Dissertation

submitted to the
Combined Faculty for the Natural Sciences and Mathematics
of the Heidelberg University, Germany

for the degree of
Doctor of Natural Sciences

Put forward by

M.Sc. Nina Isabelle Niebuhr
Born in Marburg, Germany

Oral examination: 17.12.2019

Biological Dose Accumulation in Image-guided Radiotherapy

Referees: Prof. Dr. Joao Seco

Prof. Dr. Oliver Jäkel

Biological Dose Accumulation in Image-guided Radiotherapy

Dose accumulation (DA), the computation of the total delivered 3D dose distribution D_a of a fractionated radiotherapy treatment using daily patient imaging, is getting established in clinical practice. Although the aim is to estimate the achieved biological outcome of the treatment, the principles of biological effect estimation are currently not followed consistently in the process. In this thesis, the biological effect accumulation approach and total biological dose (bEQD) were derived as a biologically consistent DA workflow. Clinical relevance of bEQD and its dependence on individual workflow aspects were investigated in data from three patient cohorts. It was found that D_a systematically underestimates the obtained biological effect, which can be avoided by the use of bEQD. Results showed that this is strongest for late-responding organs at risk (OAR) with low α/β values in dose gradient regions around the target that are prone to organ motion. bEQD to D_a deviations occurred locally, in so-called hotspots, showing individual cases of high difference magnitude but only small statistical impact. Hotspots of $bEQD - D_a$ deviation around 4 Gy in bladder and rectum were found in patients treated for prostate carcinoma. Hypofractionation increased these deviations strongly up to 8 Gy and also showed clinically relevant deviations in dose-volume analysis. Dose-response correlation in standard fractionation showed only little impact on the DA approaches. Workflow uncertainties are dominated by those from deformable image registration, which are in the same range as the difference between bEQD and D_a. bEQD should be considered in the application of treatment adaptation, especially to avoid damage to OARs in individual cases.

Biologische Dosisakkumulation in der bildgeführten Strahlentherapie

Dosisakkumulation, die Berechnung der summierten 3D-Strahlendosisverteilung D_a einer fraktionierten Strahlentherapie Behandlung mittels täglicher Bildgebung, wird derzeit Teil klinischer Routine. Obwohl das Ziel ist, das biologische Resultat der Behandlung abzuschätzen, werden mathematische Modelle zur Berechnung des biologischen Effektes bisher nicht konsistent in der Dosisakkumulation berücksichtigt. In dieser Arbeit wurde ein Effektakkumulations-Ansatz sowie die gesamtbiologische Dosis (bEQD) eingeführt, um einen biologisch konsistenten Prozess der Dosisakkumulation zu ermöglichen. Die klinische Relevanz der bEQD sowie deren Unsicherheiten und Abhängigkeit von einzelnen Workflow Schritten, wurde anhand von drei Patienten-Kohorten untersucht. Es zeigt sich, dass D_a den biologischen Effekt der Behandlung systematisch unterschätzt, was durch Anwendung von bEQD vermieden werden kann. Ergebnisse zeigen, dass diese Unterschätzung am stärksten ist für spät reagierende Gewebe mit kleinem α/β Wert, welche sich in steilen Dosisgradienten sowie in Regionen regelmäßiger Organbewegung befinden. Abweichungen zwischen bEQD und D_a treten lokal auf, in sogenannten Hotspots. Es zeigen sich Einzelfälle mit starken Abweichungen, jedoch nur geringer statistischer Einfluss. Hotspot mit 4 Gy Abweichungen in Blase und Rektum wurden in Patienten in Behandlung von Prostatakarzinomen gefunden. Hypofraktionierung erhöhte diese Abweichung auf bis zu 8 Gy und zeigte zudem klinisch relevante Abweichungen in Dosis-Volumen Untersuchungen. Die Korrelation von Dosis zu Gewebeeffekten im Rektum wurde nur geringfügig von der Akkumulations-Strategie beeinflusst. Unsicherheiten im Workflow wurden dominierten von solchen der deformierbaren Registrierung, welche ähnliche Magnituden wie der Unterschied zwischen bEQD und D_a zeigten. bEQD sollte für die Anwendung in adaptiver Strahlentherapie benutzt werden, besonders um Strahlenschäden im Normalgewebe zu vermeiden.

CONTENTS

1

INTRODUCTION

Radiotherapy (RT) aims for accuracy and precision in the delivery of ionizing radiation dose to the tumour and simultaneously sparing of the surrounding organ(s) at risk (OAR). To achieve this goal, imaging plays an indispensable role in the state-of-the-art RT workflow. The introduction of computed tomography (CT) in the 1970s enabled 3D-conformal treatment planning for the first time. The conformality of the dose delivery was further significantly improved throughout the 1980s and 1990s with the development of intensity-modulated radiotherapy (IMRT) [9] [6]. With this technique, isodose-lines could be tailored to the desired target location and shape, along with steep dose gradients that reduce the dose received by the surrounding structures.

A RT treatment is delivered in consecutive daily doses. This principle of treatment fractionation exploits the repair capability of normal tissue between fractions, while still achieving tumour cell kill. Dose therefore accumulates in the tissue over the course of therapy. While the initial CT scan prior to the treatment course, is used to set up the treatment plan, the achieved precision and accuracy of the dose delivery is strongly impaired by organ motion throughout the course of treatment, both intra- and interfractional. Due to internal organ movement as well as setup-uncertainties of the patient, the machine does no longer deliver the planned dose precisely to the targeted location on each treatment day. In consequence, the accumulating dose is increasingly washed out which can result in underdosage of the tumour and overdosage of OARs. To tackle this problem, in recent years image-guidance was established such that the setup and anatomy of the day can be monitored. Modalities used for image-guidance right before dose application are for example CT-imaging in both the kilo- and megavoltage regime as well as newly established methods using magnetic resonance imaging. This additional control step prior to RT facilitates setup corrections of the patient to ensure that the target receives the desired dose and that OARs are not being irradiated with a higher than the planned dose. Furthermore, the daily delivered dose can be estimated at voxel-level, revealing significant differences to the planned dose as reported in current literature [41] [28] [26] [56]. This enables the development of new approaches to improve the outcome of RT treatments which are a topic of current research: treatment adaptation [71] and dose-response correlation of the actually delivered dose to the treatment outcome [63] [70]. Both approaches require the calculation of the dose delivered per day as well as its computational accumulation in a common frame of reference. With the estimate of the total delivered dose on a local scale (voxel-wise), clinical decisions can be made, in case changes of the treatment plan are necessary, e.g. to spare

OARs that have already received too much dose, re-treatments can be based on how much dose has already been delivered to the patient before [5], and occurring side-effects can be correlated to the dose that they were caused by.
Nevertheless, the actual outcome of a RT treatment is not the quantity of total delivered dose but the biological effect that was achieved in the tissue: the death of tumour cells and potential damage to healthy tissue. The relationship between dose and biological effect has been a wide topic of research for a number of decades. Mathematical models try to describe the effects in tissue usually by estimation of the cell survival fraction (SF). The most commonly used model introduced in 1972 [36], which is still standard practice today [49], is the Linear Quadratic Model (LQM). Within this empirical model, the dose is linked to the SF in an exponential expression with a linear and a quadratic dose term, including tissue specific radiosensitivity-parameters.
Dose accumulation, thus, the method of computation of the voxel-wise total dose delivered to the patient over the course of fractionated treatment, by use of daily imaging, is entering clinical routine. Nevertheless, it is still discussed as a controversial topic [68] presenting new opportunities as well as challenges. This thesis builds on the observation that the workflow chain of conventional dose accumulation in current applications does not take the biological considerations into account in a consistent way. Within this work, the discussion on the rationale of dose accumulation is extended to the biological level.
This thesis tries to answer two questions: i) How can dose accumulation be performed in a biologically consistent way, i.e. to be in accordance with biological models and considerations? ii) Is it relevant to consider the accumulated total biological dose for clinical applications?
To answer the first question, the first aim of this thesis was to set up the methodology of a biologically consistent workflow for dose accumulation. Two aspects have been identified for which a critical rethinking of current applications and recent postulation is necessary in the context of biological mechanisms and models: the deformation of dose using deformable image registration (dose mapping), especially for scenarios where volumes are changing; and with this as a basis, the biological accumulation of consecutive treatment fraction doses.
Dose mapping describes the deformation of the daily delivered dose to a common frame of reference to enable voxel-wise dose accumulation. For this, deformable image registration was used and applied to the daily estimated dose distributions. The concept of dose mapping, specifically for scenarios of volume change, is a topic of ongoing debate [90]. Within this work, this process was questioned and investigated critically in a biological context. On the basis of the biologically consistently obtained voxel-wise daily doses from dose mapping, the main development of this work was to establish and investigate a biologically consistent way for effect and dose accumulation. It was found that conventional dose accumulation introduces a systematic inaccuracy in the prediction of the total biological effect. The introduced (biological) effect accumulation approach and the concept of the total biological dose

was developed to avoid this inaccuracy.
The developed methodology was investigated for three patient cohorts treated with IMRT and daily CT imaging. Each dataset, from the respective cohorts, was used to tackle certain aspects in the process to obtain the answer to the second question of this thesis: a data-based conclusion of the clinical relevance and applicability of biological dose accumulation, especially when considering uncertainties. Daily imaging and delivered dose of a cohort of 9 patients treated for prostate carcinoma at the German Cancer Research Center (DKFZ), Heidelberg, was investigated in depth to conclude on how the difference between the total biological dose and conventional dose accumulation from this data is affected by individual workflow aspects. Furthermore, uncertainties in the respective workflows were estimated. Data from a larger cohort of 251 prostate carcinoma patients from the VoxTox programme [13], Cambridge, UK, were used to find a statistical answer on the difference between both methods. Furthermore, dose-toxicity correlation on the basis of the derived total biological dose was tested in comparison to the standard approach. In addition, data from a cohort of 26 cases treated for head and neck cancers from the VoxTox programme extends the analysis to another body site.

This thesis contains five chapters following this introduction chapter. In chapter 2.1, selected fundamental concepts of radiotherapy and image-guidance are introduced as the framework of the thesis. An in depth background on specific methods used for dose accumulation and radiobiology is given in section 2.2. Materials and methods are described in chapter 3. In this chapter, the developed methodologies, that combine the concepts of dose accumulation and radiobiology, are presented: biologically consistent dose mapping, effect accumulation and the total biological dose (section 3.1). Section 3.2 describes the tools and data used in this thesis, i.e. the patient data used and the individual dose accumulation workflows. The results of the patient data analysis of biological effect and dose accumulation are presented in chapter 4. Chapter 5 discusses the developed methodology and the obtained results and concludes with the implications for the application of the presented concept (section 5.3). A short summary is given in chapter 6.

Parts of the content of this thesis have been submitted for publication in *Physics in Medicine and Biology* as a note on biological dose mapping [57] and a paper on effect accumulation and the total biological dose [58].

2

PHYSICAL AND CLINICAL FRAMEWORK OF THE THESIS

Cancer therapy searches for therapeutic agents that cause a *specific effect* on cancer cells. This means that optimally, the agent is inducing the desired effect in the target cells (i.e. cell death) while healthy cells remain unaffected. In radiotherapy (RT), this specific effect is induced by localized deposition of a certain radiation dose in the target, while the surrounding tissue optimally remains un-irradiated (or realistically, only receives low dose levels). The therapeutic agent is radiation, specifically photons in the presented work, and its *specifitcity* to kill cancer cells relies on the spatial accuracy and precision of the dose deposition as well as on the amount of dose and its delivery scheme.
In section 2.1 the physical embedding of this thesis is presented. Detailed background on the used methodological principles of dose accumulation and radiobiological are described in section 2.2.

2.1 PHYSICS OF IMAGE-GUIDED RADIOTHERAPY

The protagonist of this section is photon irradiation and its interaction with matter. Photons are not only used as the therapeutic agent, but also for imaging of the patient's anatomy. The latter is necessary to setup treatment plans but also for image guidance throughout therapy. Within this section, the physics of photon-radiotherapy and imaging with photons will be presented in section 2.1.1. How these principle are applied in RT treatment techniques is described in section 2.1.2, followed by principles of image-guided RT in section 2.1.3.

2.1.1 *Photon interaction with matter in imaging and therapy*

Physics of photon interaction with matter The attenuation of photons in matter is proportional to the incoming number of photons (i.e. the intensity I) and the thickness of the absorber x [66]:

$$-\,dI(x) = \mu I(x)dx \tag{2.1}$$

Solving the differential equation shows that the attenuation of photons in matter can be described by an exponential decay, the Lamber-Beer-law:

$$I(x) = I_0 e^{-\mu x} \tag{2.2}$$

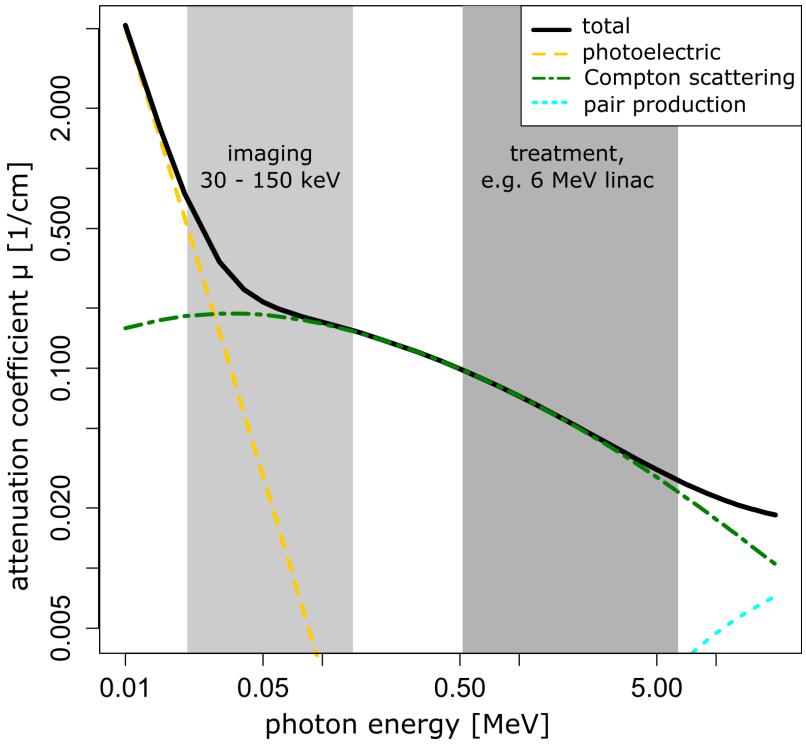


Figure 2.1. – Contribution of the photoelectric effect, Compton scattering and pair production to the total mass attenuation coefficient μ against the photon energy for the example of interaction in soft tissue (muscle). Highlighted are regions of energy ranges used for standard CT imaging and for treatment (example for a linac producing a photon energy spectrum with a maximum of 6 MeV). Created using the NIST x-ray attenuation database [29].

with the the linear attenuation coefficient μ of the absorber material which depends on the photon energy. The mass absorption coefficient is normally reported with μ/ρ where ρ is the mass density of the absorber material. Note that the above equation is only valid within a medium of homogeneous absorption properties. In an inhomogeneous medium, the quantity of μ is dependent on the location within the absorber, so that the above equation 2.2 should be written as the line integral along x:

$$I = I_0 \cdot e^{-\int \mu(x)dx} \tag{2.3}$$

The attenuation coefficient μ is characterized by the tissue's anatomical composition and can be expressed by the electron density ρ_e times the electron cross sections: $\mu = \rho_e \cdot \sigma_e(E_\gamma, Z)$, where σ_e depends on the photon energy and the charge number of the absorber material.
The manner of photon interaction in matter can be described by mainly three fundamental interactions: photoelectric (PE) effect, Compton (or incoherent) scattering and pair production. The attenuation coefficient can therefore be expressed by the sum of the individually contributing effects from those three interactions.

$$\mu(E_\gamma, Z) = \tau(E_\gamma, Z) + \sigma(E_\gamma, Z) + \kappa(E_\gamma, Z) \tag{2.4}$$

with the individual energy (E_γ) and material (Z) dependent attenuation coefficients τ, σ and κ. The contribution of the individual effects on the total absorption coefficient is shown in figure 2.1. For imaging, energy ranges between 30 - 150 keV are utilized which are characterized by an interplay of the PE effect and Compton scattering, but dominated by the latter. For the therapeutic energy range, energies in the MeV range are used for which Compton scattering dominates all other effects. Pair production, although physically possible at energies above 1.022 MeV, only gains importance for energies above 10 MeV.
Current standard RT is performed using linear accelerators (linear accelerators (linacs)). Within such a linac, the photon beam is created through bremsstrahlung of previously accelerated electrons towards a tungsten target. Thus, the resulting photon beam consists of a spectrum of energies with the maximum defined by the linacs maximum electron accelerating voltage. In common notation, for example a $6\ MV$ linac produces a photon spectrum with a maximum energy of $6\ MeV$.

Imaging with photons Photons in matter can either be transmitted through the medium, be absorbed or reflected (the latter being negligible for the here discussed energy ranges). For imaging, a delicate balance of transmission and absorption is to be found. Reasonably high transmission is favourable to achieve a high signal to noise ration and low dose to the patient. On the other hand, absorption is necessary to receive information about the traversed material, thus, to gain a higher contrast to noise ratio. In the energy ranges described above for photon imaging, both the PE effect and Compton scattering have to be taken into account. For the Compton effect, only the electron density of the material is affecting the attenuation (in addition to the dependency on the photon energy) but it is independent of the charge number Z. Thus, this gives little to no information on differences in soft tissue, since for those, electron densities are very similar. At lower energy levels, the PE starts to become more relevant. The attenuation coefficient of the PE effect has an approximate relationship to the atomic number of the absorber material of $\tau \approx Z^{3-4}$. With this, it provides critical information on the tissue composition in imaging but also leads to increased dose to the patient. Hence, for a good compromise between dose and contrast, energies between $30\ -\ 150\ keV$ are used.
For imaging with photon, e.g. in computed tomography (CT), the Lambert-Beer-law as the line integral in equation 2.3 directly describes the intensity I of photon transmission along the path from the photon source to the detector, characterized by linear attenuation coefficient μ. Through backprojection of this line integral, μ can be determined. However, the obtained value will be strongly depend on the used spectrum of photon energies of the scanner, so that instead, the relative value to that of water is used. Imaging values in CT imaging are thus given in Hounsfield Units (HU) [66] defined as:

$$HU = \frac{\mu - \mu_W}{\mu_W} \cdot 1000 \tag{2.5}$$

where μ_W is the attenuation coefficient of water. Per definition, the HU of water equals 0, while air has -1000 HU. Each HU is therefore equivalent to 0.1% of linear attenuation in water. Fatty tissue is in the range of -100 to -50 HU and other soft tissues range between 0 and 100 HU. Bone and other more dense materials range from 200 to roughly 2000 HU.
Primarily, HUs are dependent on the electron density of the tissue due to the dominating effect of Compton scattering described above. However, due to the impact of the PE effect, also the atomic charge number affects the HU value of the tissue, especially for high charged compositions of heavier materials such as bone. This effect is decreasing when using higher photon energies.

Treatment with photons The quantity of interest when using photons for cancer therapy is the dose [66]:

$$D = \frac{d\bar{E}}{dm} \text{ [Gy]} \tag{2.6}$$

where $d\bar{E}$ is the mean energy deposited in the mass dm given in the units of Gray, which equals Joule per kilogram. Energy and mass are both extensive quantities of the system, resulting in the fact that the dose, as the ratio, is an intensive quantity, thus independent of the system size.
In photon RT, the main contribution to dose deposition is caused by secondary Compton electrons and not directly by the photon radiation. The absorption of energy is caused by the Compton electrons through ionization of the absorber material based on Coulomb interaction with other electrons. When photons are entering the absorber material, a cascade of Compton electrons is induced which leads to a build-up effect of absorbed energy. The energy absorption reaches its maximum at approximately the maximum range of the secondary electrons. This range depends on the electrons energy which in turn depends on the energy of the incoming photons. Only after reaching the equilibrium of Compton electron production and photon absorption, the exponential decay described by the Lamber-Beer law (Eq. 2.3) is observed. This is shown in figure 2.2. In this equilibrium, the energy released locally from the primary uncharged particles can in first approximation be described by the above mentioned attenuation coefficient μ through the attenuation law:

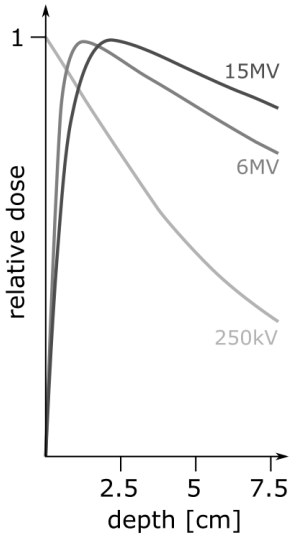


Figure 2.2. – Schematic depth dose curve of photons with different energies in water. Voltages refer to the maximum voltage used in the linac for beam creation, hence, creating the maximum of the energy spectrum in keV or MeV. Recreated from [66].

$$\Psi(E, \vec{r}) = \Phi(E, \vec{r}) E e^{-\mu(E, \vec{r})} \tag{2.7}$$

where Φ is the initial photon fluence and Ψ describes the energy fluence within the tissue at the interaction point $\vec{r}$. With this, the Total Energy Released per unit Mass (TERMA) [Gy] (for a monoenergetic photon beam) is given by:

$$T(\vec{r}) = \frac{\mu}{\rho}(\vec{r})\Psi(\vec{r}) \tag{2.8}$$

This quantity is used in dose calculation algorithms described shortly below.

2.1.2 *Techniques of photon-radiotherapy*

The depth dose curve of photons in matter depicted in figure 2.2 shows that the maximum dose deposition is reached already shortly after entering the medium, caused by the electron build-up effect described above. It is at approximately the range of the produced secondary electrons, which depends on the initial photon energy. For the used energies in RT, the range of the secondary electrons is in the order of a few centimetres. Due to the exponential decrease of the dose deposit distal to the maximum, a considerable amount of dose would be delivered also to the healthy tissue, in case the patient was treated from one direction only. Thus, in order to achieve the *specific effect* in the tumour cells, characterized by local deposition of dose and sparing of the healthy tissue, the principle of crossfire irradiation is exploited in external beam RT. This means that multiple beams from different directions are applied in such a way that an overlap is created in the target. In consequence, the dose to the healthy tissue is kept low in comparison with the dose to the target, by spreading the dose over a larger volume of healthy tissue.

IMRT and TomoTherapy In cancer therapy, the target volume is often irregularly shaped and close to adjacent organ(s) at risk (OAR). In consequence, even with the overlap of beams from several directions, the dose plan cannot cover the target without significant dose to closely surrounding structures, especially in case these targets have a concave shape. Collimators are used to shape the beam to fit the outlines of the target from the respective beam angles, called three-dimensional (3D)-conformal RT. However, as shown in figure 2.3, this does not necessarily ensure enough sparing of healthy tissue, especially closely located OARs. To improve the confromality of the dose deposition to the target, the approach of *intensity-modulated radiotherapy (IMRT)* was introduced [9]. The used principle is to modulate not only the shape, but also the intensity of the incoming photon beam based on the shape of the tumour. Beamlets (minimum individually controllable areas of a beam) that traverse through critical structures are reduced in intensity, while the intensity of beamlets that reach the target, is increased.

This leads to the fact that dose optimization for IMRT treatment planning becomes an inverse problem. The dose distribution is the given requirement, specified by the clinician or medical physicist, and parameters on fluence and beam directions need

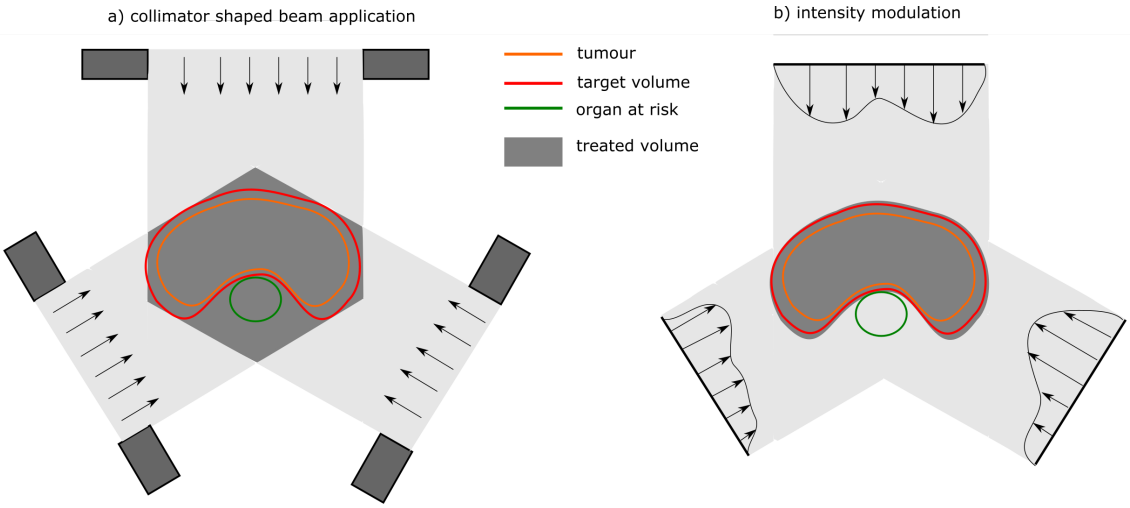


Figure 2.3. – Schematic illustration of the irradiation of a target volume with a) collimators to shape the beam according to the target projection and b) intensity modulation (IMRT) to shape the beam according to the target shape. Adapted from [66].

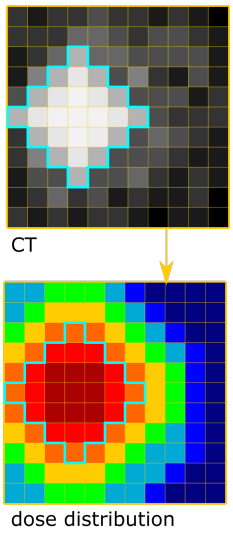
to be found by the optimizing algorithm. IMRT therefore requires CT imaging the a basis as well as computer-based dose optimization to create the treatment plan. Characteristically, IMRT dose plans show steep dose gradients around the target volume, which is advantageousness to spare the healthy tissue. Large volumes of surrounding healthy tissue also receive low levels of dose based on the used number of beam directions.

In current applications, shaping of the beam is achieved by the use of Multi Leaf Collimators (MLCs), consisting of individually movable (tungsten) leaves. A standard method of IMRT is the *step-and-shoot* (a.k.a. fixed-field) technique, where the linac is moved to the respective beam angle direction step-by-step; at each linac angle, sequentially different treatment fields are irradiated such that the summed up intensity of the fields matches the desired intensity modulation. A similar technique is given by (helical) *TomoTherapy*, where the irradiation is performed in a similar manner to CT acquisition (i.e. tomography) performed in a slice-wise rotational manner with a fan-shaped beam. This process can enable a higher number of beam direction based on the 360° rotational delivery which can be used for a high conformality in target shaping and target dose homogeneity. In turn, this can result in larger volumes of normal tissue to receive low levels of dose. In comparison, literature reports better dose homogeneity in the target volume for TomoTherapy plans, while for standard step-and-shoot IMRT, sparing of OARs was better [45] [84]. However, it should be noted that the dose distribution strongly depends on the used dose constraints, prescription dose, the optimizing algorithm and the impact of the medical physicist.

The mentioned *specific effect* achieved by RT is only partially based on the spatial conformality of the delivered dose. Other important factors are the total delivered dose and the dose delivery scheme. Standard RT is delivered in a fractionated man-

ner, meaning that low doses are delivered in a consecutive fashion of one treatment per day over a course of several days or weeks. This is based on radiobiological considerations that will be described in detail in section 2.2.2.

Dose calculation algorithms For treatment planning, the dose needs to be tailored to the structures within the patient. For this it is necessary to compute the dose absorbed in the irradiated material, conventionally based on a pre-treatment planning CT scan.
As a short example for dose calculation, the basic principle of kernel-based superposition is described (i.e. macroscopic descriptions of the physical processes of dose deposition in tissue) [2]. The computation of kernel-based dose calculation in RT is based on analytical models and is build up of three components: the primary photon beam, the absorbed energy, and the dose kernel. The final dose computation is then given by a superposition, or convolution of these three. The primary photon beam is the information of the photon spectrum of the linear accelerator, composed of the photon fluence, the energy spectrum, the photon velocity directions and the spatial distribution of the photons (for example shaped by collimators or intensity modulation). This information is necessary to calculate the energy released by the primary photons in the tissue, i.e. the TERMA, eq. 2.8. However, as described above, the main dose deposition is caused by the secondary Compton-electrons in interaction with the electrons of the tissue. This process is computed via so-called dose kernels, i.e. an energy transfer function of how energy is deposited locally. The final dose calculation is then given by the superposition of the TERMA and the dose kernel.



The electron density of the tissue influences both the calculation of the TERMA (through the path of the photons through the patient) and the dose kernel computation (absorption rate is effectively weighted by the electron density relative to that of water). The HUs described above, derived from imaging, do however not directly describe the electron density. Furthermore, the linear attenuation coefficient in the imaging energy range is different to that of the therapeutic imaging range. Therefore a calibration is necessary to convert between imaging values in HU and the necessary information of the electron density from imaging, thus exploiting the mutual physical dependency on electron density in both cases, derived experimentally. To obtain the necessary information, calibration scans of phantoms with known electron density are performed in current applications. The result is an empirical curve that links HUs to electron density, so-called *Hounsfield look-up table (HLUT)*. HLUTs are individually calibrated for each CT scanner and used image-reconstruction kernel based on a small dependency of the absolute HU number on both. An example HLUT is shown in figure 2.4.

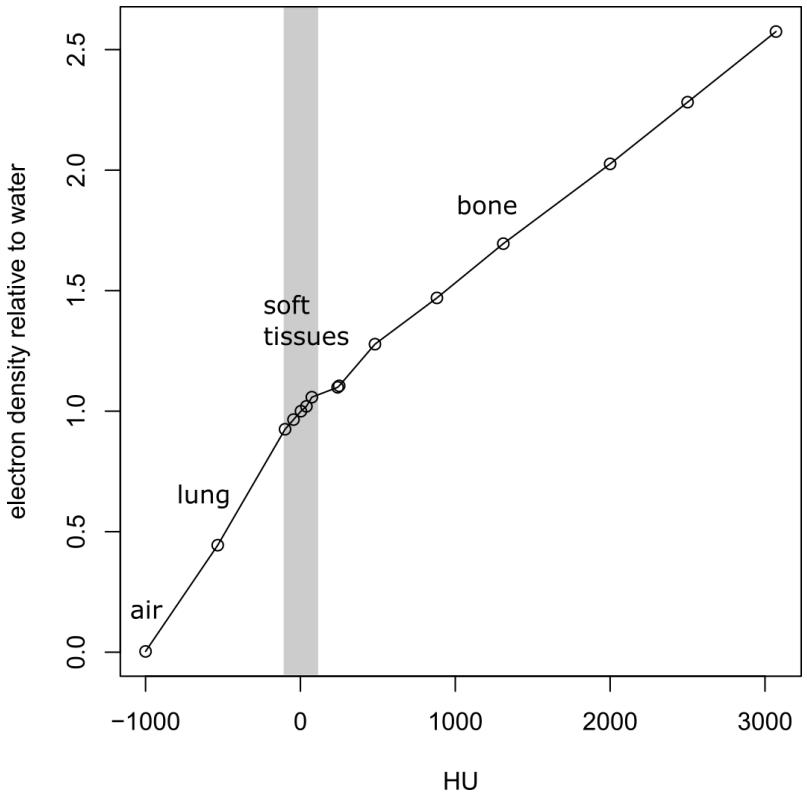


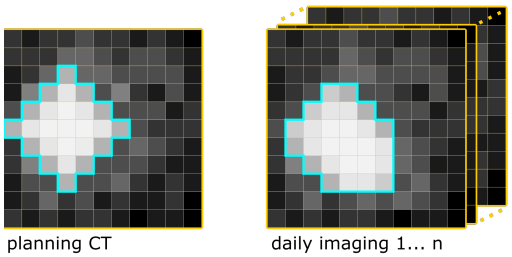
Figure 2.4. – HLUT used for treatment planning at the German Cancer Research Center (DKFZ) to convert from HU to relative electron density. Courtesy of Mona Splinter.

In addition to the kernel-based superposition approach for dose calculation, common approaches are for example given by a similar kernel-based approach based on convolution with an initial interaction function (i.e. the pencil-beam approach); the direct experimental measurement of dose using phantoms and dedicated detectors; or using Monte-Carlo simulations, thus a stochastic approach to simulate the microscopic interactions. While the latter provides the highest accuracy, kernel-based approaches offer a much faster solution which is especially important for treatment planning in IMRT and in treatment adaptation. For many applications in photon RT, kernel-based dose calculation offers sufficient accuracy.

Margins Limitations in imaging and dose delivery require the concept of margins for the target volume [35] [43]. First, the *Gross Tumour Volume (GTV)* is delineated by the clinician on the basis of imaging. For this, magnetic resonance imaging (MRI) is more and more frequently used due to the superior soft tissue contrast in comparison with CT imaging. A margin is applied to account for the microscopic spread of the tumour to surrounding tissue, which is not yet visible in imaging, forming the *Clinical Target Volume (CTV)*. To ensure tumour dose coverage, which is the foremost clinical objective, another margin is added around the CTV to account for uncertainties in the dose delivery. The thus created *Planning Target Volume (PTV)* can therefore even enclose volumes of normal tissue. The extent of these margins

is subject to the treatment optimization conducted by the treating physician and medical physicist and is usually based on dose-volume constraints (see also 2.2.2.2). The steep dose gradients in IMRT and TomoTherapy increase the importance of these margins, as small shifts in the dose field can result in severe underdosage of the tumour. Margins are therefore not necessarily smaller, but can be reduced in critical cases of nearby OARs.

2.1.3 *Techniques of image-guidance in radiotherapy*



3D imaging is the basis for RT as it enables the optimization of the 3D dose field according to the patient's anatomy. CT imaging is up to now the standard method for image-based treatment planning (section 2.1.2). Though CT imaging does not directly provide the information on electron densities, its physical principles are closely related to the electron density itself and therefore enable the easy calibration method using HLUTs.
The basic principle of CT imaging is based on two-dimensional (2D) projections of a photon beam. 360° rotational projections in a slice-wise manner (i.e. spiral) around the to be imaged object then enable the (iterative) backprojection resulting in the 3D image (in HU, section 2.1.1). Current standard CT scanners are constructed with a single photon source creating a fan-shaped beam which is opposed to a ring of detectors. The photon source is rotated around the object in a spiral shape so that 2D projections of the individual slices are projected onto the detectors. The most commonly used energy range is that of 120keVp, thus, a photon spectrum with a peak energy of 120keV. Other approaches utilize two separate energies (e.g. 80keVp and 140keVp) which enables the computation of the electron density and the effective atomic number of the material based on the dependency of μ on the photon energy. However, for standard photon RT, the HLUT-approach enables sufficient accuracy and more sophisticated approaches are only needed for RT with protons or heavy ions.
CT imaging also shows a number of limitations, with the two most important ones being the limited soft tissue contrast and the dose deposited in the patient. Both can be eliminated by the use of MRI, which is especially interesting for daily image-guidance. However, in contrast to CT imaging, imaging signals in MRI do not have a physical relationship to the required properties necessary for dose calculation (i.e. electron density). Approaches to compute synthetic CTs from MRI are therefore a prominent topic of current research, especially in MR-guided RT.
Image-guidance is entering clinical routine. This means that daily images of the patient are acquired prior to or during treatment in order to guide the treatment

process. This can be performed in a variety of manners and using different imaging modalities. Most commonly, image guidance is used for setup corrections. This means that an image is acquired prior to treatment and in treatment position which can then be checked to make sure that patient (or the target itself) is positioned correctly. Especially setup-errors but also to a large degree the motion of organs internally can impair the accurate delivery of the dose plan. The setup-correction can be based on different control points, for example: markers on the patient skin, bone position or direct location of the target. Since relative motion between the target and markers or bones can occur, the direct target location offers the best possibility to ensure correct target coverage. However, this is especially problematic using the low soft-tissue contrast of CT imaging. In case strong organ motion is detected, sometimes replanning is performed throughout treatment under consideration of a simple estimation of the summation of the planned dose, without dose accumulation of the actually delivered dose.

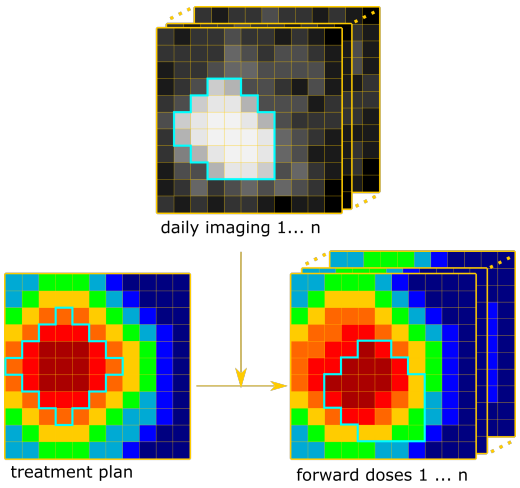


Image guidance using photon-based imaging can generally be separated in two categories: kV- and MV-based. Daily CT imaging for the presented work was performed with an in-room CT scanner on rails that enabled the acquisition of standard kV-CT scans on the treatment couch. However, often a CT or an MRI is not available in the treatment room. Image-guidance is therefore mostly based on alternative CT solution which are available with the respective treatment setup. Cone beam computed tomography (CBCT) imaging, another example for kV-CT imaging, is a standard technique used for image-guidance. In contrast to the classic CT, CBCTs only use a flat detector array opposing the beam source, which is being rotated around the patient for about 180°. The beam shape is that of a cone (hence the name). This approach results in a fast and dose reduced solution of image-guidance, but shows strong limitations in imaging contrast, accuracy and the size of the available field of view. In TomoTherpay, as described above, the treatment delivery is similar to the acquisition of a CT scan, but using different photon energies and fluences. However, the same setup can be used to acquire MV-based images (mega voltage computed tomography (MVCT)) of the patient in treatment position. Based on the mainly Compton-scattering-based interaction of the photons with matter in the MV-energy-range (figure 2.1), the signal and contrast to noise ratio is very limited for MVCT-imaging and shows even lower soft-tissue contrast than standard kV-CT. Furthermore, patients receive higher dose from imaging so that the resolution as well as the number of slices is strongly limited in clinical practice.
The daily acquisition of images of the patients anatomy is being studied for the use

in different applications in current research. Besides already commonly used direct application for setup or organ motion corrections, the aim of using daily imaging is to estimate the daily delivered dose. On the basis of repeated CT imaging, the dose estimation is based on *dose forward calculation* to compute the delivered dose in the anatomy of the day. For this, the parameters of the treatment plan that the patient was irradiated with (beam directions, intensities etc.), are used to perform the dose calculation on the basis of the image of the day, instead of the initial planning CT. Depending on the used imaging modality (CT, CBCT, MVCT), the conversion of HUs to electron densities is different based on the physical differences in the attenuation coefficient described above. With the information of the daily delivered dose, the accumulation of the dose can be performed. This information can be used to estimate the amount of dose that has been delivered to the patient on a voxel-basis[a], details of this will be given in the next section 2.2.1. Applications of this are for example treatment adaptation - changing the treatment scheme or treatment plan if necessary to ensure target coverage and normal tissue sparing; planning of re-treatments - in case of an RT treatment is applied after a previous treatment has already been finished, the dose already delivered to critical structures needs to be taken into account; and to find correlations of the delivered dose to treatment outcome - either for tumour control or for toxicities induced in the healthy tissue. The applications of the daily delivered doses are discussed in more detail with the application of dose accumulation in RT in section 2.2.1.1.
This work further focusses on dose accumulation on the basis of daily imaging in RT with photons. The quantity of interest is the delivered dose of the day that is being transformed to a reference anatomy for the purpose of accumulation of the local dose in a common frame. The next section will therefore present the details on the background of specific methods and models used in this work.

[a]voxel: 3D-volume in a regular grid, equivalent to a pixel in two dimensions

2.2 METHODOLOGICAL BACKGROUND

Embedded in the general context of image-guided RT presented in the previous sections, this work focusses on two specific aspects: dose accumulation and radiobiology. The relevant state-of-the-art in both is presented in this section. The here shown details will be used in the methodological development of this work in section 3.1.

2.2.1 *Dose accumulation*

This section describes the details of the currently used conventional dose accumulation and the standard workflow. The result of dose accumulation is the voxel-wise summation of the delivered doses for all fractions of the treatment. Multiple steps are necessary in this process that will be presented and discussed in this section. The motivation and conclusion made for the newly gained information, is presented in section 2.2.1.1 including current literature on the matter. The commonly used standard workflow that is also used within this work, is explained in section 2.2.1.2, followed by sections describing the individual steps and applications: image registration (2.2.1.3), interpolation and mapping (2.2.1.4) and accumulation (2.2.1.5).

