcome if mapping uncertainties result in cancerous regions being underdosed or OARs overdosed.

Caution is advised with plans adapted on accumulated doses. We currently recommend against daily plan adaptation based on dose accumulation in clinical settings when the mapped dose is directly incorporated in treatment optimisation. In particular, when soft-tissue contrast is limited, registration uncertainties will directly impact dose mapping [264]. Studies reporting whether accumulated dose helps to better estimate risks of toxicities are a way to demonstrate the value of the tool, for example, retrospective MRgRT dose accumulation study of Bohoudi et al. showed potential to reduce the risk of acute urinary toxicity using dose accumulation [22]. In this case, mapped and accumulated doses were used to inform treatments.

Re-irradiation

In a recent consensus, re-irradiation has been defined as "a new course of radiotherapy, either to a previously irradiated volume (irrespective of concerns for toxicity) or where the cumulative dose raises concerns of toxicity [6]", where often months or years happen between the two radiation treatments. Thus the key consideration in this scenario is the dose to critical organs that may have been previously irradiated. Re-irradiation is becoming more prevalent due to the ability of contemporary radiation therapy planning, delivery and image guidance technology to sculpt radiation dose distributions, minimising dose to previously treated critical organs/structures. The efficacy and safety of re-irradiation to tumour sites throughout the body has been demonstrated [79], [117], [120], [163], [166], [180], [192], [202], [212], however there still remains instances of high risk of catastrophic side effects due to excessive cumulative doses to critical organs/structures [70], [121], [149], [234], [255]. The risk of serious side effects is typically linked to complications that arise from maximum doses to structures ("serial" complications [49]). Notably, the cumulative maximum dose to a given organ/structure does not necessarily arise at the location of maximum dose from individual treatment courses. The concerns of toxicity from cumulative doses and whether the new course of RT overlaps with ir-radiated volume of previous course are both to be considered [6]. A lesser, but still important consideration is the less severe side effects which may impact patient quality of life such as brain necrosis and lung pneumonitis [121], [172], [255] are dependent not only on maximum organ doses but often volumetric metrics such as mean dose.

In re-irradiation, it is recommended that if the previous dose distributions are available,

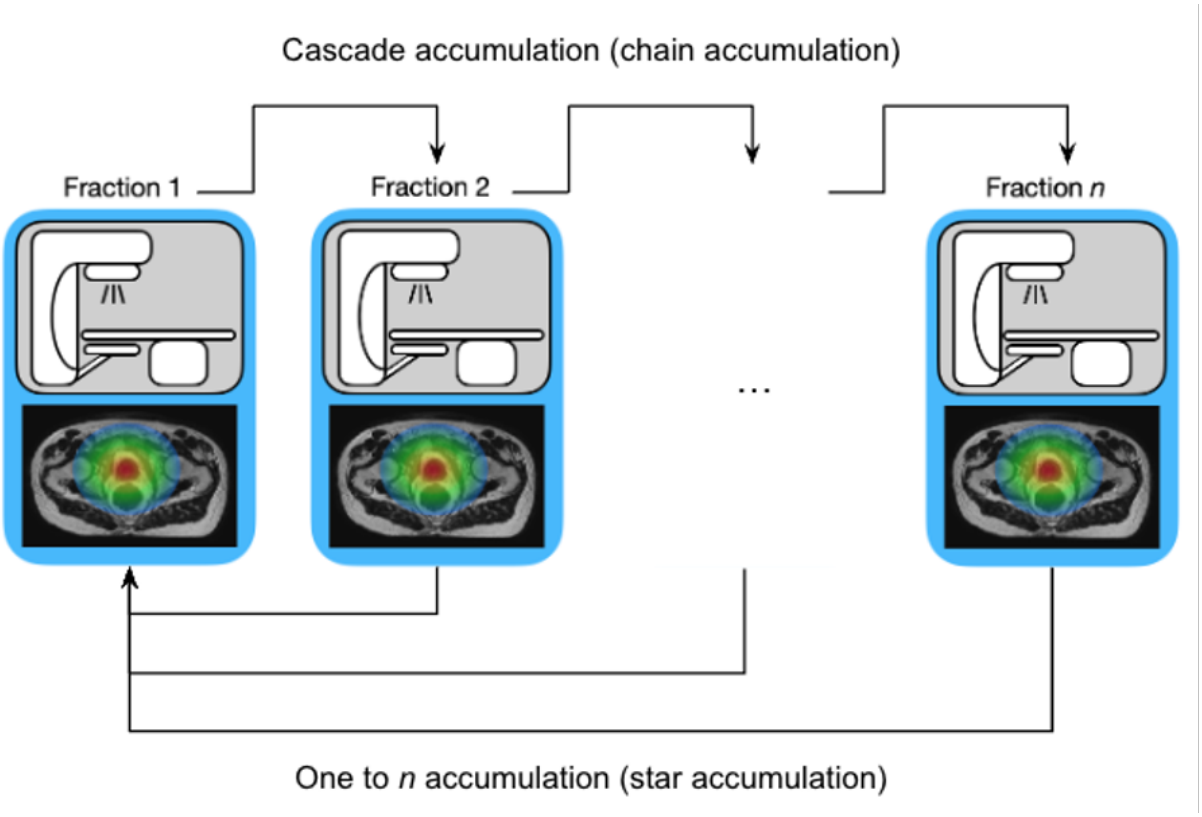

Figure S1: Illustration of the dose accumulation scenarios in ART. The "cascade" accumulation (chain accumulation) is illustrated on the top. On the button, the "one to n" accumulation also known as star accumulation is illustrated.

these are overlaid onto the new treatment planning image [6], [26], [175], [184], [209]. Dose from previous treatment(s) can be deformably mapped to a new planning image to assess delivered dose from previous treatment(s) to each spatial sub-volume of the organs/structures of interest as they appear in the new image.

The complexity of cumulative dose assessment depends on the quality of prior dose information, ranging from conservatively approximations of dose from previous treatment records in text or portal images, to dose accumulation of biologically equi-effective 3D dose [6]. From this information, the treating team can determine what dose can be safely delivered to each location of the critical organs/structures to minimise the risk of serious side effects from re-irradiation. It should be noted the importance of conversion of doses to biologically equieffective dose such as EQD2 via the linear quadratic formalism [26], [184], [209]. Interestingly, there is no consensus on the adequate order for dose mapping and radiological correction, with approaches using both orders in the literature (e.g., map then correct [175] vs correct then map [35]). Additionally, tissue recovery factors, applied to the dose in specific organs typically spinal cord and central nervous system structures may or may not be included.

The position of re-irradiation in the DMAL is highly variable depending on the organ and dose to consider. Given the critical safety issue when quantifying the previously delivered dose (and cumulative dose from all courses), it is important that a high level of caution is exercised when applying deformable dose accumulation in re-irradiation.

The position, size and shape of critical organs such as the spinal cord may be relatively constant over time-spans of months to years between treatment courses. As such, the anatomical variation is relatively small and rigid registration can be done with a relatively high degree of accuracy. Contrast this with tubular structures such as the bowel, where the filling and position will vary substantially not only between treatment courses but between each treatment session in a given treatment course. In this scenario, measurement of registration accuracy is extremely challenging and deformable dose mapping should be avoided; it is highly unlikely the accuracy of mapped dose can be determined in the instance where no clear corresponding anatomical landmarks exist. In between these two are deformable organs such as blood vessels and the oesophagus, where DIR accuracy can be improved and or assessed using contours in each image. In these cases deformable dose mapping may be considered.

Where mean doses are of importance, the variation in mean dose with different image registrations may be minimal, but likely depending on the volume of the organ; deformed mean dose to larger volume organs such as the lung and liver being more robust to uncertainties in dose mapping.

There is a high impact on patient safety from dose mapping uncertainty in organ maximum doses. To classify where high accuracy deformable dose mapping is required, one can perform a conservative estimate; in the absence of registration between the two courses, does the summation of the point maximum doses for a given organ from each individual course approach or exceed the tolerance for that organ? This is the

worst case scenario; in this case, the application of deformable dose mapping should be accompanied with a high level of rigour to maximise the accuracy of the underlying deformation, and to quantify and communicate the uncertainty. The uncertainty may be estimated via performing the assessment using multiple methods such as the conservative approach described above, and by performing rigid registration in a localised region where high doses overlap.

A further consideration is the relevance of the planned dose in the absence of real-time dose accumulation during the course of each treatment using daily volumetric imaging. The estimated dose from the treatment planning scan to relatively rigid structures such as bones and spinal cord may be very close to the delivered dose. The planned dose to organs subject to a high level of random variation day-to-day however may be a poor estimate of the actual delivered dose to the organ. Thus using the planned dose grid from the original treatment course for highly variable structures may place a ceiling on the achievable accuracy of dose accumulation in the re-irradiation scenario.

Finally, we would recommend assessing the impact of registration uncertainties on the mapped dose. Methods have been proposed, see section first-order effects in the manuscript. As these tools are not currently available in treatment planning systems, performing different registrations (by varying either algorithm parameters/initial global alignment) and comparing the doses will help identify regions where the mapped dose is (un)stable.

Response assessment

Dose-response assessment tries to establish a relationship between the dose received by an organ (or subvolume of an organ) and clinical outcome. To increase the therapeutic ratio of RT, relations of true accumulated dose delivered to tumours and organs at risk (OARs), rather than the planned pretreatment dose, should be related to clinical outcome [98].

Per-patient dose-response assessment

The dose delivered during a treatment can be accumulated between treatment fractions to account for inter-fractional anatomical variations. Bohoudi et al. [25] used such an approach to estimate the dose delivered to the bladder for 101 prostate cancer patients treated with five fractions of magnetic resonance-guided stereotactic body radiation therapy. They showed that an increase in International Prostate Symptoms Score correlated better with bladder doses accumulated after three and five fractions of prostate RT than with pretreatment planned dose. This indicates that dose accumulation could improve understanding of dose-response relationships.

Voxel-wise analysis

Another implementation of dose mapping for dose response assessment is voxel-wise analysis. Voxel-wise analysis aims at exploring and locating dose-sensitive anatomical regions for a given treatment outcome. This analysis relies on mapping dose distributions of the studied cohort into a reference anatomy, commonly using DIR [182].

Voxel-wise analysis is one of the few inter-patient dose mapping cases in current use. As anatomical variations between patients are substantial, larger deviations are expected than the per-patient approach described above. For example, this methodology has been used to identify the association between dose to the obturatorial lymph node region and treatment failure after prostate RT [249], [250], dose to the heart and overall survival after lung RT[40], [150], dose to the masseter muscle and trismus after head and neck RT [19], among others.

Both approaches are used to quantify and compare dose distributions given to patients, in a retro-spective fashion. The results of these studies are hypothesis generating and form the basis of subsequent analysis aiming at confirming or rejecting the hypothesis. In this context, the impact of registration uncertainties is limited for a given patient.

9.1.2 Supplementary B: List of discussed articles

Table S1: List of discussed articles on DIR and dose accumulation to identify current challenges during the online meetings and via email.

Ref.	Aim	Highlights, issues and challenges	Keywords
Articles discussed during the online meetings			
Bosma et al., 2021 [27]	Influence of DIR algorithm, dose warping strategy, and signal-to-noise ratio on accumulated motion-compensated dose for MR-guided radiotherapy of prostate cancer.	The chosen dose warping strategy, DDM vs. EMT, affected the accumulated dose more than the DIR algorithm employed or the signal-to-noise ratio of the images.	DDM/EMT motion-compensated dose accumulation (MCDA) Prostate
Li et al., 2014 [136]	Comparison of dosimetric difference between DDM and EMT for 4D dose accumulation in lung SBRT.	Differences between DDM and EMT of 11 % and 4 % were observed for PTV and ITV minimum dose, respectively.	DDM/EMT
Rosu et al., 2005 [199]	Investigation of dose grid size and dose interpolation method on dose accumulation in a deforming lung anatomy.	The interpolation method alone did not yield clinically significant changes in dose accumulation results. These effects were larger in high dose gradient regions for larger dose grid sizes.	Dose grid size effects (Lung)
Niebuhr et al., 2021 [171]	Based on daily CT images an assessment of biological dose accumulation was performed.	The usage of bEQDd (total biological dose) in terms of dose accumulation, copes with a potential systematic inaccuracy predicting biological effects, in particular, serial OAR in regions with dose gradients and for hypofractionation.	Biological & physical dose accumulation Re-irradiation
Nix et al., 2021 [175]	Development of a software tool for use with commercial treatment planning systems, taking radiation biology and anatomical changes into account, by comparing three dose summation approaches.	Re-irradiation cases require fractionation correction to meaningfully assess cumulative doses and reduce the risk of unintentional OAR overdose. DIR can add clinically relevant information in selected cases, especially for significant anatomical change.	Re-irradiation Pelvic Dose summation Radiobiologically

Continued on next page

Table S1 – continued from previous page

Ref.	Aim	Highlights, issues and challenges	Keywords
Kainz et al., 2022 [106]	Presentation of a new and practical QA method, the DVH overlap technic, for DIR-based dose accumulation.	In order to assess DIR for dose accumulation, in which the ICE was low, the DVH overlay technique can be used.	QA Inverse consistency
Paganelli et al., 2018 [181]	Enabling a patient-specific assessment of registration quality and subsequently improved personalised treatment.	Patient-specific evaluation of DIR with both geometric and dosimetric methods were reviewed; challenges in clinical use, guidelines, and validation of deformable dose propagation were discussed.	QA DIR Patient specific validation
Palma et al., 2020 [182]	Review article on how to implement voxel-wise analysis to explore local anatomical radiosensitivity for a given outcome.	As DIR is often used to map the dose distributions of multiple patients to a common reference to analyse a given treatment outcome, several considerations are introduced in this article. In particular, different methods to evaluate inter-patient DIR are presented.	VBA Image-based data-mining Outcome modelling
Chetty and rosu-Bubulac, 2019 [45]	Review of currently available DIR and dose accumulation methods as well as their related clinical application.	Overview article of DIR used for dose accumulation, mainly focused on the registration aspects.	DIR Methods of DA
Garcia-Molla et al., 2015 [69]	Validation of a CT-CBCT ART DIR algorithm by using POIs, and development of a method of encountering the results affecting the dose by using IC.	The algorithm is useful for head and neck patients, however clinical decisions for plan adaptation need to be ruled with respect to the variation of the results.	ART Head and neck Validation
Qin et al., 2018 [189]	Investigation of the differences of clinical-relevant doses in intensity-homogeneous organs warped by purely image-based and biomechanical model based DIR (BM-DIR).	For intensity-homogeneous organs BM-DIR could be useful, because these organs are affected with larger deformation, shrinkage.	Biomechanical model Dose warminguncertainty
Roussakis et al., 2015 [201]	Workflow demonstration and evaluation for the interfractional deformable image registration followed by dose accumulation.	Under clinically realistic scenarios the dependability of a CT-to-CT commercial DIR-based dose warping and a workflow to assist clinical personnel to assess the results was shown.	QA, uncertainties head and neck
Bohoudi et al., 2021 [25]	Identification of bladder dose parameters based on the accumulated treatment dose and investigation of the prediction of the accumulated treatment dose via the first three fraction doses.	In the prediction of urinary symptoms the accumulated dose was more specific than the planned dose, also the prediction of the accumulated dose via the first three fraction doses showed a positive correlation.	Accumulated dose prediction Toxicity prediction
Ziegenhein et al., 2018 [274]	Development of a real-time dose reconstruction via energy/mass transfer mapping.	The developed method makes real-time dose reconstitution via energy/mass transfer. Clinically applicable with the caveat that more computing power is required.	EMT Real-time dose reconstruction
Chao et al., 2012 [41]	Development of a retrospectively reconstruction method of the delivered dose by using DIR in relation to the anatomical changes.	The method provides verification of the treatment dose with potential use in ART.	Verification treatment dose ART Irradiation total body

Continued on next page

Table S1 – continued from previous page

Ref.	Aim	Highlights, issues and challenges	Keywords
Rigaud et al., 2019 [194]	Description of current main DIR applications, their methods and principles and the evaluation.	DIR should be evaluated with geometrical and dosimetric metrics appropriately.	DIR QA
Additional discussed articles			
Yin et al., 2009 [266]	Introduces a new similarity measure, the sum of squared tissue volume difference (SSTVD), which accounts for density changes between CT images, thus ensuring mass is preserved by the registration.	Over the registration pairs the SSTDV methods results showed a smaller average landmark error than the SSD method.	Sum of squared tissue volume difference SSTVD
Frederick et al., 2022 [66]	Demonstration of a volumetric image-guided radiotherapy and dose accumulation approach to derive planning target volume (PTV) margins for a patient population.	The derived PTV margins were 3 and 5 mm and ensured that in 90% of the patients at 98% of the CTV 95% of the prescribed dose was delivered.	PTV margins Oropharyngeal cancer Breast cancer
Kim et al., 2021[116]	Review for external beam radiation therapy and brachytherapy dose summation strategies.	For clinical use DIR is still in development regarding accuracy and efficiency. From the effective biological dose a composted DVH of the biological dose summation (EBRT dose map and the IGBT fractions) can be generated but an accurate radiobiological parameter is tissue dependant. The limitations of these methods of approximating dose accumulation needs to be considered.	Dose summation Biologic weighted dose maps

9.2 A multi-institutional comparison of retrospective deformable dose accumulation for online adaptive MR-guided radiotherapy

9.2.1 Supplementary A: Material and Methods

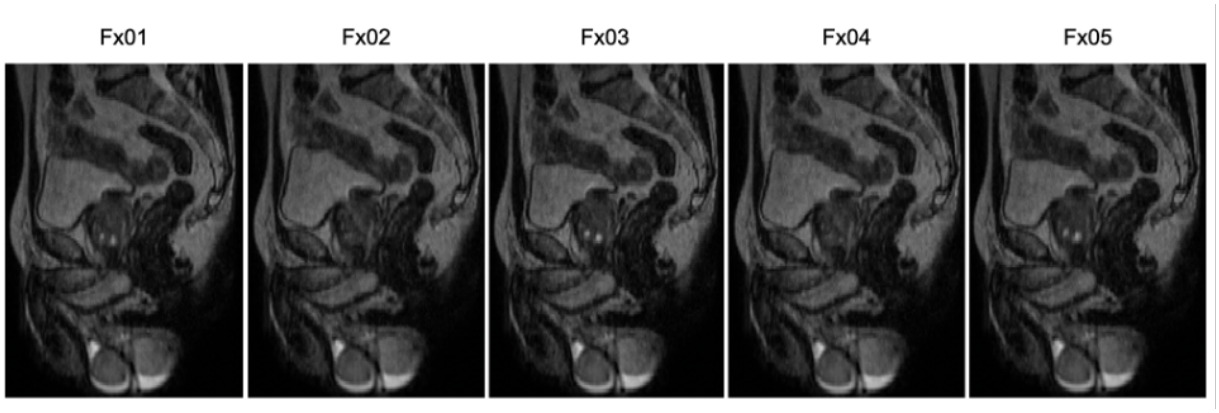

Figure S2: Illustration of the five deformed images, which resulted from the FEM experiment, in a representative sagittal slice.

*Table S2: Specific MR images, dose files, and SBRT treatment information for the gold standard (GS), and the clinical cases; (I) prostate 1, (II) prostate 2, (III) cervix, (IV) liver, and (V) lymph node. Abbreviations; TR: repetition time TE: echo time. * MR image acquisition type: 4D, a MidP image was created for planning, and a motion-averaged 3D MR image from the 4D MR image was used for DIR [185].*

Specific information	MR Image												Dose				SBRT Treatment	
Cases	MR Image type for DIR	TR [ms]	TE [ms]	Pixel Bandwith [Hz/Px]	Flip Angle [°]	Acquisition Matrix	Number of Frames	Rows	Columns	Pixel Spacing [mm]	Slice Thickness [mm]	Spacing between Slices [mm]	Number of Frames	Rows	Columns	Pixel Spacing [mm]	Fractions	Dose [Gy]
Gold standard	3D	1535	277	740	90	268 x 268	150	480	480	0.83 x 0.83	2	2	100	150	217	3 x 3	5	7.25
(I) Prostate 1	3D	1535	277	740	90	268 x 268	150	480	480	0.83 x 0.83	2	2	100	150	217	3 x 3	5	7.25
(II) Prostate 2	3D	1535	277	740	90	268 x 268	150	480	480	0.83 x 0.83	2	2	100	150	217	3 x 3	5	7.25
(III) Cervix	3D	1400	150	694	90	497 x 497	125	672	672	0.8 x 0.8	2	2	85	188	217	3 x 3	5	7
(IV) Liver	3D*	4.3	2.071	775	50	256 x 256	84	256	256	1.64 x 1.64	2.38	2.38	67	190	217	3 x 3	5	7
(V) Lymph node	3D	1535	277	740	90	268 x 268	300	480	480	0.83 x 0.83	2	1	150	251	325	2 x 2	5	6

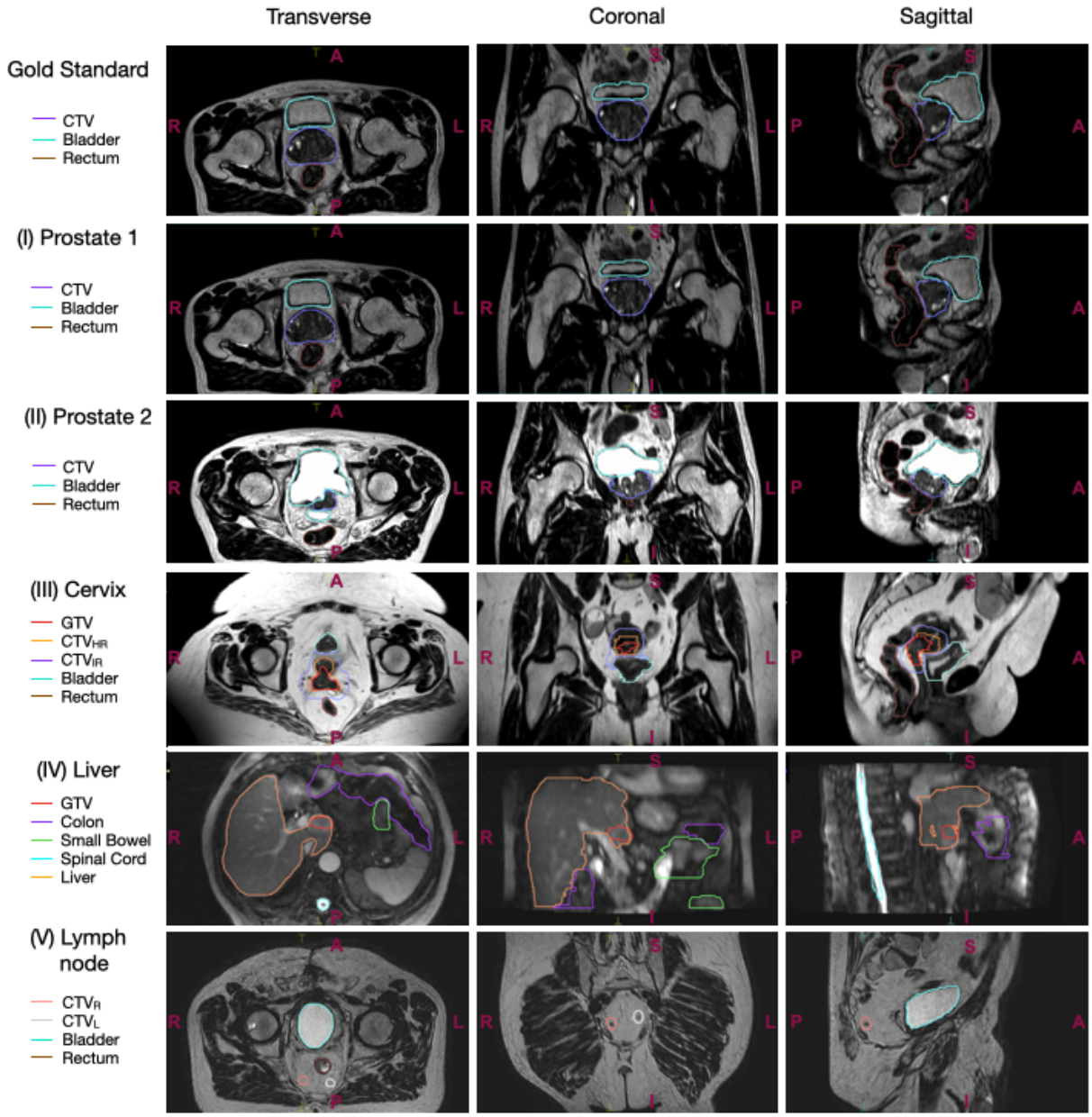

Figure S3: Illustration of the golds standard (GS), and the clinical cases; (I) prostate 1, (II) prostate 2, (III) cervix, (IV) liver, and (V) lymph node. The illustration shows one slice of transverse, coronal and sagittal view of the first fraction (Fx01) of each case. The different contours per case show the offline recontoured contours of Fx01. Note that for the GS, prostate 1 fraction 1 was used as base image. Abbreviations; CTV: clinical target volume, GTV: gross tumor volume, CTV$_{HR}$: High risk clinical target volume, CTV$_{IR}$: Intermediate risk clinical target volume, R: right, L: left, P: posterior, A: anterior, I: inferior and S: superior.

Table S3: Clinical dosimetric criteria (CDC) results, fraction-based and deformable dose accumulation (DDA) of the institutes for the gold standard (GS) and the clinical cases; (I) prostate 1, (II) prostate 2, (III) cervix, (IV) liver, and (V) lymph nodes. The respective CDC for the individual institutes DDA were calculated by ProKnow by the use of offline re-contoured structure set of fraction one (Fx01). Abbreviations; CTV: clinical target volume, GTV: gross tumor volume, CTV_{HR}: High risk clinical target volume, CTV_{IR}: Intermediate risk clinical target volume, CTV_L: left clinical target volume, CTV_R: right clinical target volume, Sum: summed dose.