2.2.1.1 *Motivation and usage of dose accumulation*

The impact of organ motion on delivered dose Organ motion, setup uncertainties and change in the tumour volume can cause deviations between planned and delivered dose. Especially in precise treatment techniques such as IMRT (section 2.1.2), the steep dose gradients within the normal tissue will most likely cause daily variations of local voxel doses, even if organ motion is at a minimum.
With the increasing use of daily imaging, more studies emerge, that report on the differences between the planned and delivered doses. A trend that was observed throughout literature was a generally good tumour coverage despite interfractional organ motion, indicating sufficiently large treatment margins and recommending the use of image-guidance (section 2.1.3). On the other hand, significant differences in the dose to the OARs of bladder and rectum were reported. Kupelian et al. [41] performed an early study in 2006 on the delivered dose in prostate TomoTherapy using daily MVCT imaging. They found substantial differences in the doses to bladder and rectum with V_{2Gy} (volume receiving at least 2 Gy) for the rectum varying from 0.1 to 67.3 cc and for bladder from 0.3 to 36.8 cc, while the $D_{95\%}$ (dose to 95% of the structure's volume, see section 3.2.3) for the prostate only varied between 1.79 and 2.2 Gy. This was also found by Hatton et al. [28] who performed twice weekly CBCT scans for comparison of the planned to the delivered dose-volume parameters. While CTV coverage shown by the $D_{95\%}$ was only marginally impaired, they found statistically significant difference for the volumes receiving 40 Gy, 60 Gy and

70 Gy in both bladder and rectum for the majority of observed cases. Godley et al. [26] similarly found only a maximum of around 3% deviation of the planned to delivered $D_{95\%}$ for the CTV in prostate RT while the V_{70Gy} of bladder and rectum was found to be up to twice the value of the planned dose distribution. Nassef et al. [56] measured the mean dose standard deviation to the bladder to be 6.9 Gy (maximum 18.1 Gy) and rectum to be 2.0 Gy (maximum 4.2 Gy) in prostate IMRT. Dose differences along various dose-volume measures ($V_{50Gy} - V_{75Gy}$) varied from -10% to +7% for the rectum and -4% to +26% for the bladder in 75% of cases.
It should generally be noted that for all these reports, the delivered dose was either higher or lower than the planned dose without indications on a systematic trend. A systematic trend was, however, observed by Shelley et al. [70] who found that the equivalent uniform dose (EUD) (section 2.2.2.2) was systematically lower for delivered than planned dose to the rectum in prostate TomoTherapy. They concluded that observed toxicity is therefore linked to smaller than expected dose levels of the rectum. The dataset used in their analysis comprises a discovery cohort of patients of the VoxTox programme. Parts of the used data in this thesis were taken from the same programme, comprising a later acquired cohort with the same treatment objectives, described in the materials and methods sections 3.2.2 and 3.2.4. Scaife et al. [65] reported on data from the same study and also found that daily variation of the rectal dose were considerable even though the median position of the structure was close to the planning scan position.
For head and neck RT, Nobnop et al. [61] found that parotid doses increased over the course of therapy by more than 10% for 3/10 patients (range 13-43%) based on volume shrinkage. Dose to the spinal cord was only marginally increased. Here again, there was very little impact on the target dose.

Applications of dose accumulation Dose accumulation is now entering clinical practice to enable different new workflows and investigations. First, the knowledge of the actually delivered dose allows for **treatment adaptation** by adapting the dose plan during the course of therapy, in addition to simple setup corrections. In case of overdosage of the normal tissue, it might be necessary to change the plan in order to spare certain areas even more. On the other hand, lower than expected dose in normal tissues might allow for dose boosts of the tumour volume. This might also be necessary in case of underdosage of the tumour, though as described above, good tumour coverage was reported in the majority of publications on delivered dose. The aim of treatment adaptation is to further improve the clinical outcome of the treatment in terms of both tumour control and sparing of healthy tissue. It is therefore especially important to keep in mind the biological impact of dose on the irradiated tissue even though it is the quantity of the delivered dose itself that is usually the focus of observations. The rationale for this is tackled in the main developed method in this work presented in section 3.1.
While treatment adaptation requires the knowledge of the dose of the already delivered fractions during the treatment course for decisions on the next fraction, the total

delivered dose becomes important whenever patients will receive a **re-irradiation** after a previously completed treatment. Boman et al. [5] investigated in a recent study the delivered dose for purpose of re-treatment. They compared the simple summation of dose based on rigid registration of the images (image registration techniques see section 2.2.1.3) with a biological summed dose on a simple EQD_2-based rescaling approach (equivalent dose in 2 Gy fractions (EQD_2), see section 2.2.2.2) and found differences ranging from -14 Gy to +5 Gy. Here again, the dose already received by the normal tissue is of high importance for the setup of a new treatment plan in order to reduce the risk for side effects.
Another field of application of dose accumulation is the estimation of the delivered dose *during* treatment with the use of time-resolved imaging. For example Dolde et al. [19] recently published a study on dose calculation based on four-dimensional (4D)-MRI (3D + time) for the study of the interplay effect in gated treatment of pancreatic cancer. This ***intrafractional* dose accumulation** is not investigated in this work. The applicability of the here developed methods (section 3.1) for the use in intrafractional dose accumulation will, however, be discussed in chapter 5. Its application is of special importance in lung cancer treatment as well as for strategies such as gating, investigated in the mentioned publication.
Finally, in order to further improve the understanding of the dose-response relationship in RT, the accumulated delivered dose is being studied for **correlation of dose to outcome**. Due to the yet limited number of completed trials that include daily (or weekly) imaging of the patient and detailed record of the treatment outcome, the study of delivered dose for this application is yet at an early stage. The Quantitative Analyses of Normal Tissue Effects in the Clinic (QUANTEC) reports from 2010 (specifically Marks et al. [47] and Jaffray et al. [31]) underline the lack of studies relating delivered dose to outcome for the quality of the reported dose to toxicity correlations and state that more accurate predictions *"[...] require models that relate the true accumulated dose to clinical outcome"* [31]. First publications already predict better dose-to-outcome correlations when using the accumulated delivered instead of the planned dose. Since dose accumulation facilitates the analysis of the delivered dose on a voxel-scale, there is a number of studies that report on local dose deposition investigated with the use of dose surface maps (DSMs) (more information on this technique also exploited in parts of this thesis will be given in section 3.2.4). Shelley et al. [70] investigated DSMs of the rectal wall and found stronger correlations of the delivered dose to rectum toxicity than for planned dose. This dataset including the used toxicity data will also be analysed in this work. Wortel et al. [81] found that investigated side effects in prostate RT showed significantly different dose distributions on delivered DSMs in patients, and they report doses with differences up to 10 Gy to patients without reported toxicity. Similar results on DSMs for the bladder were reported by Palorini et al. [63].
The various steps involved in the process of dose accumulation (that will be described in the next sections) also add to the uncertainty budget of the process. Therefore, a fundamental question in the rationale of dose accumulation is whether the newly

gained information is still advantageous when taking into account all uncertainties, compared with the simple way of only considering the planned dose. This will be discussed in more detail in section 5.3 on the basis of the results obtained in this work.

2.2.1.2 *Workflow overview*

The standard workflow for conventional dose accumulation, which is also used in this work, is schematically shown in figure 2.5. Section 3.2.3 describes how this workflow is implemented.
In the planning stage (figure 2.5 **a**), the patient receives a treatment planning CT on which dose calculation and plan optimization is based. The physics of dose calculation on the CT image is explained in more detail in the background section 2.1.2. The dose plan is then created, based on the prescription dose to the tumour as well as on dose limiting constraints for OARs. The patient will then receive the dose in a course of daily repeated treatment fractions (number of fractions: n), each delivering the same dose plan. Image-guidance is nowadays used routinely for setup corrections or even translational corrections based on the target position. More information on image-guidance is given in section 2.1.3.
For each of the delivered fractions, the workflow chain required for the ultimate dose accumulation is performed (figure 2.5 **b**). In total, four main components are part of the dose accumulation workflow before the actual summation of the dose can be performed: repeated imaging (section 2.1.3), dose forward calculation (section 2.1.3), image registration (section 2.2.1.3) and dose mapping (section 2.2.1.4). Based on this combined information, the daily delivered doses d_i (where $i = 1...n$ is the fraction number index) are given on a voxel scale and can be summed up to give the total (conventionally) accumulated dose D_a (section 2.2.1.5). Adaptation of the treatment might be performed based on the information obtained throughout the course of therapy given by the respective accumulated dose up to then. As described above, the total accumulated dose can later be used in case of re-treatment or correlation of dose to the treatment outcome.

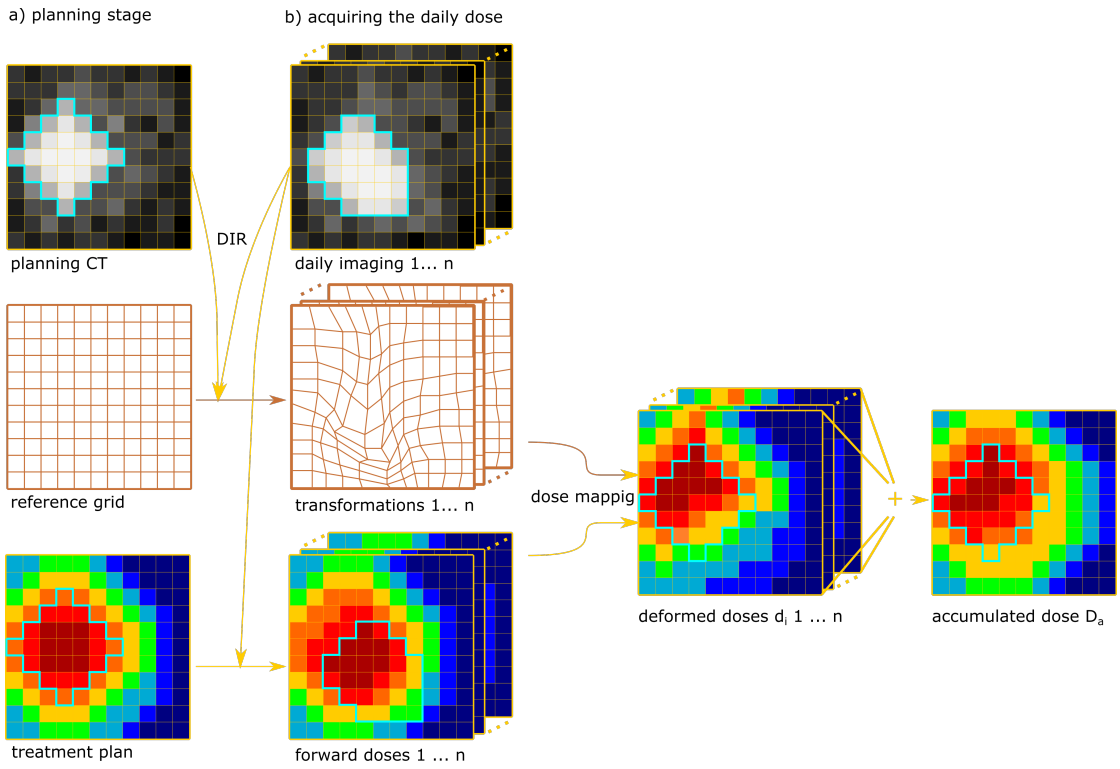


Figure 2.5. – The standard workflow for conventional dose accumulation from daily imaging and deformable image registration (DIR). In the planning stage (**a**), the patient receives a planning CT on which the treatment plan is based. In stage **b**, the daily dose is acquired for all n treatment fractions: daily patient imaging, DIR (to register the daily image to the planning CT), and dose forward calculation on the daily image based on the treatment plan. Transformations from DIR and the forward doses are then used to derive the deformed doses $\mathbf{d_i}$ in the reference frame which are then accumulated to the total dose $\mathbf{D_a}$.

2.2.1.3 *Deformable image registration*

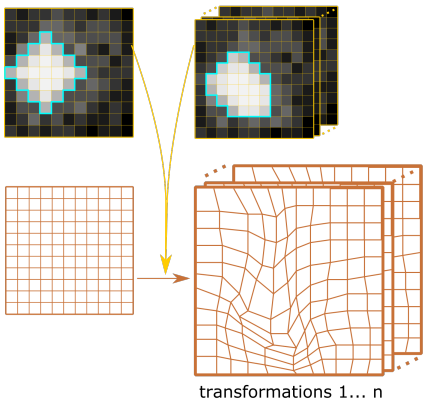


Image registration is the process of finding the geometric transformation that describes (via a deformation vector field (DVF)) the translation of grid center points in the moving image (to be deformed) to match the same ones in the reference frame of the fixed image. Thus, the information obtained from image registration is the DVF that results in (ideally) identical positions of anatomical points in the two registered images. Registration is performed between patient images obtained at different time points here. A good overview of image registration background and application in image registration can be found in the 2017 review by Brock et al. [11].

A registration works on a uniform discretized voxel grid of the image. In common application, the grid center points of the voxels are registered. The voxel value (from imaging or dose) is thus given as a point measure assigned to the volume of the voxel. Voxelization therefore is a spatial average of the volume values from imaging or dose (a discussion on this principle is given in 5.2).

Figure 2.6 illustrates an example registration between two images as the DVF from the regular fixed image to the new location of these points in the moving image. The deformed grid (that is also shown in the overview illustrations throughout these chapters) serves the purpose of visualization of the deformation, however, it should be noted that it is the DVF and not the deformed grid that is the result of deformable image registration (DIR) and the basis for interpolation explained in the next section.

Three types of registration are being distinguished. (A) Rigid registration - refers to a tranformation that is solely represented by translation and rotation while preserving the distance between all points. It can therefore only represent translations and rotations of the entire patient geometry without reflecting internal organ motion and deformations. (B) Affine registration - in addition to rigid registration, affine registration includes scaling and shearing which preserves parallels lines in the grid but not the distance between points. (C) Deformable registration - assigns individual displacement vectors to each voxel (center point) that can move in all three dimensions. Point distances are therefore not preserved. For the following, this work only focusses on DIR which is necessary to map organ motion and deformations. This information is crucial to enable the voxel-wise accumulation of dose from each treatment fraction.

A registration algorithm contains **three main components**: A similarity metric, a transformation model and an optimization algorithm. The details of this are not used or discussed in this work, but in the following, a few aspects are summarized

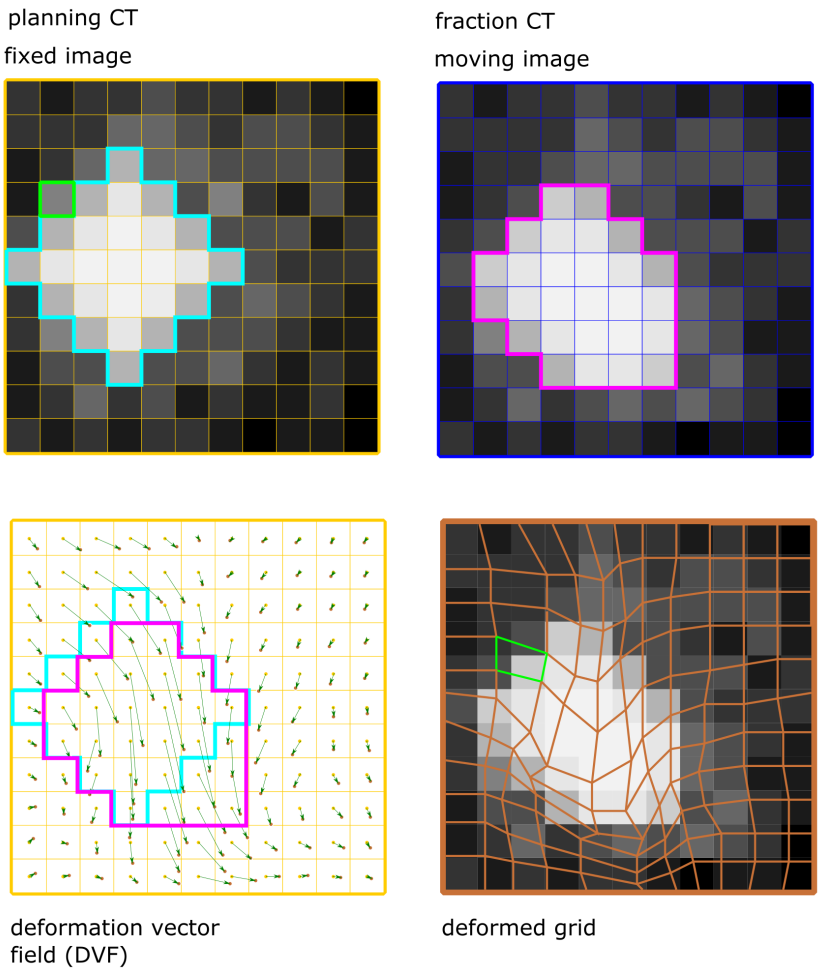


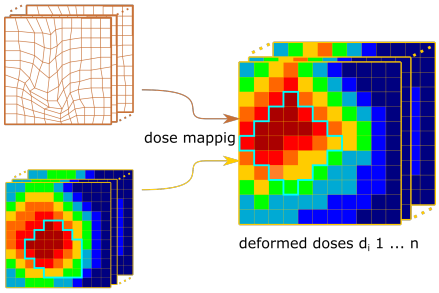
Figure 2.6. – Visualization of DIR between the fixed image (e.g. planning computed tomography (CT)) and the moving image (e.g. fraction CT). The DVF is the result of the registration and points from the regular reference grid of the fixed image to the new location of the respective points after registration. The deformed grid is an illustration of this deformation but does not represent the information that the algorithm is generating and only the serves the purpose of deformation illustration. The depicted deformation is a rather strong example. The green enclosed voxel will be viewed in detail in figure 2.7 when explaining the interpolation step for value mapping.

for further understanding of methods and results:
One of the main components of a registration algorithm is the **similarity metric**, a measure for the similarity between the two images. The registration algorithm tries to maximize or minimize this metric in order to best map the two source images (optimization step). A metric is either geometry- or intensity-based. A geometry-based metric tries to match positions of specific points identified on both images (point matching) or maximize the overlap of two structure contours (surface matching). Manually setting landmarks or contours is labour-intensive but ensures that deformations are anatomically realistic. Furthermore, this technique is more robust to imaging noise and using different modalities. Deformation in landmark-free regions are extrapolated making these regions error prone. Furthermore the manual interaction of setting landmarks is subject to uncertainties. Biomechanical modelling would be required in landmark-free regions which is not commonly implemented in currently used software. On the other hand, intensity-based metrics measure the similarity between the two images based on the grayscale voxel value information. Thus, no manual work in setting landmarks is required but there is also no guarantee that the deformation is anatomically realistic. For the sake of completion but without further details, commonly the following intensity-based similarity metrics are used: sum of squared differences (SSD), mean squared difference (MSD), correlation coefficient (CC) or mutual information (MI). In short, all of these metrics assume a relation between the voxel values in both images, from equivalent (SSD, MSD), over linear (CC) to statistical (MI). Improving the outcome of intensity-based algorithms for example requires quality control of the images, e.g. by intensity corrections.
Usually, applications are either intensity- or geometry-based, but some solutions combine both approaches as the here used ANACONDA algorithm ([77], section 3.2.3).
The minimization or maximization of the metric is performed iteratively by an **optimizer**. Commonly, a registration contains several stages, each of which is iteratively optimized and going from low to fine voxel resolutions. In addition, the optimizing algorithm is often given constraints and regularizations to ensure that the resulting transformation is anatomically realistic. Smoothness constraints are often utilized to avoid high frequencies, vortices or singularities in the deformation. This, however, can wrongly represent sliding organ surfaces.
In this work, different **transformations** are used in the DIR process that are being compared in their results for dose accumulation. More information on the used software is given in section 3.2.3 where the above mentioned components of the used registration algorithms are described shortly. Additionally to clinically applied standard software, two transformations were performed based on the *demons* and *b-spline* algorithm, respectively. The b-spline algorithm is point-based and uses polynomials (bases functions) to connect weighted control points in the grid. This description offers a smooth deformation solution. The demons algorithm works on intensity gradients, optimization is here based on a gradient descent method and

additional smoothing is achieved via a Gaussian smoothing approach.
Deformable image registration is necessary whenever the voxel-wise dose information of the day needs to be mapped to the reference frame in order to enable dose accumulation and its further applications described above. Another application of DIR is the propagation of contours to reduce the labour intense tasks of contouring in clinics. Furthermore, DIR is used to create synthetic CT images from other modalities such as magnetic resonance imaging (MRI) or cone beam computed tomography (CBCT). An approach to do this is for example described by Kraus et al. [40], where the planning CT is deformably registered to daily MRI scans to create the electron density map of the day for dose calculation (see section 2.1.3).
DIR is implemented to register two images with essentially the same geometry but with deformations. However, it is often the case that organ motion also includes **volume changes** (i.e. growth or shrinkage of a structure). For example by tumour growth or shrinkage, bladder or rectal filling changes, lung motion etc. In these cases new material is added to the system or material leaves the system. The problem is then ill-posed for the registration algorithm. While the algorithm might eventually come up with a geometric solution that can be satisfactory, its underlying mechanism, especially when it comes to the biological basis, needs to be addressed and implemented in a way to be in accordance with biology. This topic will be handled in detail in section 3.1.1 for structures undergoing volume changes based on different causes.
Volume changes will be visible in the deformed grid or DVF as areas of higher grid point density (shrinkage) or lower grid point density (expansion). A measure for the voxel-wise change in volume, based on the DVF is the determinant of the Jacobian matrix of the DVF. This parameter can also be used for quality assurance of the DIR. A Jacobian larger than 1 refers to volume growth, while a number below 1 refers to shrinkage. Negative values indicate areas where the registration is folding in on itself and is therefore physically and anatomically unrealistic. While this parameter can be an indicator for registration quality, it is no guarantee that the transformation is correct.
Quality assurance (QA) of DIR is a topic of ongoing research and debate. While other metrics exist in addition to the Jacobian determinant, this field is highly limited by the missing knowledge of the underlying ground truth of the voxel-wise deformation in a patient setup. Qualitative registration QA is recommended and usually performed by clinicians on the deformed data set. Quantitative approaches, for example, measure the alignment of specific points in the images. For this, the target registration error (the average residual between the points) can be calculated. There will, however, always be an underlying uncertainty in the identification of the two respective points and this metric has only limited informative value for areas in between points. Similarly, organ contours can serve for purpose of control of the registration quality. The Dice similarity coefficient or the mean distance to agreement can be used to quantify the amount of overlap of the structures. It should be noted that whenever one of these metrics is used to guide the registration process (in for

example geometry-based algorithms), it is not reasonable to use the same metric for purpose of QA.
Nevertheless, a **quantification of the registration uncertainty** for the individual voxels is not possible with either of these metrics. There are several approaches proposed in literature for quantification of uncertainties that require model-based approaches of error sampling. This will be discussed in section 5.2 along with the obtained results on the accuracy of registration and accumulation.
The registration and resulting DVF are only the basis to create the deformed (mapped) images. In addition, strategies are required to transform one image to another based on the information from the DVF. For this, interpolation is necessary which is being addressed in the next section.

2.2.1.4 *Interpolation and mapping*



Transformation and interpolation are essentially two parts of the registration process. The aspects described in this section are therefore still a part of the overall DIR process to derive the final deformed image.
Registration provides the information on where points in the grid have moved to at another time point. Mapping extracts the new voxels values in the fixed frame based on the information of the location in the moving frame. In the deformed grid, the displaced grid center points do not necessarily directly match voxel center points of the regular uniform grid of the undeformed original (moving) image. Therefore, interpolation is required to obtain the value of the respective grid center points (i.e. voxels).
The result of the registration is a DVF that contains the information where points from the fixed image have moved to in the moving image. The deformed grid (as a visualization) therefore shows how the information from the first to the second timepoint has taken place. Interpolation is used to derive the values of the individual voxels in the reference frame at this second time point. Different interpolation strategies are proposed in literature which will be discussed below as well as in section 5.2. Once the point value is computed, it is "pulled-back" to the regular grid of the reference frame, i.e. the value is assigned to the respective voxel (center point) that it is linked to via the DVF. Figure 2.7 illustrates this process as a follow up on registration shown in figure 2.6.
This method is commonly referred to as "value-pulling". There are essentially two directions for interpolation, thus "value pulling" or "value pushing". The second one requires computations on the basis of the unregular deformed grid which is much less time and memory efficient than the first method. "Value pulling" is used to deform the image to a common reference frame (i.e. the planning CT). Commonly, "value pushing" is used when the information from the planning scan (grey values

or dose) is transformed into the anatomy of the day (e.g. generation of synthetic CTs). An example of the application of both methods in different aspects of the same workflow is given by Glitzner et al. [25]. In this work, only the first option is required to compute the daily deformed doses in the same frame of reference.

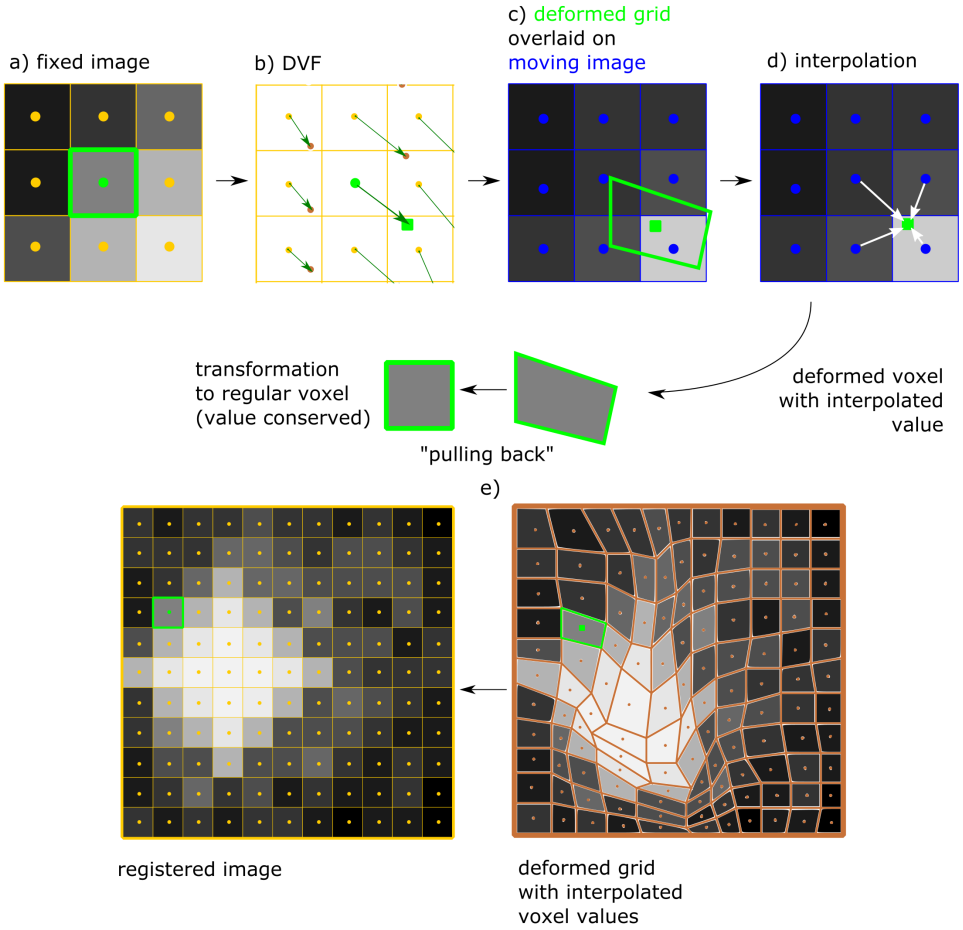
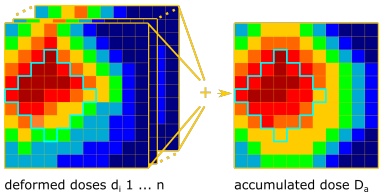


Figure 2.7. – Illustration of the mapping process on the basis of DIR shown in figure 2.6. The regular grid (a) of the fixed image is linked to the moving image via the DVF (b). c) deformed grid voxel overlayed on moving image grid (for visualization). d) Interpolation computes the value of the mapped voxel as a weighted average of the neighbouring voxel values in the moving image which is then assigned to the volume and location of the voxel it is linked to in the fixed image ("value pulling") (e). Deformation grid only for visualization: Size and shape of the deformed voxel will be backprojected to the regular voxel shape (e), resulting in a volume shrinkage in this case, while the value of the voxel is conserved in the process. Detailed analysis in section 3.1.1.

A standard method for interpolation that is also used in this work is **trilinear interpolation**. For this, the nearest neighbouring voxels in the grid of the undeformed moving image are identified to retrieve the value of the respective points in the deformed grid (see figure 2.7). The point value is calculated as a weighted mean of the values of the neighbour points where weighting is based on the distance to the point of interest. It can be shown mathematically that this process is essentially

a volume-weighted average of the regular voxels of the undeformed grid that contribute to the deformed voxel. In figure 2.7 c), in the deformed voxel overlayed onto the gray values of the moving image (highlighted in green), averaging of the gray values included in this voxel can be interpreted accordingly.
Rosu et al. [64] have shown that for areas of large volume changes, this basic trilinear interpolation method is not sufficient to map the volumes correctly since voxel values might be neglected that still contribute to the voxel of interest. They base their analysis mainly on lung tissue, where this problem has increased importance. In their solution, they propose sub-voxelization of the grids in order to push the problem back to scenarios with less strong volume changes. Before mapping, the voxel are sub-divided into smaller voxels that are being used for interpolation. The interpolation of these sub-voxels in then again the standard trilinear interpolation. After interpolation, the corresponding sub-voxel are merged again to respective larger voxel. According to Rosu et al., this method can impact the outcome of the interpolation mainly in dose gradient regions. This method is also shortly studied in this work.
Interpolation can be performed on the gray values of the image (in Hounsfield Units (HU)) to obtain the deformed anatomical representation, but also based on the voxel-wise dose information. For the following, the latter will be referred to as *dose mapping* (also known as dose warping or dose deformation). Dose mapping is essential for dose accumulation to compute the daily deformed doses.
The mapping of dose can either be based on the dose value or the deposited energy. Within this work, the used implementations are dose-based. The reasoning and implications of this will be discussed later in section 5.2. Moreover, dose mapping can be performed in a way to conserve the voxel value of the irregularly deformed voxel when being pulled back to the reference grid (which may include a change in volume) as depicted in figure 2.7 e). It can also be implemented in such a way as to scale the voxel value according to the change in volume. The question of how the interpolation should be implemented comes back to the biological basis of the deformation, especially when volumes are changing. This is a field of ongoing debate which will be tackled in this work in detail in section 3.1.1 by supplementing the debate with the underlying biology.

2.2.1.5 *Accumulation*



Once the daily deformed doses d_i are obtained from DIR (transformation + interpolation, i.e. mapping), the calculation of the conventionally accumulated dose D_a is given by the summation over d_i:

$$D_a = \sum_{i=1}^{n} d_i \tag{2.9}$$

with i denoting the fraction number and n the total number of fractions. This is possible since all daily dose distributions are based on the same reference grid through

registration and voxels ideally link to the same locations in the patient's anatomy. However, uncertainties in the dose mapping process will naturally be forwarded to the accumulated dose. From error propagation, the (voxel-wise) uncertainty of D_a with respect to the individual errors in d_i (which are not related) is given by:

$$\Delta D_a(\Delta d_i) = \sqrt{\sum_{i=1}^{n} (\Delta d_i)^2} \tag{2.10}$$

2.2.2 *Radiobiology*

For the success of a radiotherapy (RT) treatment, the fractionation concept needs to be based on the biological response to radiation. Specifically, the interplay between the total dose and dose per fraction for tumour as well as for normal tissue. Since the early beginnings of RT, various models to describe the biological radiation response mathematically have been proposed. An overview of these models can for example be found in the 2016 review by Bodgi et al. [4] and in the 2019 review by McMahon [49].
Radiation response is usually reported as the (cell) survival fraction (SF) based on experimental observations. For these experimental approaches, the quantification of the SF with respect to dose is assessed in clonogenic cell assays. A cell is considered "dead" when it looses its ability to form colonies (therefore, the cell should rather be called "sterile" than dead in strict terms). In a simple clonogenic assay setup, an initial number of cells N(0) is irradiated with a dose d. The number of colonies that are formed after irradiation gives the number of initial cells that have not been killed (i.e. sterilized) by the radiation, N(d). The cell survival fraction is thus given by $SF(d) = N(d)/N(0)$. It was generally observed that the SF follows an exponential decrease with dose while details of this decrease are described by different approaches. The most common and robust approach is the descripton of the SF by a linear-quadratic relationship to dose within the exponential, the Linear Quadratic Model (LQM). As the basic model used up to date and within this work, it is explained in more detail in section 2.2.2.1 below. In the following, the development and hypothesis behind other biological models described in literature are shortly described to be related to the LQM.

The **target theory** as a concept was already hypothesised in the early radiobiological experiments in the 1920s. The basis of this theory is that exposure to radiation is essentially seen as hitting the cells (targets) with random projectiles. The probability density function to hit the target was described by a Poisson distribution, with the number of hits being directly proportional to the applied dose. In order for radiation induced cell death to occur, it was hypothesised that *"[...] it is necessary that its [the cell's] sensitive zone absorbs a minimal number s of quantas"* by Curie [16] (referred to as *quantum radiobiology*), so that cell kill was described by a simple exponential function: $SF(D) = e^{-D/D_0}$ (with D_0 being the mean dose that leads to

one hit per target). However, advances in experimental radiobiology later revealed a shouldered exponential response curve. The target theory was therefore further adapted through the following decades from the initial single-hit, single-target model to a single-hit, multi-target model. In this approach, a certain number of targets m within one cell need to be hit in order to induce cell kill [20]:

$$SF(D) = 1 - (1 - e^{-D/D_0})^m \tag{2.11}$$

Cell survival is therefore linked to a lack of hits in the sensitive region. Nevertheless, experimental observations showed discrepancies to the curves provided by the above model, especially in low and high dose regions.
Follwing the target theory, the LQM was introduced by Kellerer and Rossi in 1972 [36]. The description via a linear and a quadratic term linked to dose showed much better agreement to the observed SF curves.
Opposing one basic assumption of the target theory is the actual number of hits that cells receive during irradiation: For a dose of 1 Gy in photon therapy, an estimated 40 double strand breaks (DBS) in the DNA occur, accompanied by around 1000 single strand breaks (SSB) and 10 000 base damages. Thus, the cell must be capable of **DNA repair** that results in radioresistance and subsequent survival rather than the simple approach of a lack of hits. Therefore, later approaches in the 1980s tried to include the repair capability and different radiosensitivities of cells in the concept and interpretation of the LQM - for example the Repair–MisRepair-, the Lethal-Potentially-Lethal-, and the saturated-repair-models. However, there is a number of contradictory findings against these models including the nature of DNA repair kinetics and consequences in cells as well as the basics of direct and indirect radiation damage to DNA.

2.2.2.1 *The linear quadratic model and its variations*

"The Linear-Quadratic (LQ) model, which is to date the most frequently used model in radiobiology and RT, dominates all the other models by its robustness and simplicity." [4]

The linear-quadratic approach first proposed by Kellerer and Rossi [36] provides a good fit to experimental data for the commonly used dose regimes. Figure 2.8 shows a comparison of the above mentioned models based on the target theory with the LQM-based survival fraction.
The survival fraction for a single fraction of dose d is given by:

$$SF_{LQM}(d) = e^{(-\alpha d - \beta d^2)} \tag{2.12}$$

The parameters α and β represent tissue specific quantities. The shoulder region of the survival curve (on a semi-logarithmic scale) is represented by the linear dose

term and the according α value while the linear region is characterized by the β term.

For **consecutive fractions**, it was shown that fractions of equal dose d are isoeffective, so that the LQM for fractionated treatment is given by a power-to-n law [27]:

$$SF_{LQM}(d,n) = e^{(-\alpha d-\beta d^2)^n} = e^{(-\alpha n d-\beta n d^2)} = e^{(-\alpha D-\beta d D)} \tag{2.13}$$

with n being the number of fractions and $D = nd$ being the total treatment dose. This is based on the fact that after sufficient time (6-8 hours), DNA repair is completed and cells are either dead (repair was not possible) or can be treated as fully repaired cells without dose memory.

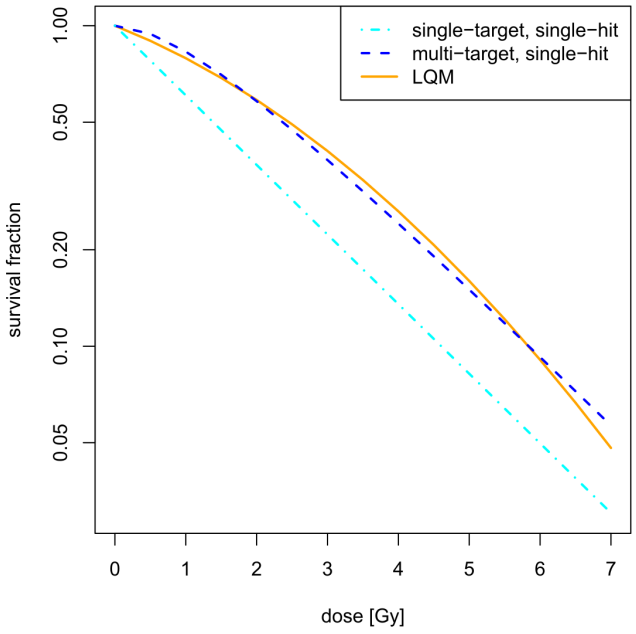


Figure 2.8. – SF curve comparison for the different described models. Exemplary values taken from [49]. For the single-target and the multi-target model, D_0 was set to 2 Gy. For the multi-target model, m was set to 2. For the LQM: $\alpha = 0.2$, $\alpha/\beta = 6$. It can be seen that the multi-target, single-hit model and the LQM are very similar for the chosen parameters, underlying that both models can provide good empirical fits to the data. However, in detailed analysis the multi-target model failed to represent the SF curve of experiments especially in the low and high dose regions.

While the LQM provides a simple application (due to the low number of variables) and good representation of experimental findings, it needs to be noted that it is still an empirical model and its **biologically based interpretation** remains unsolved. It was initially set up by Kellerer and Rossi based on the assumption of a two-level cell kill (or inactivation) process: either due to a single particle hit (α term) or coming from a first damage to a second one resulting in kill (β term). In later explanations, this was interpreted as a death following either an unrepairable double strand break, or the succession of single strand breaks. A mathematical explanation on the other hand follows the assumption that any type of response can in principle be modelled using polynomials from a Tailor expansion. The LQM in this sense only shows that the response of cells to ionizing radiation might be already adequately characterized by the linear and quadratic terms. With the nature of radiation response with respect to the LQM still unknown, its

use and interpretation needs to be treated with caution. This is especially valid for the underlying **tissue-dependent α and β parameters**. Conventionally, the α/β [Gy] value is used and reported for clinical applications instead of the individual parameters. Its determination is usually based on cell culture experiments of the above described clonogenic assays. This involves a number of uncertainties since the tissue structure and environment also plays a role in term of its radiosensitivity. The estimation of the α/β value in patient analysis on the other hand is endpoint-specific and includes a range of assumptions on the tissue and including uncertainties in the overall process. Therefore, reported α/β values often show a wide range, especially when derived from patient data (see also table 3.2 in section 3.2.3 for the applied ranges in this thesis). For example Williams et al. [78] investigated on the α/β value of prostate cancer. They found that for the 95% confidence interval, the results of α/β ranged from 3.7 Gy to infinity. By increasing the range of included dose per fraction schemes (including boost therapy) this interval shrunk to a range from 0.9 to 4.8. They concluded that the spatial and temporal variance of the delivered dose between patients that could not be accounted for, had a large impact on the resulting uncertainty range. It should be noted that also the patient variability (genetics and way-of-life background) impacts the radiosensitivity of tissue substantially. Thus, the α/β value for the same tissue type can vary between patients. Furthermore, organs can show substructures which can potentially have different α/β values which is not reflected in the common approach to assign one value for the whole organ. In summary, this means that there is possibly a natural range of α/β values for the same tissue type.
It should be noted that the underlying individual α and β values play an important role in the description of the LQM curve, as shown in figure 2.9 a). The impact on the results in this work are discussed in section 5.2.
Despite the missing biological explanations, the LQM provides a good basis for clinical applications and interpretations of fractionated treatment. **Fractionation** is based on the concept that tissue can be spared from damage when using consecutive low doses. While this is true for both the tumour and normal tissue, the difference in repair capability, which is higher in normal tissue, results in the effect that normal tissue benefits from fractionation to a larger extent than tumour tissue (dependent on the choice of the fractionation scheme as outlined below). The α/β dependency and the impact of the fractionation scheme is shown in figure 2.9 b).