Case	Clinical dosimetric criteria (CDC)				Fraction based						Accumulated							
					Fx01	Fx02	Fx03	Fx04	Fx05	Sum	A	B	C	D	E	F	Mean	Std
Gold Standard	CTV	V34.4Gy	$\geq$ 99.0	[%]	99.8	83.4	99.9	97.5	100.0	96.1	86.3	93.1	92.6	92.9	95.8	92.0	92.1	3.1
	Bladder	$D5cm^3$	< 37.0	[Gy]	36.9	37.8	34.4	35.6	35.6	36.1	34.6	34.0	34.2	34.2	33.0	33.8	34.0	0.5
		V28.0Gy	< 15.0	[%]	10.9	15.9	8.7	6.1	8.0	9.9	8.6	8.0	7.9	7.9	8.1	7.7	8.0	0.3
		V32.0Gy	< 20.0	[%]	7.5	12.0	4.9	3.9	5.1	6.7	4.7	4.1	4.1	4.1	3.2	3.9	4.0	0.5
	Rectum	$D1cm^3$	$\leq$ 38.0	[Gy]	36.3	27.6	35.9	35.8	37.7	34.6	33.0	34.0	34.1	34.0	35.3	33.9	34.1	0.7
		V28.0Gy	$\leq$ 15.0	[%]	12.1	1.4	14.8	12.4	17.3	11.6	9.8	10.2	10.3	9.8	12.5	9.8	10.4	1.1
		V32.0Gy	$\leq$ 20.0	[%]	8.5	0.3	10.6	8.4	13.0	8.2	3.5	4.8	4.9	4.5	7.6	4.2	4.9	1.4
(I) Prostate1	CTV	V34.4Gy	$\leq$ 99.0	[%]	100.0	99.7	100.0	100.0	100.0	99.9	100.0	100.0	100.0	99.9	100.0	100.0	100.0	0.0
	Bladder	$D5cm^3$	< 37.0	[Gy]	34.7	31.7	36.0	36.3	36.3	35.0	33.4	35.8	34.0	31.9	36.5	33.8	34.2	1.7
		V28.0Gy	< 15.0	[%]	7.6	2.8	10.5	5.5	12.6	7.8	7.6	11.3	8.1	5.9	16.1	7.5	9.4	3.7
		V32.0Gy	< 20.0	[%]	4.5	1.2	6.8	3.7	8.4	4.9	3.8	7.1	4.5	2.5	10.5	4.1	5.4	2.9
	Rectum	$D1cm^3$	$\leq$ 38.0	[Gy]	37.3	37.6	36.6	37.7	37.2	37.3	36.4	35.8	36.5	36.7	36.5	36.3	36.4	0.3
		V28.0Gy	$\leq$ 15.0	[%]	15.5	15.7	12.9	16.0	10.9	14.2	14.0	14.0	14.6	14.5	14.7	13.9	14.3	0.4
		V32.0Gy	$\leq$ 20.0	[%]	11.8	12.2	9.2	12.3	6.5	10.4	9.6	9.4	10.4	10.4	9.6	9.0	9.7	0.6
(II) Prostate2	CTV	V34.4Gy	$\leq$ 99.0	[%]	99.8	100.0	99.5	100.0	100.0	99.9	100.0	100.0	100.0	100.0	97.1	100.0	99.5	1.2
	Bladder	$D5cm^3$	< 37.0	[Gy]	36.8	36.6	36.6	36.7	36.4	36.6	36.3	36.2	36.3	36.3	36.5	36.2	36.3	0.1
		V28.0Gy	< 15.0	[%]	30.4	31.9	18.4	23.0	20.5	24.8	22.8	22.4	22.7	25.6	30.0	22.7	24.4	3.0
		V32.0Gy	< 20.0	[%]	22.6	23.0	12.9	16.0	14.9	17.9	14.8	14.1	14.7	16.8	21.5	14.5	16.1	2.8
	Rectum	$D1cm^3$	$\leq$ 38.0	[Gy]	33.6	35.3	36.1	36.8	36.0	35.6	34.7	34.8	34.0	35.0	32.8	34.9	34.4	0.8
		V28.0Gy	$\leq$ 15.0	[%]	6.6	9.8	11.5	16.7	9.1	10.7	9.0	8.4	8.1	10.7	5.5	9.5	8.5	1.7
		V32.0Gy	$\leq$ 20.0	[%]	3.0	5.2	6.1	11.3	4.8	6.1	4.7	4.6	3.7	5.5	2.2	5.2	4.3	1.2
(III) Cervix	GTV	D98%	> 35.0	[Gy]	35.8	35.5	34.4	35.6	34.4	35.1	34.9	35.2	34.8	34.7	33.8	34.6	34.7	0.5
	CTV_{HR}	D90%	> 33.0	[Gy]	33.6	33.7	33.2	33.4	28.2	32.4	33.0	33.0	32.9	33.0	32.9	33.0	33.0	0.1
	CTV_{HR}	D98%	> 25.0	[Gy]	32.0	32.0	31.6	31.8	25.7	30.6	31.0	30.9	31.0	30.9	31.1	31.2	31.0	0.1

Continued on next page

Table S3 – continued from previous page

Case	Clinical dosimetric criteria (CDC)				Fraction based						Accumulated							
					Fx01	Fx02	Fx03	Fx04	Fx05	Sum	A	B	C	D	E	F	Mean	Std
	CTV_{IR}	D98%	> 14.5	[Gy]	24.9	24.4	24.6	25.1	18.2	23.4	25.3	25.4	25.2	25.7	23.8	25.2	25.1	0.7
	Bladder	$D2cm^3$	< 27.5	[Gy]	26.7	27.8	27.0	27.3	26.0	27.0	26.9	26.9	26.9	26.4	27.0	26.7	26.8	0.2
	Rectum	$D2cm^3$	< 21.5	[Gy]	16.5	19.4	14.3	21.3	15.9	17.5	17.4	17.1	17.3	19.3	16.5	17.7	17.6	0.9
	Sigmoid	$D2cm^3$	< 21.5	[Gy]	21.4	20.7	22.0	20.9	22.6	21.5	21.1	22.1	20.9	23.5	21.7	21.1	21.7	1.0
(IV) Liver	Colon	Dmax	< 32.0	[Gy]	21.7	19.5	21.0	18.3	20.1	19.7	22.6	23.5	22.9	19.7	17.1	22.8	21.4	2.5
	Duodenum	Dmax	< 30.0	[Gy]	21.8	20.7	24.0	32.4	18.4	23.9	23.7	22.5	23.0	17.7	16.6	22.8	21.1	3.1
	Small bowel	Dmax	< 30.0	[Gy]	11.4	16.4	14.9	15.5	11.9	14.7	12.7	10.2	11.0	10.2	11.3	11.0	11.1	0.9
	Spinal cord	Dmax	< 22.5	[Gy]	5.2	6.4	5.3	5.7	6.0	5.9	5.4	5.5	5.4	5.1	5.3	5.4	5.4	0.1
	Stomach	Dmax	< 30.0	[Gy]	31.8	28.2	33.1	32.3	30.6	31.1	28.9	29.8	29.3	26.3	26.6	29.8	28.5	1.6
(V) Lymphnode	CTV_L	$D0.5cm^3$	< 42.0	[Gy]	41.5	41.8	41.2	41.9	41.8	41.7	40.8	39.3	40.9	38.2	40.1	41.5	40.1	1.2
		D98%	> 30.0	[Gy]	31.8	32.8	33.2	30.5	31.8	32.1	37.4	33.6	39.6	33.1	37.5	38.6	36.6	2.7
	CTV_R	$D0.5cm^3$	< 42.0	[Gy]	41.3	41.1	41.5	41.9	42.5	41.8	41.0	40.6	40.6	37.4	41.1	41.7	40.4	1.5
		D98%	> 30.0	[Gy]	30.5	31.8	31.8	32.8	33.2	32.4	39.4	37.7	39.5	31.9	38.9	40.0	37.9	3.0
	Bladder	$D0.5cm^3$	< 32.0	[Gy]	8.9	10.7	12.8	8.0	7.6	9.8	9.0	9.0	9.0	9.1	9.3	9.0	9.1	0.1
	Rectum	$D0.5cm^3$	< 32.0	[Gy]	16.6	11.3	13.8	16.2	16.3	14.4	14.2	14.8	14.4	14.1	18.7	14.8	15.2	1.8
	Sigmoid	$D0.5cm^3$	< 32.0	[Gy]	0.7	0.7	1.0	5.6	2.0	2.3	1.6	1.4	1.8	1.2	1.5	1.5	1.5	0.2

Table S4: Overview of the software used by the institutes (A-F) for deformable image registration (DIR) and dose accumulation. DVF: deformation vector field, DDM: direct dose mapping, EMT: Energy mass transfer.

Institute	A	B	C	D	E	F
Software (Version)	Monaco Research ADMIRE (3.44.2) for DIR and Slicer 3D (4.11) for dose mapping/ accumulation	MIM (6.8.5)	Raystation (8B) ANACONDA	MIM (7.0.6)	In-house developed dual-force demons deformable registration	RRTracker (v4.0)
DIR algorithm	Hybrid intensity / structure based	Mixed / hybrid DIR	Hybrid intensity / structure based	Contour based Deformable	Intensity based	Normalized gradient fields / structure based
DIR Contour guidance	Yes	Yes	Yes	Yes	No	Yes
DIR transformation	DVF	DVF	DVF	DVF	DVF	DVF
Contour propagation transformation	DVF	DVF	DVF	DVF	DVF	DVF
Dose transformation	DVF	DVF	DVF	DVF	DVF	DVF
Dose method	DDM	DDM	DDM	DDM	DDM	EMT

Deformable image registration / Contour propagation / Dose accumulation software:
Institute A used Monaco ADMIRE Research (Elekta, Stockholm, Sweden) for deformable image registration and contour propagation. The software used an Efficient Inverse-Consistent Diffeomorphic Image Registration Method [83]. Case-specifically selected contours at the first fraction were used for guidance. The offline re-contoured contours of Fx02-Fx05 were propagated to Fx01 during DIR by using the corresponding DVF. For the dose mapping and accumulation 3D slicer [63] an open software tool was used. The doses of each fraction were mapped according to the DVF to the first fraction of the treatment. The accumulated dose was calculated by weighting the deformed doses by $20\,\%$.

Institute B: Hybrid deformable image registration involves registrations using contours as well as image intensity. For contour-based deformable registration, MIM iteratively minimizes the signed distance differences between the two images, as measured from the surface of contour pairs. For intensity-based registration, intensity differences be-

tween the two images are minimized using a constrained, free-form deformable registration based on intensity. For the hybrid mode, the influence of the contour surface matching decreases with increasing distance from the contour surface. As part of DIR workflow, the MRI of Fx01 was selected as reference to which all other images were deformably registered, using an appropriate set of contours as well as image intensity. The DVF generated for the DIR was used to deform the associated dose for each case. A dose accumulation was then performed by equally weighting each of the fractional doses included in the registration to achieve the prescribed dose. No changes in the DIR protocols were made during the study. However, two different versions of MIM were used; with the last dataset being processed with version 7.2.7.

Institute C: We employed a hybrid intensity and structure-based DIR approach called ANAtomically CONstrained Deformation Algorithm (ANACONDA) integrated in the RayStation 8B treatment planning system. This algorithm computes a deformation vector field by optimizing both image similarity and alignment of controlling ROI structures segmented a priori on the registered images. The optimization includes regularization terms to control smoothness and invertibility as well as penalize generation of large ROI deviations. The optimization employs a nonlinear limited memory solver applied sequentially over 10-, 5- and 2.5-mm resolutions, with result of each resolution used as the starting point for the next level. We employed default settings in RayStation to generate each DIR and used the deformation vector field to map the dose for each fraction back to the first fraction image. Accumulation was performed by weighting each fractional dose by 1/total number of fractions and then performing a summation of all fractional doses.

Institute D used the contour-only deformable registration for this study. The registration starts from an image intensity-based rigid registration and followed by a deformable registration that iteratively minimizes the signed distance between the surfaces of contour pairs in the two images. The default parameters in MIM were used for both the clinical and gold standard cases without any modification. However, individual contours used in each registration may vary. Details are specified in the corresponding figures.

Institute E: The daily MR images (Fx02, . . . , Fx05) were deformably registered to the first fraction MR image (Fx01) using a dual-force demons deformable registration [242]. The 'Demons' algorithm introduced demons force from thermodynamics into image domain for deformable registration [233]. The demons force is driven by image gradients and image intensity differences and is updated iteratively and successively over entire image domain followed by a Gaussian-type kernel to smooth the force vectors and to propagate the forces to areas with low image gradients [269]. The dual-force demons algorithm was first developed for CT-CT deformable registration and was later validated for MR-MR deformable registration [71]. For this application, the algorithm parameters were set to allow only small deformation. The resulting deformation vector fields were used to deform the daily dose to first fraction and the deformed dose was then scaled to a fraction dose for accumulation [129], [138], [262].

Institute F: The employed dose accumulation solution relied on a variational registra-

tion algorithm, estimating displacements as the minimizer of a cost function composed of three terms: an image similarity term, a contour matching term and a regularization term. In summary, the image similarity term uses normalized gradient fields [80] in order to match similar contrast patterns between the registered images, whereas the contour matching term penalizes misalignments between any particular contour pairs defined by the user on the images. Finally, the regularization term constrains the estimated deformations to be spatially smooth, ensuring the well-posedness of the minimization problem. Further details regarding the method can be found here [28]. The deformations estimated by the DIR algorithm were subsequently used to warp each daily dose distribution onto the reference grid, via energy-per-mass transfer [136], with the warped doses being added-up in order to generate the accumulated dose.

Data analysis:
The DSC is defined by

$$DSC\left(X,Y\right)=\frac{\left(2|X\cap Y|\right)}{|X|+|Y|}, \tag{7}$$

where X,Y are the two volume sets to be compared. AAPM TG132 [34] indicates a DSC tolerance of 0.8 - 0.9, except for very large or very small volumes. In addition, the SDSC, was calculated with a threshold value of $2\,\mathrm{mm}$. The SDCS quantifies the similarity between two surfaces by measuring the overlap and mismatch between them. Third, HD was determined, which is defined as the greatest distance of a contour set X to the closest point of the other set Y [95]

$$HD\left(X,Y\right)=\max\left(h\left(X,Y\right),h\left(Y,X\right)\right), \tag{8}$$

where h is defined by

$$h\left(X,Y\right)=\max\min\|x-y\|, \tag{9}$$

$x\in X$, $y\in Y$.
Due to the outlier sensitivity of HD, the HD95% was calculated over the histogram of all distances given points on X and Y by excluding the $5\,\%$ highest outliers [198].

9.2.2 Supplementary B: Results

Table S5: Statistical values of dose accumulation errors selected for the gold standard (GS) case for the participating institutes. Values are in [Gy] and are presented individually for CTV (prostate), bladder, and rectum. Abbreviations; CTV: clinical target volume, IQR: interquartile range.

Institute	Volume of interest	Mean	Standard deviation	5th percentile	95th percentile	Median	IQR
A	CTV (Prostate)	0.09	0.24	-0.24	0.50	0.05	0.16
B		0.30	0.46	-0.18	1.31	0.10	0.45
C		0.30	0.45	-0.18	1.22	0.10	0.49
D		0.28	0.42	-0.18	1.20	0.10	0.44
E		0.39	0.85	-0.52	2.22	0.16	0.65
F		0.27	0.40	-0.18	1.08	0.11	0.44
A	Bladder	0.40	0.76	-0.58	2.02	0.16	0.71
B		0.41	0.67	-0.50	1.73	0.22	0.73
C		0.47	0.72	-0.49	1.89	0.24	0.81
D		0.12	0.74	-0.85	1.51	0.00	0.46
E		0.68	1.84	-1.40	4.38	0.18	1.05
F		0.40	0.67	-0.63	1.68	0.24	0.74
A	Rectum	0.21	0.43	-0.33	1.02	0.02	0.47
B		0.37	0.51	-0.20	1.44	0.12	0.62
C		0.32	0.53	-0.26	1.45	0.05	0.59
D		0.23	0.44	-0.26	1.20	0.00	0.40
E		2.92	2.29	0.03	7.31	2.84	3.86
F		0.25	0.50	-0.40	1.22	0.11	0.49

Table S6: Gold standard (GS) contour propagation evaluation metrics, DSC (Dice similarity coefficient), SDSC (surface DSC) with 2 mm threshold, and HD95% (Hausdorff distance 95%) for the contours and deformed fraction. G: contour guidance, if checked the contour was used as guidance by the institute. DSC printed in red if < 0.8. Abbreviations; $Femur_L$: left femur, $Femur_R$: right femur, CTV: clinical target volume, GTV: gross tumor volume.

			A				C				D				E				F			
			DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G
Gold Standard	Bladder	Fx02toFx01	0.99	0.42	5.27	x	0.99	0.47	5.07	x	0.97	0.43	4.49	x	0.98	0.38	5.36		0.98	0.65	3.74	x
		Fx03toFx01	0.99	0.42	5.27		0.99	0.49	5.16		0.98	0.42	4.25		0.98	0.40	2.64		0.99	0.69	4.01	
		Fx04toFx01	0.99	0.42	5.24		0.99	0.49	5.03		0.97	0.41	4.53		0.96	0.33	5.41		0.99	0.61	4.18	
		Fx05toFX01	0.99	0.42	5.26		0.99	0.49	5.24		0.97	0.41	4.45		0.98	0.38	2.58		0.99	0.64	4.00	
	Rectum	Fx02toFx01	0.98	0.56	3.40	x	0.98	0.63	3.38	x	0.97	0.58	3.13	x	0.97	0.54	3.49		0.95	0.66	2.75	x
		Fx03toFx01	0.99	0.57	3.39		0.99	0.66	3.37		0.95	0.55	3.27		0.98	0.56	3.43		0.94	0.70	3.05	
		Fx04toFx01	0.98	0.56	3.41		0.98	0.64	3.37		0.95	0.55	3.36		0.94	0.51	3.69		0.94	0.63	2.75	
		Fx05toFX01	0.99	0.56	3.39		0.99	0.66	3.31		0.96	0.57	3.16		0.98	0.56	3.50		0.93	0.62	3.20	
	Sphincter	Fx02toFx01	0.97	0.55	3.62		0.98	0.66	3.40		0.97	0.59	3.22	x	0.96	0.54	3.59		0.91	0.63	3.23	
		Fx03toFx01	0.99	0.57	3.64		0.98	0.67	3.56		0.92	0.53	3.95		0.98	0.57	1.26		0.91	0.64	3.33	
		Fx04toFx01	0.97	0.56	3.66		0.98	0.66	3.36		0.95	0.56	3.23		0.92	0.50	3.80		0.90	0.66	4.10	
		Fx05toFX01	0.99	0.58	3.63		0.98	0.68	3.32		0.97	0.58	3.30		0.97	0.57	3.56		0.89	0.58	3.29	
	$Femur_L$	Fx02toFx01	1.00	0.50	4.24	x	1.00	0.62	4.12		0.99	0.52	3.98		0.98	0.48	4.31		0.97	0.70	2.73	
		Fx03toFx01	1.00	0.49	4.24		1.00	0.62	4.15		0.89	0.20	5.31		0.99	0.49	4.32		0.98	0.83	3.20	
		Fx04toFx01	1.00	0.50	4.23		1.00	0.62	4.13		0.87	0.18	7.24		0.95	0.41	4.75		0.95	0.58	3.29	
		Fx05toFX01	1.00	0.50	4.24		1.00	0.62	4.13		0.88	0.19	5.09		0.99	0.48	4.31		0.97	0.70	2.75	
	$Femur_R$	Fx02toFx01	1.00	0.54	4.03	x	1.00	0.64	3.90		0.99	0.54	4.00		0.98	0.51	4.10		0.95	0.57	3.47	
		Fx03toFx01	1.00	0.53	4.04		1.00	0.64	3.86		0.89	0.27	4.39		0.99	0.53	1.03		0.97	0.72	3.09	
		Fx04toFx01	1.00	0.54	4.02		1.00	0.65	3.86		0.88	0.20	5.87		0.95	0.47	4.25		0.98	0.77	2.72	
		Fx05toFX01	1.00	0.54	4.03		1.00	0.65	3.88		0.91	0.30	4.57		0.98	0.52	4.11		0.95	0.57	3.28	
	CTV	Fx02toFx01	0.97	0.45	4.35		0.99	0.57	4.07	x	0.95	0.44	3.72	x	0.97	0.44	4.35		0.97	0.65	3.72	x
		Fx03toFx01	0.99	0.47	4.31		0.99	0.57	4.09		0.97	0.48	3.76		0.98	0.49	4.35		0.98	0.72	3.28	
		Fx04toFx01	0.97	0.46	4.30		0.99	0.57	3.97		0.94	0.38	4.15		0.96	0.41	4.36		0.96	0.64	3.62	
		Fx05toFX01	0.97	0.45	4.34		0.99	0.57	4.08		0.96	0.44	3.97		0.97	0.44	4.61		0.96	0.61	3.46	
	GTV	Fx02toFx01	0.92	0.69	2.87		0.93	0.74	2.74		0.90	0.68	2.88		0.91	0.67	2.84		0.89	0.77	2.98	x
		Fx03toFx01	0.93	0.69	2.97		0.95	0.77	2.72		0.91	0.70	2.78		0.95	0.70	2.82		0.92	0.84	2.72	
		Fx04toFx01	0.92	0.68	2.90		0.94	0.77	2.94		0.90	0.68	2.77		0.88	0.62	3.24		0.90	0.83	2.76	
		Fx05toFX01	0.92	0.68	2.88		0.94	0.73	2.67		0.91	0.68	2.80		0.91	0.66	2.93		0.83	0.64	4.02	

Table S7: (I) Prostate 1 contour propagation evaluation metrics, DSC (Dice similarity coefficient), SDSC (surface DSC) with 2 mm threshold, and HD95% (Hausdorff distance 95%) for the contours and deformed fraction. G: contour guidance; if checked the contour was used as guidance by the institute. DSC printed in red if < 0.8. Abbreviations; $Femur_L$: left femur, $Femur_R$: right femur, CTV: clinical target volume, GTV: gross tumor volume.

			A				C				D				E				F			
			DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G
(I) Prostate 1	Bladder	Fx02toFx01	0.99	0.74	3.10		0.98	0.73	3.38		0.91	0.54	4.31		0.70	0.23	17.50		0.98	0.58	5.81	
		Fx03toFx01	0.99	0.75	3.05	x	0.99	0.79	3.29	x	0.97	0.64	3.53	x	0.96	0.55	4.12		0.99	0.65	4.31	x
		Fx04toFx01	0.99	0.74	3.06		0.98	0.73	3.26		0.94	0.51	4.19		0.85	0.25	8.02		0.98	0.67	4.13	
		Fx05toFx01	0.99	0.75	3.11		0.99	0.80	3.22		0.97	0.64	3.73		0.97	0.58	2.59		0.99	0.63	4.56	
	Rectum	Fx02toFx01	0.97	0.84	1.45		0.95	0.83	2.81		0.86	0.54	5.50		0.83	0.46	6.85		0.81	0.35	8.99	
		Fx03toFx01	0.98	0.87	2.33	x	0.96	0.86	3.08	x	0.93	0.73	2.91	x	0.91	0.70	4.27		0.92	0.60	3.66	x
		Fx04toFx01	0.96	0.84	2.33		0.88	0.69	4.83		0.88	0.61	6.58		0.86	0.58	14.17		0.86	0.53	15.15	
		Fx05toFx01	0.96	0.84	2.71		0.93	0.76	3.09		0.89	0.63	4.06		0.86	0.56	17.34		0.88	0.51	14.94	
	Sphincter	Fx02toFx01	0.91	0.66	2.28		0.90	0.76	2.42		0.92	0.68	2.94		0.84	0.54	3.13		0.82	0.49	4.80	
		Fx03toFx01	0.96	0.75	1.73		0.97	0.81	2.89		0.95	0.77	2.72	x	0.87	0.60	2.79		0.89	0.54	3.55	
		Fx04toFx01	0.90	0.61	2.87		0.89	0.61	2.81		0.84	0.48	3.21		0.86	0.50	3.59		0.86	0.40	5.10	
		Fx05toFx01	0.88	0.67	2.71		0.92	0.76	2.20		0.92	0.71	3.31		0.81	0.39	4.07		0.83	0.46	4.39	
	$Femur_L$	Fx02toFx01	1.00	0.90	2.18		0.98	0.93	2.24		0.77	0.26	10.68		0.97	0.89	1.17		0.95	0.61	4.14	
		Fx03toFx01	0.99	0.90	2.17	x	0.98	0.93	2.25		0.91	0.63	4.67		0.97	0.88	2.30		0.95	0.57	4.26	
		Fx04toFx01	1.00	0.90	2.18		0.99	0.93	2.18		0.83	0.36	10.37		0.97	0.89	1.32		0.95	0.62	4.15	
		Fx05toFx01	0.99	0.90	2.18		0.96	0.89	2.35		0.89	0.51	4.67		0.95	0.81	2.62		0.94	0.50	4.13	
	$Femur_R$	Fx02toFx01	0.99	0.92	2.10		0.99	0.95	2.10		0.45	0.12	24.20		0.98	0.91	2.16		0.97	0.64	4.25	
		Fx03toFx01	0.99	0.92	2.10	x	0.98	0.94	2.13		0.91	0.56	3.71		0.97	0.90	2.24		0.96	0.62	4.12	
		Fx04toFx01	0.99	0.92	2.10		0.98	0.94	1.91		0.84	0.40	6.85		0.98	0.90	2.23		0.96	0.63	4.11	
		Fx05toFx01	0.99	0.92	2.11		0.96	0.93	2.22		0.80	0.34	9.72		0.95	0.88	2.06		0.96	0.62	4.08	
	CTV	Fx02toFx01	0.93	0.54	4.11		0.98	0.80	2.89		0.94	0.55	3.74		0.93	0.48	4.25		0.96	0.58	3.82	
		Fx03toFx01	0.93	0.53	3.65		0.98	0.83	2.70	x	0.95	0.64	3.42	x	0.92	0.46	3.59		0.96	0.60	3.89	x
		Fx04toFx01	0.95	0.65	3.34		0.98	0.78	2.93		0.92	0.51	4.36		0.94	0.57	3.75		0.96	0.57	3.70	
		Fx05toFx01	0.93	0.49	4.07		0.98	0.81	2.78		0.95	0.65	3.23		0.92	0.45	4.30		0.96	0.58	4.08	
	GTV	Fx02toFx01	0.65	0.45	4.21		0.82	0.77	3.48		0.70	0.53	5.22		0.67	0.47	4.26		0.66	0.37	5.29	
		Fx03toFx01	0.76	0.64	3.14		0.81	0.76	2.75		0.84	0.73	3.02		0.75	0.60	3.58		0.74	0.52	4.12	x
		Fx04toFx01	0.92	0.94	2.10		0.91	0.93	2.24		0.92	0.94	2.04		0.93	0.95	2.06		0.91	0.75	2.75	
		Fx05toFx01	0.75	0.57	3.64		0.82	0.72	2.77		0.78	0.57	4.28		0.76	0.55	3.72		0.74	0.45	4.26	

Table S8: (II) Prostate 2 contour propagation evaluation metrics, DSC (Dice similarity coefficient), SDSC (surface DSC) with 2 mm threshold, and HD95% (Hausdorff distance 95%) for the contours and deformed fraction. G: contour guidance; if checked the contour was used as guidance by the institute. DSC printed in red if < 0.8. Abbreviations; $Femur_L$: left femur, $Femur_R$: right femur, CTV: clinical target volume, GTV: gross tumor volume.