Within the LQM, the α/β reflects the tissues response to radiation. Tissues with a low α/β (1-5 Gy) are known as late responding tissues due to a low proliferation rate (like kidneys, heart or the rectum). On the other hand, a larger α/β value refers to early or acute responding tissue (like skin and bone marrow). Figure 2.9 a) illustrates the LQM-based SF curves for different α/β values. It can be seen that the cell kill per unit dose is increased for lower α/β values based on the increased impact of the β term and its quadratic relation to dose. Based on this, figure 2.9 b) shows that tissues with low α/β values benefit more from fractionation with a high number of fractions and low dose per fraction because the high dose regime is

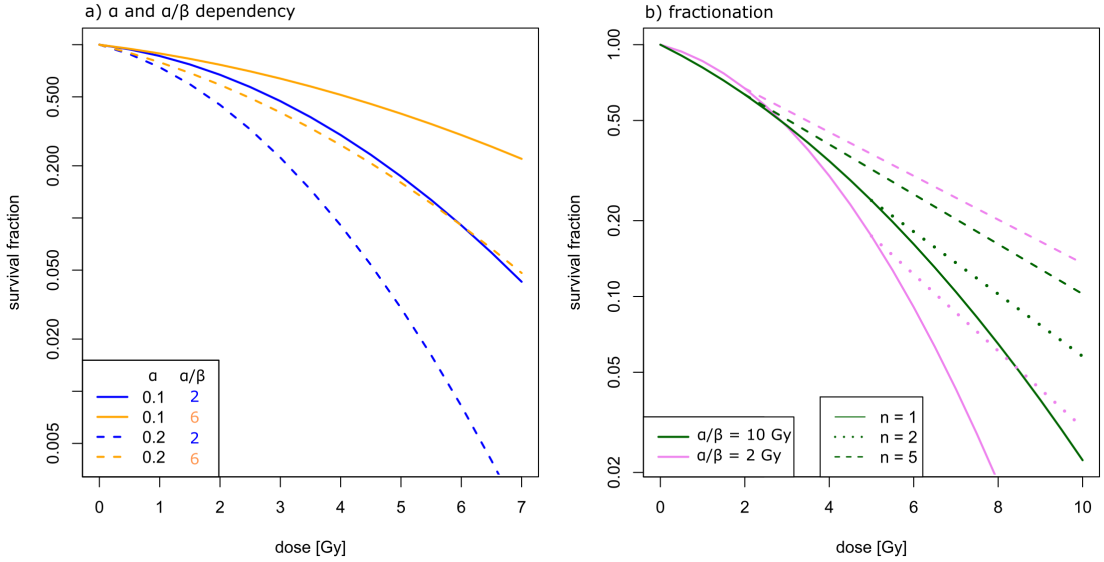


Figure 2.9. – The cell SF according to the LQM for a) different underlying α and α/β values and b) the fractionation dependency. a) shows that the cell kill per unit dose is increased for lower α/β (and same α). The underlying α value determines the bending of the curve (extent of the shoulder). b) Visualization of the fractionation principle (green: $\alpha = 0.1$, violet: $\alpha = 0.19$), n = number of consecutive fractions. For n=1 and n=2, the curve with higher α/β value shows a higher SF, showing that a low number of high dose fractions is advantageous for normal tissue with high α/β while achieving higher cell kill in tumours with lower α/β (principle of hypofractionation). However, in case the tumour (as conventionally expected) has a higher α/β values than the normal tissue, a higher number of low dose fractions is advantageous: for n = 5, the upper curve now is based on the lower α/β value, using the advantage of lower cell kill for lower doses due to the shoulder region, justifying standard fractionation schemes.

cut while for tissue with a high α/β, the SF curve is almost linear to the dose per fraction and therefore benefits less from decreasing the fraction dose.
It was conventionally assumed, that due to the interpretation of cancer being a volume of fast and uncontrolled cell proliferation, tumours have a high α/β value (commonly applied tumour $\alpha/\beta = 10$ Gy). On the other hand, slow proliferating normal tissues shows lower values (around 3 Gy). This is the basis for standard fractionation in usually 34-37 fraction of around 2 Gy that optimized normal tissue sparing based on its recovery and the total treatment time.
However, different strategies should be applied for cases where the α/β value of the tumour is lower than that of the surrounding normal tissue. In the case of prostate RT, recent works pointed out that the α/β value of prostate carcinoma might be low compared to the common standard (e.g.: [1], [10], [53], [78]). Miralbell et al. [53] for example suggest $\alpha/\beta = 1.4$ Gy. For these cases, hypofractionation is proposed, thus lowering the number of fractions while increasing the fraction dose. The beneficial impact of this can be seen in figure 2.9 b) where for higher doses per fraction, the tissue with higher α/β value (in this case the normal tissue) will be spared more (higher survival). Details on how to convert different fractionation schemes are given below in the next section.

The **recommended dose range** for using the LQM is between 1 to 5 Gy per fraction [34] while some other recommendation go up to 10 Gy. Experimental observations indicate that the LQM might not be appropriate outside of this range, underlining the missing biological interpretation of the model. For high dose regions, while the LQM shows continuous bending, experimental observation indicate a rather straight line. Different approaches are suggesting how to account for the observed discrepancies, each including at least one more tissue dependent parameter. The derivation of three instead of two parameters from the limited information in experimental setups (especially in the high dose region where the SF decreased to below 10%) provides a big challenge up to the point that these models did not find clinical applications up to date. One of these is the linear-quadratic-cubic model that compliments the LQM with an additional cubic parameter and the tissue specific constant γ.
On the other side of the scale, some experimental observations [48] indicated a low-dose hypersensitivity for doses below 1 Gy. It was found for some mammalian cell lines, that the region below about 10 cGy shows a much steeper SF slope than the LQ-fits for higher doses suggest. This slope is followed by an increased resistance again. An explanation for this is a sort of repair-inducing dose threshold at about 10 cGy that only then causes certain repair mechanisms to take action. For the low dose region, another modification of the LQM was therefore suggested, called the Induced Repair model. Essentially, the Induced Repair Model combines two versions of the standard LQM: for low doses, it represents the LQM with parameters α_s and β and for high doses the LQM with α_r and β. α_s therefore indicated the steeper slope at lower doses. The low-dose hypersensitivity might become important for treatments that irradiate large volumes of the patients anatomy with relatively low doses. This is for example the case for intensity-modulated radiotherapy (IMRT) where large volumes of normal tissue is irradiated with doses below 1 Gy. However, this topic is still in debate with only very limited experimental proof.
The basic form of the LQM does not consider time-linked parameters. In case of consecutive fractions where fractions are delivered before cell repair is completed (multiple fractions per day or continuous irradiation), more factors need to be considered in the use of the LQM. However, this will not be relevant within this work. Furthermore, for fast proliferation (of e.g. tumours), the significant increase in cell number also needs to be taken into account. This can for example be done by including another term in the exponential expression representing proliferation according to the specific tumour type [23]. For the analysis in this work, this factor is not considered.

2.2.2.2 *Radiobiology in clinical practice*

When planning a RT treatment, it is necessary to know which dose magnitudes can and should be applied to the human tissue. In case of tumour tissue, it needs to be known how much dose is necessary, in order to achieve *local control*: complete

inactivation of all cancer stem cells within the tumour, so that no further regrowth after therapy is possible. For normal tissue, constraints need to be applied to the planned dose field, in order to avoid normal tissue toxicities as consequences of radiation damage. Based on these numbers combined, the treatment plan is optimized to fulfil dose prescription to the tumour (usually described by dose coverage at a certain prescription dose level like the $D_{95\%}$, the dose that 95% of the structure receives) and the normal tissue constraints, usually defined by the volume V that is allowed to receive a certain dose level x, referred to as $V_{x\%/Gy}$[b].

Conversion of fractionation schemes It is often desired to change the fractionation scheme (number of fractions and dose per fraction) without changing the biological effect (*isoeffective approach*) that proved good clinical applicability. Mainly, it is the conversion from normal fractionation (standard approach of 2 Gy per fraction) to hypofractionation (lower number of fractions with doses above 2 Gy) that needs to be calculated as outlined above in section 2.2.2.1. The isoeffective conversion of fractionation schemes is a direct application of the linear-quadratic approach in clinical practice by using the LQ description of the relationship of dose to biological effect.
The **equivalent dose in 2 Gy fractions (EQD_2)** refers to the total dose of a treatment with a fraction dose of 2 Gy that is biologically equivalent (isoeffective) to a treatment delivered with fraction dose d and total dose $D = nd$ with n being the number of fractions:

$$EQD_2 = \frac{D(\alpha/\beta + d)}{\alpha/\beta + 2} \tag{2.14}$$

In clinical practice, 2 Gy per fraction is the commonly applied dose and therefore serves as a good value for comparing the new treatment scheme to the standard approach and experience. However, in principle any other number can be applied for conversion. The EQD_2 is derived from setting the effect based on the LQ approach from two different treatments to be identical, i.e. being isoeffective. This principle will be used later in the derivation of the total biological dose bEQD (see section 3.1.3).
There are other concepts in literature for the same purpose which are also mathematically and biologically equivalent and will therefore not be explained in more detail here: the biologically effective dose (BED), the extrapolated tolerance dose (ETD) or the normalized total dose (NTD).

TCP, NTCP and volume effects The interpretation of the SF in clinical practice is not straightforward since the quantification of cell survival can often not

[b]Note that two conventions exist: V_{50Gy} refers to the volume that receives at least 50 Gy while $V_{50\%}$ is the volume that receives at least 50% of the tumour prescription dose. This will be noted accordingly in the following.

directly be linked to treatment outcome due to a number of reasons outlined below, especially in normal tissue. Therefore, models are proposed that describe the relationship of dose to treatment outcome. Fundamentally, one has to distinguish between tumour and normal tissue in the description of dose response.
For the irradiation of tumours, the aim is to achieve local control as described before, which means in the classic interpretation that optimally, all cancer cells are inactivated. This is reflected in clinical terms with the **tumour control probability (TCP)** that is directly used as a tool for treatment planning in comparing different treatment strategies. The concept is based on the hypothesis that every single cancer cell can in principle generate a tumour, thus, for the success of an RT treatment, all cancer cells need to be eliminated. This is also why tumour coverage with a certain dose level is often prioritized over the risk of occurring side effects.
The standard TCP approach is based on the expression given by the LQM in first approximation that in fractionated RT, the α term alone defines the slope effectively. The initial absolute cell number is approximated by tumour volume and clonogenic cell density. The probability of cell death is approximated with Poisson statistics to derive the final TCP formula.
However, TCP is also criticized for two reasons mainly: the estimation of cell kill is model based and cannot be measured directly and secondly, not in all cases, it is necessary to inactivate all tumour cells to achieve cure (but for example only cancer stem cells [18]). Zaider and Hanin [86] propose that it is rather the *"probability of cure"* of the patient that should be aimed for rather than a certain TCP value.
The situation in **normal tissue** is even more complex. For example, the quantification of toxicity is often difficult. For many endpoints, patient reported outcome is given via a questionnaire in which the assessment is naturally subjective, thus the direct correlation of toxicity to SF is unsuitable. Furthermore, the relationship between SF and induced toxicity is not necessarily straight forward but depends on the functional structure of the organ and so called *volume effects*.
The response of normal tissue to radiation is described in so called *normal tissue complication probability (NTCP)* models. Several approaches exist. An early and widely used model was presented by Lyman in 1985 [46]. The model includes two variables, given by the dose and the volume, and is described by three model parameters: D_{50}, a uniform dose to the structure that when received by the entire volume results in a 50% probability of the certain endpoint to occur; an exponent to the volume variable, n, that describes the volume effect (see below); m, a parameter describing the steepness of the dose-response curve. This model is purely empirical.
A huge challenge in NTCP modelling is the complex architecture of normal tissue that requires the description of volumetric effects. This is also a problem in the Lyman model, since it is only valid for uniformly irradiated structures, which does not occur in actual clinical practice. A further developed form of the Lyman model was presented by Kutcher and Burman in 1989 [42] who tried to tackle this limitation by introducing the effective volume method (resulting in the so called Lyman-Kutcher-Burman (LKB) model). In their approach, the dose-volume distri-

bution (dose-volume histogram, description see section 3.2.3), is summarized into a single dose metric, the equivalent fraction volume that receives the maximum dose, which is then used within the Lyman model. This is based on the assumption that each volume fraction follows the Lyman model in the same way as the entire structure.

NTCPs are up until now a field of research and do not provide enough clinical evidence or experience to be used for treatment decision making. Moreover, the description of volume effects in and outside of NTCP modelling is a wide field of current research. The volume that receives a certain dose often determines its dose tolerance and therefore the necessary clinical **tolerance dose constraints**. For treatment planning, therefore often dose constraints to certain volumes are used as a simpler, and more robust approach instead of NTCPs. They are, for normal tissue, mostly given in the form of V_x, the volume receiving at least x Gy, as mentioned before. In many cases, clinical decisions are based on the 1991 recommendations by Emami et al. [21]. However, these are not necessarily based on clinical or experimental evidence, but rather opinion- and experience-based. In 2010, the Quantitative Analyses of Normal Tissue Effects in the Clinic (QUANTEC) reports [47] summarized the given published evidence-based experience for normal-dose constraints and provide the up to date standard. However, the reported toxicities are related to planned doses which might vary substantially from the delivered doses (see section 2.2.1.1). Furthermore, most data is based on three-dimensional (3D)-conformal treatment plans (see section 2.1.2) and not using IMRT which show substantially different dose distributions. Hypofractionation is also not covered. These problem are important aspects in current research and were also investigated in this work (discussion see section 5). Results obtained in this work were compared to the QUANTEC constraints for the planned dose and the two strategies to obtain the accumulated dose.

To tackle the volume effect question, Withers et al. [80] first introduced the concept of **functional sub-units (FSUs)**. The definition of FSUs is tolerance dose based and defines the largest volume of a structure that can be regenerated from a single clonogenic cell and on the other hand represent subunits that are independently sterilized by irradiation. The varying forms of radiosensitivity of entire organs is described by the architectural arrangement of FSUs. The physiological function of an organ is only then impaired, if either a critical number of FSUs is damaged or a FSU was exposed to a certain threshold dose. This then defines the volume of the organ that may receive certain dose thresholds without inducing organ damage.

Two types of an organ's architecture can be distinguished: ***serial* or *parallel***. For serial type organs, the damage of single FSUs (or maybe even just one) may induce organ toxicity . This means that in clinical practice, this organ is prone to side-effects when exposed to local high doses, even in small volumes, so called "hot spots". The dose distribution to the entire organ (e.g. mean dose) plays a minor role. A classical example of a serial type organ is the spinal cord or the intestines. On the other hand, parallel type organs are susceptible to radiation-induced side-effects in

cases large volumes (a certain number of FSUs) of the organ receive a certain dose level. The dose distribution within the organ becomes relevant. Lung or liver are examples for more parallel type organs. It needs to be noted that organs are never organized in a purely parallel or serial manner. It may be that for example an organ contains regions that are serial-type organized, introducing a serial-component in an otherwise parallel organ.
For the above described LKB model, the volume exponent n describes the volume effect within this NTCP model. For n close to 1, the volume effect is large, which means that the tolerance dose increases strongly with decreasing volume (more parallel), while n close to 0 describes a small volume effect given for more serial type organs. This principle was extended by the introduction of the relative seriality model [39] that includes an extra parameter s to describe the tissue architecture in this manner.
The volume effect is also used when converting an inhomogeneous dose distribution to a homogeneous dose distribution resulting in the same SF: The **equivalent uniform dose (EUD)** is defined as a generalized mean[c] used to summarize heterogeneous organ doses [59]. It is defined as a weighted mean, where the weighting is based on the tissues serial or parallel type structure by the exponent a:

$$EUD = (\sum_i v_i D_i^a)^{1/a} \tag{2.15}$$

v_i denotes the volume of the structure with absorbed dose D_i. For the exponent a, the general recommendations are rare and usually based on empirical fits, a problem that is also valid for the discussed NTCP models before. The International Commission on Radiation Units & Measurements (ICRU) report 83 [50] generally suggests: "*serial-like tissues have large positive values of* a*, and values are close to 1 for parallel-like structures*". With a close to 1, the EUD converges to the mean of the dose. Some works use the inverse of the exponent a, referred to with $n = 1/a$ also described above for the LKB-model. It should be noted that EUD is therefore as well a model-based quantity.

[c] note that the here used EUD concept is also referred to as the general EUD or gEUD

3

MATERIALS & METHODS

3.1 METHODOLOGY OF BIOLOGICAL DOSE ACCUMULATION

This work builds on the observation that conventional dose accumulation is not being used consistently with regard to radiobiological models and considerations. To overcome this inconsistency, a methodology of the accumulation of the biological effect and dose was developed in this work. This chapter builds up on the methodological background given in section 2.2 and brings together both aspects of dose accumulation and radiobiology.
As the basis for accumulation, the post-processing chain of the dose data from daily imaging (dose mapping) needs to be in accordance with biological mechanisms. Section 3.1.1 supplements current debates in literature about the rationale of dose mapping by discussing this problem based on biological considerations. The result of this section is a conclusion on how the mapping algorithm should be implemented in order to give a biologically meaningful transformation of the tissue and the dose. Based on this, biologically consistent dose accumulation is presented. This is twofold: first, the effect accumulation approach is described - the consistent use of the biological model in order to retrieve the cell survival fraction after fractionated treatment (section 3.1.2). Second, the concept of the total biological dose bEQD is presented - the total dose of the treatment that represents the mathematically correct (cell) survival fraction (SF) according to the effect accumulation approach (section 3.1.3). The chapter is summarized in section 3.1.4 to present the complete workflow of biologically consistent dose accumulation.

3.1.1 *The biological rationale of dose mapping*

The content of this section was submitted for publication in August 2019 to *Physics in Medicine and Biology* as a Technical Note as it is presented here [57].

In the previous sections, the workflow for conventional dose accumulation was presented (section 2.2.1.2). The basis of any sort of dose accumulation is the fractional delivered dose values that need to be transformed into the grid and geometry of a common frame of reference, the deformed doses the $\mathtt{d}_i$. Thus, the fundamental

prerequisite of dose accumulation is dose mapping which is based on deformable image registration. This section presents the methodological analysis of dose mapping under consideration of biological mechanisms. This serves the purpose of building the basis for the following dose accumulation. This chapter is based on a thought experiment and will hold no further data based results in the following chapters. The result of this chapter is the conclusion on how dose mapping needs to be performed in order to be biologically consistent.

The underlying principles and assumptions involved in this dose mapping procedure are the topic of an ongoing debate [31] [90] [68], especially for cases where anatomical structures change in volume over the course of treatment. Dose mapping and accumulation has uncertainties [74] [56] associated with the use of deformable image registration (DIR) as also outlined above. Two distinct cases can be identified: (i) The images to be registered represent the exact same anatomy containing only tissue deformation; (ii) the two images include changes in anatomy, for example due to tumour growth or shrinkage, changing organ fillings, or weight loss in normal tissue. While the registration problem of (i) is well-defined and, at least in principle, has an unambiguous solution, case (ii), requires the registration of two different objects and is therefore an ill-posed problem. In the following, this work focusses on the latter case. The DIR will produce a solution, which may be satisfying from a geometrical standpoint. For the biological assessment of dose mapping, however, it is necessary to consider also the biological target receiving the dose, and how doses at different time points can be added in a biologically consistent way.

Most bio-effect models are observational in nature, derived to fit time dose and fractionation dependencies of the observed responses as discussed in detail in section 2.2.2. They are by nature approximate models making certain assumptions. Biological understanding of cell organisation has been integrated into these models, and demonstrated to have clinical relevance [67] [60]. Current clinical imaging techniques do not provide a resolution at the level of individual cells nor to distinguish between viable and sterile/killed cells. We can therefore base our implementation of bio-effect models and conclusion only on the results on a voxel level which will always suffer from discretization effects (further discussed also in section 5.2).

Dose mapping based on DIR can be implemented so as to conserve dose, or so as to conserve delivered energy [90]. In a recent review on adaptive radiotherapy (RT) by Sonke et al. [71] it was stated that, in the case of volume changes of the structure *"As a consequence, however, when transforming scans and daily doses back to the planning computed tomography (CT), mass and energy are not conserved. While corrections can be applied to enforce mass and energy in such transformations yielding "physically" more correct representations, it is debatable if such corrections are also biologically more relevant."*. Here, the two existing approaches are discussed by extending the debate to include considerations at the biological level. The conclusions here are based on the basic postulation that the mapped dose and subsequent accumulation has to be in accordance with the observed biological response of tumour and normal tissues. It will be shown here, that the dose conserving approach

facilitates a biologically consistent interpretation of the relation of the accumulated dose and the biological effect. On this basis, the different effects causing volume change (for both growth and shrinkage) are analysed for DIR on a biological level and its implication for the dose mapping implementation are discussed.

Biological considerations For *incompressible soft tissues*, it is assumed that physical density remains constant over time. In biomechanics, soft tissue organ architecture is defined by cellularity, architecture of the extracellular stroma, and extracellular water content. The extracellular stromal content is the most highly conserved of these three properties. Across a wide range of physiological and pathological processes, intracellular and extracellular compartments both approximate to water density. Thus, changes in organ volume are caused by a change in voidage (cellular packing) due to reciprocal gain or loss of water and cells. Changes in organ volume are therefore directly proportional to mass [24] while the cell density can change without change in physical density. The following analysis is mainly based on this type of tissue. For *lung tissue* on the other hand, it is necessary to look at intrafractional motion. In this case, though volume changes, the mass of the object is constant while physical density changes. This scenario will be discussed briefly at the end.
Ultimately, treatment outcome is a combination of the treatment parameters as well as radiobiological parameters of the tumour and normal tissue of the patient. The argument is based on the Linear Quadratic Model (LQM) for the prediction of the biological effect given as the SF. For simplicity, the biological effect discussed here only considers cell survival as the endpoint and it does not consider changes in organ function of normal tissue which depends on more complex cellular architectures [60] [67].
In the following, both dose mapping approaches are briefly analysed with regard to the resulting mapped dose. Afterwards, the biological implications from the derived mapped doses are analysed in detail. Furthermore, the discussion of biologically consistent dose mapping is extended to four different biological causes of volume change in tissue.

Dose- vs. energy-conserving dose mapping: a thought experiment To perform dose mapping in case of mass loss or gain, two approaches can be considered: conserving the delivered dose or conserving the deposited energy. The two approaches are analysed based on a thought experiment of a structure made up of uniformly-sized volume elements (voxels). The example situation of the thought experiment, where the structure reduces in volume and mass and receives a homogeneous dose, is shown in figure 3.1, along with the transformation back to its initial state.
Conservation of dose in the context of dose mapping means that the deformation vector field (DVF) and interpolation is applied to the dose distribution, analogously to

grey values in a CT image. In this implementation, when the volume of a uniformly irradiated structure is changing (figure 3.1, a) (the smallest of which is a voxel), the dose to this volume will still be the same after the mapping process (figure 3.1, b) (the same principle as visualized before in figure 2.7). Recent works promoted the concept of energy conservation [90] [15] [89] [91] [14] for structures that change in volume. In this approach, the same transformation and interpolation is applied to the energy delivered to the structure, instead of the dose (figure 3.1, c).

For the analysis of the dose mapping processes, the following thought experiment is considered in figure 3.1: The starting point is a volume at timepoint T_1 (figure 3.1, a) that shrinks to 1/4th of its initial volume at timepoint T_2. Uniform shrinkage is assumed so that each initial voxel will also shrink to 1/4th of its volume when looking at the corresponding sub-unit at T_2. The whole structure receives a homogeneous dose of a delivered dose of $2\ Gy$ at T_2 that is then being mapped back to timepoint T_1 for the purpose of dose accumulation in the reference frame. By applying the DVF obtained from DIR between the fixed image from T_1 to the moving image in T_2, to the structure in T_2, the result is a structure of the same size as in T_1 (after applying the DVF: T_2'). We assume an anatomically correct transformation to solely focus on the underlying principles, so that for the volumes: $V_2' = V_1 = 4V_2$.
From T_1 to T_2, volume and therefore mass is reduced by a factor 4 and then virtually increased again by a factor of 4 in the mapping to T_2'. This is the direct result of biologically consistent mapping by assigning water equivalent density at all stages. A direct mapping of the structure to its initial volume and location is a prerequisite for accumulation on a voxel scale in the reference frame represented by timepoint T_1.
For the dose conserving mapping approach shown in figure 3.1 b, the dose in the entire volume in T_2' remains the same as in T_2. Since mass was virtually increased in the mapping, the delivered energy is also (virtually) increased by a factor 4, since energy, in contrast to dose, is an extensive quantity of the system. In contrast, when conserving the delivered energy (figure 3.1 c), the total energy after mapping is the same as in the structure at T_1, so $E_2' = E_2$. Since dose is defined as energy per unit mass, the dose within the structure is reduced to 1/4th of its initial value.

Biological implications of dose- and energy-conserving dose mapping The biological effect as predicted by the LQM is solely linked to dose, not to delivered energy. Dose is an intensive quantity of the system and is therefore the same for each mass element (i.e. cell) in one voxel within the limits of discretization. For the presented scenario in figure 3.1, all cells in the highlighted yellow region receive 2 Gy at timepoint T_2. Thus, the cell survival fraction resulting from the irradiation at T_2 is given by the LQM for $d_2 = 2\ Gy$. When conserving the dose (figure 3.1, b), the survival fraction within the mapped volume is the same for both the original cells from T_2 and virtually added volume (water or cells) in T_2'. In the energy conserving approach (figure 3.1, c), the dose is decreased by the factor increase in mass. Therefore, also the dose to each cell is decreased and the SF within the mapped volume

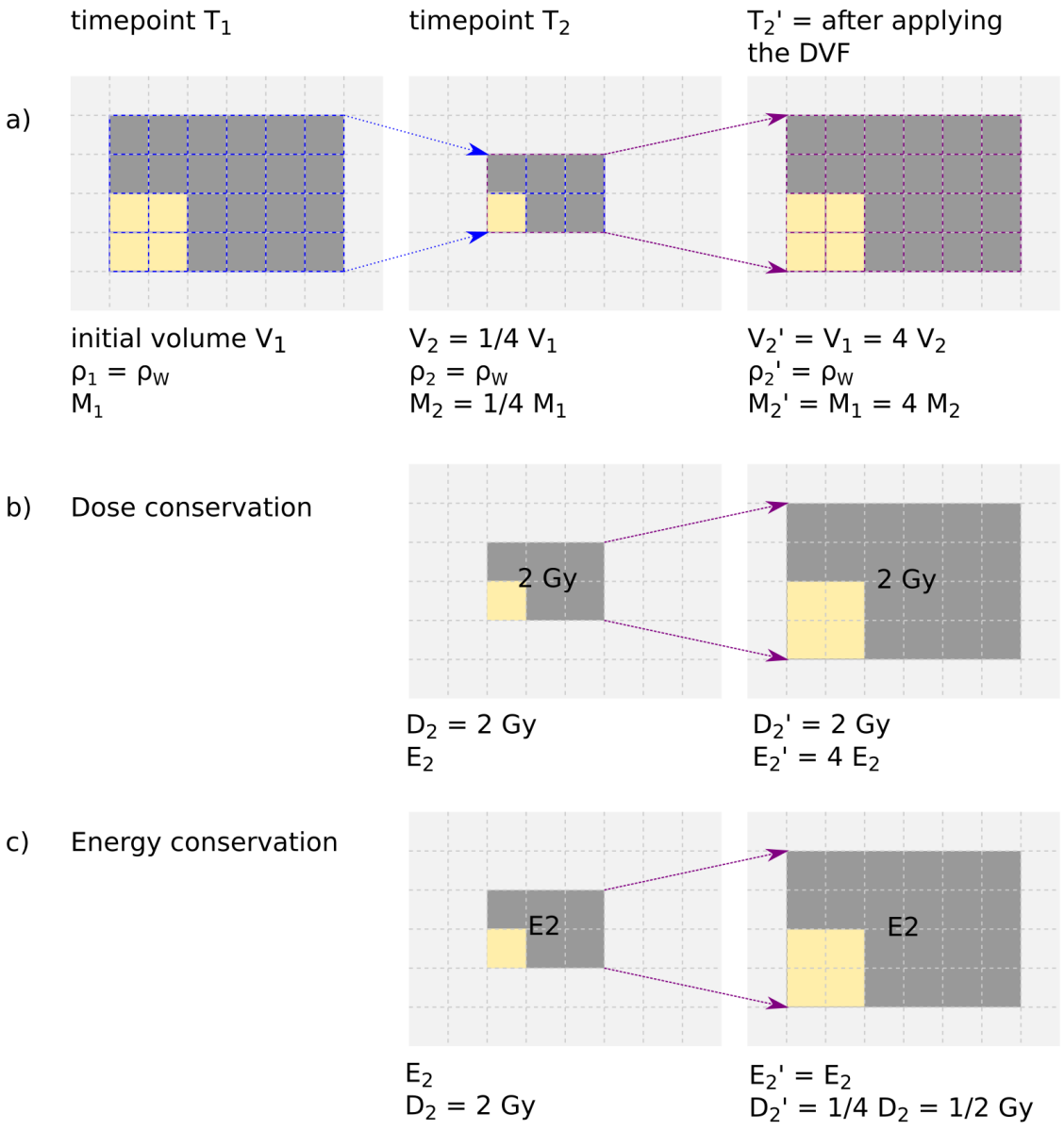


Figure 3.1. – Example case of a structure shrinking to 1/4th of its initial volume from timepoint T_1 to T_2. Density is assumed to be approximately water equivalent (ρ_W) at all times, therefore volume is directly proportional to mass. a) Deformation and resulting mapped structure (T_2'). b) Dose conservation within the entire structure that receives homogeneous dose. Deposited energy is scaled according to the factor change in mass. c) Energy conservation implies that the dose for the entire structure is reduced to 1/4th of its initial value according to the increase in mass. For cases of volume increase, the calculation can be performed analogously.

will be that of the LQM for 1/2 Gy:

$$SF(T_2) = LQM(2\ Gy);$$
$$SF(T_2')_{dose-conserved} = LQM(2\ Gy);$$
$$SF(T_2')_{energy-conserved} = LQM(1/2\ Gy)$$

In summary, the dose conserving approach conserves the SF at timepoint T_2 in the mapping process while altering the deposited energy to the overall structure. Conserving the energy that was delivered by the treatment device, on the other hand, results in altering the biological effect imposed at timepoint T_2 by reducing the dose received by the individual cells.
A requirement for biologically consistent dose mapping is to conserve the survival probability of cells in fraction x. This is to conserve the biological impact of the irradiation of each fraction and enable to calculate the overall survival fraction of the treatment schedule in the reference frame. As the process of DIR along with dose mapping is not a physical process, but a mathematical projection of one state to another, this work suggests to adjust and interpret the implementation of this process with our knowledge about the biological system. The structure change between T_1 and T_2 is a physical process that results in a mass change. Thus, the structures in T_1 and T_2 do not represent the same physical system. In the attempt to map the structure from T_2 to T_1, a non-closed biological system is simulated, so the implementation should not try to conserve a closed system in the process. Whenever volume (i.e. mass) changes, so does the absorbed energy. In this sense, the aim of dose mapping should not be to conserve what energy has been delivered by the treatment device to one of these systems but rather which effect has been obtained in the biological tissue. Thus, the question needs to be answered, what energy would have been needed to be deliver to the structure in the reference frame, in order to cause the same biological effect as was obtained in the structure at fraction x. This is achieved by mapping the dose that was delivered to the structure in T_2 in a dose conserving-manner, while the energy of the system is scaled with the change in volume induced by the mapping process. Thus, one must not demand energy conservation but equality of the ratio of change in energy and mass which is given by the conservation of dose.
Recently, Zhong and Chetty [90] published an opinion piece in which they state that energy conservation is violated in the process of dose mapping for scenarios where structures undergo volume changes. In their conclusion, they propose energy conservation for dose mapping. In this thesis, it is agreed that delivered energy is not conserved in this process as presented above, but disagreed with the conclusion, and argued that energy conservation is not legitimate in instances where volumes are changing while the physical density is conserved. This is in accordance with a short reply by Hugo et al. [30] to the mentioned opinion statement above. Two different states of an imaged patient's anatomy along with the delivered dose do not

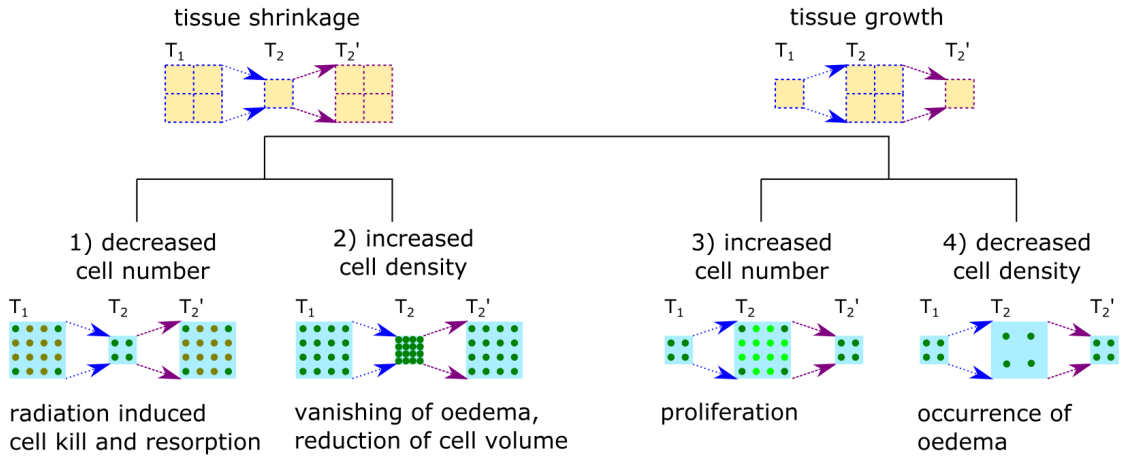


Figure 3.2. – Scenarios of volume change in biological (un-compressible) soft tissue. In the top row, the physical implementation is depicted analogously to figure 3.1, referring to tissue shrinkage and growth. In the bottom row, possible biological mechanisms of the individual cases are shown. Both cells and extracellular space have approximately water equivalent density so that a change in cell density does not induce a change in physical density. For the cases of change in cell number, the result of the mapping process is that cells are virtually added or removed, while a change in cell density is given by a virtual increase/decrease in water content. In tissue, the depicted scenarios usually occur in a mixed manner.

represent the same closed physical system. The very fact that volumes change (loss or gain of mass) underlines this. This is in-line with the above conclusion that it is dose, not energy that needs to be conserved during the mapping process for a volume of homogeneous dose even if this volume changes in size.

Biological consistency of the dose-conserving mapping method in different scenarios There are four main processes that drive volume change in soft, non-compressible tissue, presented in figure 3.2. Using dose conservation in the mapping process, a consistent relation between mapped dose and biological effect can be achieved. This will be discussed separately for the four cases that are depicted as well as for the case of lung tissue:

Scenario 1 - *Volume decrease due to decreasing cell number* (radiation induced cell kill and resorption): Cells that are killed and resorbed from T_1 to T_2 will be virtually added in the projection from T_2 to T_2' ($V_2' = V_1$) and assigned with the dose of the voxel they contribute/ are shifted to. This ensures that the dose delivered to the original vital cells present in T_2 is not diluted and the SF of the cell cohort at T_2 is conserved within the volume V_2'. It needs to be noted that all virtually added cells represent dead cells (since they were resorbed as such from T_1 to T_2). In the accumulation process, this SF needs to be taken into account as will be discussed in detail in section 3.1.2. Thus the dose conserving dose mapping approach conserves the SF at each fraction as well as the absolute number of surviving cells for accumulated doses. It has to be noted that this argument is also valid, in case not all cells that were killed in the first fraction are resorbed prior to the subsequent fraction,

but remain in the volume as sterile cells.
Scenario 2 - *Volume decrease due to increasing cell density* (reduction of cell volume or vanishing of an oedema): The dose to a cell within a voxel of x Gy is also x Gy, independent of the absolute number of cells within that voxel (cell density). The increase of volume from T_2 to T_2' causes a decrease of cell density back to its original value (as in T_1) by virtually adding water to the system. Whether the virtually added mass is made up of dead cells (first case above) or water that receives the delivered dose (i.e. additional virtually delivered energy) is irrelevant as long as in the accumulation process, the survival fractions from previous fractions are considered analogously to the argument above. Dose conservation ensures that dose to individual cells is not diluted.
Scenario 3 - *Volume increase due to increasing cell number* (proliferation): In the projection of T_2 to T_2', volume is decreased again, virtually removing cells from previous proliferation. When applying the LQM, we assume that repair has been finished before the next treatment fraction is delivered (for standard fractionation of one fraction per day). Thus, cells can be considered being either sterilized/killed or completely repaired and therefore without dose memory and in the latter case, will behave like un-irradiated cells. Thus, it is not necessary to distinguish previously irradiated and newly proliferated cells in the volume decreasing step, which would not be possible on the basis of the imaging data. By conserving the dose, the SF of the cell cohort is conserved accordingly. Note that here again, for the consistent accumulation of the biological effect, the survival fraction from previous fractions needs to be considered. Regarding the absolute cell number, the increase by proliferation between the two fractions has to be considered. This can be done by adding a proliferation term in the LQM as explained before. It needs to be questioned whether dose (or effect) accumulation on the initial volume is still reasonable for scenarios where volume increases over the course of therapy due to proliferation since additional tissue that might be present at the end of the treatment compared with the initial state is being neglected in the analysis of the accumulated delivered dose in the reference frame.
Scenario 4 - *Volume increase due to decreasing cell density* (e.g. increase in water content due to oedema): In this case the volume at T_2 is increased compared with T_1 and is mapped back to the initial volume by the mapping process. Virtually, water is removed from the system and the same argument applies as in the previous case that conservation of dose also conserves the SF of the cells within the volume decreased in the mapping process.
Scenario 5 - *Volume change due to physical density change* (intrafractional lung motion, not illustrated in the figure): This case is different to the four cases discussed above since physical density changes and the motion occurs on a small time scale. For the overall structure of the lung, mass is conserved in this case. This holds true as long as the entire structure is being observed and mapped (no migration of lung tissue outside the field of view that is being mapped). In this case, both energy- and dose-conserved mapping are equivalent.

In conclusion, in all scenarios of tissue shrinkage and growth, the dose conserving mapping approach facilitates a consistent implementation of the biological effect in terms of SF as well as absolute cell number, independent of the cause of volume change. In biological tissue, volume change is caused by various processes happening simultaneously that could for example be described by combinations of the described scenarios. Since all discussed causes of volume change can be treated in the same way in the mapping process, combinations of these are also permissible.
In conclusion of this analysis, dose mapping should be performed in a dose conserving manner. This way of implementation is used throughout the workflow of this thesis and will not be further analysed but used as a given standard. A short discussion on implementation consideration is presented in section 5.2 .

3.1.2 *The effect accumulation approach*

Fractionation is a key principle in RT to open the therapeutic window, i.e. to achieve higher differences between the cell kill in tumour and normal tissues. The basis of this principle is the non-linear relationship between dose and biological effect. As described above in section 2.2.2, the LQM represents a good empirical model to describe this biological effect by an exponential decrease in the cell survival fraction with dose. As given in equation 2.12, within the exponential expression, the biological effect is linked to a linear and a quadratic dose term. With this, the described shoulder-shape of the cell survival curve over dose offers the chance to take advantage of the difference in radiosensitivities of tumours and normal tissues by fractionation.
The advantages achieved by fractionation are, however, reduced by fact that between fractions, organs will move. As described in section 2.2.1.1 this can strongly impair the delivered dose per fraction and lead to underdosage of the tumour or overdosage of normal tissues. On a local scale, the dose within each voxel will change from day to day, even for small organ motion amplitudes. Above, the principle of dose mapping and accumulation is discussed to retrieve the total delivered dose of the treatment course that includes these daily variations on a voxel scale (section 3.1.1). Even though the daily dose variations are in fact summarized in the total conventionally accumulated dose D_a (equation 2.9), in this process, the information on the daily dose variations is lost. Instead, this information is reduced to the mean dose per fraction $d = D_a/n$.
As stated above, the relationship of dose and biological effect is non-linear, manifested in the β-term in the exponential expression, which is quadratic to dose. Thus, using the averaged fraction dose d in the LQM equation 2.12 will lead to a different SF outcome than using the individual fraction doses d_i. Thus, a mathematical flaw is introduced by neglecting the non-linear relationship by first performing a linear operation on the dose information.
The effect of this inaccuracy is visualized in figure 3.3 that shows cell survival curves

according to the LQM for three consecutive treatment fractions. Each colour represents the same total dose but build up of different fraction doses that will in consequence each result in a different SF at the end of the three fractions. The blue curve shows three equal fraction doses. As can be seen, this leads to the highest cell survival fraction in total. Thus, the total biological effect of the treatment does not only depend on the total dose, but on each delivered fraction dose throughout the treatment.