			A				C				D				E				F			
			DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G
(II) Prostate 2	Bladder	Fx02toFx01	0.99	0.82	2.72		0.98	0.81	3.43		0.95	0.59	3.94		0.95	0.56	5.14		0.97	0.62	4.42	
		Fx03toFx01	0.99	0.83	2.71	x	0.98	0.83	2.81	x	0.94	0.53	5.30	x	0.92	0.52	10.35		0.98	0.66	4.12	x
		Fx04toFx01	0.99	0.83	2.71		0.99	0.86	2.80		0.96	0.62	3.19		0.93	0.55	7.17		0.98	0.64	4.23	
		Fx05toFx01	0.99	0.83	2.74		0.98	0.79	3.05		0.91	0.42	8.94		0.85	0.40	12.49		0.97	0.59	4.17	
	Rectum	Fx02toFx01	0.98	0.86	2.38		0.84	0.73	15.17		0.87	0.51	4.62		0.81	0.52	12.79		0.87	0.43	5.54	
		Fx03toFx01	0.97	0.85	1.81	x	0.97	0.87	2.47	x	0.82	0.47	7.05	x	0.83	0.55	11.25		0.81	0.40	10.50	x
		Fx04toFx01	0.98	0.87	0.97		0.97	0.89	2.45		0.78	0.44	7.32		0.93	0.71	2.62		0.89	0.51	3.37	
		Fx05toFx01	0.98	0.86	1.08		0.96	0.86	2.56		0.78	0.39	9.65		0.74	0.38	15.12		0.84	0.40	9.81	
	Sphincter	Fx02toFx01	0.81	0.51	3.36		0.78	0.55	3.71		0.82	0.59	2.87		0.79	0.40	3.90		0.85	0.51	4.27	
		Fx03toFx01	0.79	0.59	3.35		0.78	0.68	3.51		0.83	0.62	3.67	x	0.79	0.46	3.70		0.84	0.55	3.83	
		Fx04toFx01	0.73	0.49	5.10		0.73	0.49	5.43		0.77	0.52	5.12		0.70	0.34	5.44		0.64	0.35	5.77	
		Fx05toFx01	0.75	0.51	5.60		0.76	0.55	5.87		0.72	0.51	4.05		0.68	0.27	6.55		0.72	0.37	5.19	
	$Femur_L$	Fx02toFx01	0.99	0.90	2.18		0.98	0.93	2.20		0.14	0.07	31.85		0.96	0.88	1.66		0.91	0.39	5.79	
		Fx03toFx01	0.99	0.90	2.18	x	0.98	0.93	2.20		0.12	0.06	34.02		0.97	0.89	1.55		0.97	0.60	4.60	
		Fx04toFx01	0.99	0.90	2.18		0.97	0.93	2.23		0.21	0.08	27.35		0.96	0.87	2.34		0.93	0.47	4.89	
		Fx05toFx01	0.99	0.90	2.19		0.93	0.66	2.78		0.04	0.03	40.26		0.92	0.61	2.92		0.90	0.36	5.49	
	$Femur_R$	Fx02toFx01	0.99	0.90	2.14		0.90	0.50	4.11		0.59	0.16	18.36		0.89	0.49	4.20		0.83	0.20	5.65	
		Fx03toFx01	0.99	0.91	2.14	x	0.96	0.90	2.32		0.37	0.09	25.70		0.96	0.86	2.40		0.96	0.55	4.68	
		Fx04toFx01	0.99	0.91	2.14		0.97	0.94	2.20		0.65	0.14	12.68		0.97	0.88	2.27		0.96	0.57	4.91	
		Fx05toFx01	0.99	0.91	2.13		0.95	0.89	2.32		0.34	0.08	24.02		0.95	0.85	2.43		0.94	0.46	4.74	
	CTV	Fx02toFx01	0.93	0.56	2.76		0.98	0.85	2.61		0.89	0.31	4.58		0.92	0.52	3.25		0.95	0.55	4.33	
		Fx03toFx01	0.92	0.51	3.51		0.98	0.85	2.74	x	0.93	0.56	3.63	x	0.90	0.49	4.04		0.95	0.55	3.46	x
		Fx04toFx01	0.92	0.51	3.18		0.98	0.85	2.65		0.92	0.55	3.81		0.90	0.45	3.82		0.94	0.53	3.96	
		Fx05toFx01	0.91	0.51	4.10		0.98	0.83	2.70		0.92	0.55	3.88		0.89	0.44	3.93		0.94	0.54	4.27	
	GTV	Fx02toFx01	0.43	0.49	3.76		0.49	0.42	4.79		0.60	0.55	4.03		0.23	0.61	4.82		0.19	0.31	5.21	
		Fx03toFx01	0.43	0.46	4.06		0.40	0.30	6.07		0.16	0.25	6.85		0.26	0.52	5.25		0.23	0.28	4.88	x
		Fx04toFx01	0.19	0.53	6.80		0.16	0.49	8.06		0.14	0.43	8.17		0.12	0.38	7.57		0.06	0.45	10.12	
		Fx05toFx01	0.73	0.76	3.38		0.35	0.35	7.00		0.01	0.14	10.21		0.23	0.50	4.41		0.27	0.30	5.27	

Table S9: (III) Cerfix contour propagation evaluation metrics, DSC (Dice similarity coefficient), SDSC (surface DSC) with 2 mm threshold, and HD95% (Hausdorff distance 95%) for the contours and deformed fraction. G: contour guidance; if checked the contour was used as guidance by the institute. DSC printed in red if < 0.8. Abbreviations; GTV: gross tumor volume, CTV_{HR}: High risk clinical target volume, CTV_{IR}: Intermediate risk clinical target volume.

			A				C				D				E				F			
			DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G
(III) Cervix	Bladder	Fx02toFx01	0.97	0.80	1.32		0.97	0.76	2.01		0.86	0.23	5.87		0.93	0.71	2.93		0.97	0.62	3.14	
		Fx03toFx01	0.97	0.70	1.98	x	0.97	0.70	2.10	x	0.82	0.62	5.63	x	0.93	0.67	2.59		0.97	0.51	3.02	x
		Fx04toFx01	0.97	0.75	1.37		0.97	0.79	2.04		0.87	0.84	4.14		0.94	0.73	2.77		0.97	0.56	3.08	
		Fx05toFx01	0.97	0.70	1.84		0.97	0.68	1.94		0.77	0.25	6.17		0.92	0.67	2.70		0.96	0.27	2.09	
	Rectum	Fx02toFx01	0.97	0.64	1.94		0.97	0.61	1.84		0.76	0.23	7.35		0.90	0.58	2.59		0.93	0.69	3.12	
		Fx03toFx01	0.97	0.57	1.98	x	0.97	0.56	1.84	x	0.79	0.59	4.34	x	0.90	0.57	2.39		0.94	0.62	3.09	x
		Fx04toFx01	0.97	0.67	1.94		0.97	0.67	1.80		0.79	0.78	4.78		0.92	0.65	2.22		0.92	0.62	3.18	
		Fx05toFx01	0.97	0.07	2.00		0.97	0.09	1.88		0.69	0.11	7.78		0.92	0.08	2.21		0.93	0.32	3.09	
	Sigmoid	Fx02toFx01	0.95	0.40	1.87		0.95	0.38	1.83		0.59	0.20	11.49		0.81	0.30	6.72		0.80	0.34	6.27	
		Fx03toFx01	0.96	0.43	1.09	x	0.96	0.46	1.52	x	0.60	0.47	9.30	x	0.82	0.39	4.09		0.82	0.39	4.28	
		Fx04toFx01	0.97	0.54	1.50		0.96	0.54	1.45		0.74	0.58	4.57		0.86	0.50	2.47		0.85	0.44	3.08	
		Fx05toFx01	0.96	0.16	1.62		0.96	0.19	1.47		0.58	0.12	6.91		0.84	0.16	4.13		0.82	0.15	3.86	
	GTV	Fx02toFx01	0.90	0.97	2.68		0.88	0.97	2.92		0.66	0.49	6.90		0.88	0.84	2.86		0.92	0.65	3.51	
		Fx03toFx01	0.86	0.97	4.60		0.85	0.97	5.07		0.82	0.41	3.28	x	0.87	0.85	4.39		0.88	0.66	4.11	
		Fx04toFx01	0.89	0.97	2.58		0.89	0.97	2.46		0.92	0.55	2.34		0.87	0.85	2.76		0.89	0.64	3.40	
		Fx05toFx01	0.85	0.97	3.20		0.84	0.98	3.71		0.66	0.31	6.50		0.84	0.83	3.80		0.70	0.64	6.42	
	CTV_{HR}	Fx02toFx01	0.89	0.97	3.38		0.88	0.98	3.54		0.76	0.51	6.03		0.89	0.85	3.76		0.95	0.68	3.18	
		Fx03toFx01	0.85	0.97	5.24		0.85	0.98	4.99		0.85	0.55	3.78	x	0.86	0.85	4.62		0.94	0.69	2.66	x
		Fx04toFx01	0.88	0.97	2.40		0.89	0.98	2.48		0.92	0.50	3.20		0.88	0.91	3.10		0.93	0.62	3.40	
		Fx05toFx01	0.55	0.96	11.28		0.56	0.98	11.32		0.53	0.40	12.66		0.56	0.89	11.20		0.81	0.69	7.56	
	CTV_{IR}	Fx02toFx01	0.92	0.96	5.05		0.90	0.98	4.46		0.86	0.37	6.52		0.90	0.71	5.91		0.92	0.45	4.65	
		Fx03toFx01	0.91	0.98	5.77		0.91	0.98	6.10		0.91	0.37	4.22	x	0.90	0.76	5.77		0.91	0.49	5.16	
		Fx04toFx01	0.94	0.98	2.55		0.94	0.98	2.72		0.95	0.48	3.72		0.94	0.83	2.66		0.93	0.50	4.24	
		Fx05toFx01	0.72	0.98	11.10		0.72	0.98	11.15		0.70	0.31	12.55		0.73	0.80	10.58		0.72	0.47	11.10	

Table S10: (IV) Liver contour propagation evaluation metrics, DSC (Dice similarity coefficient), SDSC (surface DSC) with 2 mm threshold, and HD95% (Hausdorff distance 95%) for the contours and deformed fraction. G: contour guidance; if checked the contour was used as guidance by the institute. DSC printed in red if < 0.8. Abbreviations; $Kidney_L$: left kidney, $Kidney_R$: right kidney, GTV: gross tumor volume.

			A				C				D				E				F			
			DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G
(IV) Liver	Stomach	Fx02toFx01	0.95	0.86	3.17		0.78	0.65	4.57		0.51	0.18	13.58		0.70	0.43	8.01		0.75	0.25	7.30	
		Fx03toFx01	0.95	0.83	2.50	x	0.77	0.61	5.09	x	0.59	0.22	10.43		0.66	0.39	11.78		0.75	0.24	7.29	x
		Fx04toFx01	0.95	0.84	3.27		0.78	0.66	4.33		0.61	0.27	10.51		0.70	0.47	9.92		0.76	0.26	7.41	
		Fx05toFx01	0.94	0.81	3.18		0.77	0.61	5.38		0.62	0.22	9.53		0.68	0.45	12.84		0.75	0.25	9.42	
	Spleen	Fx02toFx01	0.96	0.89	2.74		0.82	0.66	3.14		0.70	0.26	7.42		0.80	0.57	5.16		0.82	0.23	5.80	
		Fx03toFx01	0.93	0.76	3.81		0.79	0.56	4.83		0.68	0.24	8.79		0.78	0.59	5.27		0.78	0.17	6.96	
		Fx04toFx01	0.96	0.88	2.68		0.81	0.64	3.75		0.75	0.28	5.23		0.81	0.68	4.90		0.82	0.22	6.25	
		Fx05toFx01	0.95	0.87	2.74		0.81	0.61	4.39		0.70	0.35	8.56		0.76	0.61	16.31		0.82	0.24	6.04	
	Spinal cord	Fx02toFx01	0.87	0.83	1.69		0.80	0.80	9.53		0.64	0.51	7.28		0.83	0.80	3.27		0.79	0.47	4.48	
		Fx03toFx01	0.87	0.27	4.89	x	0.53	0.27	7.08		0.51	0.26	8.24	x	0.51	0.27	7.71		0.50	0.13	8.39	
		Fx04toFx01	0.88	0.81	0.83		0.83	0.81	2.62		0.68	0.49	4.67		0.84	0.81	2.77		0.81	0.51	3.70	
		Fx05toFx01	0.88	0.80	0.80		0.81	0.78	2.88		0.60	0.46	6.50		0.81	0.79	2.95		0.79	0.49	3.56	
	Small bowel	Fx02toFx01	0.82	0.72	19.92		0.65	0.30	17.53		0.60	0.20	16.83		0.68	0.33	18.61		0.58	0.09	19.45	
		Fx03toFx01	0.85	0.72	7.34	x	0.65	0.25	17.27		0.61	0.16	15.66	x	0.61	0.28	20.38		0.62	0.09	22.67	
		Fx04toFx01	0.85	0.69	5.84		0.55	0.22	19.06		0.52	0.12	18.41		0.60	0.29	17.69		0.54	0.07	22.00	
		Fx05toFx01	0.84	0.67	6.22		0.61	0.29	14.73		0.51	0.18	15.10		0.61	0.34	19.50		0.40	0.09	21.57	
	Pancreas	Fx02toFx01	0.80	0.64	6.44		0.93	0.87	3.20		0.57	0.26	8.69		0.87	0.73	4.49		0.81	0.24	8.71	
		Fx03toFx01	0.80	0.62	5.65		0.93	0.87	3.26	x	0.63	0.35	8.01	x	0.84	0.68	6.06		0.80	0.22	7.98	
		Fx04toFx01	0.82	0.67	4.86		0.95	0.91	2.65		0.73	0.37	7.03		0.88	0.77	4.12		0.83	0.29	8.20	
		Fx05toFx01	0.76	0.54	6.21		0.94	0.87	2.65		0.58	0.32	11.06		0.82	0.63	5.57		0.77	0.24	10.52	
	Liver	Fx02toFx01	0.90	0.60	4.96		0.85	0.74	2.74		0.76	0.19	9.62		0.82	0.47	6.15		0.83	0.15	7.14	
		Fx03toFx01	0.85	0.63	18.47		0.85	0.75	3.66	x	0.73	0.20	10.68		0.80	0.42	8.71		0.79	0.14	14.18	
		Fx04toFx01	0.91	0.67	5.00		0.85	0.74	3.39		0.75	0.15	9.64		0.83	0.60	4.67		0.83	0.18	7.00	
		Fx05toFx01	0.91	0.73	3.90		0.85	0.77	3.06		0.75	0.18	10.26		0.83	0.65	4.69		0.84	0.19	3.77	
	$Kidney_R$	Fx02toFx01	0.82	0.69	3.01		0.84	0.69	10.61		0.69	0.26	13.26		0.89	0.76	10.26		0.87	0.22	11.15	
		Fx03toFx01	0.84	0.70	2.18	x	0.84	0.77	4.81		0.66	0.22	11.96	x	0.87	0.80	3.73		0.82	0.24	7.46	
		Fx04toFx01	0.84	0.77	2.21		0.87	0.79	3.87		0.71	0.29	9.33		0.87	0.81	4.46		0.85	0.29	7.03	
		Fx05toFx01	0.84	0.76	2.12		0.87	0.82	2.96		0.71	0.27	8.60		0.88	0.86	2.74		0.86	0.29	6.25	
	$Kidney_L$	Fx02toFx01	0.83	0.77	3.44		0.81	0.80	19.71		0.67	0.30	12.61		0.87	0.75	10.27		0.86	0.24	14.59	
		Fx03toFx01	0.85	0.66	2.30	x	0.88	0.78	3.75		0.80	0.36	6.37	x	0.88	0.76	5.30		0.87	0.22	6.24	
		Fx04toFx01	0.85	0.78	2.31		0.92	0.86	3.26		0.85	0.53	4.78		0.92	0.87	2.88		0.92	0.30	2.66	

Continued on next page

Table S10 – continued from previous page

			A				C				D				E				F			
			DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G
		Fx05toFx01	0.85	0.81	2.06		0.91	0.91	2.48		0.79	0.38	6.63		0.91	0.88	2.68		0.92	0.31	5.38	
	GTV	Fx02toFx01	0.71	0.45	5.29		0.79	0.85	2.32		0.26	0.26	12.00		0.49	0.31	5.75		0.76	0.49	4.45	
		Fx03toFx01	0.73	0.48	5.49		0.79	0.78	3.12	x	0.44	0.45	8.53	x	0.60	0.35	5.29		0.74	0.32	5.74	x
		Fx04toFx01	0.75	0.52	6.90		0.79	0.82	2.38		0.61	0.42	5.04		0.55	0.32	7.57		0.76	0.45	4.40	
		Fx05toFx01	0.72	0.49	7.72		0.78	0.81	2.75		0.32	0.20	11.50		0.52	0.29	7.18		0.73	0.33	4.59	
	Duodenum	Fx02toFx01	0.79	0.79	5.23		0.92	0.89	2.07		0.53	0.38	14.04		0.78	0.73	15.15		0.76	0.33	73.26	
		Fx03toFx01	0.81	0.78	2.54	x	0.87	0.87	2.90	x	0.52	0.30	10.70	x	0.78	0.68	4.64		0.73	0.32	34.02	x
		Fx04toFx01	0.81	0.78	2.61		0.91	0.88	2.89		0.53	0.28	9.73		0.84	0.80	2.70		0.76	0.33	33.68	
		Fx05toFx01	0.81	0.78	2.58		0.87	0.90	2.52		0.46	0.21	12.18		0.81	0.76	3.25		0.72	0.32	13.28	
	Colon	Fx02toFx01	0.86	0.72	3.91		0.58	0.24	18.47		0.55	0.19	16.36		0.60	0.28	16.10		0.44	0.07	63.04	
		Fx03toFx01	0.87	0.76	6.16	x	0.46	0.17	23.35		0.60	0.17	13.96	x	0.54	0.20	19.12		0.45	0.06	65.06	
		Fx04toFx01	0.86	0.73	6.53		0.46	0.16	24.36		0.55	0.16	15.30		0.51	0.19	20.52		0.37	0.04	64.42	
		Fx05toFx01	0.87	0.75	5.22		0.48	0.15	20.16		0.60	0.19	13.07		0.51	0.20	19.85		0.40	0.05	64.62	

Table S11: (V) Lymph node contour propagation evaluation metrics, DSC (Dice similarity coefficient), SDSC (surface DSC) with 2 mm threshold, and HD95% (Hausdorff distance 95%) for the contours and deformed fraction. G: contour guidance; if checked the contour was used as guidance by the institute. DSC printed in red if < 0.8. Abbreviations; CTV_L: left clinical target volume, CTV_R: right clinical target volume.

			A				C				D				E				F			
			DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G	DSC	SDSC	HD95% [mm]	G
(V) Lymph node	Bladder	Fx02toFx01	0.99	0.99	0.62		0.96	0.95	1.43		0.69	0.14	11.01		0.77	0.63	10.64		0.98	0.53	1.79	
		Fx03toFx01	0.99	0.99	1.04	x	0.98	0.98	1.19	x	0.85	0.52	6.08	x	0.96	0.91	1.90		0.98	0.55	3.09	x
		Fx04toFx01	0.99	1.00	1.02		0.99	0.99	1.17		0.84	0.43	7.11		0.97	0.92	1.82		0.98	0.53	3.13	
		Fx05toFx01	0.99	0.99	1.03		0.98	0.98	1.22		0.92	0.58	3.45		0.97	0.93	1.64		0.98	0.54	3.01	
	Rectum	Fx02toFx01	0.98	1.00	0.70		0.97	0.99	1.24		0.78	0.54	6.67		0.89	0.96	1.68		0.96	0.63	2.61	
		Fx03toFx01	0.98	1.00	0.74	x	0.96	0.99	1.01	x	0.76	0.50	6.68	x	0.87	0.82	2.55		0.95	0.61	2.62	x
		Fx04toFx01	0.98	0.99	0.82		0.97	0.99	1.26		0.70	0.44	8.19		0.86	0.95	6.08		0.95	0.60	2.74	
		Fx05toFx01	0.98	0.99	0.79		0.97	0.98	1.27		0.79	0.60	6.23		0.89	0.97	1.27		0.95	0.60	2.50	
	Sigmoid	Fx02toFx01	0.76	0.88	51.75		0.64	0.62	52.19		0.26	0.29	55.91		0.61	0.60	54.68		0.96	0.32	2.61	
		Fx03toFx01	0.63	0.80	52.98	x	0.40	0.62	61.45		0.39	0.55	58.96	x	0.37	0.58	66.65		0.95	0.40	2.62	
		Fx04toFx01	0.76	0.88	50.16		0.51	0.52	51.46		0.31	0.22	11.27		0.55	0.56	25.83		0.95	0.27	2.74	
		Fx05toFx01	0.80	0.89	22.98		0.53	0.59	56.27		0.52	0.50	12.43		0.59	0.63	52.89		0.95	0.32	2.50	
	CTV_L	Fx02toFx01	0.72	1.00	1.37		0.89	1.00	0.79		0.46	0.84	2.90		0.58	0.95	2.22		0.68	0.91	1.75	
		Fx03toFx01	0.40	0.92	3.44		0.87	1.00	0.77	x	0.00	0.44	7.09		0.27	0.89	3.67		0.33	0.86	4.21	x
		Fx04toFx01	0.67	1.00	2.04		0.91	1.00	0.90		0.00	0.41	7.27		0.53	0.94	2.49		0.53	0.88	2.48	
		Fx05toFx01	0.70	0.98	1.34		0.55	0.99	0.94		0.00	0.00	n.a.		0.69	0.96	1.27		0.51	0.86	2.55	
	CTV_R	Fx02toFx01	0.82	1.00	1.01		0.91	1.00	0.87		0.00	0.22	6.98		0.52	0.75	3.18		0.75	0.90	1.90	
		Fx03toFx01	0.83	0.97	1.20		0.87	1.00	0.88	x	0.58	0.84	2.27		0.22	0.49	4.26		0.74	0.92	2.46	x
		Fx04toFx01	0.82	1.00	1.15		0.49	1.00	1.02		0.02	0.39	7.04		0.31	0.50	3.91		0.72	0.93	1.83	
		Fx05toFx01	0.82	0.98	1.17		0.55	1.00	1.04		0.00	0.00	n.a.		0.62	0.84	2.64		0.69	0.94	2.40	

Table S12: Summary of the ANOVA for repeated measures with post-hoc test for DSC analysis. Abbreviations CTV: clinical target volume, GTV: gross tumor volume, CTV_{HR}: High risk clinical target volume, CTV_{IR}: Intermediate risk clinical target volume, CTV_L: left clinical target volume, CTV_R: right clinical target volume, A-F are the Institutes, N/A: not applicable, n.s.: not statistically significant.

Cases & contours		ANOVA for repeated measures with post-hoc test		
		Global hypothesis	Pairwise comparison	
		p-Value	Combination of pairs	p-Value
Gold Standard	Bladder	<0.05	A C A F C F D E	n.s.
			all other combination pairs	<0.05
	Rectum		A D A F C D C F E F	<0.05
			all other combination pairs	n.s.
	Sphincter		A F C F D F E F	<0.05
			all other combination pairs	n.s.
	$Femur_L$		A D C D D E D F	<0.05
			all other combination pairs	n.s.
	$Femur_R$		A D C D D E	<0.05
			all other combination pairs	n.s.
	CTV		C D C E C F	<0.05
			all other combination pairs	n.s.
	GTV		C F	<0.05
			all other combination pairs	n.s.
(I) Prostate 1	Bladder	<0.05	A E	<0.05
			all other combination pairs	n.s.
	Rectum		A D A E A F	<0.05
			all other combination pairs	n.s.
	Sphincter		all combination pairs	n.s
	$Femur_L$		A D C D D E D F	<0.05
			all other combination pairs	n.s.

Continued on next page

Table S12 – continued from previous page

Cases & contours		ANOVA for repeated measures with post-hoc test		
		Global hypothesis	Pairwise comparison	
		p-Value	Combination of pairs	p-Value
	$Femur_R$		A D C D D E D F	<0.05
			all other combination pairs	n.s.
	CTV		A D A E C F D E	n.s.
			all other combination pairs	<0.05
	GTV	0.78	N/A	
(II) Prostate 2	Bladder	<0.05	A D A E C E E F	<0.05
			all other combination pairs	n.s.
	Rectum		A D A E A F C D	<0.05
			all other combination pairs	n.s.
	Sphincter	0.90	N/A	
	$Femur_L$	<0.05	A D C D D E D F	<0.05
			all other combination pairs	n.s.
	$Femur_R$		A D C D D E D F	<0.05
			all other combination pairs	n.s.
	CTV		A D A E D E	n.s.
			all other combination pairs	<0.05
	GTV	0.21	N/A	
(III) Cervix	Bladder	<0.05	A D C D D E D F	<0.05
			all other combination pairs	n.s.
	Rectum		A C A F C F E F	n.s.
			all other combination pairs	<0.05
	Sigmoid		A C E F	n.s
			all other combination pairs	<0.05
	CTV_{HR}	0.70	N/A	

Continued on next page

Table S12 – continued from previous page

Cases & contours		ANOVA for repeated measures with post-hoc test		
		Global hypothesis	Pairwise comparison	
		p-Value	Combination of pairs	p-Value
	CTV_{IR}	0.99	N/A	
	GTV	0.26	N/A	
(IV) Liver	Colon	<0.05	C D C E C F D E	n.s.
			all other combination pairs	<0.05
	Duodenum		A E	n.s.
			all other combination pairs	<0.05
	$Kidney_L$		A E D E D F	n.s.
			all other combination pairs	<0.05
	$Kidney_R$		C E C F E F	n.s.
			all other combination pairs	<0.05
	Pancreas		A E A F E F	n.s.
			all other combination pairs	<0.05
	SmallBowel		A C A D A E A F	<0.05
			all other combination pairs	n.s.
	SpinalCord	0.59	N/A	
	Spleen	<0.05	C E C F E F	n.s.
			all other combination pairs	<0.05
	Stomach		C F	n.s.
			all other combination pairs	<0.05
	Liver		C E C F E F	n.s.
			all other combination pairs	<0.05
	GTV		A C A F C F D E	n.s.
			all other combination pairs	<0.05
(V) Lymph node	Bladder	<0.05	A D C D D F	<0.05
			all other combination pairs	n.s.
	Sigmoid		A C A D	<0.05
			all other combination pairs	n.s.