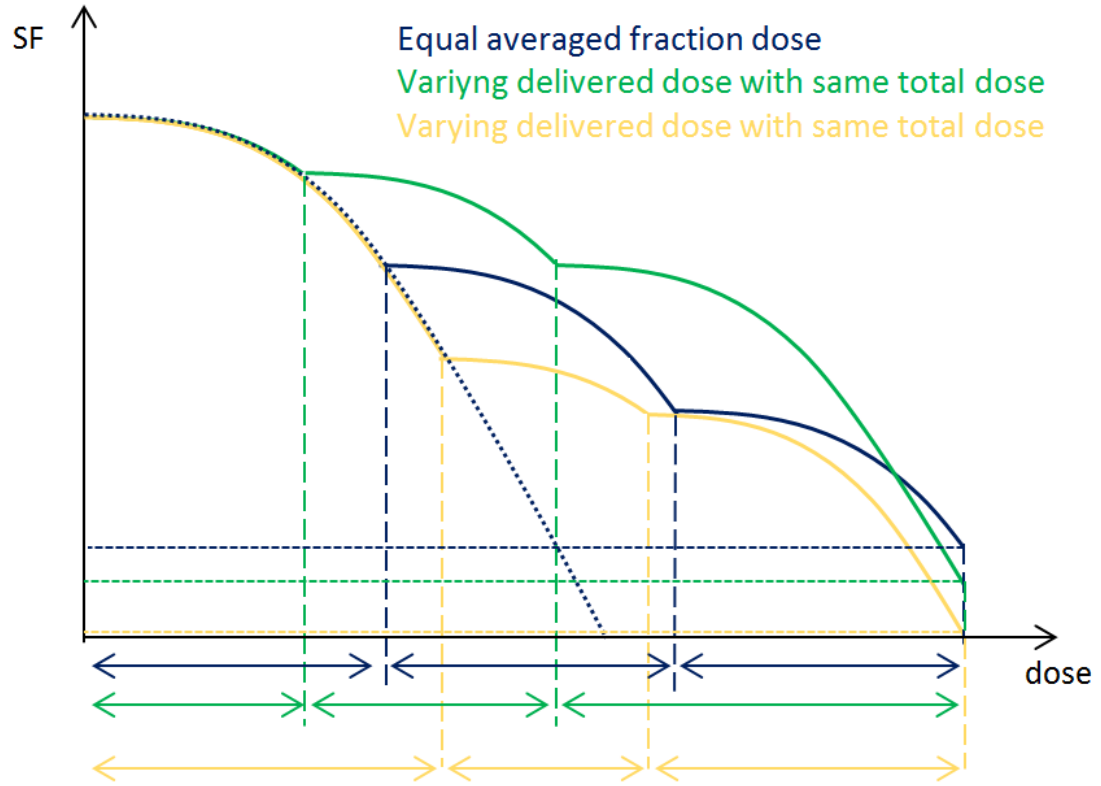


Figure 3.3. – SF according to the LQM for three consecutive fractions. Each colour shows the SF for the same total dose, but delivered in different fraction doses. It can be seen that the total SF depends on the individual fraction doses which shows the inaccuracy in using the conventionally accumulated dose for biological effect prediction in case of varying fraction doses. Reprinted from [58].

To overcome this inaccuracy, the steps of effect prediction and accumulation need to be interchanged. In the following, the principle of mathematically correct effect and dose accumulation is derived and compared to conventional dose accumulation and effect prediction. The following derivations along with parts of the results presented in chapter 4 were submitted for publication in *Physics in Medicine and Biology* [58]. The magnitude of the difference between the shown exemplary curves in figure 3.3 is topic of the results of this work in chapter 4.

In section 2.2.2, the biological effect after n fractions of dose d and total dose D_a was described, which is obtained from the LQM by a power-to-n-law resulting in the total SF of n fractions of:

$$SF(d) = e^{(-\alpha D_a - \beta n d^2)} \tag{3.1}$$

In the following SF(d) will denote the SF of n fractions of dose d, called the *dose accumulation approach* (the conventional workflow).
The flaw in this principle for use in actual treatment scenarios is the fact that the dose d is changing in each fraction. However, the power-to-n-law is only valid to use for a constant fraction dose d. This law is applied in common practice by using $d = D_a/n$ for all fractions, thus the average of the daily dose d derived from conventional dose accumulation.
In order to retrieve the mathematically correct total survival fraction, it is necessary to calculate the biological effect of each fraction dose *prior* to accumulation.
The effect of a single fraction dose d_i is given by equation 2.12. When accumulating the effects of multiple fractions (i=1...n), then the accumulated effect is given by the product of all fraction effects with doses d_i:

$$SF(d_i) = e^{(-\alpha d_1 - \beta d_1^2)} \cdot e^{(-\alpha d_2 - \beta d_2^2)} \cdot ... \cdot e^{(-\alpha d_n - \beta d_n^2)} = e^{(-\alpha D_a - \beta \sum_{i=1}^{n} d_i^2)} \tag{3.2}$$

A visualization of the two approaches to estimate the final SF of the treatment are depicted in figure 3.4, following up figure 2.5 of the conventional dose accumulation workflow. The workflow principles follow the same path until after the transformation from DIR is applied to the forward calculated daily doses (the steps until this point are discussed in section 3.1.1). Thus, both approaches build up on the same transformation and use the same deformed doses d_i.

Comparing $SF(d)$, called the *dose accumulation approach* representing the conventional workflow, with $SF(d_i)$, the *effect accumulation approach* postulated in this work, the difference between both is given solely for the β-term. In other words: by performing dose accumulation, the amount of cell kill due to the radiation of previous fractions is not adequately taken into account and for all following fractions, the accumulation of any further fractions is based on a wrong number for cell survival. Using Jensen's inequality for a convex function [33], both approaches can be directly compared. Generally, Jensen's inequality states:

$$f(\sum_{i=1}^{n} \lambda_i x_i) \leqslant \sum_{i=1}^{n} \lambda_i f(x_i) \tag{3.3}$$

where λ_i is a non-negative parameter with $\sum_{i=1}^{n} \lambda_i = 1$. This is given when using λ as a normalizer for the number of fractions n, so that all $\lambda_i = 1/n$. The definition of the *effect* is given by the expression within the exponential of the LQM, which represents a convex function and can be directly inserted in the inequality:

$$\alpha \sum_{i=1}^{n} \frac{1}{n} d_i + \beta (\sum_{i=1}^{n} \frac{1}{n} d_i)^2 \leqslant \sum_{i=1}^{n} \frac{1}{n} (\alpha d_i + \beta d_i^2) \tag{3.4}$$

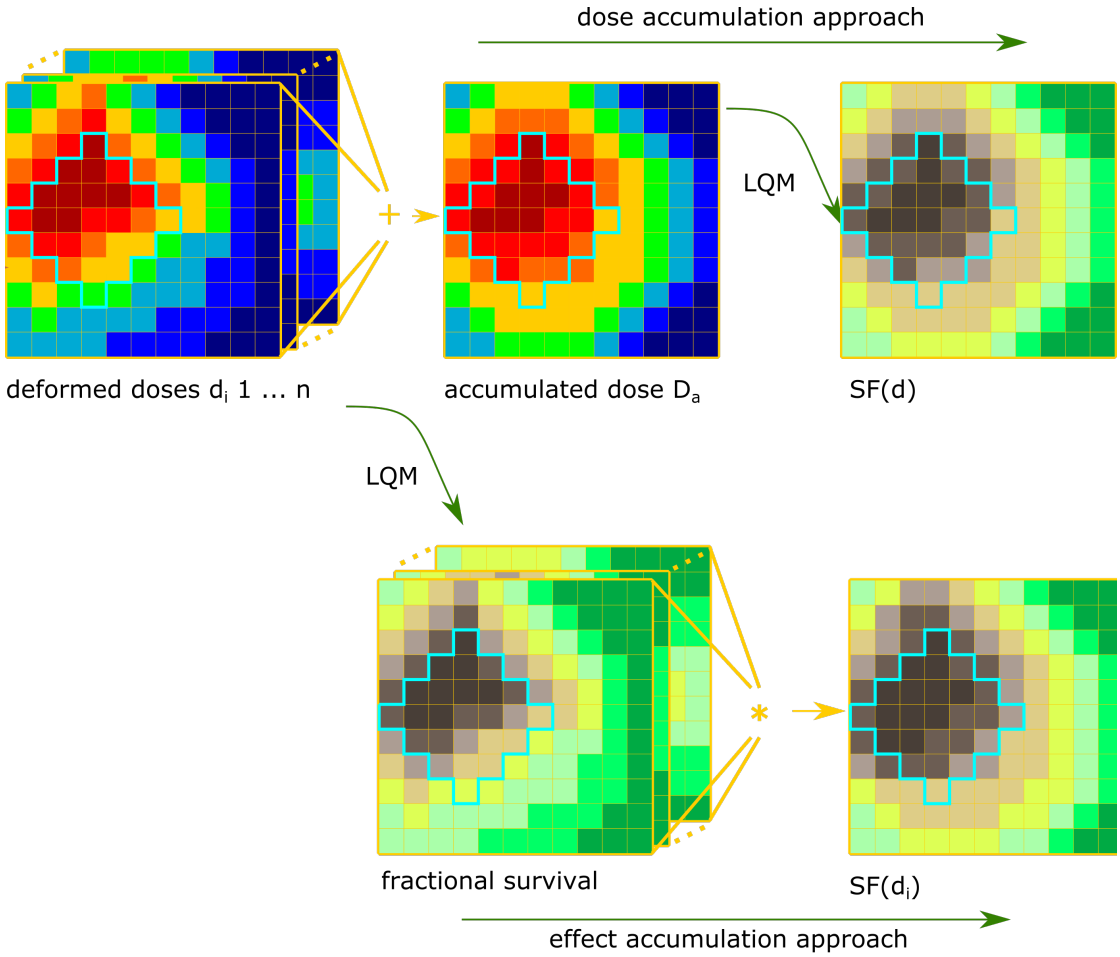


Figure 3.4. – The dose accumulation and effect accumulation workflows in comparison, following up on figure 2.5.

This can be rearranged to:

$$\alpha D_a + \beta n d^2 \leqslant \alpha D_a + \beta \sum_{i=1}^{n} d_i^2 \tag{3.5}$$

For the negative exponential of this expression this yields:

$$e^{(-\alpha D_a - \beta n d^2)} \geqslant e^{(-\alpha D_a - \beta \sum_{i=1}^{n} d_i^2)} \tag{3.6}$$

Comparing to equations 3.1 and 3.2, the above expression directly compares the survival fractions resulting from either the dose or the effect accumulation approach. It can be seen that when using conventional dose accumulation, the survival fraction of the total treatment is systematically overestimated:

$$SF(d) \geqslant SF(d_i) \tag{3.7}$$

Going back to figure 3.3, indeed the delivery of equal fraction doses (as assumed when using the conventional dose accumulation approach) will always lead to a higher survival rate than any delivery of varying doses with the same total dose. This might become especially important for normal tissues, since the survival rate after dose accumulation is assumed to be higher than what is actually predicted by the biological model. Therefore, this work will focus on the impact of the inaccuracy in biological effect estimation in normal tissue.

3.1.3 *The total biological dose bEQD*

For interpretation and clinical applications, the SF is an inconvenient quantity. Thus, following up on the previous chapter the question is, what would be the total dose delivered to the tissue that represents the mathematically correct cell survival fraction of the treatment? With the estimate of a total dose that consistently considers the LQM, common comparisons of dose constraints and prescription doses for the total dose of the treatment can be based on a biologically consistent workflow.
In order to answer this question, the total achieved biological effect needs to be conserved in the conversion process. For this purpose, the concept of the equivalent dose (described in section 2.2.2.2) was adapted to convert from a treatment of constant dose d, to a treatment with varying fraction doses d_i while being biologically isoeffective (instead of conversion between fractionation schemes). For this, the estimates of the fraction doses d_i are used, to find a total dose that results in the same biological outcome.

The total biological effect of a treatment with constant fraction dose d [Gy] is given by [82]:

$$E_d/\alpha = bEQD(1 + \frac{d}{\alpha/\beta}) \tag{3.8}$$

Here, the total biological dose (bEQD) is the total dose of a treatment with constant fraction doses d that results in the total biological effect E_d. This effect is now to be equal to the total effect after application of the effect accumulation approach (equation 3.2), that is, the biological effect after varying fraction doses d_i:

$$E_{di}/\alpha = \sum_i d_i + \frac{\sum_i d_i^2}{\alpha/\beta} = D_a + \frac{\sum_i d_i^2}{\alpha/\beta} \tag{3.9}$$

To resolve the question for bEQD, E_d and E_{di} have to be equal, so that the *total biological dose* is given by:

$$bEQD(d_i) = \frac{\alpha/\beta D_a + \sum_i d_i^2}{\alpha/\beta + d} \tag{3.10}$$

This total biological dose bEQD is an alternative to conventional dose accumulation (D_a) that is in contrast in accordance with the underlying biological model by taking into account the varying fraction doses as well as the tissue radiosensitivity. In the following work, bEQD will be compared to the conventionally accumulated dose D_a. For a direct comparison, the constant fraction dose d in equation 3.10 is set to be the same as the mean fraction dose of the treatment, given by: $d = D_a/n$. This way, the difference between both parameters will directly reflect the systematic inaccuracy caused by the method of conventional dose accumulation in contrast to effect accumulation.

A visualization of the principle of the bEQD embedded in the workflows described above can be seen in figure 3.5, following up on figures 2.5 and 3.4. As for the effect accumulation approach, it uses the deformed doses d_i. Applying the LQM on the total biological dose will result in the same SF as given by the effect accumulation approach (note that the change in number of fractions needs to be taken into account).
Comparing bEQD and D_α can be performed analogously to the above derivation of the difference between the effect from dose accumulation and effect accumulation using Jensen's inequality (equation 3.3). From this, it follows that the total biological dose bEQD is systematically higher than the conventionally accumulated dose D_α:

$$bEQD \geqslant D_\alpha \tag{3.11}$$

which is in accordance to the fact that the effect accumulation approach results in a systematically lower SF (equations 3.6 and 3.7).

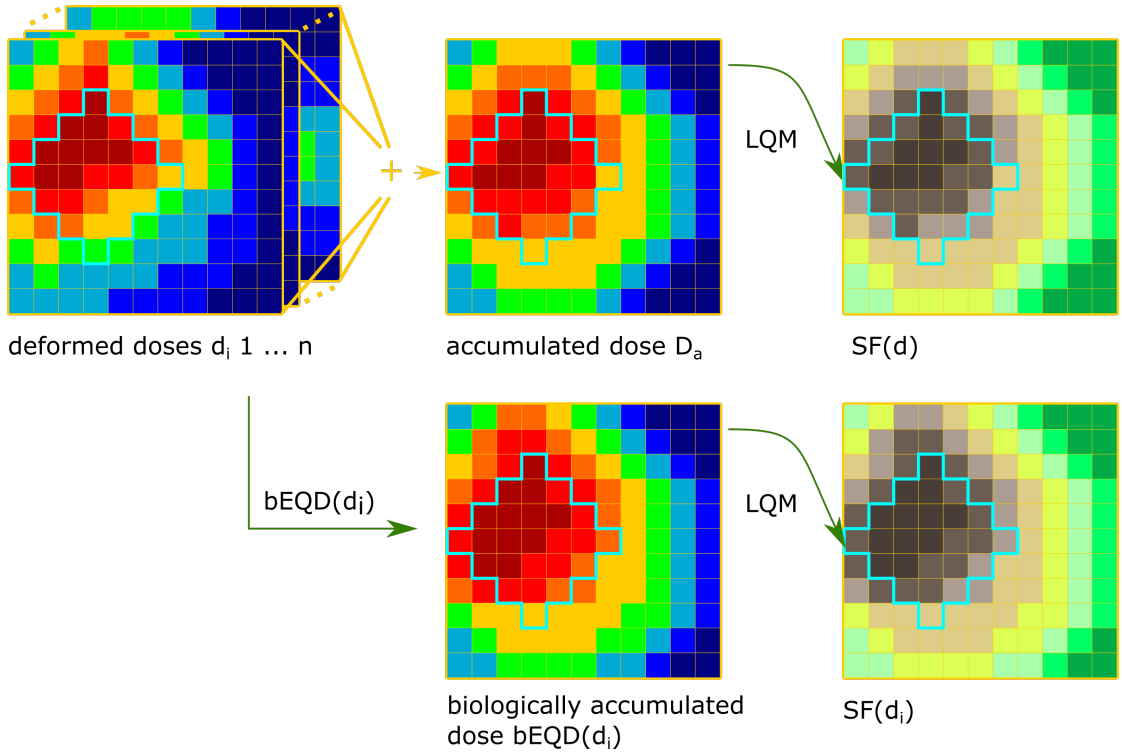


Figure 3.5. – Illustration of the effect accumulation workflow (top) in comparison with the biologically equivalent bEQD workflow (bottom row), following up on figure 3.4.

Uncertainties from dose mapping are also inherited by the bEQD. The uncertainty of bEQD with respect to the uncertainties in the respective d_i and with constant conversion parameter d is given by:

$$\Delta bEQD(\Delta d_i) = \frac{\sqrt{\sum_i (\alpha/\beta \Delta d_i + 2 d_i \Delta d_i)^2}}{\alpha/\beta + d} \tag{3.12}$$

3.1.4 *Summary: the biological workflow*

Figure 3.6 summarizes the complete biologically consistent workflow aspects of dose mapping and accumulation, derived in this work. For correct dose mapping, computation should be based on the dose, instead of energy (to be biologically correct) nor the mapping of the SF directly (to reduce uncertainties). For the accumulation, however, computations need to be based on the SF directly by following the effect accumulation approach to derive the mathematically correct survival fraction; or by calculation of the total biological dose bEQD.
Within this work, biologically consistent dose mapping is treated as an already given requirement. This is because common applications, like the ones used here, are implemented in a way to conserve the gray value of the voxel, given as dose in case dose distributions are being mapped. Thus, the common approach of dose mapping is implemented in a dose-conserving manner, though these implementation were not based on biological considerations. A change in this concept was recently postulated and is being used more commonly by now. However, in this work, a conclusion was assessed on a biological level and it was derived that the commonly used approach needs to be utilized, instead of changing it to energy-conservation. Therefore, the following chapters do not focus on the dose mapping aspect but on the dose accumulation approaches of bEQD and D_a. In contrast to dose mapping, the application of conventional dose accumulation is commonly performed in the described biologically systematically wrong manner. The magnitude of the deviation between bEQD and D_a is the focus of this work in theoretical analysis as well as in patient data. The impact of different aspects of effect accumulation (e.g. fractionation scheme, α/β, uncertainty assessment) were analysed in detail as well as statistical a assessment of the total dose deviation and dose-toxicity correlation based on bEQD.

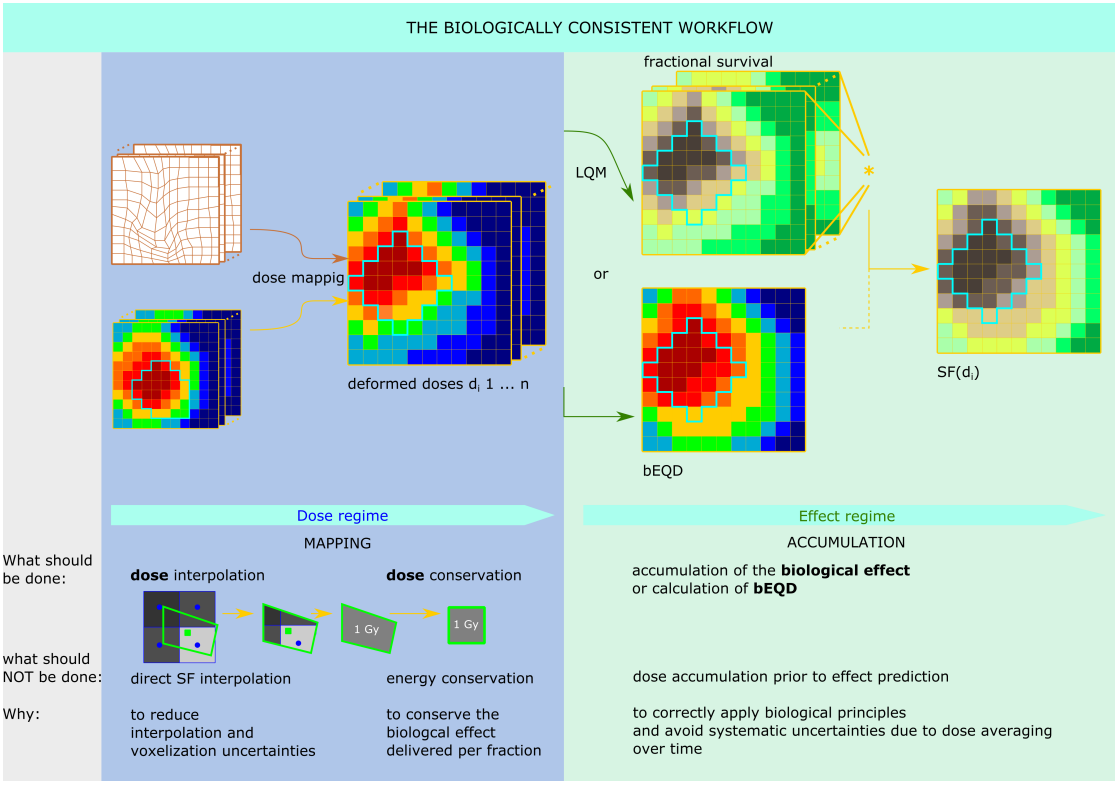


Figure 3.6. – Summary of the biologically consistent dose accumulation using either the effect accumulation approach or bEQD. In the first part of the data-post-processing, dose mapping, computations should be based on dose (not energy). For accumulation, computations have to take into account the non-linear relation of dose to the biological effect.

3.2 DATA & WORKFLOWS

This section describes the tools and data used for analysis in this work. Systematic investigations were based on a theoretical analysis of the dependency between dose variation magnitude and its impact on the different dose accumulation strategies (total biological dose (bEQD) and D_a), described in section 3.2.1. An overview of the used patient data is given in section 3.2.2. Sections 3.2.3 and 3.2.4 describe the analysed patient cohorts and the details on how the methodologies presented in section 3.1 are implemented for the respective workflows along with all tools and methods used in this process.

3.2.1 *Theoretical analysis*

The first part of investigation of biological dose accumulation was performed on a theoretical fictional voxel undergoing certain magnitudes of daily dose variations (as also submitted for publication [58]). This analysis was performed to be able to conclude on the magnitude of impact from effect accumulation and bEQD in comparison to conventional dose accumulation to assess resulting difference magnitudes and focus points of further investigations. All computations were performed in **R**[a]. Results on the theoretical analysis are presented in section 4.1.
The simulated fractionation scheme was chosen analogously to the used German Cancer Research Center (DKFZ) patient data (see next section 3.2.3): standard fractionation of 76.5 Gy in 2.25 Gy per fraction; moderate hypofractionation with 63 Gy in 3.0 Gy per fraction (equivalent dose to 75.9 Gy with an α/β-value of 1.4 Gy). Four different scenarios were chosen: a voxel in a high and in a medium dose region, for standard and for hypofractionation, respectively, giving the following fictional voxel mean doses: standard high = 2.25 Gy x 34 fractions, standard medium: 1.5 Gy x 34 fraction, hypo high: 3.0 Gy x 21 fractions, hypo medium: 2.25 Gy x 21 fractions. The investigations focused on the dependencies of the deviation between bEQD and D_a (see section 3.1.3) on the magnitude of dose variations, the assigned α/β-value, mean dose and the fractionation scheme. Fractional voxel doses were chosen manually to guarantee the desired mean dose and dose variation magnitude. The dose variation magnitudes, given as the standard deviation are chosen to be 0%, 12.5%, 25% and 50% with respect to the mean dose, where bEQD with 0% dose standard variation is equal to the conventionally accumulated dose D_a.

[a] **R** Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org. It is used in this work for statistical computation and generation of graphics and plots. Version 3.3.1 was utilized, scripts were written using RStudio Version 0.99.903.

3.2.2 *Patient data overview*

Three different patient cohorts are investigated: two cohorts of patients treated for prostate carcinoma (one acquired at DKFZ, one acquired for the VoxTox programme, Cambridge, UK) and one dataset of patients treated for head and neck (H&N) cancers (VoxTox programme). While the DKFZ pelvis cohort and VoxTox H&N cohort were given as three-dimensional (3D) data, for the VoxTox rectum data, dose surface map (DSM) of the rectum alone were available and analysed. An overview of the used cohorts and their focus within this work is shown in table 3.1 along with the data source. The main focus is set on the pelvic data. The pelvic region is prone to sudden and undirected motion caused by gas in the intestine that can potentially impair the planned dose distribution significantly.
Each cohort is used in this work to focus on specific aspects in the investigations of biological dose accumulation. The DKFZ pelvis data (section 3.2.3) was used for an in-depth analysis of the used strategies and workflows for the individual cases: detailed comparison to planned dose (γ-index analysis, dose-volume analysis, comparison to current dose-constraint recommendations), treatment schemes (normal vs. hypofractionation), uncertainties (deformable image registration (DIR), α/β-value and interpolation) and volumetric analysis (local vs. dose-volume histogram (DVH), whole organ volumes vs. organ walls). The VoxTox rectum cohort (section 3.2.4) constitutes a larger dataset and was therefore used to conclude on statistical outcome. Based on the rare toxicity data available from this cohort, dose-toxicity correlation was performed in comparison to published analysis on this data that was using conventional dose accumulation. The VoxTox H&N data (section 3.2.4) were used to compare the results from the pelvis region with another body site showing different motion patterns. It also serves for comparison between daily imaging strategies.

Table 3.1. – Overview of the analysed patient data cohorts and their focus in this work. 1: DKFZ, Heidelberg. Ethics approval S-380/2017 by the Ethics Committee of the Medical Faculty of Heidelberg, Germany; 2: Cambridge University Hospital NHS Foundation Trust, UK, Ethics approval East of England - Essex Health Research Committee (13/EE). Intensity-modulated radiotherapy (IMRT); computed tomography (CT); mega voltage computed tomography (MVCT)

Cohort	**DKFZ pelvis**	**VoxTox Rectum**	**VoxTox H&N**
Site	Prostate carcinoma	Prostate carcinoma	Various H&N cancers
Data type	3D data (pelvis)	DSM (rectum)	3D data (H&N)
Analysis focus	Impact of treatement and post-processing strategy	Statistics and dose response modelling	Comparison to other cancer sites
Imaging modality	In-room kV-CT	MVCT	MVCT
Treatment modality	IMRT - step-and-shoot	IMRT - TomoTherapy	IMRT - TomoTherapy
Source	DKFZ[1]	VoxTox programme[2]	VoxTox programme[2]
number of cases	9	251	26 (investigated)
Results sections	general: 4.2.1; fractionation: 4.2.2; local vs. volume: 4.2.3; diff. to plan: 4.2.4; uncertainties - DIR, α/β, interpolation: 4.2.5	general & diff. to plan: 4.3.1; toxicity corr.: 4.3.2	general: 4.4

3.2.3 *The DKFZ data and workflow*

An overview of all workflow steps and acquired results is given in figure 3.7. The results from the "DKFZ pelvis" cohort are presented in section 4.2.
The investigation on biologically consistent dose accumulation with the DKFZ pelvis cohort were based on data from a clinical study by Bostel et al. [8] (Ethics approval S-380/2017 by the Ethics Committee of the Medical Faculty of Heidelberg). The data had been acquired previous to this work. Within this thesis, the post-processing of the data for conventional and biological dose accumulation was set up, in parts in a collaboration with the department of Medical Image Computing at DKFZ, described below. Additional analysis regarding fractionation, volumetric analysis and uncertainties was performed.

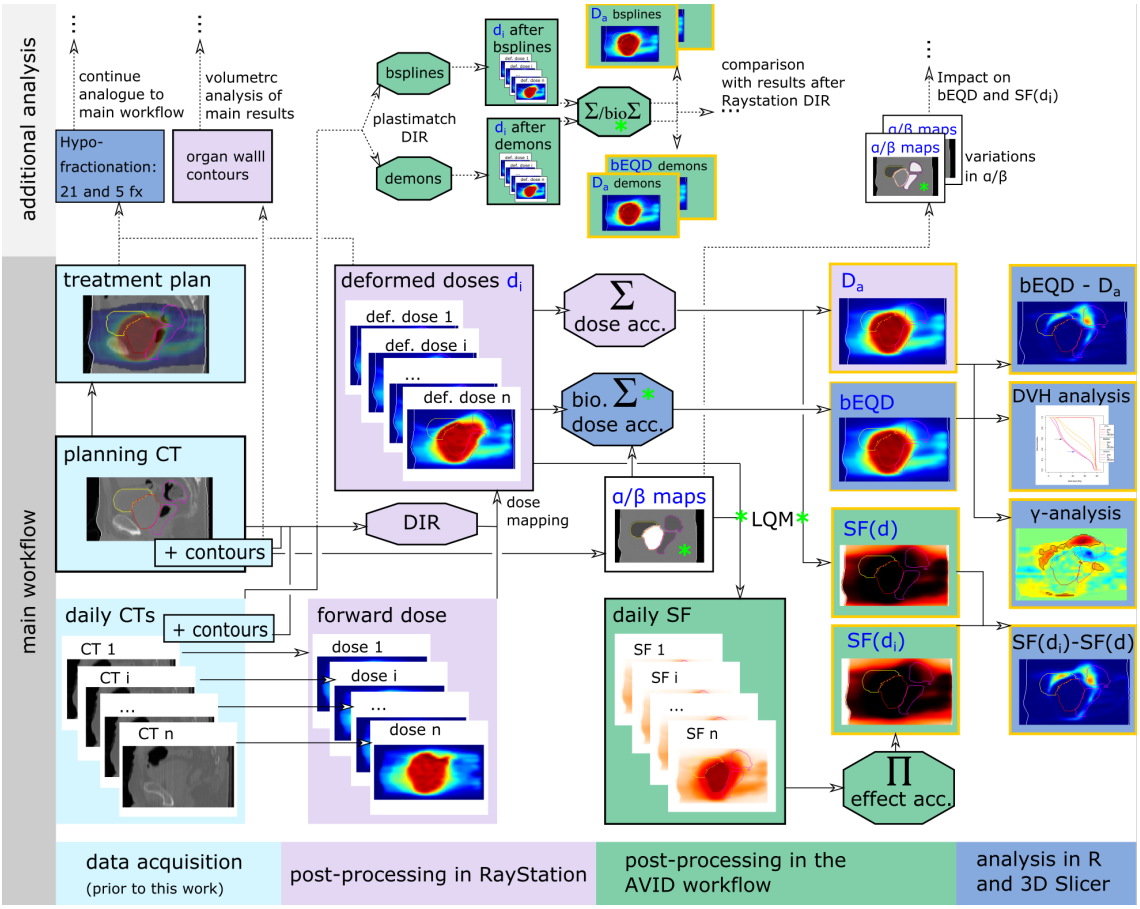


Figure 3.7. – Workflow followed for data acquisition, post-processing and analysis with the DKFZ pelvis cohort. Boxes framed black represent data in the reference geometry of the planning CT. Boxes framed yellow are final results that are being presented in the results section 4.2.

Patient cohort "DKFZ pelvis" The cohort analysed in this work consists of nine patient cases that were treated for prostate carcinoma at DKFZ. Acquisition of the data was done prior to this work for analysis of the daily delivered dose and focus on re-treatment strategies. A first analysis of this cohort was recently published by Bostel et al. [8]. In cooperation with the clinicians in charge of the study, the patient data were retrospectively used for analysis of biological dose accumulation in this work. The study was approved by the Independent Ethics Committee of the Medical Faculty of the University of Heidelberg, Germany (S-380/2017).
Patients were scored with intermediate-risk prostate cancer stage T1c to T2b. 6 MV IMRT step-and-shoot treatment (background see section 2.1.2) was performed with 9 coplanar fields (Artíste linear accelerator, Siemens, Erlangen, Germany). Treatment plans were created using the **RayStation**[b] treatment planning system (TPS). The Clinical Target Volume (CTV) was defined by the prostate gland with an additional 7 mm margin for the Planning Target Volume (PTV). For standard fractionation,

[b]The **RayStation**® is a treatment planning system by RaySearch Laboratories (Stockholm, Sweden). Version 6.1.1.2 was utilized for computations in this work.

a prescription dose of 76.5 Gy was chosen, delivered in 34 fractions of 2.25 Gy. Three of the nine patients were treated for the prostatic fossa after surgery with 34 fractions of 2 Gy. Dose constraints followed the Quantitative Analyses of Normal Tissue Effects in the Clinic (QUANTEC) recommendations [51].

Data acquisition Daily CT-imaging with an in-room CT scanner (Primatom; Siemens OCS, Malvern, USA) was performed in diagnostic quality in treatment position prior to each treatment session (background on imaging can be found in section 2.1.3). The CT-scanner itself is approved for treatment planning scans.
For each fraction, the target and organ(s) at risks (OARs) were delineated by board-certified radiation oncologists. For analysis of the dose to the OARs, the whole organs of bladder and rectum was used in the standard workflow. Further analysis for this work was performed for the bladder and rectum walls. Based on the former delineated organs, the organ wall contours were created in an approximate manner by an isotropic subtraction of a 3 mm smaller organ delineation to obtain organ wall contours (performed in RayStation).

Computation of the daily deformed doses For computation of the treatment plan as well as forward calculated of the daily doses (see section 2.1.1), the CT scans were imported in the RayStation TPS. An in-house Hounsfield look-up table (HLUT) (DKFZ-HLUT) was used for dose calculations on the basis of CT Hounsfield Units (HU) (see also section 2.1.1).
The RayStation TPS was used to perform all basic computations on the acquired data to compute the daily deformed doses d_i and the conventionally accumulated dose D_a. The following steps were performed after data acquisition during patient treatment: dose forward calculation on the basis of the daily image using the dose plan, DIR, dose deformation and dose summation.
For DIR, the *hybrid intensity and structure based* option was chosen which is based on the ANACONDA (anatomically constrained deformation algorithm) algorithm [77]. The algorithm is both intensity-based and additionally geometry-based by the organ contours (region of interests (ROIs)) in the target and moving image (called *controlling ROIs*). The objective function to this algorithm consists of four metrics in a linear combination that is to be minimized: image similarity (measured with a correlation coefficient), smoothness of the grid (grid regularization), a shape based regularization and a penalty term (the latter two referring to the *controlling ROIs* in the images). Background information on DIR can be found in section 2.2.1.3. For all patients and fractions, the contours of the bladder, the rectum and the PTV were selected as *controlling ROIs*. DIR was performed with the planning CT as the target image which is then used as basis for dose accumulation (reference frame).

Computation of α/β-maps Both variations to compute the total biological effect ((cell) survival fraction (SF)) require the underlying information of the α/β-

value of the individual tissues. For this, α/β-maps were created based on the delineated organ structures in the reference frame of the planning CT. From the contours contained in the structure-sets, binary maps for the individual organs were created using the in-house developed *VoxelizerTool*, an implementation in the RTToolbox [88]. In this tool, voxels within the delineated contour are assigned a value of 1, while everything else is set to 0. The binary organ maps were then imported into **3D Slicer**[c] and by using the built-in Python framework, α/β-maps were created containing the values for the CTV, bladder and rectum as well as all other tissue in one map in the geometry of the reference frame.

Since the α/β-values of the individual organs show a range of uncertainty, the impact of this uncertainty is studied for the result of the total biological effect and the total biological dose. This was done by assigning a range of different values to the tissues based on reported ranges in literature. Details on the α/β-value and literature references can be found in section 2.2.2.1. For the ROIs used in this work, the range of studied values is shown in table 3.2. Clinical standard values[d] shown in bold are used as standard values assigned in all cases expect for analysis with focus on α/β-variation. α was set to the constant value of $0.1\ \mathrm{Gy}^{-1}$ and the variation in α/β was achieved by only changing β. To study the dependency on α, results were also compared to those given from $\alpha = 0.2\ \mathrm{Gy}^{-1}$ with the same α/β ratios.

Table 3.2. – α/β-values used for the respective tissues for computation of $\mathrm{SF}(d)$, $\mathrm{SF}(d_i)$ and bEQD, based on literature reported values (section 2.2.2.1). Bold: clinical standard values

CTV	**1.4**, 10
rectum	2, **3**, 4, 5
bladder	**5**, 10
all else	**3**

Accumulation Conventional dose accumulation was performed in the RayStation system. D_a and the individual d_i were exported from the RayStation anonymized for further processing. Calculation of the total survival fraction from conventional dose accumulation and from effect accumulation were performed using the **AVID**[e] framework. Within this work, the AVID framework was extended to calculate the survival fractions from the two respective approaches as described in section 3.1.2. The biological effect based on the accumulated dose D_a is using equation 3.1. The total effect from effect accumulation based on the individual d_i is obtained using equation 3.2. The required data were: the deformed doses d_i as output from the RayStation and the computed α/β-maps.

[c] **3D Slicer**: an open source software platform for medical image informatics, image processing, and three-dimensional visualization. Version 4.5.0-1 was used [22] along with the SlicerRT extension Version 0.17.1.; https://www.slicer.org/

[d] personal communication with PD Dr. Juliane Hörner-Rieber, radiation oncologist at the University Hospital Heidelberg

[e] **AVID**: Analysis of Variations in Interfractional RaDiotherapy. An automated software framework to allow handling of large datasets, in-house development at DKFZ.

bEQD as described in section 3.1.3 was computed using the same d_i and α/β-maps. Calculations of bEQD were performed in R. bEQD was then compared to the conventionally accumulated dose D_a as output from the RayStation.

Hypofractionation The same patient cohort was used to restrospectively calculate hypofractionated treatment plans. Two hypofractionation strategies were chosen[f]. First, the in-house standard approach for moderate hypofractionation was used, given by 3 Gy per fraction. Using an α/β-value of 1.4 Gy for the prostate (to be consistent with further analysis), the standard equivalent dose in 2 Gy fractions (EQD_2)-based conversion was performed (section 2.2.2.2). This resulted in a treatment scheme of 21 fractions of 3 Gy (total dose 63 Gy, equivalent to 75.9 Gy). For patients treated for the prostatic fossa, 21 fractions of 2.7 Gy were chosen. For strong hypofractionation, a number of five treatment fractions was fixed. Using the same conversion principle, this resulted in a treatment scheme of 6.8 Gy per fraction (6.1 Gy for the fossa) over five fractions, giving a total of 34 Gy (equivalent to 76.4 Gy).

Hypofractionated plans were calculated by rescaling of the initial standard fractionated dose plans, performed in R. Dose constraints to the OARs were converted according to EQD_2 with the respective α/β values. For all plans, the desired constraints were fulfilled after rescaling. Rescaling of the dose for further analysis was performed on the deformed doses $D-i$ after export from the RayStation.
In order to simulate realistic treatment conditions, the first 5 and 21 fractions, respectively, for each patient were rescaled and analysed. Optimization of new treatment plans with hypofractionation doses was not performed since this would potentially impair comparability of the results, as will be discussed later. The workflow and analysis using the rescaled deformed doses was performed analogously to the described standard workflow.

Dose-volume histograms (DVH):
In a *cumulative* DVH, all voxels that receive at least a certain dose x correspond the according percentage volume on the y-axis. Each point on the curve can be calculated by:

$$DVH(D) = 1 - \frac{1}{V}\int_0^{D_{max}} \frac{dV(D)}{dD} dD \tag{3.13}$$

For example, the point on the DVH-curve at an x-value of 50 Gy represents the volume including all voxels that received at least 50 Gy. This point would be referred to as V_{50Gy}. Analogously, $D_{95\%}$ describes the dose to 95 % of the volume. $D_{95\%}$ therefore describes the dose coverage of the structure while $D_5\%$ describes the high dose areas within the volume. It should be noted that spatial information is lost in the computation of a DVH and all dose-volume parameters are reduced to a volumetric summary.

[f] after personal communication with PD Dr. Juliane Hörner-Rieber, radiation oncologist at the University Hospital Heidelberg

Analysis The survival fraction after dose and effect accumulation were compared as the absolute and relative cell survival differences. Analogously, the absolute dose difference between bEQD and D_a was analysed. The differences were investigated on a voxel-scale. All computations were performed in R. Visualization of the results was realized using **MITK**[g]

γ-index analysis: This method was initially described by Winiecki et al. [79]. It analyses dose difference as well as spatial displacement between two points, usually given for a voxel in a reference frame compared with another dose distribution. For all points $\vec{r}_c$ in the reference frame γ is calculated by:

$$\gamma(\vec{r}_c, \vec{r}_m) = \sqrt{\frac{|\vec{r}_c, \vec{r}_m|^2}{DTA^2} + \frac{|D(\vec{r}_c) - D(\vec{r}_m)|^2}{\Delta D^2}} \tag{3.14}$$

with the distance between the analysed points given by $|\vec{r}_c, \vec{r}_m|$, the dose difference by $|D(\vec{r}_c) - D(\vec{r}_m)|$ and DTA (distance to agreement) and ΔD (dose difference) representing scaling factors chosen as the $\Delta D/DTA$ criterion. The final $\gamma(\vec{r}_c)$-value is given by the minimum of $\gamma(\vec{r}_c, \vec{r}_m)$. The γ-criterion is passed for values below 1. In common practice, a γ-passing rate is reported, represented by the percentage of volume with $\gamma < 1$. Clinical standard for a passed γ-analysis is a γ-passing rate of 95%.

In addition to this local analysis, **Dose-volume-histograms** and dose-volume (DV)-parameters were compared between planned dose, conventionally accumulated and biologically accumulated dose. Cumulative dose-volume histograms were created in R (see box for more information on DVHs). DV-parameters were computed: for the CTV the $D_{2\%}$, $D_{5\%}$, $D_{95\%}$ and $D_{98\%}$; for bladder and rectum the V_{30Gy}-V_{75Gy} (standard fractionation). For moderate and strong hypofractionation, the respective V_{xGy} were converted using the EQD_2 with $\alpha/\beta = 5$ Gy for the bladder and $\alpha/\beta = 3$ Gy for the rectum. Dose constraints for bladder [75] and rectum [51], given by the QUANTEC reports, were analysed for the planned and the accumulate doses. These constraints were also converted to hypofractionation for comparison but it should be noted that these are not recommended for use. The analysis was based on the organ volumes defined by the contours in the reference frame. Additionally, the analysis of the whole organ volumes of bladder and rectum were compared with the dose values to bladder- and rectal-wall alone.

Parts of the results are presented in the form of Boxplots which show: the median dose as the central thick line, the interquartile ranges given by the coloured box, the minimum/maximum values ± 1.5 times the interquartile range as error bars and outliers as circles. Boxplots were created in R.

Comparison of bEQD and D_a with the planned dose, respectively, was performed in a **γ-Index analysis** (see box for more information). Passing rates for a 3%/3 mm

[g] **MITK**: The Medical Imaging Interaction Toolkit. An open source software for medical image processing [62]; Version: ITK 4.7.1, VTK 6.2.0, QT 5.4.1. www.mitk.org.

criterion were investigated. A lower dose cutoff threshold was set to 10%. γ-analysis was performed in 3DSlicer using the SlicerRT extension. Plots for visualization were created in MITK. In contrast to the standard process, the analysis was adjusted in two ways: 1) Usually, the analysis is given for the entire imaged volume of the patient. Hot spots of a failed γ-criterion in smaller volumes can be hidden in the large analysed volume. Therefore, the γ-analysis was performed for the individual OARs of bladder and rectum in this work. 2) The γ-value does not report on the direction in which the dose in the compared voxels differ. While for OARs, lower doses than planned are advantageous, a γ-value above 1 can be obtained which reduces the passing rate nevertheless. Therefore, in this work for all voxels with lower delivered than planned dose in the OARs, the γ-value was multiplied by -1. All voxels with $\gamma < 1$, including negative values are rated as passed.

DIR Uncertainties The analysis of uncertainties in conventional and biological dose accumulation inherited from uncertainties in DIR was performed in two ways: (I) In section 3.1.3, it is described how errors from DIR propagate mathematically to D_a and bEQD respectively. Since a direct quantification of errors in DIR is a not yet solved problem, as described in detail in section 2.2.1.3, a simplified approach was tested. A 3 mm error in DIR was assumed to each single voxel in the image. The value was chosen according to reported values in literature [3]. Within this 3 mm distance from each voxel, the highest dose difference to the center voxel was computed. This represents a worst case estimate for an error in the mapped voxel dose. This value was then used in the calculation of the respective uncertainties of D_a and bEQD as described before. Calculations were performed in R.
(II) A direct comparison of the outcome of biological dose accumulation for different registration algorithms was performed. For this purpose, DIR and dose mapping was integrated into the above mentioned AVID workflow using the **plastimatch**[h] DIR algorithm. The demons as well as the b-spline algorithms were used, respectively. The registrations are based on a three- and four-staged parameter file, respectively that can be found in the appendices B. No further guidance of the registration was performed (no landmarks or structure contours guidance).
Interpolation of dose values for dose mapping was based on trilinear interpolation. Background on DIR and interpolation is described in sections 2.2.1.3 and 2.2.1.4. For comparison, the sub-voxelization interpolation method by Rosu et al. [64] was investigated which is implemented in the AVID framework. Additional data required for the computation in AVID were the planning CT, the daily CTs and the forward calculated doses that are to be deformed. The results for D_a and bEQD were compared to those following the standard RayStation based workflow.

[h] **plastimatch**: an open source software for image computation with focus on image registration, Version 1.6.4. www.plastimatch.org

3.2.4 *The VoxTox programme*

The VoxTox programme was initiated to test the hypothesis *"that delivered dose provides a better predictor of toxicity than planned dose"* [13] by Neil Burnet and colleagues at the University of Cambridge (Cambridge University Hospital NHS Foundation Trust, UK, Ethics approval East of England - Essex Health Research Committee (13/EE)). Daily MVCT scans within the TomoTherapy unit used for treatment were acquired for the purpose of patient positioning guidance in the standard treatment workflow. These daily acquired scans were then used within the VoxTox programme to calculate the delivered dose of the day, perform conventional dose accumulation and correlate the result to the collected patient-reported toxicity. The objective of the collaboration project was to implement the derived biological dose accumulation (as described in section 3.1.3) within the workflow pipeline of the VoxTox programme for both pelvis DSM analysis and H&N cases. Statistical conclusions on the magnitude of difference between conventional and biological dose accumulation were aimed for. The impact of biological dose accumulation in comparison with conventional dose accumulation and planned dose in dose-toxicity correlation for the rectum was investigated.

Based on the large amount of data, one objective of VoxTox programme was to facilitate an automated workflow of data post processing for the correlation of dose to treatment outcome. Implementation and automation of steps included data transfer, auto-contouring (of the rectum), dose re-calculation on the MVCTs, conventional dose accumulation and dose-volume reporting, including automated control of the entirely pipeline.
The acquired data had three main advantages compared with the DKFZ cohort described above: a much larger number of cases is available; data exist for two different entities: prostate and H&N radiotherapy (RT); toxicity reports are available (minimum of 2 years follow up time, 4 year median for prostate cases). Together with the colleagues from Cambridge, a cooperation was initiated within this work to exploit the potential of this highly valuable dataset also for purposes of biological dose accumulation.

Patient cohort "VoxTox rectum" The **cohort** used for analysis in this work comprised of 251 prostate cancer cases from the study consolidation cohort for which daily dose data in the form of dose-surface maps (DSMs) was made available. A previous study by Shelley et al. [70] was based on 109 cases as a discovery cohort (different cohort) and more details on the data and the post-processing workflow can be found in the given publication. The workflow steps and choice of investigated dose measures was based on this work in order to compare the results of the bEQD analysis to the existing conclusions. Patients were treated with TomoTherapy (see section

2.1.2) in either 37 fractions with a total dose of 74 Gy (136/251) or with moderate hypofractionation of 20 fraction with a total dose 60 Gy (115/251)[i]. For analysis of the hypofractionated cases, the dose was rescaled to the equivalent of 37 fractions with the standard biologically effective dose (BED) approach (see section 2.2.2.2). Daily imaging was performed using MVCT. Calculation of the daily doses followed a Cambridge in-house developed ray-tracing algorithm [72] (more information on dose calculation is given in the background section 2.1.1). The contours of the rectum were generated automatically (due to the large amount of data) on all daily MVCT scans and manually on the planning CT scans. MVCT scans did not cover the same volume in all fractions but were sometimes shifted towards head or feet direction. Follow-up toxicity data were not available for all 251 cases for all endpoints. An overview of the available toxicity data and incidence rates is given in table 3.3. The total biological dose (bEQD) approach was implemented into the VoxTox workflow to compare the results for the conventionally and the biologically accumulated doses. The workflow was implemented in **Matlab**[j]. Figure 3.8 illustrates the workflow and where it was complimented for purpose of this work.

Table 3.3. – Available rectum toxicity data analysed in this work. G1 and G2 refer to grade 1 and grade 2 toxicity scaling as reported in [70].

endpoint	incidence rate [%]	available cases
proctitis (G1)	17.7	136
proctitis (G2)	13.2	136
rectal bleeding (G1)	33.1	136
rectal bleeding (G2)	11.8	136
bowel bother (G4)	39.1	138

The provided workflow contained the following: automated segmentation of the rectum contours; generation of dose-surface maps (dose surface maps (DSMs)) of the rectal wall along with conventional dose accumulation; extraction of dosimetric features. **DSM generation** was based on algorithms described in literature by Buettner et al. [12] and Murray et al. [55]. A DSM is a two-dimensional unfolding of a structure, in this case the rectum, considered as a cylinder. For this, the rectum wall was virtually cut on the posterior side of the center of mass and represented as a two-dimensional (2D)-image. The height is in this case defined by the number of slices (3 mm thickness on the planning DSM and 6 mm on the daily mega voltage computed tomography (MVCT) DSMs) and the width was scaled to equal the height. Daily MVCT scans were restricted in length (number of slices) and did not necessarily match the planning DSM height. In these cases, missing data were substituted with planned DSM data. As a dose summarizing measure, the equivalent uniform dose (EUD) was calculated for the planned and both accumulated dose versions with a = 11.11 (assuming a serial-type structure of the rectum, see

[i] Recruitment details for the used data from the CHHiP trial can be found here: https://www.icr.ac.uk/our-research/centres-and-collaborations/centres-at-the-icr/clinical-trials-and-statistics-unit/clinical-trials/chhip

[j] **Matlab**: Release R2016a, The MathWorks, Inc., Natick, Massachusetts, United States

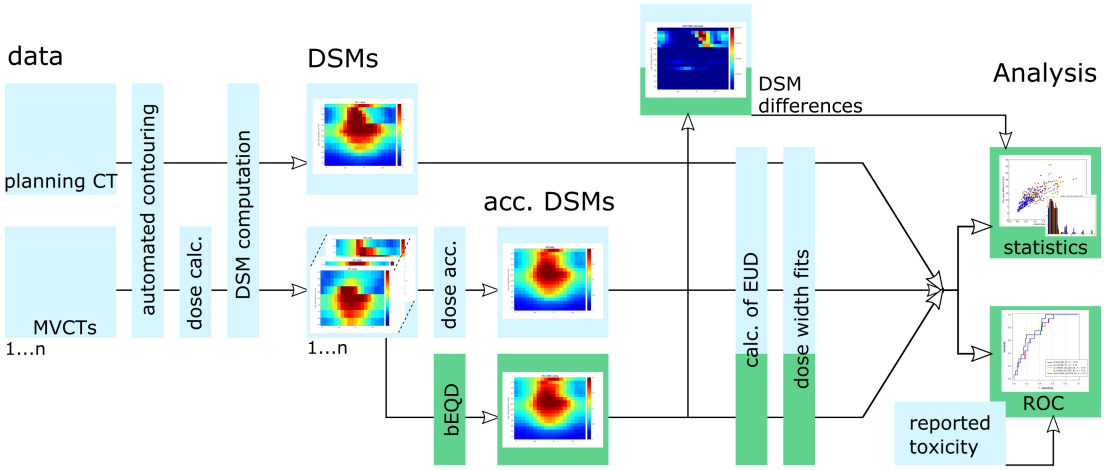


Figure 3.8. – Workflow of the analysis of the VoxTox rectum DSM data. Blue boxes show workflow and data provided by the VoxTox group, green boxes indicate complimentary steps added for bEQD analysis within this work. For calculation of the bEQD, α/β-values of 2 Gy, 3 Gy and 4 Gy were used for the rectum, analogously to the investigations for the DKFZ pelvis cohort, and results were compared respectively. 3D data from prostate TomoTherapy (planning CT and daily MVCTs) was used. MVCT rectum contours were generated automatically. Daily DSMs were used to generate the accumulated DSM dose D_α and additionally the code was adapted to calculate the bEQD on the same basis. No image registration was performed to map the images on a voxel scale, this is solely based on the contouring of the rectal wall on the daily scans. The equivalent uniform dose (EUD) and dose widths were computed for all DSMs. Results were analysed statistically and using receiver operator characteristics (ROC) analysis to investigate correlations of dose to toxicity.

equation 2.15 and details in section 2.2.2.2). Furthermore, the **'dose-width'** was fitted to the DSMs based on the method reported by Buettner et al. [12]. For this, an ellipsoidal-fit was performed based on a certain iso-dose line. Dose data were therefore transformed to binary maps for certain dose thresholds and an ellipse was fitted to the largest central cluster. The maximum lateral extend relative to the total DSM width of certain iso-dose levels was used as the respective dose-width. The dose-width was calculated for iso-dose levels of 30, 40, 50, 60, 65 and 70 Gy.

The VoxTox rectum DSM cohort allowed for **statistical conclusions** on the impact of biological dose accumulation due to the 251 analysed patient cases in contrast to the low number of cases from the German Cancer Research Center (DKFZ) pelvis cohort. As described above in section 3.1.3, the difference between bEQD and D_α is systematic and both accumulated doses built up on the same daily deformed dose data. Therefore, statistical tests for significant difference between these results is neither valid nor expedient. The comparison between both methods was thus done by descriptive statistics of dose measures computed in R.

ROC curve analysis: The Receiver operator characteristics (ROC) curve was calculated to determine correlations between dose and treatment outcome. The ROC curve itself is a measure for separability between classes. Here, the two classes being compared are the ones reporting the certain toxicity, and the ones without. The ROC curve is plotted as sensitivity vs. 1-specificity. The sensitivity is a synonym for "true positive rate" given by: true positives / (true positives + false negatives). 1-specificity is given by the false positive rate: false positive / (true negative + false positive).
The measure reported to quantify magnitude of correlation is the area under the curve (AUC), which is given as a number between 0 and 1. For an AUC of 0.5, the correlation of input and output is deemed to be purely by chance, thus, there is no correlation. AUC values differing by greater or lesser amounts from 0.5 can provide stronger of weaker evidence for correlation.

Correlation of rectum toxicity to delivered dose (D_α and bEQD) was done by performing a **ROC-curve analysis**. The result of this is to assess whether individual dose-volume parameters can be predictors for different reported toxicities. This analysis was performed for comparability to the already published results from this VoxTox cohort by Shelley et al. [70]. The analysis was performed in **SigmaPlot**[k].

Patient cohort "VoxTox H&N" Within this thesis, the investigation of the head and neck (H&N) cases serve the purpose of comparing results in the pelvis region to another body site to be able to estimate the impact of biological dose accu-

[k] **SigmaPlot** is a scientific data analysis and statistics software (Version 12.0), Systat Software, Inc., San Jose California USA, www.systatsoftware.com.

mulation in radiotherapy (RT) in a more generalized way. Additionally, the dataset enabled investigations of three-dimensional (3D) data, like for the DKFZ data, but using different applications in the workflow. Therefore, this section is not part of the central analysis of this work but for complimentary analysis and conclusions.
For the H&N **cohort**, about 200 cases where available that were treated for various cancer sites in the head and neck region. However, for more detailed analysis, the focus of investigations was put to dose to brainstem and spinal cord. As will be discussed later (section 5.2), this is because of the rather serial than parallel nature of these organs for which the bEQD is more relevant. Therefore, a test sub-cohort of patients with dose plans that might indicate increased doses in the respective regions was studied, resulting in 26 cases for investigation (6 cases of potentially increased brainstem dose, 23 cases for spinal cord).
The workflow followed for dose accumulation in the H&N data is essentially the same as presented in section 2.2.1.2 for conventional dose accumulation and described above for the DKFZ cohort. Patients were treated with TomoTherapy and underwent daily image-guidance. In comparison with the methods described for the DKFZ workflow, main differences were: a) the used modality for daily imaging and b) the registration algorithm. For a), MVCT imaging was used on a daily basis to estimate the daily delivered doses. As described for the DSM data, dimensions of the MVCT were restricted in length (number of slices) compared with the planning scan and might have varied from day to day. Due to the use of MVCT, dose-recalculation on the daily images was required using the same algorithms as described above [72]. For b), the used image registration algorithm is based on the **Elastix**[1] software package [38][69]. Within the Elastix implementation, mutual information between the moving and fixed image are maximized, using stochastic gradient descent methods. A first affine registration was followed by a cubic-bspline transform (background on deformable image registration (DIR) see section 2.2.1.3). More details in the appendix B.
The provided workflow pipeline worked on the already deformed daily doses (DIR + interpolation) and using the manually generated organ structure sets containing the contours. The data post processing pipeline within the VoxTox programme was set up in Python. The provided dose accumulation pipeline was complimented by integrating the effect accumulation and bEQD formalism for this work. The assigned α/β values for the respective tissues were: spinal cord - 2 Gy, parotids - 3 Gy, brainstem - 2.5 Gy, brain - 2.5 Gy, Clinical Target Volume (CTV) - 10 Gy, all else - 3 Gy ([34] and values as used within the clinical workflows of the VoxTox programme). Analysis was performed on the 3D data and for volumetric parameters of the structures of interest using the organ contours.

[1] **Elastix:** open source software toolkit for image segmentation and registration. http://elastix.isi.uu.nl/

4

RESULTS

This chapter is organised by the analysed data: theoretical fictional voxel analysis (section 4.1), the DKFZ pelvis cohort (section 4.2), the VoxTox rectum dose surface map (DSM) cohort (section 4.3) and the VoxTox head and neck (H&N) cohort (section 4.4). An overview of the data can be found in the materials & methods chapter in table 3.1. The results are partially submitted for publication as indicated in the respective sections [58].

4.1 THEORETICAL PREDICTIONS

The results as presented in this section are partially submitted for publication [58]. The impact of the dose per fraction and the assigned α/β value on the survival fraction is shown in figure 4.1 for a fictional voxel receiving conventional and hypofractionated irradiation dose with different daily dose variation magnitudes. It can be seen that the absolute (cell) survival fraction (SF) decreases strongly for lower α/β. Lines that correspond to a higher underlying α and β but for the same ratio α/β show lower SF . The differently dotted lines indicate the impact of the fractional dose variation as the dose standard deviation (std). The higher the dose variation magnitude, the lower the survival fraction according to the effect accumulation approach. For the conventional dose accumulation approach, there is no difference since only the mean dose is considered ($d = D_a/n$), which is the same in all cases. The absolute difference between the effect accumulation approach $SF(d_i)$ and the conventional approach $SF(d)$ is higher for higher α/β, while the relative difference increases strongly for lower α/β. For 12.5% std, the relative SF difference is low (below 10% difference) for α/β above 3 Gy, but can reach values up to 20% for α/β below 3 Gy. At 25% std, these values already increase to 20-80% and even higher for 50% std where the relative difference can reach 100%. In case of hypofractionation (figure 4.1, right side), the SF curve is steeper and tissue with higher α/β is less affected (note the different axis but that at 1.4 Gy, both curves show the same SF). However, the relative SF difference is very similar and only slightly increased compared with the standard fractionation case.

Figure 4.2 shows the results on total biological dose (bEQD) for the analysis of the same fictional voxel calculations as above. The emphasis here is put more onto the impact of the delivered dose and fractionation scheme involved. Again, three dependencies of bEQD can be seen: bEQD changing with the dose variation magnitude, given as the dose standard deviation (std), the underlying α/β-value and the

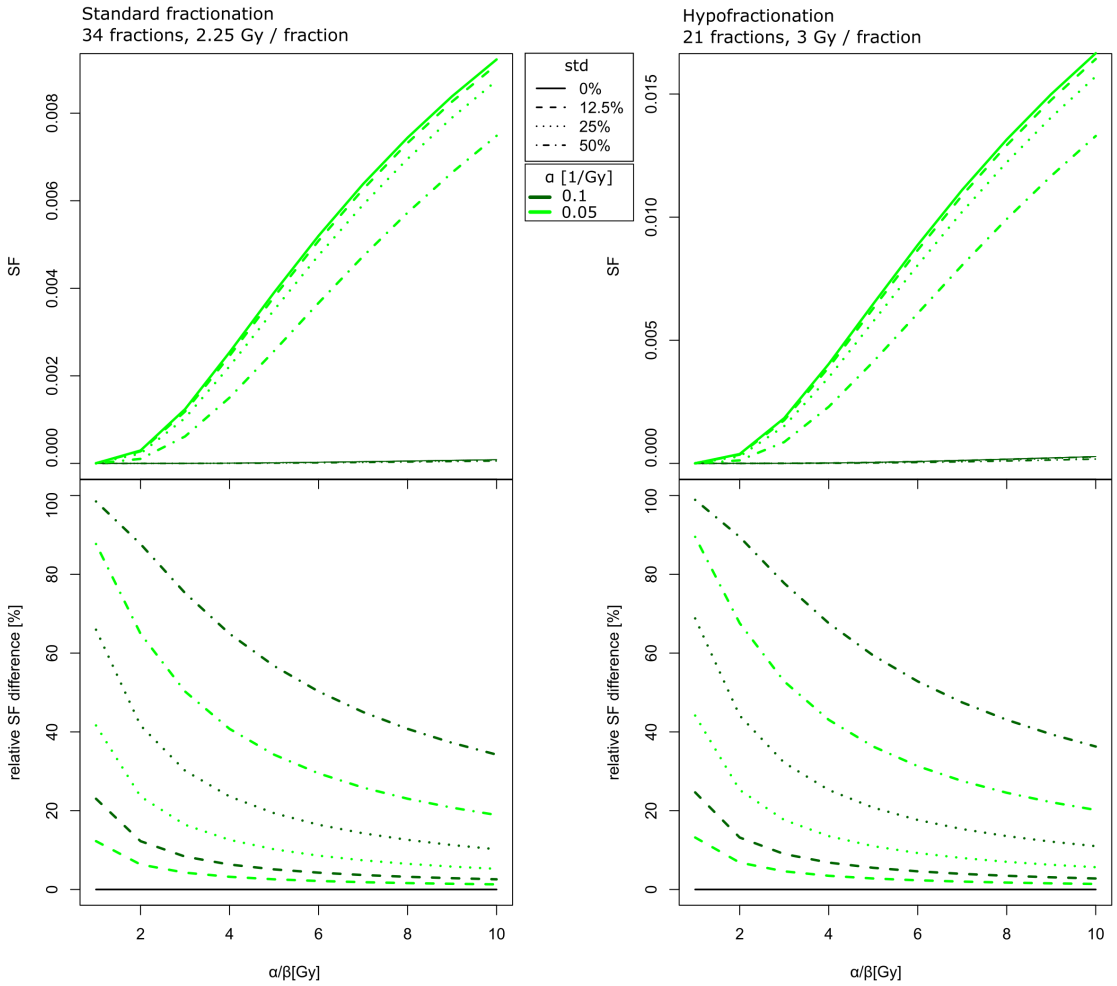


Figure 4.1. – Theoretical analysis of the SF based on the effect accumulation approach $SF(d_i)$ (Eq. 3.10) (top row) and the relative SF difference to conventional accumulation ($SF(d)$, Eq.2.9), bottom row, for a fictional voxel receiving standard (left) or hypofractionated (right) irradiation with different magnitudes of dose variation. Both fractionation schemes are biologically equivalent at 1.4 Gy (equivalent dose in 2 Gy fractions (EQD_2) -based conversion, section 2.2.2.2). The dose variation magnitude is given as the standard deviation of the fractional doses d_i in percent with respect to the mean dose d. Note that the $SF(d_i)$ with $std(d_i) = 0\%$ is the same as $SF(d)$. Light green curves are calculated with $\alpha = 0.05\ Gy^{-1}$, dark green curves with $\alpha = 0.1\ Gy^{-1}$.

fractionation scheme (number of fractions and the mean delivered dose d). The difference to the base line referring to a std of 0%, so that bEQD and D_α are equivalent, visualizes the total dose difference between both approaches. The difference between

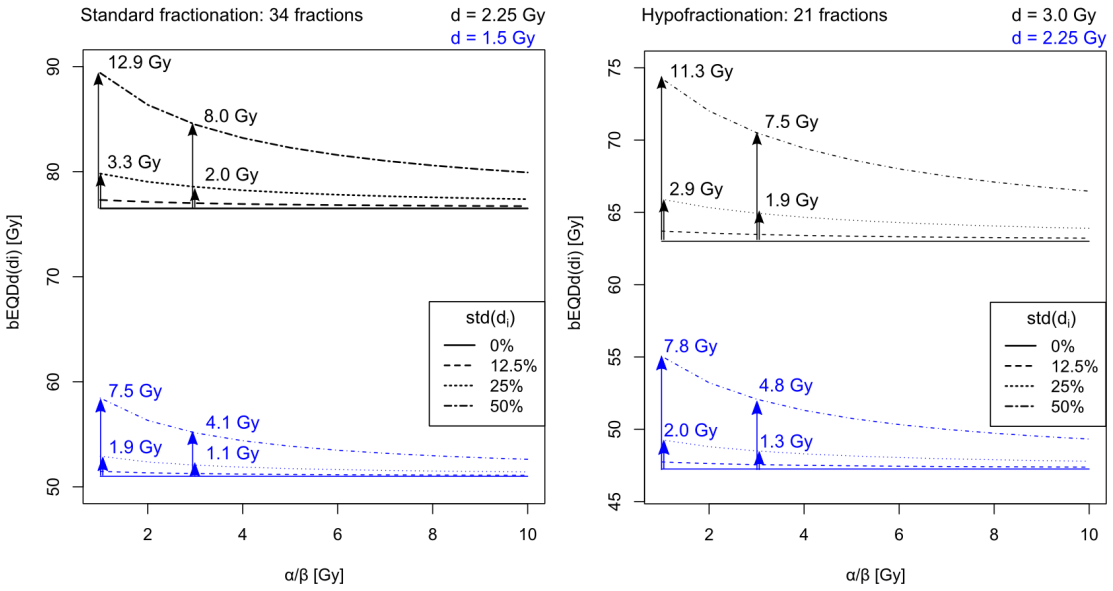


Figure 4.2. – Biologically accumulated dose bEQD for a fictional voxel receiving standard (left) or hypofractionated (right) irradiation with different magnitudes of dose variation. The dose variation magnitude is given as the standard deviation of the fractional doses d_i in percent with respect to the mean dose d. Note that bEQD with $\mathrm{std}(d_i) = 0\%$ equals D_α. Black curves represent the voxel in the respective high dose regions of 2.25 Gy (34 fractions) and 3.0 Gy (21 fractions) that are biologically equivalent at 1.4 Gy (EQD_2 -based conversion, section 2.2.2.2). Blue lines represent voxels in the respective fractionated treatment in a medium dose regime of 1.5 Gy or 2.25 Gy. Numbers in the plot are given to show dose difference to the 0 % dose variation case (D_α). For less than 12.5 % dose variation, the difference remains below 1 Gy.

bEQD and D_α strongly depends on the dose variation magnitude. At a standard deviation below 12.5% of the mean physical dose, the difference between bEQD and D_α remains below 1 Gy. For high dose variation magnitudes of 50% with respect to the mean physical dose, the difference between the two accumulation methods can reach values of up to 12.9 Gy.

It can also be seen that the lower the α/β-value used in bEQD, the higher the difference between bEQD and D_α. For example the highest impact in the given example can be seen for a standard fractionated treatment in the high dose regime with the highest dose variation magnitude where the difference of bEQD ($\alpha/\beta = 1$ Gy) to bEQD ($\alpha/\beta = 10$ Gy) is 9.5 Gy. This dependency induces an uncertainty in the calculation of bEQD based on the uncertainty of the α/β-value which is analysed in the patient data results.

The curves in figure 4.2 show that for the same number of fractions, the higher the mean dose, the higher the difference between bEQD and D_α. Generally, the absolute differences are higher for the standard fractionated case. This is because the difference between D_α and bEQD accumulates over the course of treatment and

scales with the total delivered dose. Furthermore, it also scales with the dose per fraction d. Therefore, the difference between D_α and bEQD relative to the total delivered dose is higher for the hypofractionated case for the respective high and low dose regimes (e.g. 2.25 Gy x 34 fractions compared with 3.0 Gy x 21 fractions).

4.2 COHORT DKFZ PELVIS

The results for the DKFZ pelvis cohort show the analysis of three-dimensional (3D) data sets. Details on data and acquisition are given in section 3.2.2. Parts of the results as presented in this section are submitted for publication [58] and further complimented with additional details and analyses.
The first part of the analysis in section 4.2.1 focusses on the deviations found in the 3D data for the SF and total dose between conventional and biologically consistent accumulation. The impact of the choice of the fractionation scheme (standard fractionation, moderate and strong hypofraction) on the difference between both methods to accumulate the total delivered dose is shown in section 4.2.2. Section 4.2.3 investigates the dose to volume relation, given by condensing the 3D results into commonly used dose-volume parameters as well as by analysing the dose to organ walls (bladder and rectum) rather than total organs. A comparison of bEQD and D_α to the planned dose is shown in section 4.2.4. Finally, section 4.2.5 shows the results for uncertainty assessment in the accumulation process from deformable image registration, the assigned α/β values and the use of different interpolation algorithms.

4.2.1 *Survival fraction and total dose difference in 3D data*

Figure 4.3 shows the results of the comparison of the survival fraction from the dose and the effect accumulation approach for the nine analysed patient cases. As shown theoretically, the SF from effect accumulation is systematically lower than the results after conventional dose accumulation. Largest deviation were found in all cases in the gradient regions around the target. The magnitude is between 35 - 40% in 6/9 cases and reaches 50.7% for case 6) in a distinct hot spot. Areas of SF deviation are mainly located in the organ(s) at risk (OAR) in bladder and rectum and the space between those, while within the Clinical Target Volume (CTV), the difference remains marginal.

The analysis of the nine patient cases analysed with respect to their bEQD-D_α difference are shown in figure 4.4. As shown theoretically, bEQD is systematically higher than D_α. In all cases, larger areas with pronounced deviations in total dose between bEQD and D_α were found in the dose gradient regions around the target volume with values around 1-3 Gy, analogously to the SF difference. More distinct deviation hot spots could be found for 6 out of 9 patients within the bladder and

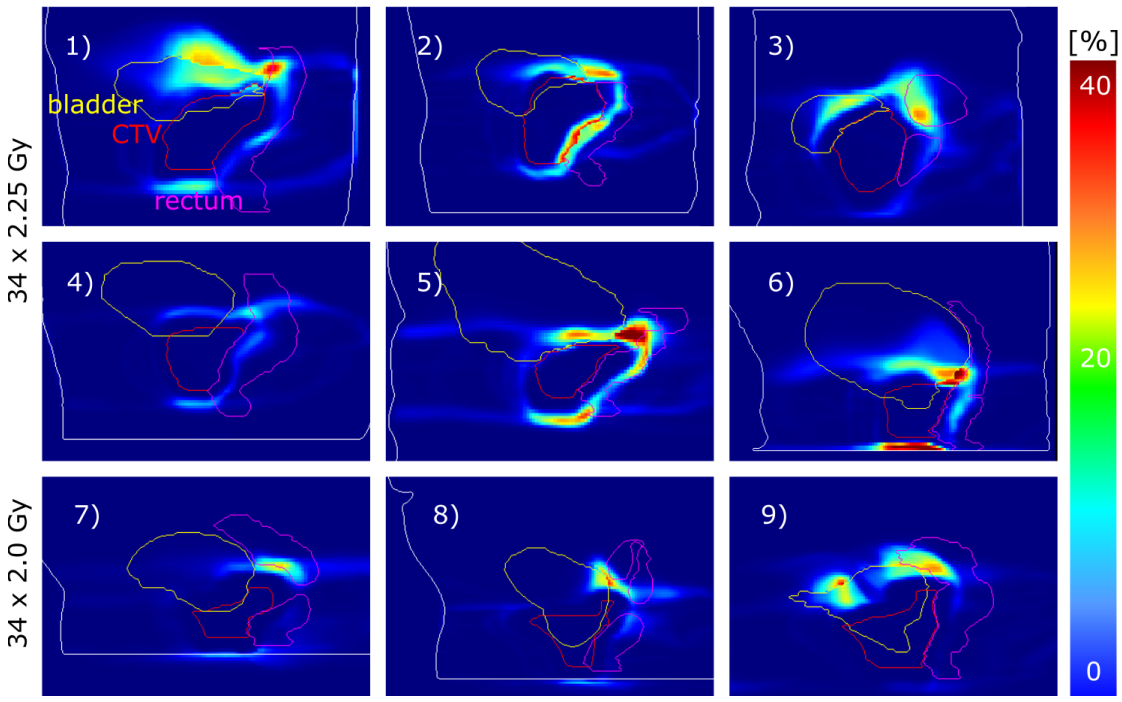


Figure 4.3. – Relative SF difference as calculated from the dose accumulation approach $SF(d)$ and the effect accumulation approach $SF(d_i)$. Presented as $(SF(d) - SF(d_i))/SF(d)$ in [%] for the nine analysed patient cases calculated with $\alpha = 0.1\ 1/Gy$ for all tissues and an α/β-map with CTV = 1.4 Gy, rectum = 3 Gy, bladder = 5 Gy, all other tissue = 3 Gy, assigned by changing the according β value. As mathematically predicted, $SF(d_i)$ was systematically lower than $SF(d)$. Each picture shows the sagittal slice with the highest SF deviation of the respective patient. Cases 1)-6) represent patients treated for prostate carcinoma with the whole prostate marked as CTV ; cases 7) - 9) show cases of the prostatic fossa with a slightly lower prescription dose of 2.0 instead of 2.25 Gy per fraction. All cases were treated and imaged for 34 fractions. For seven of the nine cases (except cases 4) and 7)), deviations were 34.8 – 50.7 % in distinct hot spots with maximum difference of 50.7% in case 6). Only case 4) showed deviations below 18% in the entire volume.

the rectum with maximum deviations around 3.3 – 4.9 Gy in regions of high motion amplitudes. The maximum difference was given in case 6 with 4.9 Gy. For only one patient (case 4), the deviations were below 1.4 Gy in the entire monitored volume. A region of repetitive high deviation amplitude could be found for all patients between the delineated organs above the apex of the CTV, often coinciding with the rectal and bladder wall.

Naturally, total SF deviation hot spots shown in figure 4.3 coincided with hot spots of total dose difference shown in figure 4.4. Both figures are qualitatively very similar with only little differences. The magnitudes of these deviations depended on the total dose delivered in the respective voxel along with the dose variation magnitude. Table 4.1 summarizes the maximum deviations for both the SF approaches and the dose accumulation approaches. It can be seen that both magnitudes scale very similarly from 0 - 5 Gy and from 0 - 50%, respectively. Areas of higher deviation are a little broader for the SF comparison.

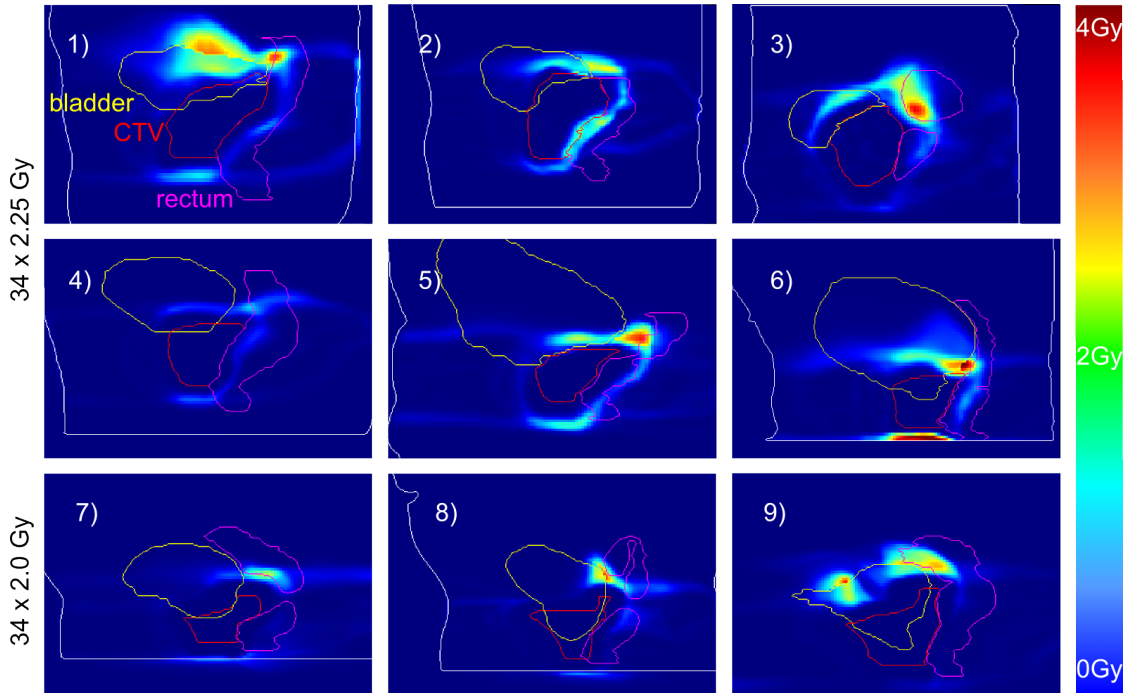


Figure 4.4. – Total dose deviation of the biologically accumulated dose and the conventionally accumulated dose: bEQD – D_α [Gy], same cases and slices as shown in figure 4.3. As mathematically predicted, bEQD was systematically higher than D_α. For six of the nine cases (1), 3), 5), 6), 8), 9)), deviations were 3.3 – 4.9 Gy in distinct hot spots with maximum difference of 4.9 Gy in case 6). Only case 4) showed deviations below 1.4 Gy in the entire volume.

Table 4.1. – Maximum deviations between the effect and the dose based methods: SF: $(SF(d) - SF(d_i))/SF(d)$ in [%]; dose: $bEQD - D_\alpha$ [Gy], for the 9 cases shown in figures 4.3 and 4.4.

case:	1	2	3	4	5	6	7	8	9
SF [%]:	36.5	39.4	35.3	17.9	35.8	50.7	26.5	34.8	35.2
dose [Gy] :	3.3	2.5	3.9	1.4	3.4	4.9	2.2	3.5	3.4

4.2.2 *The impact of fractionation schemes*

Figure 4.5 gives an overview of all 9 patient cases for all 3 fractionation schemes for the analysed total dose difference. Difference magnitudes strongly increased when increasing the dose per fraction while lowering the number of fractions. Highest differences were found for strong hypofractionation in 5 fractions where 7 of 9 cases show difference hot spots above 4.3 Gy with a maximum of 8.4 Gy (case 9). The difference was above 8 Gy in 3 cases. Highest difference hot spot for the 21-fraction scheme was 6.8 Gy for case 6) with 5 cases showing difference hot spots above 5 Gy. For comparison, at 34 fractions (see section 4.2.1) the maximum hot spot was at 4.9 Gy. Hot spot areas are broader for the 21-fraction scheme in 7 cases and decreasing in size but for most cases increasing in magnitude, when reducing the fraction number to 5. In one case, the magnitude of difference decreased when going from 34 to 21 fractions and for two cases when going from 21 to 5 fractions.

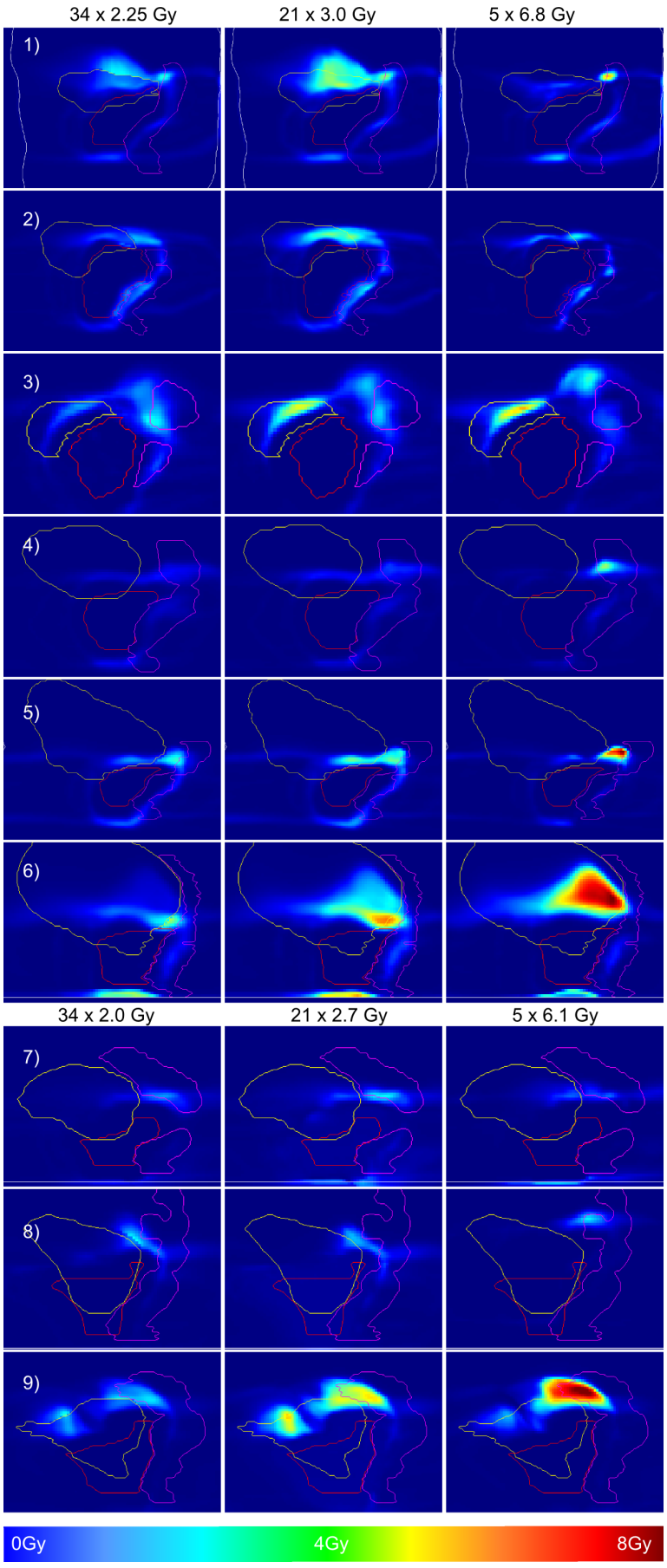


Figure 4.5. – Total dose deviation (bEQD – D_α [Gy]) for the nine analysed patient cases (same as in figure 4.4) using different fractionation schemes: 34 fractions of 2.25 Gy / 2.0 Gy as used for treatment, 21 fractions of 3.0 Gy / 2.7 Gy (moderate hypofractionation), 5 fractions of 6.8 Gy / 6.1 Gy (strong hypofractionation). The lower 3 cases show patients treated for the prostate fossa. Schemes were converted using the standard EQD formalism with prostate $\alpha/\beta = 1.4$ Gy to be biologically equivalent. Note that the dose difference color scaling is different to previous figures (0 – 8 Gy). Differences were generally increasing when going to higher fraction doses for the majority of cases as predicted in figure 4.2. Differences were highest for 7/9 cases for strong hypofractionation (above 4.3 Gy) with a maximum of 8.4 Gy (case 9). At 21 fractions, case 6) showed the highest difference hot spot with 6.8 Gy.