Continued on next page

Table S12 – continued from previous page

Cases & contours		ANOVA for repeated measures with post-hoc test		
		Global hypothesis	Pairwise comparison	
		p-Value	Combination of pairs	p-Value
	Rectum		A C A F C F	n.s.
			all other combination pairs	<0.05
	CTV_L		A D C D D E D F	<0.05
			all other combination pairs	n.s.
	CTV_R		A D A E C D D F	<0.05

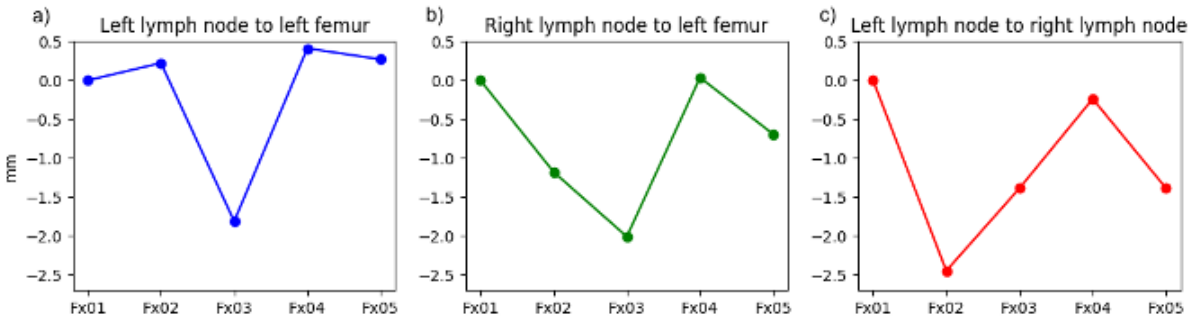

Figure S4: Differences in magnitude vectors a) left lymph node to left femur (mm) b) right lymph node to left femur (mm) and c) left lymph node to right lymph node (mm) of fractions 2-5 to fraction 1. The magnitude vectors were calculated from the x, y and z coordinates of the vectors a) left lymph node to left femur (mm) b) right lymph node to left femur (mm) and c) left lymph node to right lymph node (mm) which were calculated from the center of mass of the left lymph node, right lymph node, and left femur.

Table S13: Summary of the intraclass correlation coefficient (ICC) and the ANOVA for repeated measures with post-hoc test. Abbreviations CTV: clinical target volume, GTV: gross tumor volume, CTV_{HR}: High risk clinical target volume, CTV_{IR}: Intermediate risk clinical target volume, CTV_L: left clinical target volume, CTV_R: right clinical target volume, A-F are the Institutes, N/A: not applicable, n.s.: not statistically significant.

Cases & contours		Reliability			ANOVA for repeated measures with post-hoc test		
					Global hypothesis	Pairwise comparison	
		ICC	95% Confidence Interval		p-Value	Combination of pairs	p-Value
Gold Standard	CTV	1.00	1.00	1.00	<0.05	all combination pairs	<0.05
	Bladder	1.00	1.00	1.00	<0.05	A B	n.s
						A F	n.s
						B F	n.s
						all other combination pairs	<0.05

Continued on next page

Table S13 – continued from previous page

Cases & contours		Reliability			ANOVA for repeated measures with post-hoc test		
					Global hypothesis	Pairwise comparison	
		ICC	95% Confidence Interval		p-Value	Combination of pairs	p-Value
	Rectum	0.99	0.99	0.99	<0.05	all combination pairs	<0.05
(I) Prostate 1	CTV	1.00	1.00	1.00	0.24	N/A	
	Bladder	0.99	0.99	0.99	<0.05	all combination pairs	<0.05
	Rectum	0.99	0.99	0.99	<0.05	all combination pairs	<0.05
(II) Prostate 2	CTV	1.00	1.00	1.00	<0.05	A C C F all other combination pairs	n.s n.s <0.05
	Bladder	1.00	1.00	1.00	<0.05	D E all other combination pairs	n.s <0.05
	Rectum	0.98	0.98	0.98	<0.05	all pair of combination	<0.05
(III) Cervix	GTV	1.00	1.00	1.00	<0.05	D F all other combination pairs	n.s <0.05
	CTV_{HR}	1.00	1.00	1.00	<0.05	all pair of combination	<0.05
	CTV_{IR}	1.00	1.00	1.00	<0.05	all pair of combination	<0.05
	Bladder	1.00	1.00	1.00	<0.05	B D all other combination pairs	n.s <0.05
	Rectum	1.00	1.00	1.00	<0.05	A C all other combination pairs	n.s <0.05
	Sigmoid	1.00	1.00	1.00	<0.05	A E C E all other combination pairs	n.s n.s <0.05
(IV) Liver	GTV	0.92	0.92	0.92	0.33	N/A	
	Colon	1.00	1.00	1.00	<0.05	all pair of combination	<0.05
	Duodenum	1.00	1.00	1.00	<0.05	all pair of combination	<0.05
	Small bowel	0.99	0.99	0.99	<0.05	all pair of combination	<0.05
	Spinal cord	0.99	0.99	0.99	0.22	N/A	
	Liver	1.00	1.00	1.00	<0.05	A D all other combination pairs	n.s <0.05
(V) Lymph node	Bladder	1.00	1.00	1.00	0.05	N/A	
	Rectum	1.00	1.00	1.00	<0.05	A F all other combination pairs	n.s <0.05
	Sigmoid	1.00	1.00	1.00	<0.05	A E A F E F all other combination pairs	n.s n.s n.s <0.05
	CTV_L	0.84	0.84	0.85	<0.05	all pair of combination	<0.05
	CTV_R	0.72	0.71	0.73	<0.05	A E all other combination pairs	n.s <0.05

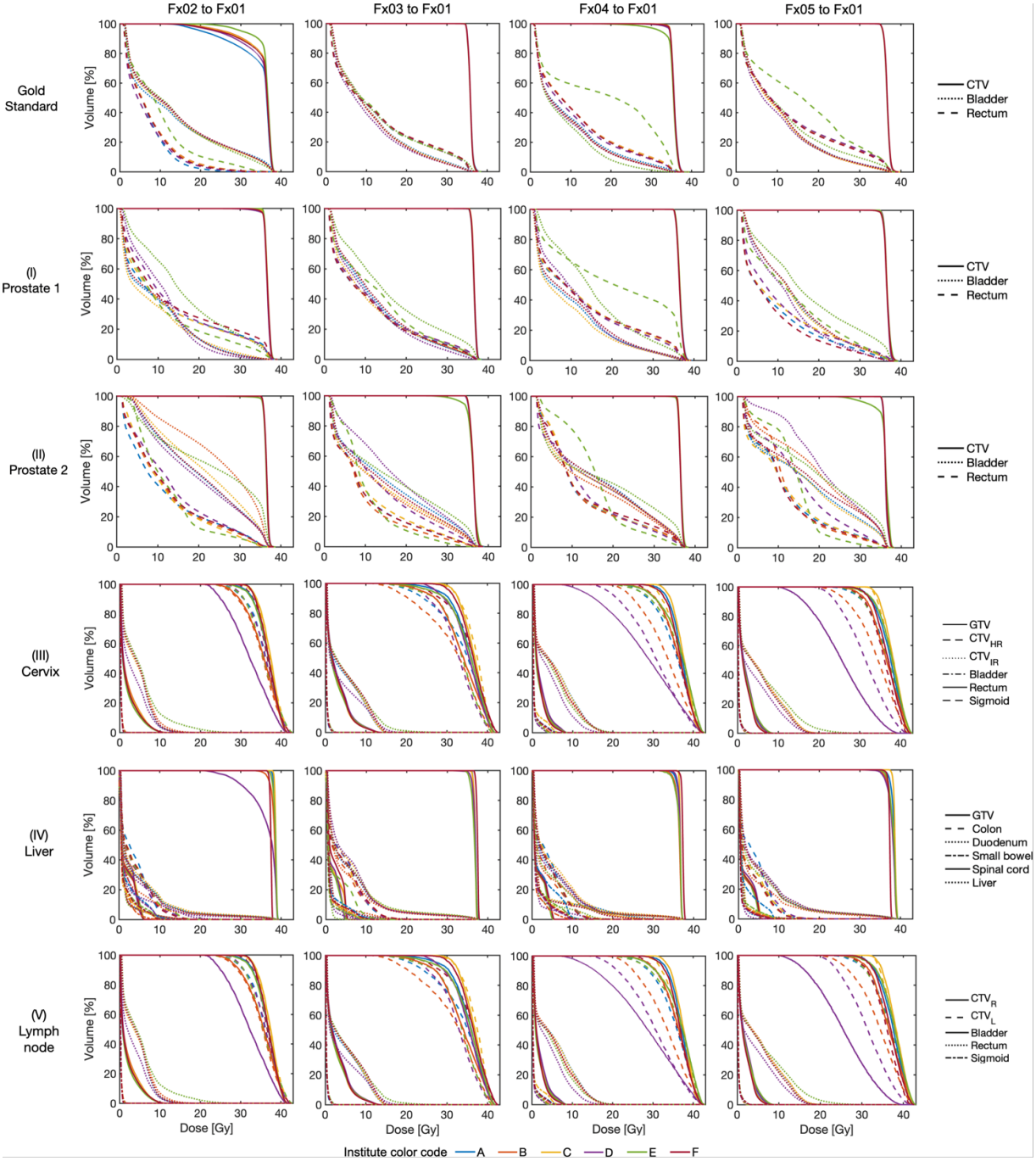

Figure S5: Gold standard and clinical cases resulted DHVs of the deformed dose for fraction two (Fx02) till fraction five (Fx05) to fraction one (Fx01) by the use of offline re-contoured structure set of fraction one (Fx01). The colors illustrate the different institutes (A-F) and the line type illustrates the clinically relevant contours. Note: Institute B did not provide mapped dose files for cervix and prostate 1. Abbreviations; CTV: clinical target volume, GTV: gross tumor volume, CTV_{HR}: High risk clinical target volume, CTV_{IR}: Intermediate risk clinical target volume, CTV_L: left CTV, CTV_R: right CTV.

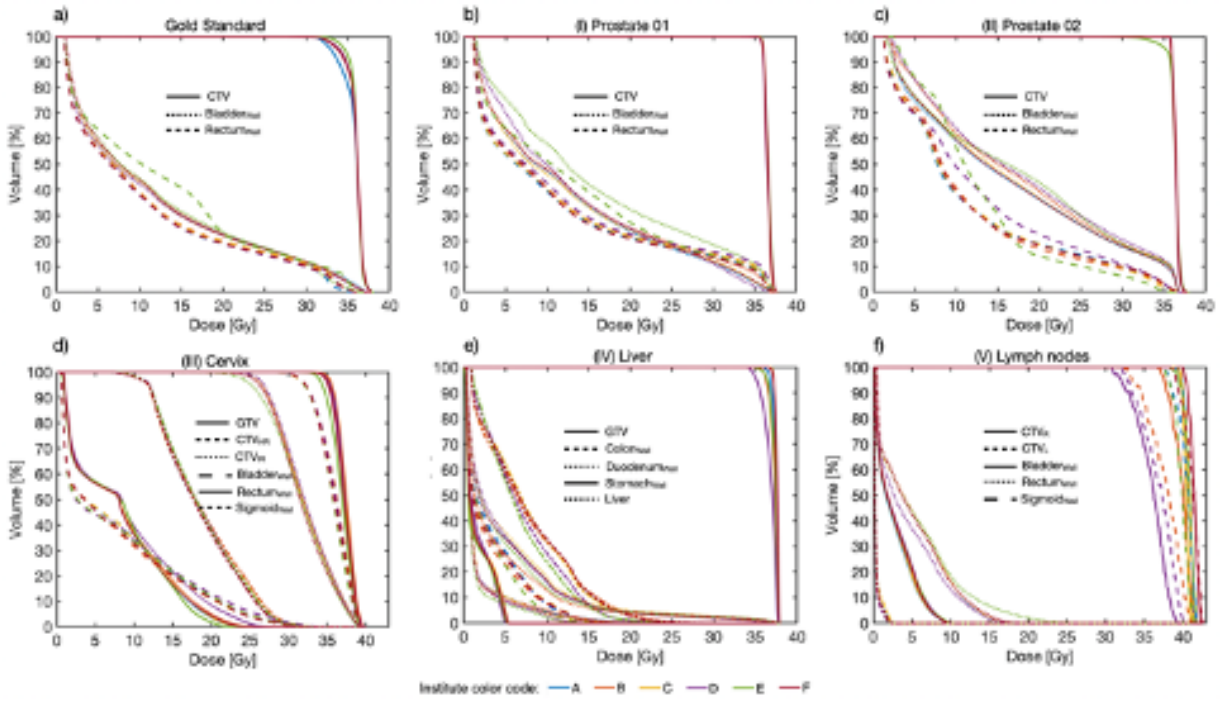

Figure S6: Results of the accumulated DVH per tumor volume and wall contours and institutes; for the cases gold standard (a), (I) prostate 1 (b), (II) prostate 2 (c), (III) cervix (d), (IV) liver (e) and (V) lymph node (f). Wall contours for rectum, colon, and duodenum, likewise for bladder and stomach were created using negative margins of 3 and 4 mm. The institutes (A-F) are presented in different colors, the case-related contours are presented in different line styles. Contours of fraction one were used for the respective DVH calculation.

Table S14: Clinical dosimetric criteria (CDC) results of deformable dose accumulation (DDA) of the institutes for the gold standard (GS) and the clinical cases; (I) prostate 1, (II) prostate 2, (III) cervix, (IV) liver, and (V) lymph nodes. Wall contours for rectum, colon, and duodenum, likewise for bladder and stomach were created using negative margins of 3 and 4 mm. The respective CDC for the individual institutes DDA were calculated by Monaco by the use of offline re-contoured wall structure set of fraction one (Fx01).