4.2.3 *Local vs. volumetric analysis*

DVH analysis Dose comparison using dose-volume histograms (DVHs) revealed only small differences between the two accumulation methods for standard fractionation although distinctive hot spots of increased total dose difference are visible for all patient cases in the 3D data (figure 4.4). The DVHs of cases 3) and 9) are shown in figure 4.6. For case 3), in comparison to figure 4.4, the distinct hot spot within the volume of the rectum is visible as only a slight difference in the DVH of the rectum. There is no difference of D_α and bEQD for the bladder or the CTV , although for the bladder there was a higher difference to the planned dose. The visible difference between the DVH curves of D_α and bEQD for the rectum were similar to the depicted case 3) for two other cases, while 6 cases showed lower difference. For

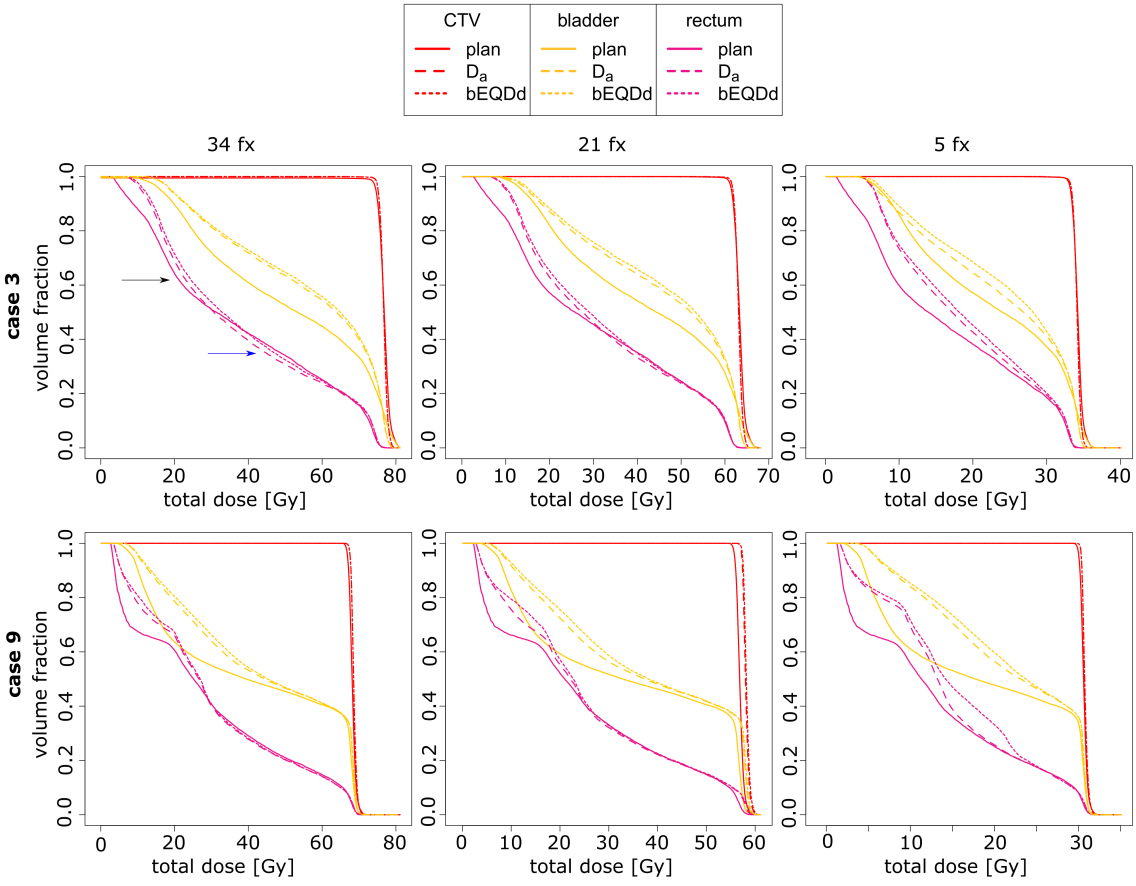


Figure 4.6. – Cumulative DVHs of case 3) and 9) (see figures 4.4, 4.5 and 4.8), showing the planned dose, D_α and bEQD for the three investigated fractionation schemes. For case 3) the distinct hot spot difference between bEQD and D_α is visible only as a small difference in the curves for the rectum in the DVH in the mid dose region, while no clear difference can be seen for CTV and bladder. Black arrow: dose regime where bEQD increases the overdosage between planned and delivered dose further, as also shown by the black arrow in figure 4.8. Blue arrow: The increase in the delivered dose calculation by bEQD is still below the planned dose (seen also figure 4.8). For case 9) the strong difference hot spot in the rectum is also visible as a stronger difference in the DVH for hypofractionation with 5 fractions (fx).

hypofractionation, differences were slightly more distinct for case 3) based on the overall higher difference reported in the previous section. For case 9), the difference hot spot in the rectum for strong hypofractionation also shows a stronger difference in the DVH . The DVHs for the bladder showed generally lower differences between the two accumulation methods than those for the rectum. For both bladder and rectum, there was a distinct difference to the planned dose, similar to the depicted case in a total of three cases while the other 6 cases showed only small differences. For all cases, the dose coverage of the tumour was similar as in the depicted DVH. More information and analysis on the difference between planned and delivered dose is given in the next section.

DVH metrics An overview of the comparison of dose-volume (DV) parameter as the difference of these as DV(bEQD) - DV(D_α) for all patients is given in figure 4.7 (standard fractionation) for: $D_{2\%}$, $D_{5\%}$, $D_{95\%}$, $D_{98\%}$ of the tumour (CTV), V_{30Gy} - V_{75Gy} for bladder and rectum, and mean dose and equivalent uniform dose (EUD) for all. The results are shown for the entire organs of bladder and rectum, as well as for bladder and rectal wall.
The systematic difference between bEQD and D_α can also be seen for all dose metrics. The median difference (thick black line) is however close to 0 in all cases, while the range differs. Generally, the organs walls of bladder and rectum showed lower differences compared with results for the whole organs.
For both bladder and rectum the difference in mean dose is around 0.2 to 0.3 Gy (median) with a range of about 1 Gy. The EUD is lower in comparison, only around 0.1 Gy difference, along with a smaller range of about 0.2 - 0.3 Gy. For the bladder, all V_{xGy} metrics show differences below 0.5% volume, decreasing with increasing dose threshold, with the exception of a few outliers, along with a small range of 0.5%. These are slightly higher for the rectum where the median is round 0.5%. V_{50Gy} shows a higher range compared with the other metrics of about 1.5%. The difference is increasing for the mid dose range and decreasing again for higher dose thresholds. The CTV shows generally very low differences below 0.2 Gy for mean dose and EUD and below 1 Gy for $D_{98\%}$ and $D_{95\%}$. The difference in $D_{5\%}$ and $D_{2\%}$ is negligible.
Generally, the differences in the DV-metrics were increased for moderate and strong hypofractionation. The analogue figures to 4.7 can be found in appendix B, figures B.1 and B.2. The above described trends can be seen analogously for the other fractionation schemes: for the bladder, the V_{xGy}-differences are decreasing with increasing dose threshold. For the rectum, the mid dose regimes show the highest differences with a median close to 1% for moderate and strong hypofractionation, with a range of $\pm 2\%$. In one case, a difference of 7% in V_{21Gy} for the rectum in strong hypofractionation was found and 4% for the bladder in V_{23Gy}. More details will be given in section 4.2.4 for comparison to the planned dose.

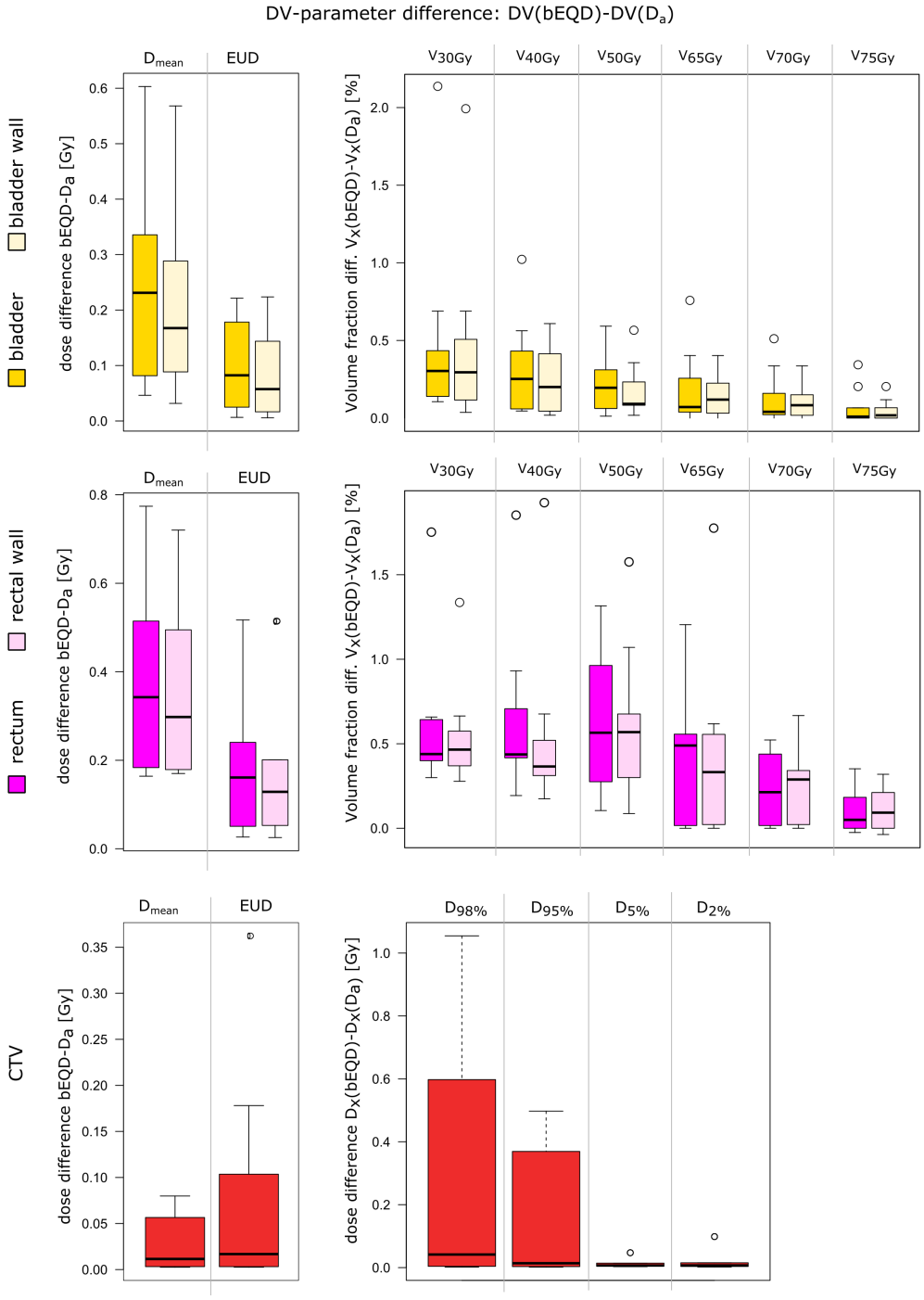


Figure 4.7. – Comparison of DV parameters for standard fractionation as the difference between DV(bEQD) and DV(D_a) of the structures of interest of bladder, bladder wall, rectum, rectal wall and tumour (CTV) for various dose metrics. A comparison to the planned dose will be given in figure 4.9 for the same dose metrics. The results for hypofractionation can be found in appendix B.

4.2.4 *Comparison to the planned dose*

Gamma-analysis In the Gamma-analysis for the 3mm/3% criterion, the passing rate was lowered for all cases and all volumes when using bEQD instead of D_a by an average of 1.0% for the bladder and 1.6% for the rectum with a maximum of 3.0%. The passing rate was below 95% for the rectum in all cases and in 5/9 cases for the bladder. The gamma analysis is shown for three cases in figure 4.8. The main influence of bEQD compared with Da was pronounced as an increase of the γ value in regions of already failed γ criterion, as shown for the maximum values in figure 4.8 (red). The deviation hotspots between bEQD and D_a coincided with increased γ values in regions of already failed γ criterion ($\gamma \geqslant 1$), indicating a more severe deviation to the plan than calculated with D_a (black arrows) in four cases in the bladder volume and in five cases in the rectum volume. However, regions of high deviation between bEQD and D_a (figure 4.8 left) did not necessarily coincide with overdosed regions (red) in the γ analysis (figure 4.8 right). For example, in case 3) the strongest total dose deviation hot spot coincided with an underdosed region (blue in γ analysis), effectively only lowering the magnitude of the underdosage (blue arrows) but not adding to the failing rate. A similar scenario is observed in a total of two bladder and two rectum volumes.

DVH metrics Figure 4.9 shows the same dose volume metrics for standard fractionation as before in the comparison between both accumulation schemes, now for the comparison to those derived from the planned dose (for both bEQD and D_a). The same analysis for hypofraction can be found in appendix B in figures B.3 and B.4. More information and analysis on the difference between planned and delivered dose for conventional dose accumulation can be found in a preceding study [8]. The median difference to the planned dose is close to 0 Gy in all cases. The range shows much higher magnitudes than the difference between both accumulation strategies. For the majority of cases, it can be seen that the computed dose metrics for the organ walls are smaller than that of the entire organs. This can either mean that they are closer to 0 Gy difference to the planned dose or show a higher negative difference.
The range of difference is around 10 - 12 Gy for the *bladder* for the mean dose and the EUD . While the mean dose is mostly higher for the accumulated results than the planned dose, the EUD is very close to or lower than planned. For the V_{xGy} dose-volume metrics of the bladder, the ranges generally tend to shift towards negative values for higher dose thresholds, especially for the bladder wall. The ranges are higher for the whole bladder (12 - 22% absolute range) than the bladder wall (8 - 12 % absolute range).
For the *rectum*, difference of the mean dose for the whole organ is at 1.5 Gy (median) but shifted towards 0 Gy (median) when only considering the rectum wall. The EUD varies uniformly around 0 Gy difference to the plan. For the V_{xGy} metrics, the range and median are similar for all with an increase in outliers in both directions towards

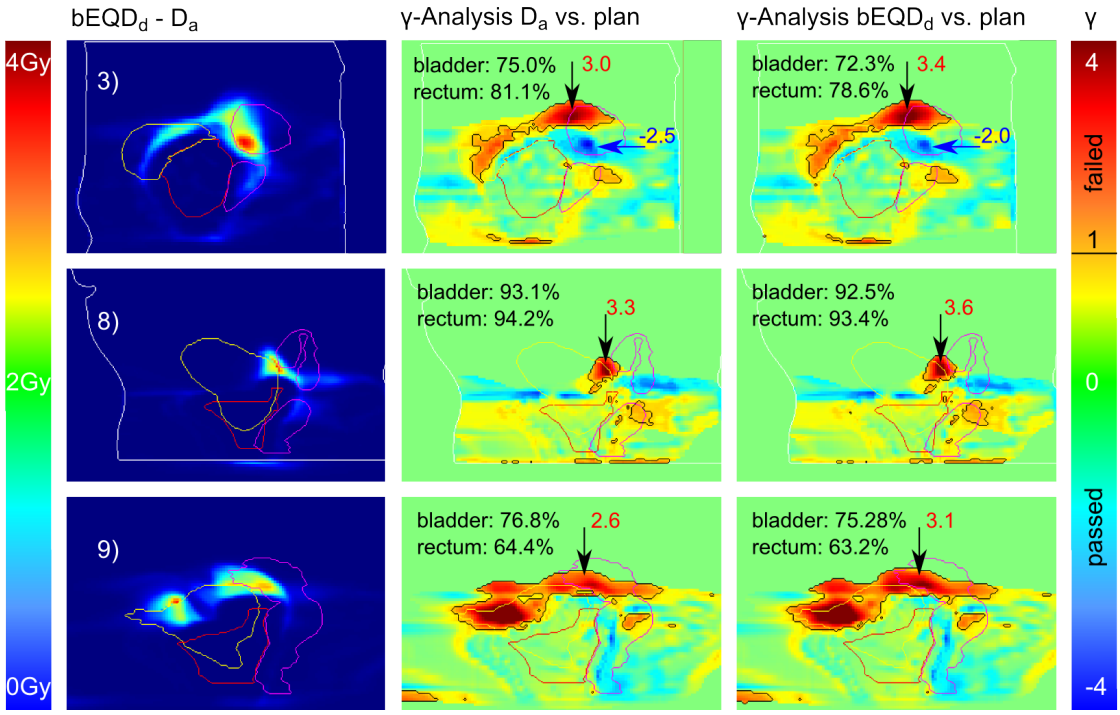


Figure 4.8. – γ-analysis for three cases (3, 8 and 9). Left: $\mathrm{bEQD} - \mathrm{D_a}$ as depicted in figure 4.4. Center and right: γ-analysis of the conventional accumulated dose ($\mathrm{D_a}$, center) and the biological dose (bEQD, right) compared to the planned dose $\mathrm{D_p}$, respectively (for standard fractionated treatment). Percent γ passing rates for a 3 mm/3% criterion for bladder and rectum are given. Failed γ-criterion is marked in orange and red colors and separated by a black line. Black arrows highlight regions of failed γ-criterion, enhanced in the case of using bEQD compared with $\mathrm{D_a}$. These regions coincide with total dose deviation hot spots shown on the left. Blue arrows in case 3) show an area where the total dose deviation hotspot coincided with a negative γ-value region, thus increasing γ (less negative) without changing the passing rate outcome.

higher dose thresholds. Again, the difference between whole organ and organ wall is more pronounced than for the bladder. This is also true for the difference between bEQD and $\mathrm{D_a}$. In comparison to the bladder results, for the mid dose thresholds deviations between the bEQD and $\mathrm{D_a}$ based differences can be seen.
The CTV dose difference range of the mean dose is limited to ± 1 Gy and a little more towards negative values for the EUD. The range of the investigated $\mathrm{D_x}$ is strongly decreasing for the metrics of low volume (referring to the highest dose regions), while those for large volumes (referring to dose homogeneity) show absolute ranges of about 10%. $\mathrm{D_{98\%}}$ was lower by 1 Gy (median) for $\mathrm{D_a}$ and bEQD compared with the planned dose.
Especially for the rectum, the *impact of the fractionation scheme* in the difference to the planned dose was clearly visible. As a summary of figures 4.9, B.3 and B.4, figure 4.10 shows the DV-difference between the accumulation methods as well as the difference of the respective DV to the planned dose for the rectum for all fractionation schemes. It can be seen that the median as well as range of DV-difference is increasing for hypofractionation. It needs to be noted that based on the conversion, the high dose thresholds for hypofractionation are already above the CTV planning

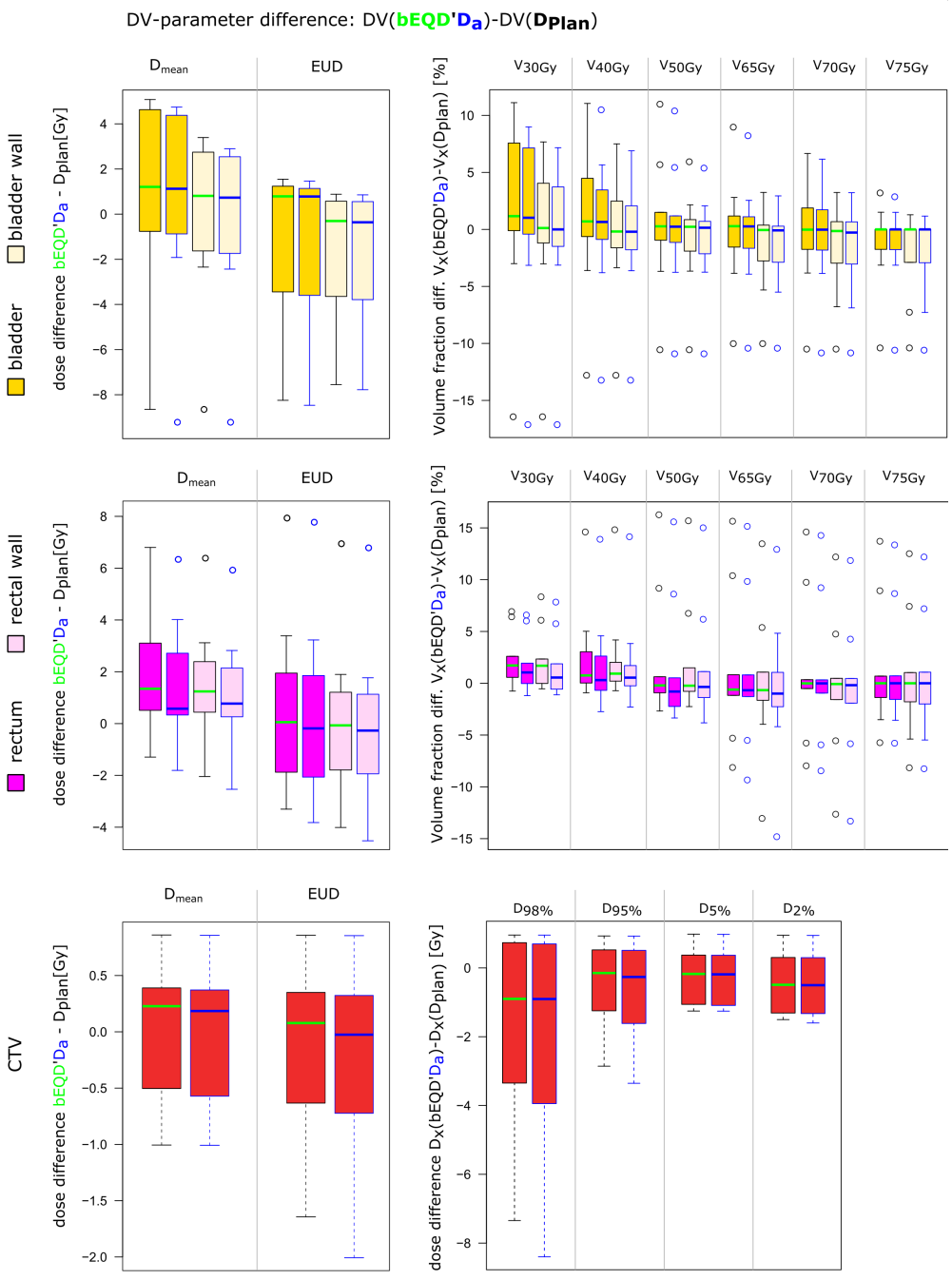


Figure 4.9. – Comparison of DV parameters as the difference between either DV(bEQD) or DV(D_α) to DV(D_{plan}), respectively. Shown for the structures of interest of bladder(B), bladder wall (BW), rectum (R), rectal wall (RW) and tumour (CTV) for various dose metrics. The absolute difference between the same parameters derived from bEQD and D_α is furthermore shown in figure 4.7.

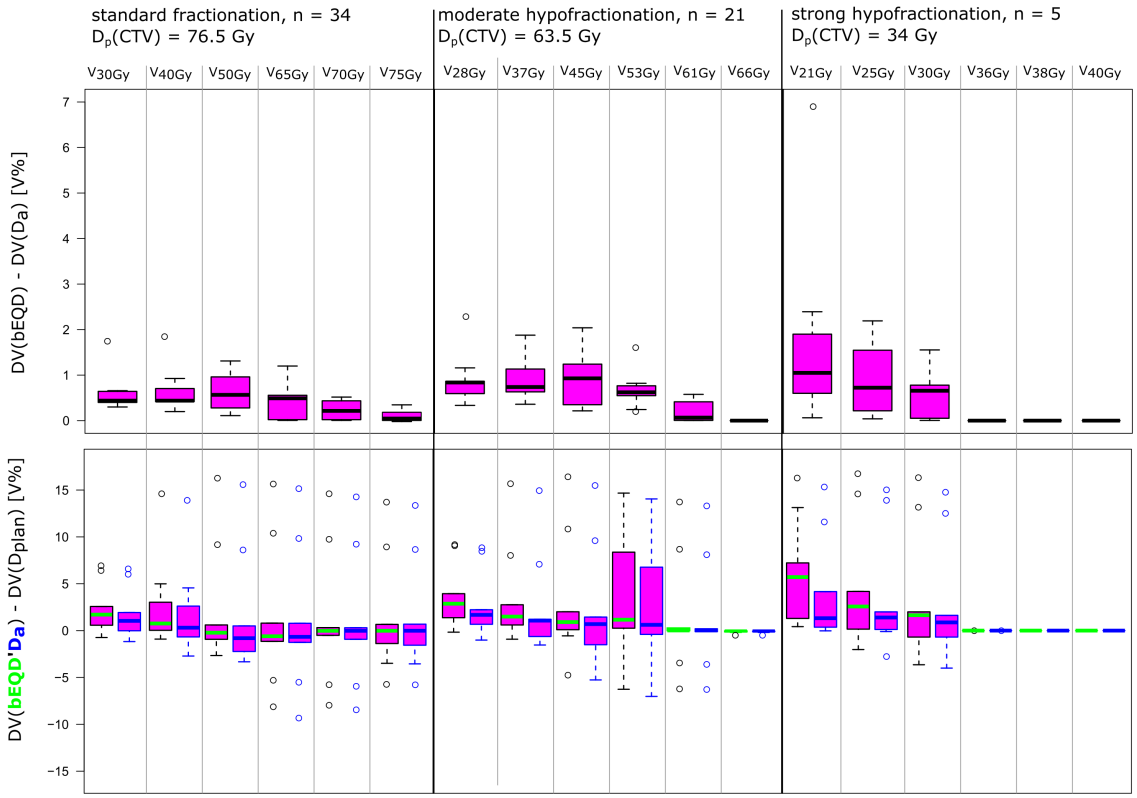


Figure 4.10. – Overview of the DV-analysis in the rectum for all investigated fractionation schemes. Top row: absolute difference in the respective DV-metric between bEQD and D_α. Bottom row: respective difference to the planned dose. Complementing figures 4.7 and 4.9.

dose. The bottom row depicts the respective differences to the planned dose. As can be seen, this is especially increased for strong hypofractionation. Furthermore, an increasing difference between bEQD and D_α can be observed, especially for the median dose differences.

For an evaluation of the derived total doses with current dose constraint recommendations found in the Quantitative Analyses of Normal Tissue Effects in the Clinic (QUANTEC) reports, figure 4.11 shows the planned dose, bEQD and D_α for all cases and V_{xGy} metrics along with the recommended constraints. As noted before, the dose conversion for hypofractionation works in favour of the normal tissues that are assigned with a higher α/β than the tumour, so that the highest dose threshold values in normal tissue are above the CTV planned dose. It can be seen, that especially for the rectum, dose constraints are violated for 1 - 2 cases for the high dose thresholds. The violation is already given for D_α, the effect is slightly increased in all cases for bEQD. In appendix B, figure B.5, the same results are shown for $V_{25\%}$, $V_{50\%}$, $V_{75\%}$ and $V_{100\%}$ (volume receiving x% of the planning dose) for a better comparison of the fractionation schemes. As can be seen there, the trends are equal in all cases but the difference between bEQD and D_α is generally increased for hypofractionation.

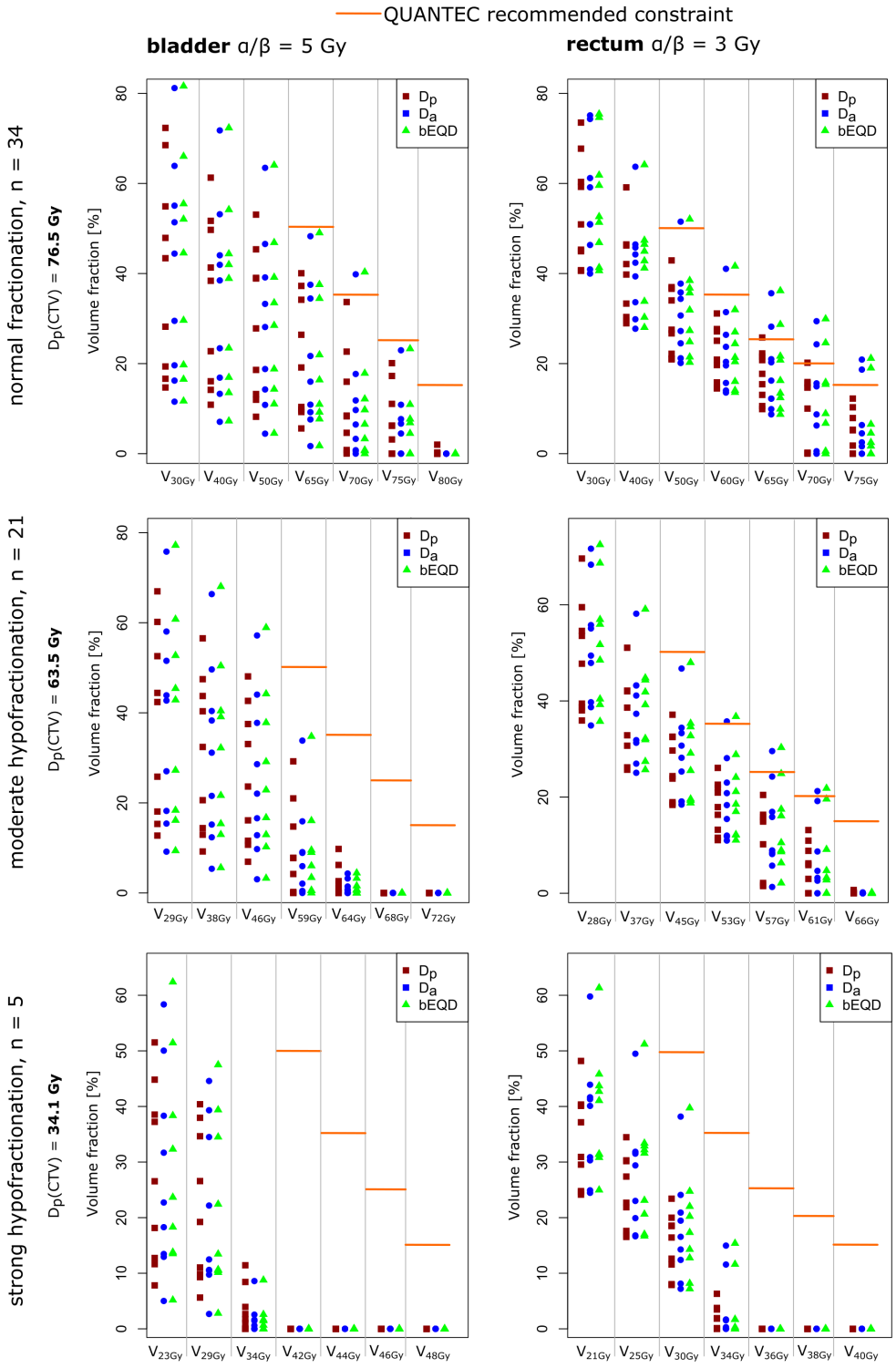


Figure 4.11. – V_{xGx} comparison between planned (D_p), accumulated (D_a) and biological accumulation dose (bEQD) with respect to QUANTEC dose constraint recommendation for rectum and bladder. The respective V_{xGy} and dose constraints were converted for hypofractionation according to the EQD_2 approach.

4.2.5 *Uncertainty assessment*

Worst case DIR error estimation Image registration is required in order to perform either strategy of dose accumulation (bEQD and D_α). The impact of a faulty registration on the daily dose was estimated as the highest dose difference to that of each respective voxel within a 3 mm distance. This daily dose error was propagated to D_α (eq. 2.10) and bEQD (eq. 3.12), respectively. Figure 4.12 depicts the resulting error estimation for cases 1-3). Generally it can be seen that the magnitude of uncertainty is higher for ΔbEQD than ΔD_α. Furthermore, the volume showing differences above 1 Gy is higher for the dose mapping error estimation than the difference between the two dose accumulation methods (left). In the analysis of all patients, ΔbEQD showed maximum values around 9 Gy while the maximum of ΔD_α was at 5 Gy. Uncertainties are located in the gradient regions resulting from the neighbouring voxel deviation estimation. The magnitude of ΔD_α is similar to that of (bEQD $- D_\alpha$), while that for ΔbEQD is higher. Regions of highest dose mapping error estimations coincide for ΔD_α and ΔbEQD but do not coincide with strongest hot spots in (bEQD $- D_\alpha$).

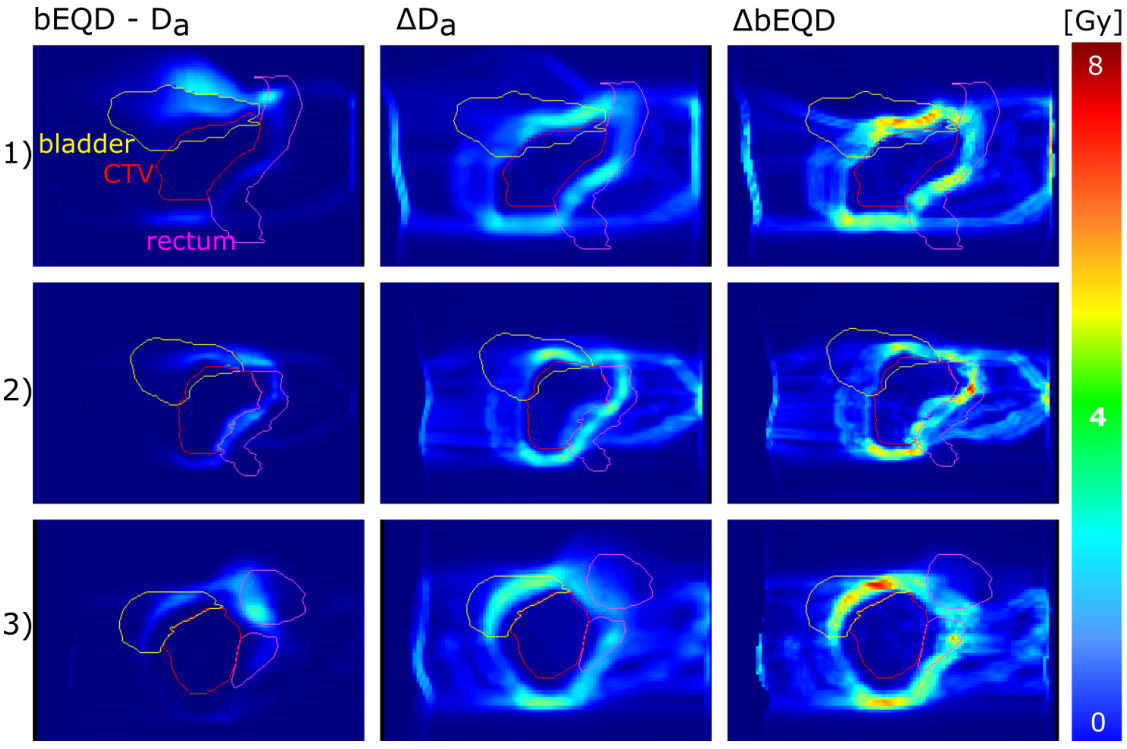


Figure 4.12. – Worst case dose mapping error estimation for cases 1-3). A 3 mm deformable image registration (DIR) registration is assumed for each voxel, the maximum dose difference within that range to the voxel of interest is calculated as the error. Error propagation on bEQD and D_α is calculated, respectively. Left - Total dose difference bEQD $- D_\alpha$ as shown in figure 4.4 for comparison; center - ΔD_α, eq. 2.10; right - ΔbEQD, eq. 3.12.

Different DIR implementations For the main analysis of this work, registration was performed in the RayStation system which uses the ANACONDA algorithm and is guided by the structure contours. For comparison, plastimatch was used within the AVID workflow, either based on the b-splines or the demons algorithm, respectively, but without landmark or contour guidance. For information on DIR algorithms, see section 2.2.1.3. Figure 4.13 shows the results of the total dose difference $\mathrm{bEQD} - \mathrm{D_a}$ based on the three different DIR approaches for three example cases. As can be seen, the results differ substantially, especially for the b-spline-based registration. The magnitude of difference is similar for the RayStation and the demons based DIR with larger volumes of high difference in the RayStation examples. The b-spline-based results show both higher difference magnitude and larger volumes of higher difference. In all cases, high differences are located mainly in the dose gradient regions.

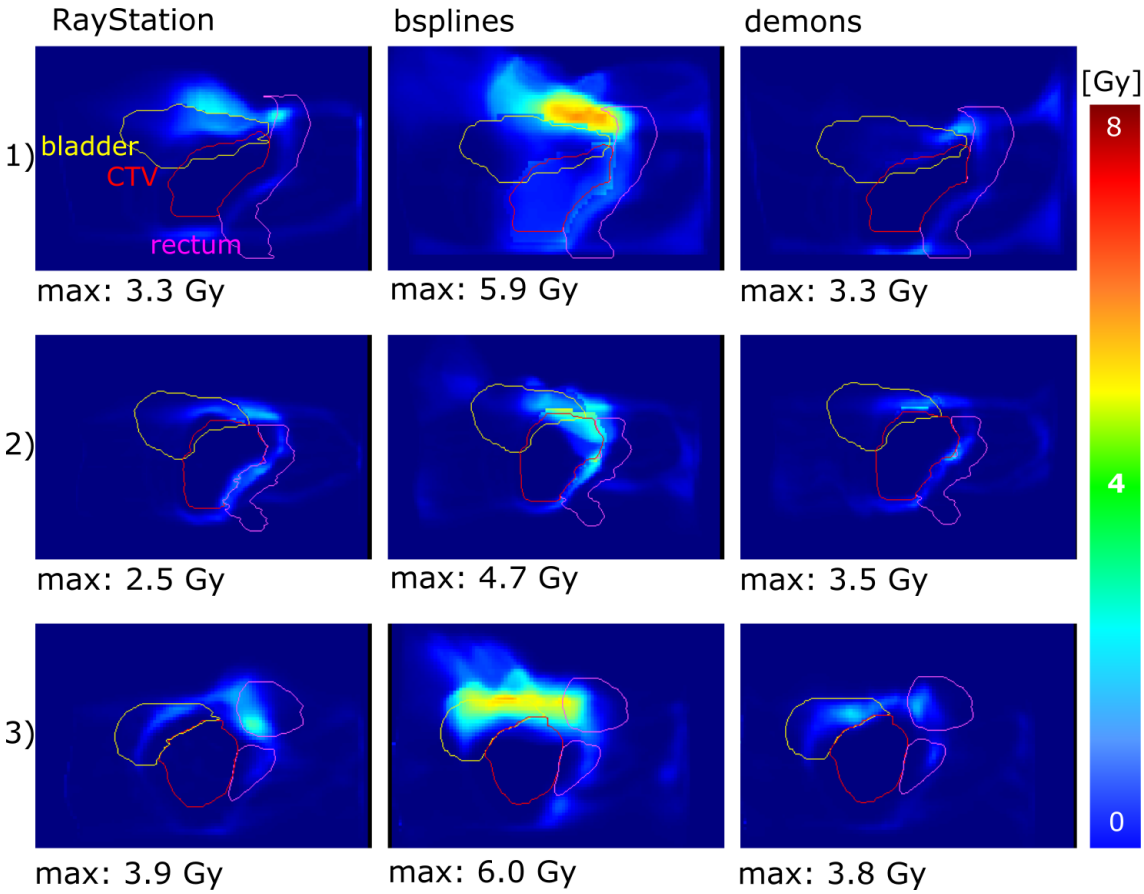


Figure 4.13. – Example comparison of cases 1-3) for the total dose difference $\mathrm{bEQD} - \mathrm{D_a}$ [Gy] for three different underlying DIR implementations: left - RayStation ANACONDA algorithm; center - plastimatch using b-splines; right - plastimacth using demons. Note that the scale is changed compared with previous figures. Case 1) is highlighted in detail in figure 4.14.

For case 1), b-spline DIR , the difference within the CTV is unexpectedly high. Therefore, this case was investigated in more detail in figure 4.14. Here, the total dose differences are again shown in the upper row for the demons- and the b-spline-based results along with the planning computed tomography (CT), which serves as the reference frame (fixed image). For the example fraction CTs, it can be seen that both the bladder and the rectum volume increased. While the bladder volume is similar in all fraction CTs, the rectal volume is especially increased in fractions 1 and 22. The mapped CTs show in all cases that the bladder volume was not correctly mapped in the upper region. B-splines shows generally large deformations while those for the demons algorithm are small. In fractions 1 and 22, b-splines shows that the grid is folding in on itself. This can also be seen on the right periphery of the demons registration which is folding in voxels from outside the dose field. In fraction 22, the b-splines algorithm clearly failed to find an anatomically plausible deformation field, resulting in the fact that the CTV dose was entirely mapped out of the CTV contours, which results in the deviation in this area in the total dose difference. For fraction 21, both the b-spline and the demons algorithm show anatomically realistic deformations on the deformaed CTs and qualitative assessment of the DIR would not reveal any substantial differences. Nevertheless, the two resulting mapped dose distributions are substantially different in the bladder and upper rectum volume.

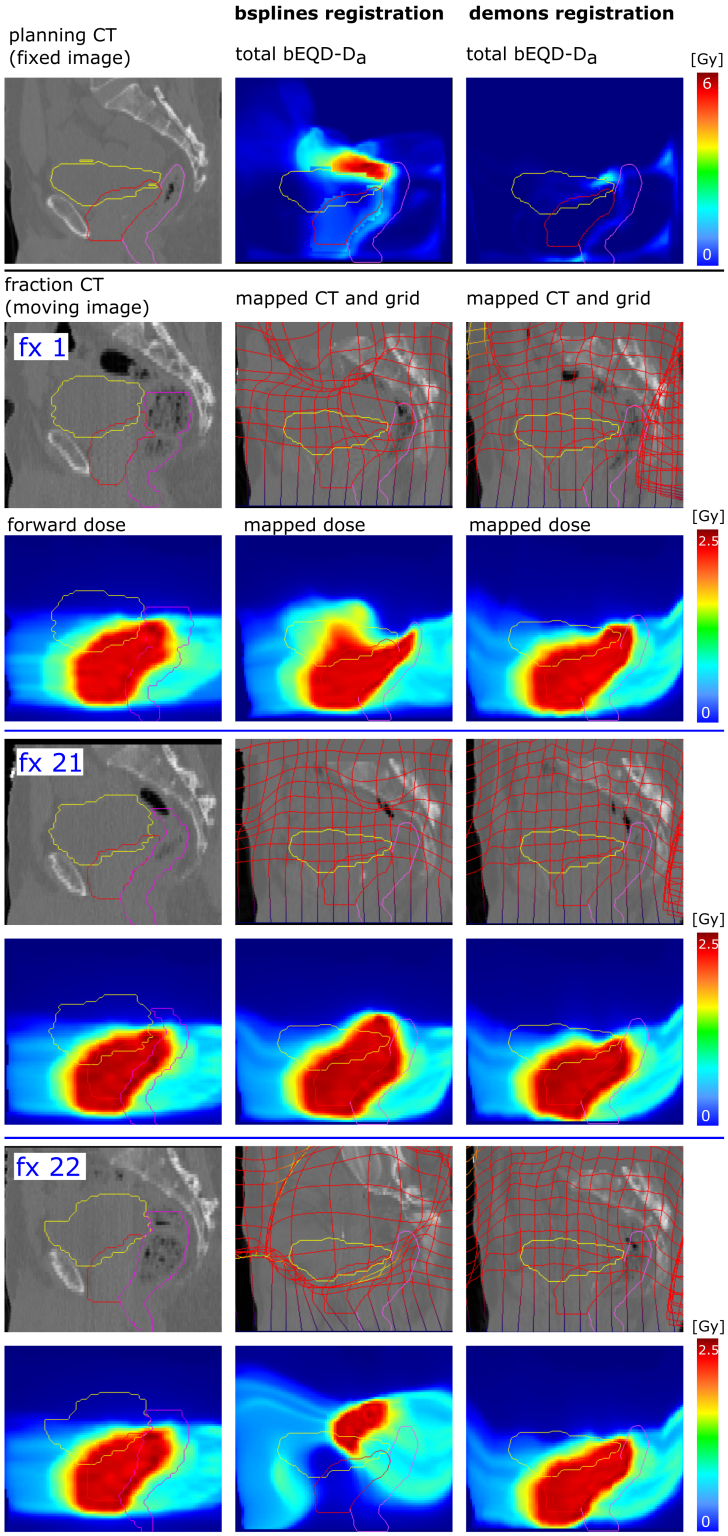


Figure 4.14. – Three example fractions of registered CT and mapped dose of case 1), using b-splines and demons registration, respectively. Top left: planning CT and contours for reference; Top center and right: Total dose difference $\mathrm{bEQD} - \mathrm{D}_{\alpha}$ (as shown in figure 4.13), as the result of the below shown DIR-based dose mapping for b-splines and demons registration, respectively. Below - Left: fraction CTs and contours along with the forward calculated doses. Center: b-spline registration results for the mapped CTs (deformation grid corresponding to deformed voxels of the fixed image that are being pulled back to the regular grid, thus enclosing the imaging/dose values they "pull back", structures show the reference structures of the planning CT), along with the resulting mapped doses. Right: results from demons registration, same as for center. The anatomically unrealistic deformation of fraction 22, b-splines, is the reason for the unexpected total dose difference within the CTV shown in the top row.

Uncertainties from α/β In contrast to conventional dose accumulation, biological dose accumulation depends additionally on the assigned α/β-values of the respective tissues. Generally, as reported above, the lower the α/β-value of the tissue, the higher is bEQD. Since α/β-values for tissues suffer from an uncertainty of their own, this will be passed on to the uncertainty in the calculation of bEQD. Based on the range of recommended values in literature (see table 3.2 materials & methods section 3.2.3), bEQD was calculated for the patient cases with different underlying α/β-maps. As an example, figure 4.15 shows the impact of changes

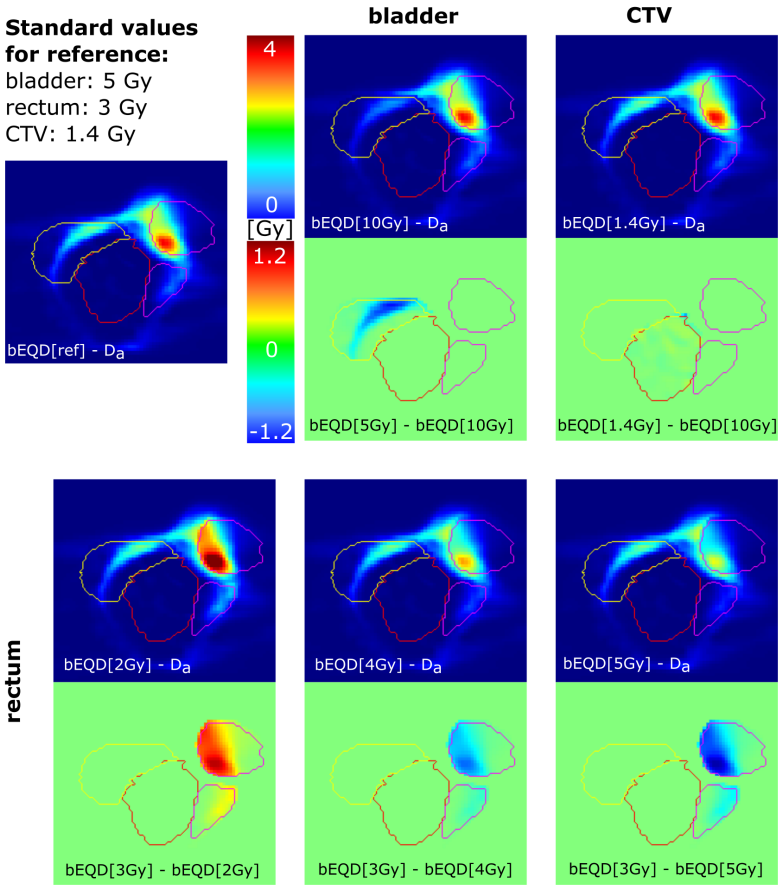


Figure 4.15. – Impact on bEQD of assigning different underlying α/β values to the respective structures according to ranges of recommendations in literature. Analysis shown for case 3), separately for the respective structures in comparison to the reference results in the upper left.

in α/β-values for case 3) for the individual structures. Highest difference between bEQD computed with different α/β-maps coincided with regions of highest differences to D_a. Changing the α/β-value for the rectum from 3 to 4 Gy resulted in a maximum decrease of 0.6 Gy (0.2 – 1.3 Gy) in bEQD averaged over all patients. Lowering α/β from 3 to 2 Gy led to a maximum increase in bEQD of on average 1.0 Gy (0.4 – 1.3 Gy). For the bladder, the difference between the lowest recommended α/β-value of 5 Gy to the highest of 10 Gy resulted in a maximum difference of 0.9 Gy (0.3 – 1.5 Gy) on average. In the CTV, changes were generally lower due to the overall lower dose variation magnitude. Assigning the standard α/β-value for

tumours of 10 Gy instead of the more recently promoted low α/β-value of around 1.4 Gy changed bEQD by a maximum of 0.7 Gy (0.0 – 1.7 Gy) on average.
The bEQD solely depends on the ratio α/β, so that the individual values do not affect the outcome, when looking at the same ratio. This is different for the SF itself. Figure 4.16 shows the impact of changing the individual α and β, while keeping the same ratio. It can be seen that the magnitude of absolute dose difference between the dose and the effect accumulation approach is not affected by the choice of α. However, the region of high difference is shifted towards the outer dose field for higher α (figure 4.16, a). The relative SF difference is decreased by approximately a factor 1/2 for assigning the half as high α values (figure 4.16, b). As for the absolute difference, the difference in the individual accumulation approaches is similar (figure 4.16, c). The magnitude is high in the regions of low dose (up to 25%) and becomes very small for higher doses as well as for regions where the dose goes to 0 Gy.

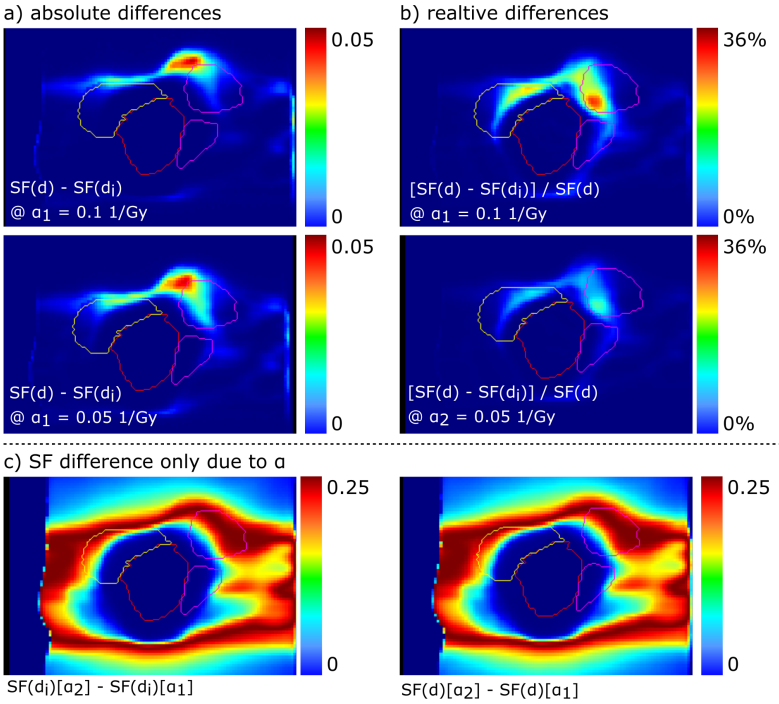


Figure 4.16. – Assessment of the impact of the choice of the individual α and β on the resulting SF from $SF(d)$ and $SF(d_i)$, respectively. Example for case 3). a) absolute SF difference between the dose and the effect accumulation approach, based on either $\alpha = 0.11/Gy$ (used standard) or $\alpha = 0.051/Gy$. β was changed accordingly to keep the same ratio of α/β in both. b) relative SF difference for the same parameters. c) SF difference for the individual $SF(d_i)$ and $SF(d)$.

Uncertainties from interpolation As a comparison to standard trilinear interpolation of the dose, the sub-voxelization method of Rosu et al. [64] (see also section 2.2.1.4) was applied within the AVID framework. Figure 4.17 shows the results for the difference between both interpolation methods for case 3). The range of difference is the same for both bEQD and D_a. Larger differences can generally be seen for the demons based registration in this case while the total dose difference of $bEQD - D_a$ was lower than for the b-splines-based DIR in this case. Difference are not systematic but can be positive or negative in magnitude, ranging between -2 to +4.4 Gy. Highest absolute difference are located in the gradient regions around the target but also at the beam penumbra.

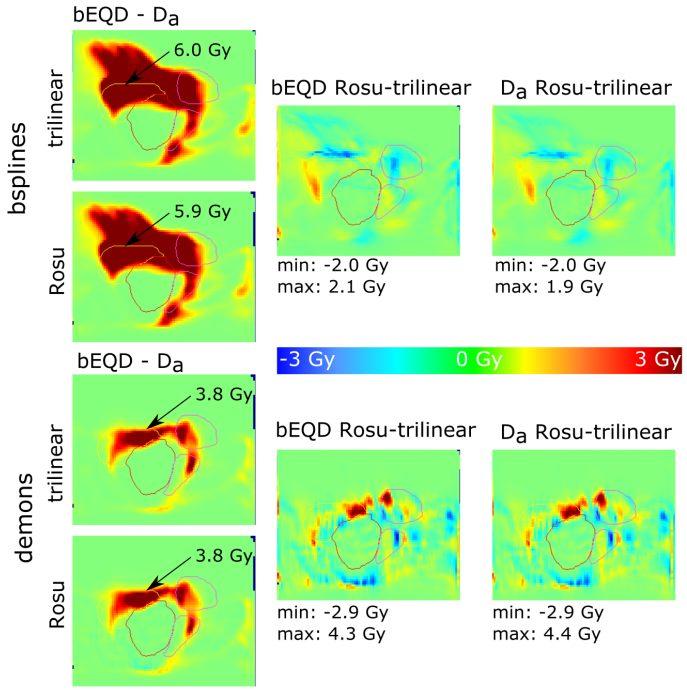


Figure 4.17. – Comparison of the mapped dose using different underlying interpolation algorithms: classic trilinear interpolation and the sub-voxelization method introduced by Rosu et al. [64]. Shown for case 3) for both b-splines- and demons-based DIR . Left - Total dose difference $bEQD - D_a$ [Gy] for reference with the same colour scale (same results as in figure 4.13). Center - difference of the bEQD results between the Rosu-interpolation and trilinear interpolation. Right - Same difference for D_a.

4.3 COHORT VOXTOX RECTUM

The VoxTox rectum cohort compliments the above investigation of the German Cancer Research Center (DKFZ) pelvis cohort by i) local dose analysis based on DSMs of the rectal wall; ii) allowing for statistical conclusions based on the much higher number of 251 cases (section 4.3.1); iii) dose-toxicity correlation based on the available patient-reported outcome data (section 4.3.2).
Since the data and workflow was made available through a cooperation with the VoxTox group in Cambridge, UK, the bEQD approach was implemented into the existing workflow. Investigation of the individual workflow steps, as done for the DKFZ pelvis cohort, was not feasible within this work.

4.3.1 *Total dose deviations on DSM*

The results of the dose calculated on the rectum surface as DSMs were computed for conventional dose accumulation D_a and biological dose accumulation bEQD, analogously to the analysis of the DKFZ pelvis cohort. Figure 4.18 a) shows the difference $bEQD - D_a$ for the maximum and the mean dose deviation for all 251 investigated cases. bEQD was calculated with different underlying α/β values to investigate the impact on the total dose deviation. The systematic deviation between bEQD and D_a was also confirmed for this cohort, with bEQD being systematically higher than D_a. The magnitude of maximum dose difference ranges from 0 - 3.3 Gy with the majority of cases below 1.5 Gy. For the mean dose difference, values reach a maximum of 0.3 Gy with the majority of cases below 0.15 Gy. A trend can be observed that for higher maximum dose difference, the mean dose difference is higher as well. As was also shown theoretically and for the DKFZ data, the lower the α/β, the higher the dose difference (highest difference for the calculations based on $\alpha/\beta = 2$ Gy). Using the same α/β of 3 Gy as for the DKFZ pelvis cohort, magnitudes decreased to 2.5 Gy (max. difference) with an average of 0.74 Gy and 0.25 Gy (mean difference) with an average of 0.08 Gy. Figure 4.18 a) shows the boxplot statistics for this case as well. For an α/β of 4 Gy, the difference decreased to below 2 Gy (max. difference) and 0.2 Gy (mean difference).

Figure 4.18 b) shows the same data as in a) as the respective difference to the planned dose (D_p). For the maximum dose difference, a range from 0 - 37 Gy can be seen. The mean dose difference ranges from 0 - 3.5 Gy. Again, the impact of the assigned α/β value can be seen with a generally higher impact on the points located at higher maximum and mean dose difference. D_a and bEQD show systematically higher mean values than the planned mean dose. In comparison to the difference to the planned dose, the difference between the accumulation methods is small. Nevertheless, the additional difference to the mean planned dose induced by the use of bEQD (instead of D_a) is clearly visible in figure 4.18 b) and it reaches up to 25 %. A histogram of the additional shift in mean dose by bEQD compared with D_a is

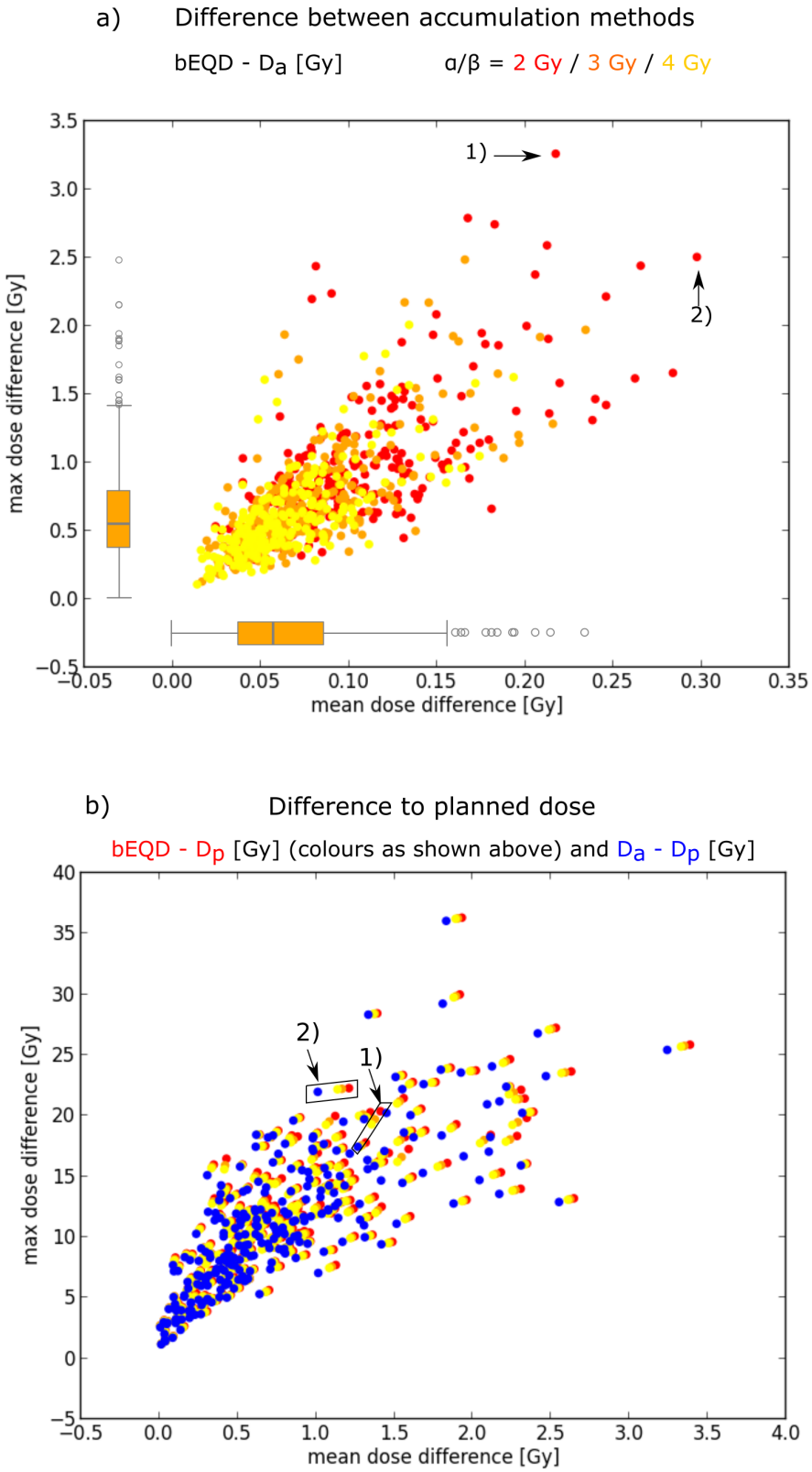


Figure 4.18. – Overview of the mean and the maximum dose difference in the DSM data between a) the two accumulation methods as $\mathrm{bEQD} - \mathrm{D}_\alpha$ (boxplots shown for difference with and $\alpha/\beta = 3$ Gy) b) the individual D_α and bEQD and the planned dose D_p. The results from all 251 patient cases are shown. bEQD calculation is depicted for different underlying α/β values of the rectum as indicated by the colours. Highlighted cases 1) and 2) refer to the highest maximum (3.3 Gy) and highest mean (0.3 Gy) dose difference, respectively, shown in detail in figure 4.20.

shown in figure 4.19. A peak around an additional 6% can be seen and a substantial amount of cases above 10%.

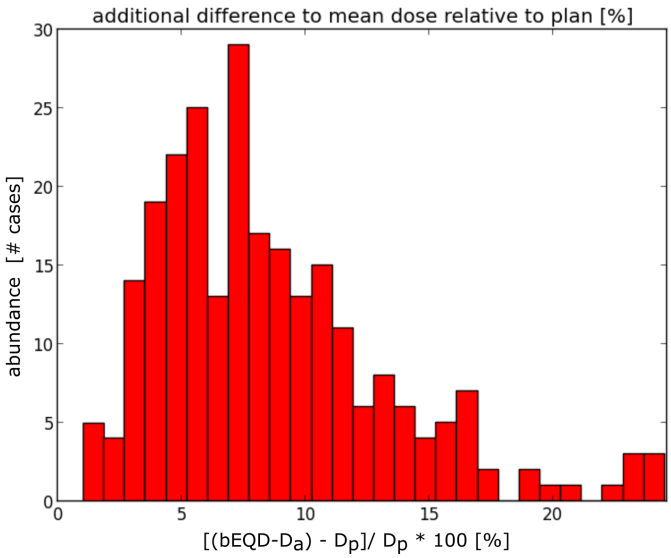


Figure 4.19. – Additional mean dose difference relative to the planned dose induced by the difference in bEQD to D_a in [%].

As an example, the two cases showing the highest maximum dose difference (1) 3.3 Gy) and the highest mean dose difference(2) 0.3 Gy), respectively, are visualized in figure 4.20, as indicated as well in figure 4.18. Mainly, the highest differences between both accumulation methods as well as highest differences to the planned dose are located in the peripheral superior dose region. Case 1), showing the highest maximum dose difference, is characterized by the distinct difference hotspot of about 3.3 Gy, while in the rest of the area, the difference remains marginal. This is similar for the difference to the planned dose. For case 2), which shows the highest mean dose difference, the area of difference between D_a and bEQD above 1 Gy is larger than in case 1) but located as well in the peripheral superior mid dose regions and again similarly located as highest differences to the planned dose.

In the following, the DSM data were analysed with respect to the work published on this data by Shelley et al. [70]. Within this prior work, the results on the conventionally accumulated dose were investigated with respect to the planned dose. The EUD and the dose width was computed on the planned dose and the conventionally accumulated dose for a (different) discovery cohort of 109 patients. Figure 4.21 shows the results acquired in the here presented work for the statistics on dose width and EUD from the D_a as well as bEQD, compared with the planned dose, for all available 251 cases of the consolidation cohort. It can be seen that the higher the dose threshold for the dose width fit, the lower the median dose difference to the planned dose. While this is very close to 0 for a dose threshold of 30 Gy, it monotonically decreases to - 3 (relative lateral extend) at 70 Gy threshold. Along with the median dose difference shift, the range of values also tends to lower difference with increasing dose threshold. The range is generally large between ± 12 (50 Gy) to $(-35) - 5$ (70 Gy). The difference between both accumulation methods is small in comparison to the difference to the planned dose. However, it can again be seen also for the dose width that the extent is smaller for D_a than for bEQD with a difference of about 1 - 2. EUD calculated from bEQD and D_a was lower than that

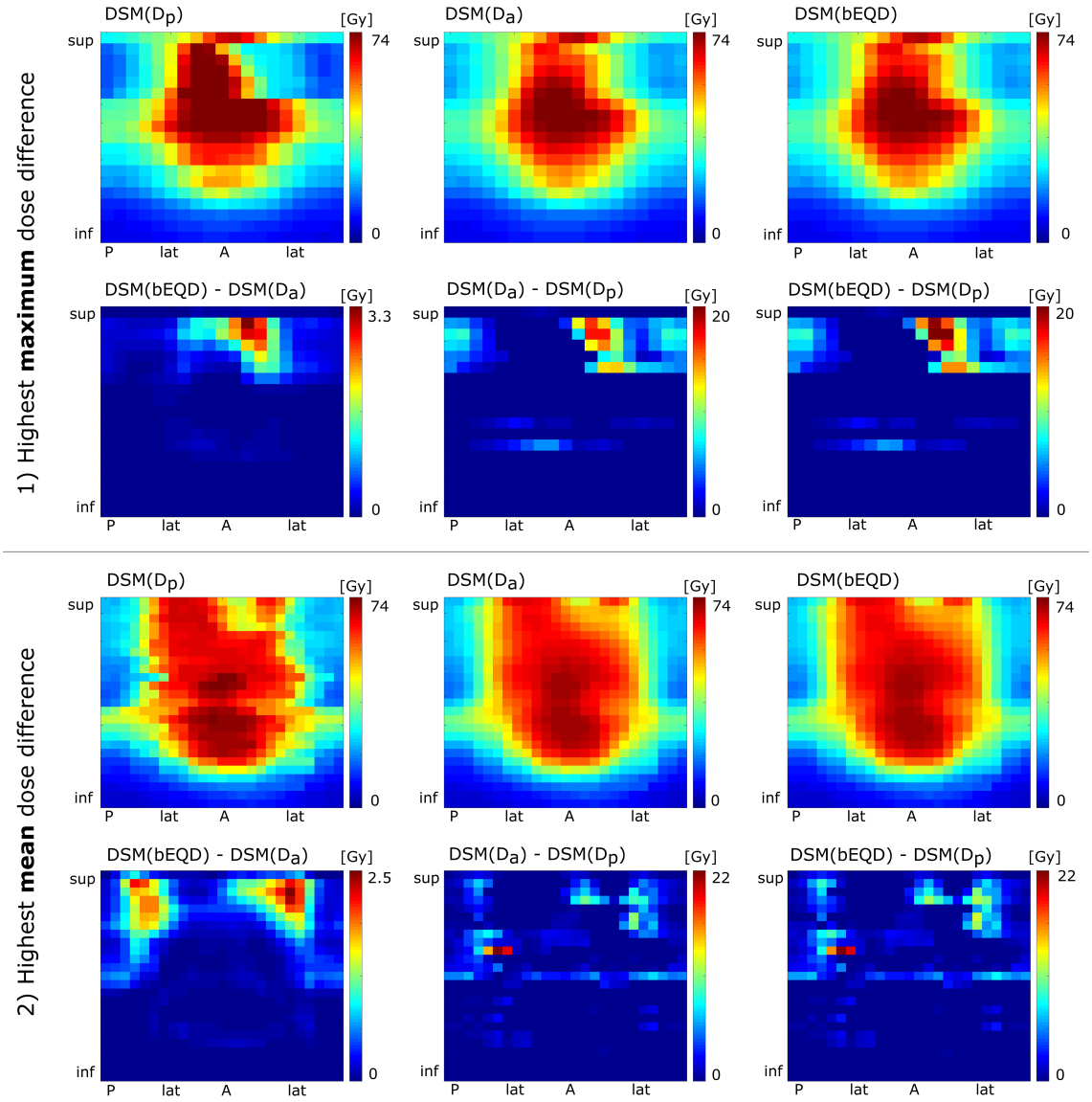


Figure 4.20. – Example cases for DSM of the case showing 1) the highest maximum dose difference of 3.3 Gy and 2) the highest mean difference of 0.3 Gy, indicated in the overview in figure 4.18, for $\alpha/\beta = 2$ Gy. Note the different scales on the colour bars. Sup: superior, inf: inferior, P: posterior, lat: lateral, A: anterior. D_p: planned dose.

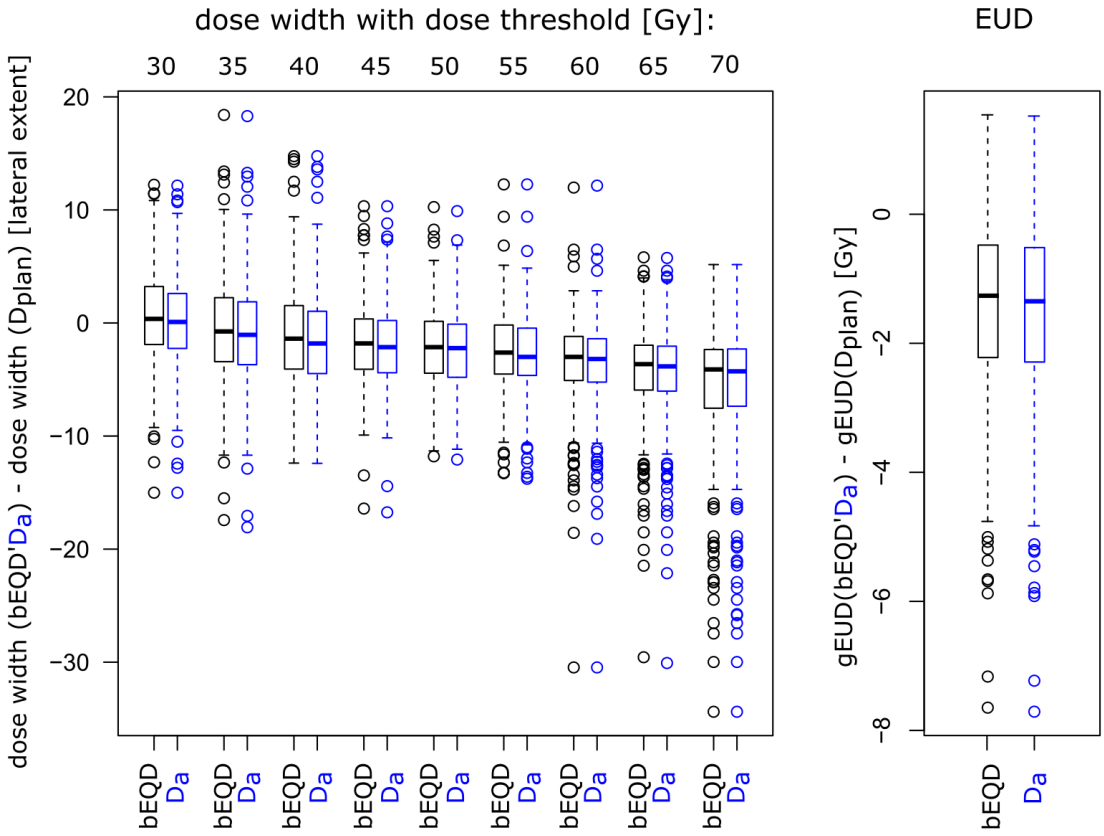


Figure 4.21. – DSM Boxplot statistics of the dose metric differences between bEQD- or D_α-based calculation and the planned dose, respectively: dose width fits based on several dose thresholds from 30 to 70 Gy; and the EUD.

based on the planned dose for the vast majority of cases by about 1 Gy (median) (note that as shown above, the mean dose was higher for the accumulated doses in all cases). The results from bEQD and D_α are only marginally different with EUD (D_α) lower by about 0.1 Gy. The same dose metrics (EUD and dose width) will be correlated to the reported toxicity in the next section.

4.3.2 *Dose-toxicity correlation*

For the correlation of dose to outcome, the above presented dose metrics of the EUD and the dose width were related to the patient-reported outcome. In section 3.2.4, the endpoints and incidence rates were reported with rectal bleeding grade 2 showing the lowest incidence of 11.8% and the highest for bowel bother grade 4 with 39.1%. For rectal bleeding grade 1, the incidence was strongly increased compared to grade 2 with an incidence rate of 33.1%. The incidence rates are shown again in figure 4.22 along with the results of the receiver operator characteristics (ROC)

analysis for the area under the curve (AUC) .

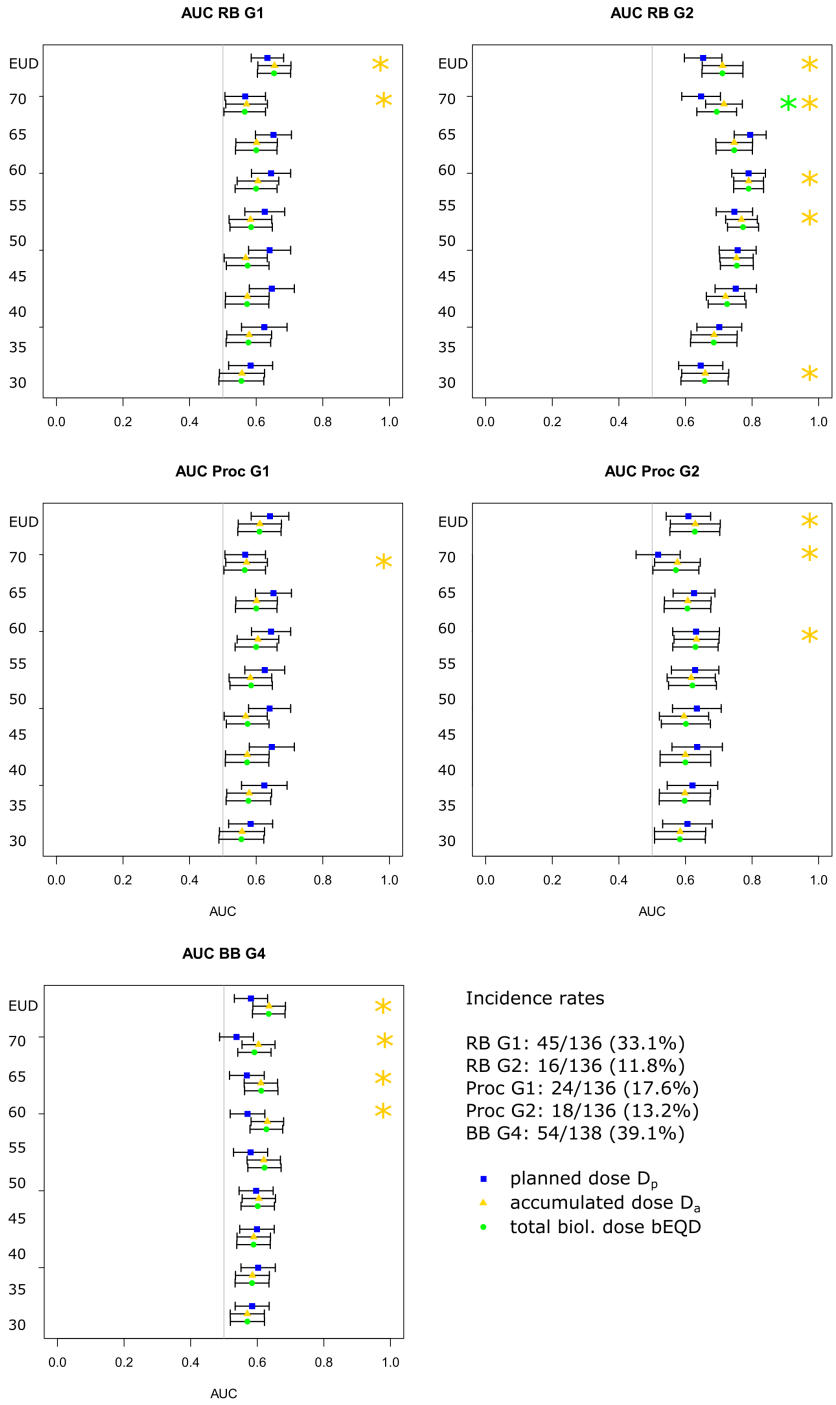


Figure 4.22. – All AUC results for the investigated endpoints: RB: rectal bleeding; Proc: Proctitis; BB: Bowel Bother; G1/2: grade 1/2. Shown for the planned dose D_p, the conventionally accumulated dose D_a and the biologically accumulated dose bEQD. Cases marked with a yellow star highlight those where AUC for the accumulated doses (bEQD and/or D_a) is higher or equal to that of the planned dose. The green star highlights the case with highest impact of using bEQD.

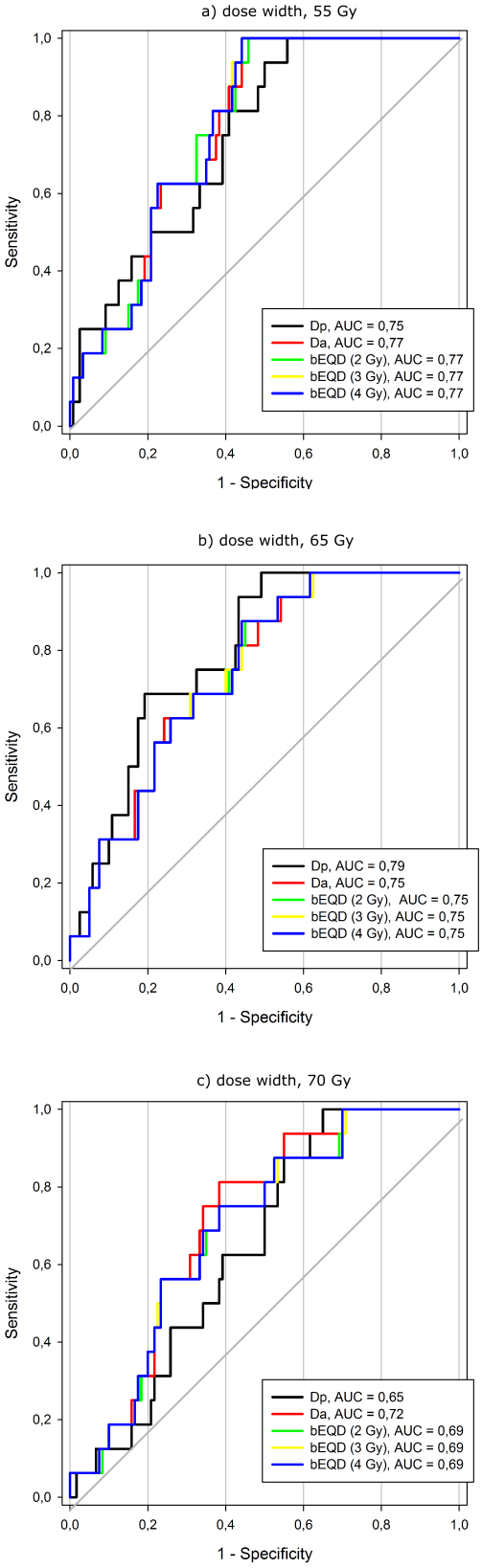


Figure 4.23. – Three example ROC curves for rectal bleeding (RB) G2. a) high AUC with that of D_α and bEQD being higher than D_p, b) AUC from D_p higher than D_α and bEQD. c) Highest difference in AUC between bEQD and D_α.

For the correlation of dose to toxicity, the same analysis was performed as reported by Shelley at al. [70] on a subcohort of the here used data. In the performed ROC analysis (see section 3.2.4) generally small differences between the planned dose and the accumulated doses bEQD and D_α can be observed. An overview of all AUC results for all investigated endpoints can be seen in figure 4.22 for the planned dose, bEQD and D_α. For the majority of cases, the AUC is close to 0.6 and shows only very little difference for the different dose metrics. Only for rectal bleeding grade 2 (RB G2), the AUC results reach values up to 0.79 with highest correlation found for the mid to high dose thresholds between 50 - 65 Gy. Cases for which the AUC is higher for the D_α and/or bEQD than for the planned dose are highlighted with a yellow star. It can be seen that for the majority of cases, actually the ROC analysis for the planned dose resulted in a higher AUC. bEQD and D_α showed a higher correlation mainly for the high dose thresholds and, for all endpoints except for proctitis grade 1, for the EUD. Three example ROC curves are shown in figure 4.23 for RB G2, dose widths with a threshold of 55 Gy and of 65 Gy which correspond to curves of high AUC, as well for 70 Gy, which shows the highest difference for the bEQD-based result, but generally lower AUC . There was no notable difference for the bEQD-based AUC results when using different underlying α/β values. Comparing the outcome of the AUC of bEQD and D_α, only very little difference can be seen which is much smaller than the already small difference to the AUC from the planned dose. The strongest difference between AUC from bEQD and D_α is shown for RB G2 (figure 4.22 and figure 4.23 c)) for a dose width threshold of 70 Gy, where the difference is at 0.03.