Case	Clinical dosimetric criteria (CDC)				Accumulated							
					A	B	C	D	E	F	Mean	Std
(Gold) Standard	$Bladder_{Wall}$	$D5cm^3$	< 37.0	[Gy]	33.0	32.7	32.8	32.4	31.4	32.6	32.5	0.6
		V28.0Gy	< 15.0	[%]	14.1	14.0	14.0	13.6	13.9	13.9	13.9	0.2
		V32.0Gy	< 20.0	[%]	9.5	9.3	9.3	8.8	6.4	9.1	8.7	1.2
	$Rectum_{Wall}$	$D1cm^3$	$\leq$ 38.0	[Gy]	33.1	34.3	34.3	34.1	35.4	34.0	34.2	0.7
		V28.0Gy	$\leq$ 15.0	[%]	12.0	12.9	13.1	11.8	13.7	12.0	12.6	0.8
		V32.0Gy	$\leq$ 20.0	[%]	6.7	8.2	8.4	7.4	9.9	7.3	8.0	1.1
(I) Prostate 1	$Bladder_{Wall}$	$D5cm^3$	< 37.0	[Gy]	37.1	33.9	32.3	31.2	35.6	30.1	33.3	2.7
		V28.0Gy	< 15.0	[%]	13.6	15.8	14.4	14.3	21.1	5.2	14.1	5.1
		V32.0Gy	< 20.0	[%]	8.5	11.3	9.5	7.3	15.7	2.5	9.1	4.4
	$Rectum_{Wall}$	$D1cm^3$	$\leq$ 38.0	[Gy]	36.5	36.4	36.6	36.8	36.6	36.4	36.6	0.2
		V28.0Gy	$\leq$15.0	[%]	15.4	16.3	16.0	16.9	16.0	15.0	15.9	0.7
(II) Prostate 2	$Bladder_{Wall}$	$D5cm^3$	< 37.0	[Gy]	35.8	36.0	35.9	36.1	36.1	35.8	35.9	0.1
		V28.0Gy	< 15.0	[%]	20.7	22.0	20.7	23.9	22.6	20.8	21.8	1.3
		V32.0Gy	< 20.0	[%]	15.2	17.0	15.4	17.3	15.8	15.2	16.0	0.9
	$Rectum_{Wall}$	$D1cm^3$	$\leq$ 38.0	[Gy]	34.9	34.9	34.3	35.1	33.1	30.1	33.7	1.9
		V28.0Gy	$\leq$ 15.0	[%]	11.8	10.5	11.1	13.2	7.7	11.6	11.0	1.8
		V32.0Gy	$\leq$ 20.0	[%]	7.8	7.3	6.8	8.5	4.4	8.1	7.2	1.5
(III) Cervix	$Bladder_{Wall}$	$D2cm^3$	< 27.5	[Gy]	26.5	27.0	26.5	26.3	26.6	26.2	26.5	0.3
	$Rectum_{Wall}$	$D2cm^3$	< 21.5	[Gy]	17.3	18.0	17.6	19.7	16.6	17.5	17.8	1.0
	$Sigmoid_{Wall}$	$D2cm^3$	< 21.5	[Gy]	21.7	21.3	21.3	22.6	22.0	21.6	21.8	0.5
(IV) Liver	$Colon_{Wall}$	Dmax	< 32.0	[Gy]	22.7	23.3	22.5	19.3	17.3	22.6	21.3	2.4
	$Duodenum_{Wall}$	Dmax	< 30.0	[Gy]	24.0	21.9	22.3	17.2	16.5	22.9	20.8	3.2
	$Stomach_{Wall}$	Dmax	< 30.0	[Gy]	29.7	30.1	30.3	26.9	28.0	30.4	29.2	1.4
(V) Lymph node	$Bladder_{Wall}$	$D0.5cm^3$	< 32.0	[Gy]	9.0	9.1	9.1	9.2	9.3	9.1	9.1	0.1
	$Rectum_{Wall}$	$D0.5cm^3$	< 32.0	[Gy]	14.5	15.0	14.8	14.3	19.2	15.2	15.5	1.8
	$Sigmoid_{Wall}$	$D0.5cm^3$	< 32.0	[Gy]	1.7	1.5	1.9	1.3	1.5	1.6	1.6	0.2

9.3 Comparison of online adaptive and non-adaptive magnetic resonance image-guided radiation therapy in prostate cancer using dose accumulation

9.3.1 Supplementary A: Material and Methods

EQD2 was calculated with the linear-quadradic model

$$EQD2 = \left(\frac{\sum_i (D_i^n \cdot V_i)}{\sum_i V_i} \right)^{\frac{1}{n}} \cdot \frac{d + \frac{\alpha}{\beta}}{2 + \frac{\alpha}{\beta}} \tag{10}$$

where D_i specifies the dose and V_i the volume at dose level i, respectively. Parameter n describes the volume effect, d the dose per fraction and the α/β ratio the dose-response relationship.
Cohen's d was calculated by

$$d = \frac{(M_1 - M_2)}{s_p} \tag{11}$$

with M_1 and M_2 being the mean values of the two groups to be compared and s_p the pooled standard deviation which is a weighted average of the standard deviations of the two groups.

9.3.2 Supplementary B: Results

Figure S7 illustrates the variability of bladder and rectum volumes for the initial treatment plan and across fractions per patient.

The two investigated DDA approaches, OA-MRgRT and conv-IGRT, yielded comparable median (range) NTCP results (cf. Supplementary Figure S8), in detail for bladder incontinence G2+ with 2.4 % (2.2 % - 3.0 %) and 2.4 % (2.3 % - 2.9 %) and for rectum late toxicity G2+ OA-MRgRT and conv-IGRT resulted in 0.4 % (0.4 % - 0.6 %) and 0.5 % (0.4 % - 0.7 %). The statistical analysis yielded no significant differences.

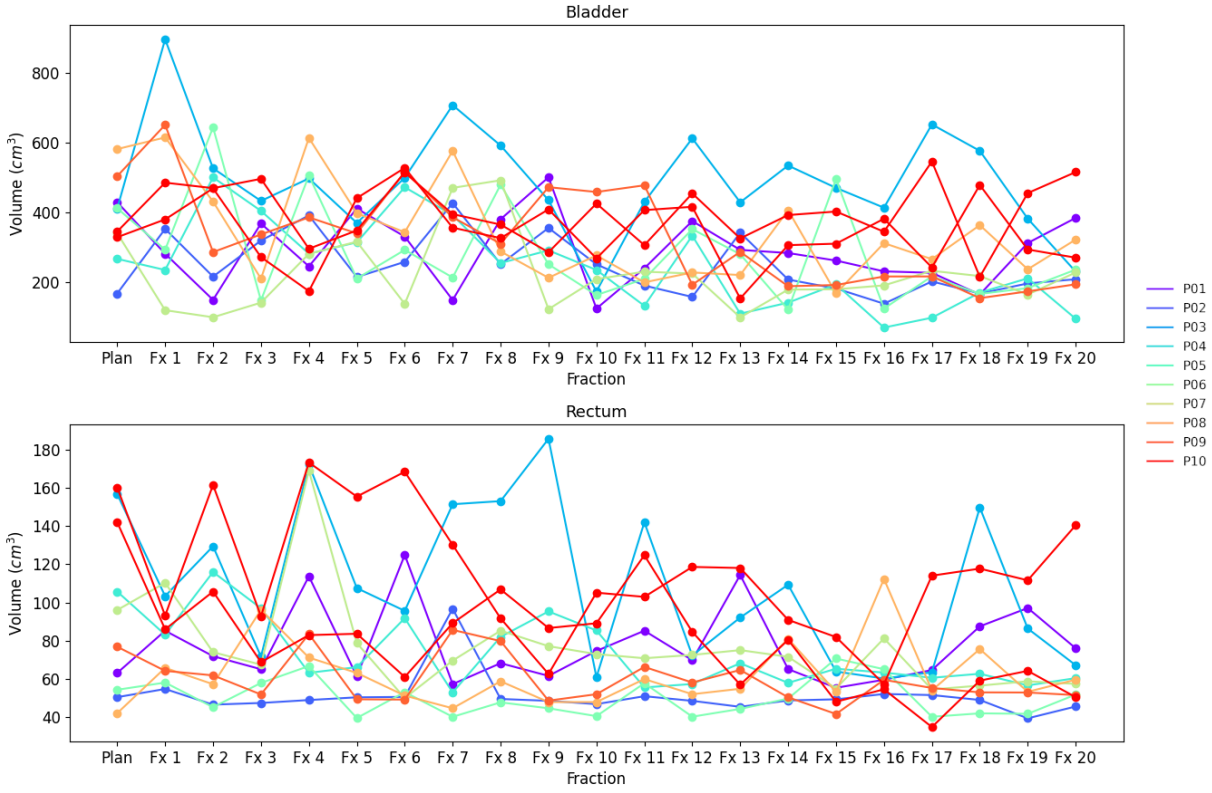

Figure S7: Variability of bladder and rectum volumes for the initial treatment plan and across fractions per patient.

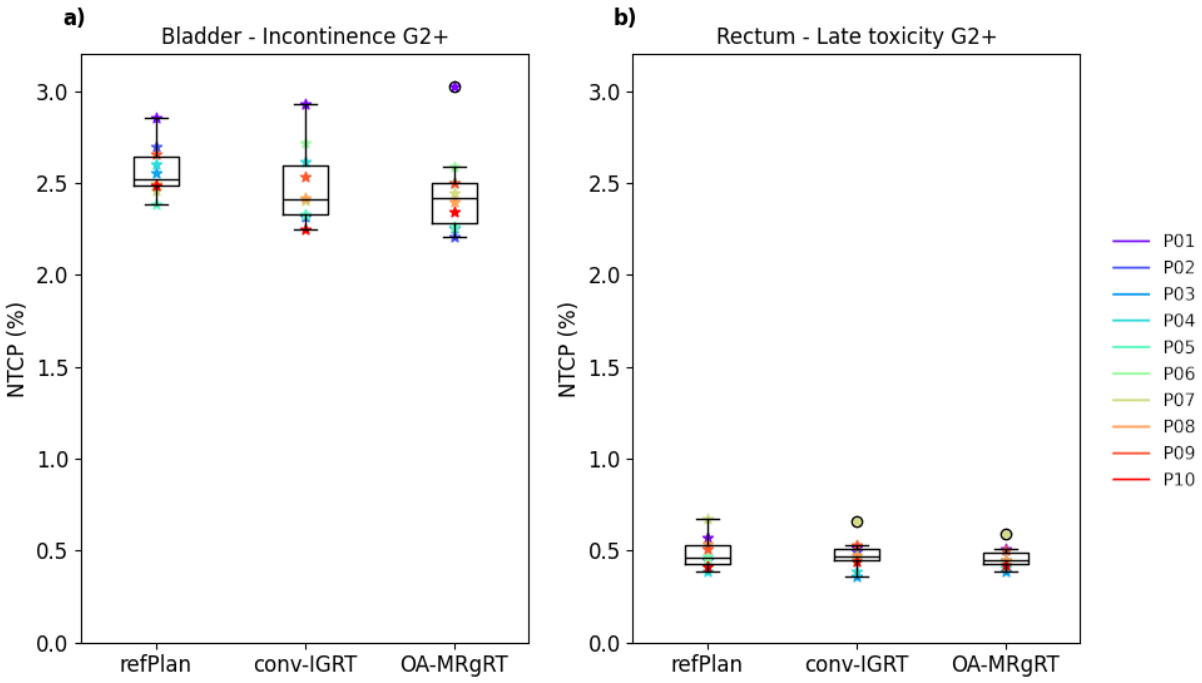

Figure S8: Boxplots showing the differences in the NTCP for a) bladder incontinence G2 and b) rectum late toxicity G2+ between refPlan, conv-IGRT, and OA-MRgRT. The boxes represent interquartile range (IQR), which is the range between the first quartile (Q1) and the third quartile (Q3). The black line inside the box is the median of the dataset. The length of the whiskers is set to 1.5 times the IQR. Individual data points beyond the whiskers are considered potential outliers and plotted as black circles. The color-coded points are the patient-individual results.

10 Acknowledgments

I have received an immense quantity of support and assistance from exceptional individuals from a variety of backgrounds during the composition of this thesis and the ongoing advancement of deformable dosage accumulation in radiotherapy.

Initially, I would like to express my sincerely gratitude to Prof. Daniela Thorwarth, my supervisor, for her unwavering support and her unwavering commitment to ensuring that I derived satisfaction from my academic pursuits.

Additionally, I would like to express my gratitude to Prof. Arndt-Christian Müller for his guidance and encouragement in this regard. I would like to express my gratitude to the Tübingen DFGPC project committee for their confidence in my ability to advance this initiative. Dr. Daniel Wegener, I am grateful for the exchange and your contributions to the DFGPC initiative. Regarding this matter, I would also like to extend my gratitude to Prof. Dr. Zips, Dr. Gani, and Dr. Böke.

MRL consortium working groups have been a fundamental component of my research. The development of my work has been substantially influenced by their insightful feedback and exceptional cooperation in the multi-center study.

Additionally, I would like to extend my gratitude to the remarkable ESTRO physics working group on dose accumulation and mapping for the stimulating, informative, and occasionally overwhelming discussions. I have acquired a wealth of knowledge from you. I am overjoyed to be a member of the community.

I would like to express my gratitude to the biomedical physics research group for the excellent collaboration and your responses to my presentations. It is also my pleasure to express my gratitude to my office buddies for the private conversations that allowed me a break from work.

Lastly, I would like to express my gratitude to my family, mother, stepfather, and grandmother for their unwavering and affectionate support, which has enabled me to successfully convey my academic research in a manner that is easily understood. For the academic discussions that have always motivated me and my brother's unwavering support in achieving the desired outcome, I am grateful to him and his wife. I would like to express my gratitude to my partner Peter for his steadfastness and composure and for his advice "Martina, just take a break" as well as for the hardships during this exciting time.

I would not be in the position I am in today without your assistance.