4.4 COHORT VOXTOX H&N

The VoxTox H&N data analysis investigates the difference between bEQD and D_a in 3D data for a different body site than the up until here analysed pelvis cases. The data is also part of the VoxTox programme so that the here proposed bEQD approach was implemented in the already existing workflow to accumulate the total dose.
In total 26 H&N cases were studied in more detail that were chosen due to an increased risk of overdosage in the brainstem and spinal cord. The brainstem was contoured in 6 cases, the spinal cord in 23 cases. Left and right parotids as well as the CTV were also studied which were contoured in all patients.
In general, in the overall imaged patient volumes the observed total dose differences were smaller than those found in the pelvic region, though the highest deviations found reached up to 8 Gy locally and even around 5 Gy locally in the spinal cord, which is higher than for standard fractionation in the pelvis regions. However, high differences were mostly localized in smaller volumes and in the peripheral superior-inferior positions of the imaged volume.
Example cases of the planning CT, a fraction mega voltage computed tomography (MVCT) and the dose difference map are depicted in figure 4.24. For all cases, the daily MVCT scans imaged smaller volumes than the planning CT scan. In the superior-inferior position, the location of these scans was not identical on all days. The difference between both accumulation methods was especially increased in the superior and posterior regions. Figure 4.25 shows two examples for the difference

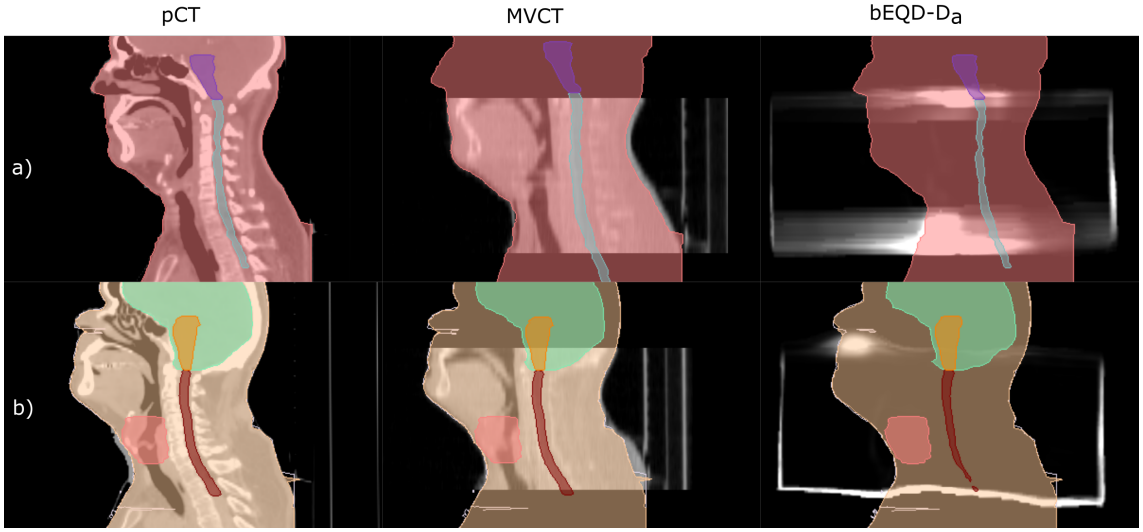


Figure 4.24. – Examples of the imaging data for two different H&N cases (planning CT, pCT, and daily MVCT) and resulting dose difference location (just qualitatively, with white showing high difference regions). Main volumes of higher dose differences were located in the superior and inferior periphery of the imaged volume. Location of the MVCT scans was not identical in all imaging fractions but shifted towards head or feet direction.

between bEQD and D_a with high differences in the regions of interest. Figure 4.26 shows the boxplot statistics of the volumes that show a dose difference higher than 1 Gy in the respective regions of interest.

For the *brainstem*, in all 6 cases, the local dose difference was higher than 1 Gy, observed in an average of $22\pm10\%$ of voxels of the brainstem volume (median of 17% with a range from 0-39% in figure 4.26). For case 1) in figure 4.25, a difference of 7 Gy locally within the brainstem can be seen. In 4 of 6 cases, the mean dose was higher than 1 Gy - by 1.6 Gy, 3.2 Gy, 1.5 Gy and 1.5 Gy, respectively.
In 21 of the 23 *spinal cord* cases, the local dose difference was higher than 1 Gy. This magnitude was seen in an average of $4.0\pm1.9\%$ of voxels of the spinal cord thus very locally. This can also be seen in the statistics in figure 4.26 where the range of difference is below 10% with a median of 4%. In both cases in figure 4.25, local maximum dose differences of 5 Gy can be seen, however in very small volumes and fast decreasing to a region of difference between 0.5 and 2 Gy. The mean dose of the spinal cord was never higher than 1 Gy. $D_{2\%}$ and $D_{5\%}$ were not different in the range of accuracy when comparing bEQD and D_a.

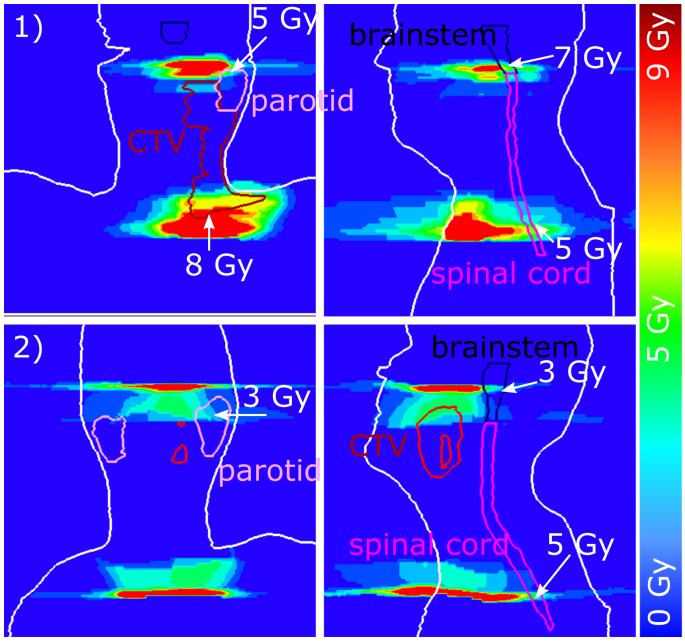


Figure 4.25. – Two examples cases of the difference $bEQD - D_a$ in Gy for the investigated VoxTox H&N cases.

Examples of high dose difference in the *parotids* can also be seen in the examples in figure 4.25. For 10 and 5 out of 26 cases of the right and left parotids, respectively, the local difference was higher than 1 Gy. The range was similar in both cases from 0 - 23 Gy. Mean doses were higher than 1 Gy in non of the cases in the parotids.
In the *CTV*, in 4 of the 26 investigated cases local differences above 1 Gy were found. Figure 4.25 shows an example of 8 Gy difference. Dose differences, as shown in figure 4.26 above 1 Gy where found in 5.3, 10.9, 3.5 and 14.9 % of the volumes for the 4 cases. The mean CTV dose was increased by more than 1 Gy in two cases by 1.7 Gy and 2.8 Gy, respectively.

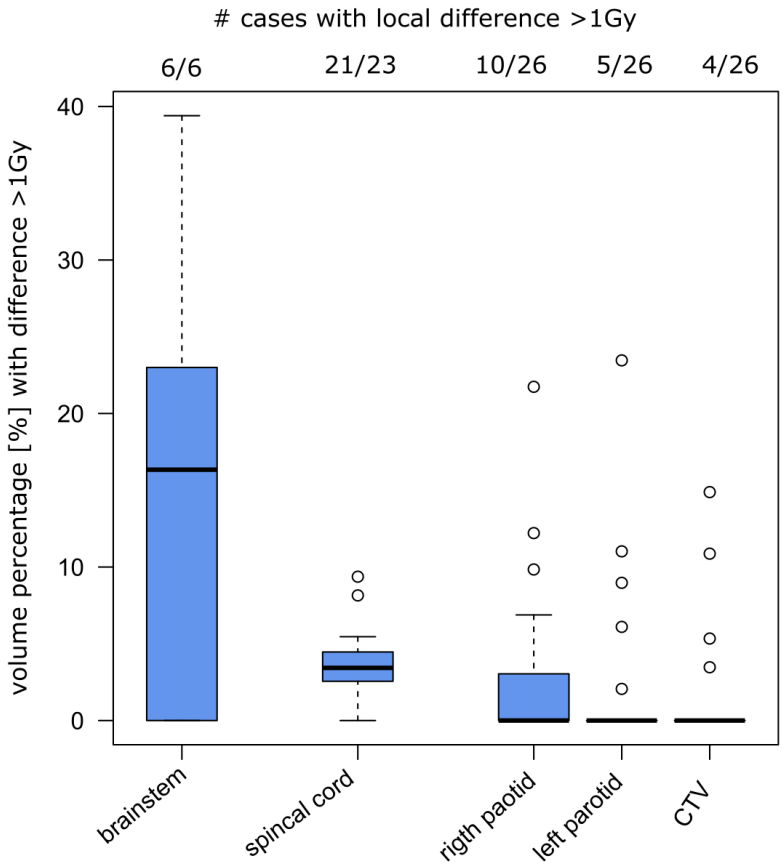


Figure 4.26. – Boxplot statistics of the volume showing a local dose difference above 1 Gy for the respective regions of interest. Note that the number of available cases of contoured structures differs.

5

DISCUSSION & CONCLUSION

Dose accumulation is getting established in clinical practice to estimate the success of radiotherapy (RT) treatments. Ultimately, however, it is the biological effect delivered in the tissue that needs to guide treatment adaptation and re-treatment and is the relevant parameter in response estimations. Section 3.1 presented the methodological workflow of **biological dose accumulation** that was derived within this work to achieve consistency of the links of individual steps to obtain a total biological effect estimation.
Two aspects have been isolated and investigated in this work, for which in current applications and research the biological effect is not adequately addressed or needs to be questioned:
i) **Dose mapping** as the basis for the computational accumulation of the total dose, needs to be implemented in a way to conserve the obtained biological effect within the tissue, as it is proposed in this work in section 3.1.1. Conventional dose mapping implementations conserve the gray value of the voxel, so that when mapping is performed on the dose distribution, a dose-conserving mapping is applied automatically. Recent postulations called for conserving the deposited energy in the mapping process, instead of the dose, in case volumes change in size [90]. In the here presented analysis of both the energy and the dose conserving approach, it was found that energy conservation violates the conservation of the obtained biological effect. It was shown here that conservation of dose is necessary for a biologically consistent dose mapping. Furthermore, this is independent of the mechanisms and direction of volume changes in the biological tissue. Remarks on implementation details are discussed in section 5.2.
ii) Conventional **dose accumulation** was found to be mathematically inconsistent with radiobiological models. It was shown that averaging of the daily delivered doses performed by using conventional dose accumulation introduces a systematic overestimation of the (cell) survival fraction (SF) and therefore an underestimation of the total dose with respect to the resulting biological effect. To overcome this inconsistency, the concept of effect accumulation and the total biological dose (bEQD) was introduced (sections 3.1.2 and 3.1.3). For effect accumulation, the steps of accumulation and effect prediction need to be interchanged. The concept of the total biological dose bEQD was derived that represents the total dose which yields the SF as obtained from the effect accumulation approach. Due to the systematic overestimation of the SF and the underestimation of the total dose by conventional dose accumulation, this effect becomes especially important for healthy tissue since

the radiation-induced effects are stronger than what was expected from the conventional approach. Thus, adaptation might become necessary at an earlier stage in order to spare the normal tissue; re-treatments would need to be optimized for lower doses in the organ(s) at risks (OARs); and in dose response-modelling, higher than expected doses would be correlated to toxicity, and side-effects might occur earlier than what conventional dose accumulation would suggest. For tumour tissue, on the other hand, the deviation between the effect- and the dose-based approaches are less important because of two reasons: as will be discussed later, the deviation between the approaches scales with the daily dose variation magnitude which is expected to be small to negligible within the clinical target region; and second, because achieving a higher cell kill than expected is advantageous for the tumour in contrast to normal tissue. Therefore, in the following, the analysis focusses on the deviations found in OARs. The investigation of the *magnitude of the differences* between the effect- and the dose-based approaches is the second aim of this thesis based on daily patient imaging data.
The discussion will continue with a general assessment of the introduced concept and derived magnitudes in comparison of the results from theory with those in the individual patient cohorts and with reports in literature in section 5.1. Afterwards, the individual investigations and workflow aspects are discussed in more detail in section 5.2. Concluding, the question of whether it is actually advantageous to perform biological dose accumulation is discussed based on the combined results and discussions with a focus on the overall uncertainty budget and possible applications of the derived approaches (section 5.3). Aspects of the here presented discussion have been submitted for publication [57] [58].

5.1 CONCEPTS, MAGNITUDES AND TRENDS

The effect accumulation approach is a simple and straightforward biologically consistent application of the Linear Quadratic Model (LQM) within the workflow of dose accumulation (introduced in section 3.1.2). Effect accumulation is based on the same data as the conventional approach and without any additional computational cost. In contrast to conventional dose accumulation, however, the value of the daily dose information does not become degraded in the process of averaging. In theoretical analysis it was found that, for strong daily dose variations, the relative SF differences between the effect accumulation approach and effect estimation after conventional dose accumulation can reach 100% relative SF difference. In patient data, relative differences around 35% were found in large volumes for the majority of investigated cases. In a single case, a hot spot of 50% SF difference was observed. These investigations were based on standard fractionation and found in tissues assigned with an α/β between 3 and 5 Gy. As will be discussed in detail in the next section, even higher deviations are to be expected for hypofractionation and lower α/β-values.

For the SF difference, the relative difference was reported while the absolute difference was studied only theoretically. Two cases have to be distinguished for the interpretation of absolute and relative differences here: High dose regions and regions of lower dose. In tumour tissue receiving generally high doses, the SF will already be so low that in good approximation, the SF difference (even when high in relative terms) will loose its importance since the absolute difference becomes irrelevant. The situation might be different when considering normal tissue that receives generally lower doses in the majority of the volume. Here, the absolute differences will be higher while the relative difference might be lower. Thus, one needs to keep in mind that a relative SF difference of e.g. 20% might be irrelevant for tumour tissue but can be significant in normal tissue. The effect from lower doses is even more pronounced in the presented analysis since normal tissue was assigned higher α/β compared with the tumour, resulting in a higher SF for a given dose.
The direct interpretation of the absolute and relative SF numbers (especially for normal tissues) and linking the quantity of SF to experience on treatment outcome, might not be common practice. Though dose itself is only linked to the outcome via models, clinical practice is often based on prescription doses and dose constraints as outlined in section 2.2.2.2. The derived concept of **the total biological dose bEQD** (section 3.1.3) describes the total dose that, when applied in the LQM, results in the mathematically correct SF as given by the effect accumulation approach. The concept of bEQD therefore enables the direct comparison of the derived biologically consistent workflow to current clinical practice of conventional dose accumulation and common dose constraints.
Effect accumulation and bEQD are biologically equivalent. It should be noted that the correction of the SF in the bEQD concept (compared to the use of D_a), as outlined in this thesis, is due to an increase in the overall dose, while the same dose per fraction as for the conventional dose accumulation approach $d = D/n$ is used (to facilitate direct comparison on the same basis of d). Thus, it needs to be considered in the calculation of the SF using bEQD - $SF(bEQD)$ - that the number of fractions is increased in the process: $n_{bEQD} = bEQD/d \geqslant n = D_a/d$. However, the calculation of $SF(bEQD)$ is unnecessarily complicated since the straight forward effect accumulation approach $SF(d_i)$ can be used instead.
Nevertheless, it might be advantageous for clinical practice to know the biologically consistent mean dose per fraction, instead of the total biological dose bEQD. Following from the above, the biologically consistent dose per fraction is *not* given by $bEQD/n$ but needs to be derived differently since n_{bEQD} is not necessarily given by an integer number and the delivery of a fraction of a dose is not feasible or well defined in clinical practice. Analogously to the bEQD approach, the ***biologically equivalent dose per fraction*** d_b can be derived from $SF(d_i) = SF(d_b)$ where $SF(d_b)$ uses equation 2.13 with the constant fraction dose d_b. The result of solving the linear-quadratic formula yields:

$$d_b = -\frac{\alpha/\beta}{2} + \sqrt{\left(\frac{\alpha/\beta}{2}\right)^2 + \alpha/\beta\, d + \sum_i d_i^2/n} \tag{5.1}$$

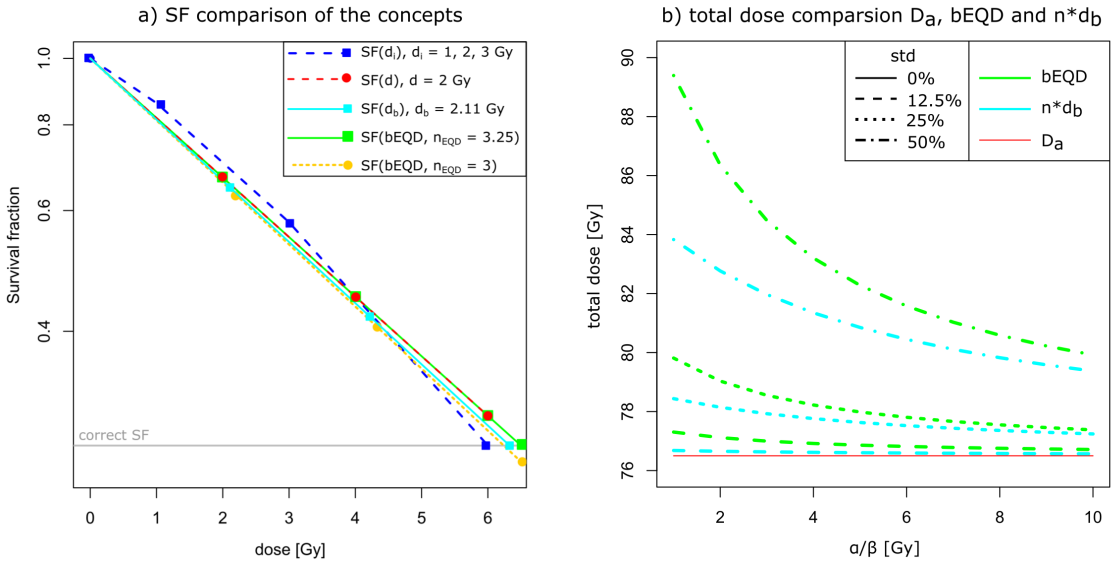


Figure 5.1. – Concept comparison: a) Example of the application of the different described approaches for a scenario where the $n = 3$ delivered fractions are: $d_1 = 1Gy, d_2 = 2Gy, d_3 = 3Gy$. $SF(d_i)$: effect accumulation approach yielding the correct SF. $SF(d)$: dose accumulation approach with $d = \overline{d_i} = 2Gy$, resulting in systematically lower SF. $SF(d_b)$: using the biological fraction dose in the LQM, n = 3, correct SF. $SF(bEQD, n_{bEQD} = 3.25)$: correct application of the bEQD concept, the correct survival fraction is achieved by delivering an extra dose on top of the three fractions, $d = 2Gy$, correct SF. $SF(bEQD, n = 3)$: wrong application of the bEQD method by assuming a constant fraction dose of $d = bEQD/n$, wrong SF. b) Comparison of the total biological dose bEQD and the biologically equivalent fraction dose d_b (as the total dose $n \cdot d_b$). Though both result in the same SF (see a), the total dose magnitudes are different. Shown for different magnitudes of fractional dose variation in % to the mean dose.

d_b is systematically larger than d, analogously to bEQD being larger than D. d_b is the constant fractions dose, that when delivered in *the same number of fractions* n as the original treatment, results in the correct SF. On the other hand, bEQD is the total dose, that when delivered in *the same mean dose per fraction* d as the delivered treatment, results again in the correct SF. For an example case, this is visualized in figure 5.1, a).

In the interpretation of bEQD and d_b, its needs to be noted that the magnitudes of bEQD and $n \cdot d_b$ are not the same (figure 5.1 b)). It was found that the latter is generally lower but follows the same trends: higher impact for hypofractionation and lower α/β values. bEQD can generally be interpreted as an increase in the total number of fractions, as outlined above. For example, in case a dose difference of 3 Gy is found in the rectum-voxel, where the mean dose per fraction was 1.5 Gy, this would mean that this spot received the biological equivalent of two more fractions than what was obtained from conventional dose accumulation. The number n_{bEQD} effectively changes for each single voxel, since the variation of the individual d_i is different in all voxels. Furthermore, n_{bEQD} is not necessarily a whole number, which is a limitation of this approach. On the other hand, the concept of d_b has the advantage, that for all voxel, the same number of fractions is equal and the

same as in the treatment plan. Only the fraction dose d_b is changed (along with the total dose $n \cdot d_b$). A direct comparison to the actual fractionation scheme that was delivered is therefore possible. In conclusion, it might be advantageous to consider d_b instead of bEQD. This concept should therefore be investigated in future work. Since lower difference to the conventional dose accumulation are to be expected than from bEQD, the same systematic deviation trend is still given and the following conclusions can in principle be also applied for d_b.
In literature, the topic of effect accumulation has been investigated only very little in three consecutive papers between 2004 to 2010. Zavgorogni [87] introduced the concept of the equivalent constant dose (ECD) to investigate the impact of fractional dose variations on the biologically effective dose (BED) (similar concept as the equivalent dose in 2 Gy fractions (EQD_2), see section 2.2.2.2). The ECD is given by:

$$ECD = -\frac{n}{2}\alpha/\beta + \sqrt{(D + \frac{n}{2}\alpha/\beta)^2 + (n\sigma_d)^2} \tag{5.2}$$

with n the number of fractions, σ_d the dose standard deviation and the total dose D. The ECD represents the total dose that provides the same SF as the given by the fractional varying doses. In comparison to the here presented approaches, this approach is identical to the above described concept of the biologically equivalent fraction dose d_b and results in the same curves as shown for $n \cdot d_b$ in figure 5.1 b). Zavgorodni reformulated the equation to include the dose standard deviation. At the time, this was advantageous since no daily dose data was available. With the now available daily doses from imaging and dose mapping, this detour in the computation process is not necessary any more and d_b provides the same, but straight forward calculation approach. Zavgorodni only estimated dose variation magnitudes that were based on setup errors of the patient. Further investigations for a theoretical patient were performed by Bortfeld et al. [7] two years later where the ECD is used under the name of the normalized total dose (NTD) for varying fraction doses. Investigations were again only based on dose variations estimated from setup errors without the use of a voxel-wise daily dose estimation from imaging and registration. In their conclusion, the impact of the NTD only becomes important for standard deviations above 10%. The concept and impact of daily dose variations on a biological basis was only studied in this manner by one other group of authors in 2010. De Xivry et al. [17] investigated the legitimacy of dose accumulation on a biological background on the basis of the ECD from Zavgorodni. The concept was tested for 10 head and neck (H&N) patients that underwent weekly computed tomography (CT) scans and treated with helical Tomotherapy. A demons based deformable image registration (DIR) was performed to map the daily forward calculated doses to the planning CT. The results were investigated on the basis of dose-volume histograms (DVHs). In their conclusion, they found no clinically relevant difference between conventional accumulation and biological accumulation in the DVH parameters. However, they state that *"Care must be taken if further clinical practices require local rather than global surrogates to adapt treatment"*. The

reported magnitudes are generally lower than the here observed values. This is however to be expected because of two reasons: as described above, the d_b approach, in comparison to bEQD results in generally lower deviations and secondly, within the H&N region, lower motion amplitudes might be observed, as will also be shown for the here investigated cohort.
The study presented here compliments the above described literature in multiple ways: the concept of the bEQD was investigated for daily imaging data, which was used by none of the above authors; the impact of hypofractionation was investigated; the analysis was performed locally in addition to volume-based, also on the basis of dose surface maps (DSMs); the uncertainties of bEQD and D_a were investigated and compared; the investigations focussed on another body site (pelvis) that regularly shows high motion amplitudes; and the impact on dose-toxicity correlation was studied. All these individual aspects will be discussed in the next section while in the following, the found magnitudes of total dose difference will be compared.
It was found that the SF fraction difference can be well described by the total dose difference in magnitude and extent as shown by comparing figures 4.3 with 4.4 and table 4.1. Thus, the focus in this thesis was put mainly to the total dose difference for an easier interpretation of the results as well as comparison of total dose metrics rather than SF values.

In summary, the **effect accumulation approach** is a straightforward solution to facilitate correct accumulation of the cell survival fraction. Therefore, this should be used instead of effect prediction after conventional dose accumulation. The derived concept of the **total biological dose bEQD** provides the total dose that predicts the same SF as effect accumulation. With this, bEQD facilitates the comparison of a biologically consistent total treatment doses to the planned dose and dose constraints. **bEQD is systematically higher than** D_a, implemented by an increase in the number of fractions. This deviation makes bEQD especially important for **investigations in normal tissue**. For a more clinically applicable approach, the **biologically equivalent dose per fraction** d_b can be used, which conserves the number of fractions but increases the dose per fractions to be biologically equivalent to effect accumulation.

Magnitudes of bEQD-D_a The **theoretical analysis** of a fictional voxel undergoing certain dose variations showed generally strong deviations when using the dose or the effect based approaches for the SF and the total dose. The strongest impacting factor on the difference was given by the daily dose variation magnitude. For both the SF and the total dose, a dose variation magnitude below 12.5% standard deviation resulted in very low impact of below 10% relative survival difference and dose differences below 1 Gy. This is in accordance with the observations by Bortfeld et al. [7] described above. On the other side, for high dose variation magnitudes, e.g. 50%, the differences can reach 100% relative SF difference and up to 12.9 Gy total dose difference. From this, it is to be expected that a strong deviation between

the two accumulation approaches will only be seen in areas of high organ motion amplitude located in dose gradient regions. In regions of either little organ motion or areas of uniform irradiation, none or only small differences are to be expected. However, the magnitude of SF and total dose difference can reach clinically significant orders of magnitude.

The magnitude of SF difference and total dose difference **in patient data** show the ranges expected from the theoretical analysis. In all cases, the systematic difference to the dose-based methods was confirmed. For the nine analysed patient cases of the **German Cancer Research Center (DKFZ) pelvis cohort**, the detected difference hotspots were around 4 Gy in five cases. These differences increased to more than 8 Gy when changing to a strong hypofractionated treatment scheme. Maximum difference for standard fractionation was at 4.9 Gy, at 6.8 Gy for moderate hypofractionation and at 8.4 Gy for strong hypofractionation. Highest differences were localized around the target volume in the dose gradient regions. The differences depended most strongly on the location and magnitude of organ motion and were therefore highly case specific. In comparison to theoretical values, even higher magnitudes could be expected when organ motion in critical dose gradient regions increases. As expected, in regions of low motion amplitude or uniform dose fields (like in the Clinical Target Volume (CTV)), the differences remained marginal. Most, though not all hot spots of total dose difference were located in the volumes of the contoured OARs of bladder and rectum. In addition, high dose deviations were observed between the three organ contours in the regions of the seminal vesicles and neurovascular bundles. The impact on these structures should therefore be investigated.

For the **VoxTox DSM** cohort with 251 patient cases and using the same rectum α/β value of 3 Gy as in the DKFZ cohort, the maximum deviations were 2.5 Gy with an average of 0.74 Gy and the highest mean dose difference for the entire area was 0.25 Gy with an average of 0.08 Gy. Thus, the observed deviations were generally smaller than for the DKFZ cohort but in a comparable range for maximum values. There are several reasons why this magnitude might be lower: i) By only looking at the dose deviation for the rectum, it is possible that other volumes of high local dose deviation (as found in bladder and other organs in the DKFZ cohort) were not considered in the analysis. ii) The analysis of the dose difference for the DSMs was not based on the use of DIR. The daily contours of the rectum were overlayed without a voxel-wise mapping. In case that daily dose deviations occurred repetitively in the same region of the rectum, shifts of the entire organ might have washed out this effect in the analysis. It would be necessary to include DIR in the analysis process to avoid this strong limitation. Furthermore, the contours on the daily images were generated automatically, which induced another source of uncertainty. This was, however, necessary to enable the analysis on the large amount of data. iii) The observed statistical difference might have been induced by different clinical protocols. The VoxTox cohort was treated with intensity-modulated radiotherapy (IMRT) in a TomoTherpay unit, while the DKFZ cohort received standard step-and-shoot IMRT.

It was found in literature that for TomoTherapy and step-and-shoot IMRT, different degrees of healthy tissue sparing and tumour dose homogeneity could be achieved (section 2.1.2). Furthermore, for the DKFZ cohort, image guidance was used for setup error correction on the basis of the target position. For the VoxTox cohort, daily imaging was used to control the status of the rectal volume and if changes were too large, remedial action was taken. The different strategies might have had an impact on the dose delivered to the rectum. In summary, it needs to be concluded that the dose deviation magnitudes depend on the clinical strategy and data post-processing. Both cohorts show, however, that the magnitudes can reach clinically significant values in individual cases, on a local scale. Section 5.2 will further discuss the outcome of the dose-volume-based results of both cohorts on a statistical basis.
The **VoxTox H&N cohort** showed cases of both lower and higher dose deviation magnitudes than in the pelvic region. In the majority of volume, the detected deviations were marginal. However, difference hotspots were detected with highly varying magnitudes of up to 8 Gy locally. For the brainstem, all 6 cases showed deviations above 1 Gy locally in large volumes (median of 17% volume). Deviations were much more localized for the spinal cord with a median volume of 4% showing deviations above 1 Gy. This is mainly due to the fact that most deviations were localized in the peripheral region of the images. A reason for this might be uncertainties in the DIR process. This is due to the fact that the daily mega voltage computed tomography (MVCT) scans did not cover the same volume in all cases and were generally shorter than the planning CT scan. This can enhance the uncertainties in the regions were not all MVCTs overlapped. However, the magnitude of impact of these uncertainties could not be quantified with the available methods (and can generally only be estimated, see below for the discussion on DIR). Smaller dose variation magnitudes might be based on overall smaller motion amplitudes than in the pelvic region, though, as just concluded before, clinical strategies impact the magnitudes as well. De Xivry et al. [17] performed ECD-based analysis for H&N cancer patients and found highest difference up to 2.6 Gy. In comparison, the magnitudes found in the here presented cohort were higher. De Xivry only analysed the data from weekly CT scans in contrast to the daily MVCT scans of the here presented cohort, which might be a reason for the different dose deviation amplitudes. The trend towards lower magnitude compared to the pelvis data is however in accordance with the presented findings.
Generally, the comparison to the cases from pelvis and H&N suggest that it is of interest to perform further investigation with a higher number of cases and different tumour sites, especially in regions that are expected to show high motion amplitudes as for example in the thorax and abdomen. The results of the magnitudes from theory further underline that difference between the two dose accumulation principles depends strongly on the local circumstances and can therefore exceed the highest here observed dose difference by a factor of two.
While the difference between the two dose accumulation strategies is based on the magnitude of dose variation from day to day, this does not necessarily coincide with

larger systematic **differences to the planned dose**. For example, the patient's bladder might have been larger during the planning CT than in any other fraction, leading to a systematic difference to the planned dose with only little day to day dose variations. This example is a likely explanation for the presented DVH of case 3) in figure 4.6 where there is very little difference between bEQD and D_α for the bladder but a higher difference to the planned dose. For the rectum, there was likely a higher magnitude of day to day motion leading to the depicted deviation hot spot and DVH difference although the total dose was lower than the planned dose. More discussion on the DVHs will be given in section 5.2. For the VoxTox cohort, the difference between the planned dose and the accumulated doses (D_α and bEQD) is very high. Generally, for the mean and the maximum dose deviations, the difference to the planned dose exceeds those between the accumulation strategies by far. Maximum difference to the plan was at 37 Gy and at 3.5 Gy for the mean rectum surface dose. Figure 4.18 b) showed however that the difference between both accumulation strategies has a visible impact on the overall difference to the planned dose and figure 4.19 showed that the mean dose increase from using bEQD instead of D_α was up to 25% of the mean dose with a peak around 6%.

In summary, the magnitudes of observed total dose and relative SF deviation are in accordance with theory and match between the three investigated patient cohorts and with the limited literature reports. The highest deviations are found in **dose gradient regions that coincide with organ motion** and therefore especially for **organs at risk surrounding the target**. Theoretically, **higher than observed deviations are possible** within realistic dose variation magnitudes. Though deviations were in the same order of magnitude, statistical differences in the total dose deviation magnitudes were observed between the investigated patient cohorts. This shows an impact of the deviation magnitudes on the method of dose delivery, the image-guidance strategy and the tumour site.

5.2 INDIVIDUAL WORKFLOW ASPECTS, UNCERTAINTIES AND RESULTS

This section discusses the individually investigated workflow details given by: Implementation remarks on biological dose mapping; the impact of α/β; the impact of the fractionation scheme; local vs. volumetric analysis (including the discussion on dose-volume (DV)-parameters); γ-analysis; DIR accuracy; interpolation; and the dose-toxicity correlation.

Implementation of biologically consistent dose mapping In addition to the discussion and conclusions given in section 3.1.1 that describes the concept of biologically consistent dose mapping, the following should be consider for its implementation:

Mean organ dose conservation: Depending on the location of growth and shrinkage, deformation of tissue within one organ may not always be uniform, which is a reason for the unknown quality of transformations on a voxel scale. It might be that parts of the organ undergo greater volume changes than others [76]. It is also not known where cells are lost or added, which is another source of uncertainty when assigning dose values by the mapping process. This is a general problem beyond the knowledge available from imaging. In many cases, organ dose is non-uniform especially in the presence of steep dose gradients in normal tissues as delivered by IMRT. As a conclusion, the mean organ dose before and after mapping may change accordingly and should therefore not be used as a dose mapping constraint.

Uncertainties from DIR: DIR is most likely the largest source of uncertainty in the dose mapping process. Instances where structures show uniform imaging intensities, but lie in steep dose gradients, are especially susceptible to dose mapping errors [74]. Errors in the mapped volume in the form $V_2' \neq V_1$ (mapped volume and reference volume), directly translate to dose errors in the overall structure. Volume changes between the reference state and during therapy at the level of the individual voxel can be estimated with the determinant of the Jacobian matrix of the deformation vector field (DVF) (see section 2.2.1.3). However, there is as yet no mechanism by which to directly measure and validate the DIR process in a physical system to calibrate in-silico modelling [68]. This is due to the anisotropic manner of deformation on a voxel scale discussed above not visible in imaging. For structures receiving inhomogeneous dose levels, constraints demanding volume conservation in the form $V_2' = V_1$ will therefore not necessarily ensure better results for the mapped dose since errors might be hidden on the voxel scale. Dose accumulation should only be performed in case the DVF from DIR are anatomically plausible (quantification of this is still a topic of ongoing research). On the other hand, volume conservation in the form $V_2' = V_2$ does not allow for dose accumulation in a common reference frame.

Interpolation: Values assigned to a voxel are given by the value interpolated at the grid centre point of that voxel in common implementations (e.g. Elastix [37]). The principle of interpolation is explained in detail in section 2.2.1.4. Thus, dose is a point value assigned to a volume and its accuracy in the mapping process depends on the used interpolation method. For the above discussion, it therefore is the best possible estimate of the interpolated dose to a voxel that needs to be conserved in the dose mapping process to calculate the SF. As tissue shrinkage or growth is always associated with cell transfer between neighbouring voxels in the mapping process (in all discussed scenarios due to volume displacement in figure 3.2), voxels in the reference frame (after mapping, T_2') may be composed of cells receiving different dose levels before mapping (T_2), which is a general problem due to discretization. The assigned dose to this voxel is interpolated. Only in the case of uniform irradiation of the voxels included in the interpolation process, is the dose to individual cells entirely conserved. However, in steep dose gradients, the shifted cells will be assigned with a different dose within the newly composed voxel due to interpolation, also in

dose conserving implementations. For a given dose gradient, this dose difference will increase with decreasing resolution of the dose grid due to the discretization. If the number of cells exchanged between voxels is small (i.e. small changes in cell number or density), the shifted cells will travel only over a small distance and the dose alteration actually received by these cells will be small as well. Furthermore, for a high resolution of the dose grid, this small dose difference will also be assigned to the shifted cells. Thus, in the limit of small changes of cell number or density and high resolution of the dose grid, the above mentioned condition of a uniform irradiation is asymptotically fulfilled on the scale of neighbouring voxels and combining these voxels, not only SF but also the absolute number of surviving cells will be maintained after dose mapping. These assumptions appear to be valid for a standard fractionated treatment at least for the case, when volume changes of the tumour or normal tissue are not too large. Both dose and SF are subject to uncertainties when being interpolated over a volume. Although SF could be mapped and interpolated as well, SF changes nonlinearly over many orders of magnitude making it more susceptible to discretization errors than dose interpolation and the use of simple interpolation techniques such as trilinear interpolation may also be questioned. Therefore, it is dose that should be mapped according to the DVF and not the SF directly. This is not based on biological considerations but on minimizing the uncertainty in the overall workflow.

In summary, in the process of dose mapping, **mean organ dose as well as volume conservation should not be used as constraints for DIR**.
Discretization and interpolation induces uncertainties in the mapping process which will be lower by **mapping the voxel dose rather then the voxel-wise SF**.

The impact of α/β In the following, the impact of α/β on the SF and bEQD is discussed generally since results in theoretical analysis (section 4.1) and in patient data (section 4.2.1) are in good agreement. In short summary: The lower the α/β, the stronger the difference in relative survival and in the total dose between the dose and the effect-based approaches. Furthermore, assigning smaller values for α and β, while the ratio α/β remains the same, results in generally lower relative differences. Lower α/β generally means that β increases with respect to α. Therefore, the β-term is weighted higher in comparison to calculations based on higher α/β. The difference between the dose and effect accumulation approach is manifested in the β-term alone, explaining the observed higher differences for lower α/β.
As presented with the basics of the LQM (section 2.2.2.1), the smaller the tissue α/β the higher the SF for a given dose. Smaller α/β are associated with slow proliferating tissues which in turns means that damage to the cell can be repaired more efficiently before the next cell cycle begins and the damaged cell might undergo mitotic cell death. On the other hand, late radiation damage might occur. For the evaluation of the SF from either dose or effect accumulation, this means that for smaller α/β, the absolute SF difference between both methods is naturally higher

than for larger α/β since the overall SF is higher. As can be seen in figure 4.1, the SF for high α/β values is already so small that differences are hardly visible on the given scale. This, however, changes when looking at the relative SF difference. Here, it was shown that for lower α/β, the relative difference is higher for high α/β. As discussed above, high relative differences are more important in medium and low dose regions, while for high dose regions, as in the CTV, a high relative difference is most likely related to very small absolute difference based on the overall SF scale.
For the computation of the SF, α was set to a constant value of 0.1 Gy and change in the α/β value was achieved by changing β only. This is reasonable for the given analysis since only the β-related quadratic dose term is affected in the comparison between the effect and the dose accumulation approach and is therefore of higher importance. The same approach is followed for example by Afsharpour et al. [1]. Changing α while keeping the same α/β of course means that β was changed by the same factor. However, as was presented in LQM background in figure 2.9, the same α/β with underlying smaller α and β results in lower cell kill (higher SF). Therefore, the same as above applies that the absolute SF difference is higher, but the relative difference is smaller.
For the comparison of bEQD with D_a, only the assigned α/β value is of interest. It was found that the difference between both total doses scales in the same way with α/β as the relative SF difference: the lower α/β, the higher the difference. Within the bEQD formula (eq. 3.10), D_a is directly factorized with α/β, so that a lower α/β leads to a smaller weighting. This trend is in accordance with the above discussed observation that the SF difference is more pronounced due to the stronger weighting of the β-term, which is essentially the same since α was kept constant.
Due to the impact of α/β on bEQD, uncertainties in the assigned value propagate to bEQD itself. Lowering the α/β for the rectum for case 3) from $3\ \mathrm{Gy}^{-1}$ to $2\ \mathrm{Gy}^{-1}$ increased the local dose difference by $1.2\ \mathrm{Gy}$. It was found that an increase from 3 to $4\ \mathrm{Gy}^{-1}$ had a lower impact than the reduction from 3 to $2\ \mathrm{Gy}^{-1}$, of only $-0.7\ \mathrm{Gy}$. Only an increase to $5\ \mathrm{Gy}^{-1}$ resulted in a similar change in total dose difference of $-1.2\ \mathrm{Gy}$. Thus, tissues with low α/β are prone to higher uncertainties inherited from the α/β uncertainty. The uncertainty scales with the dose variation magnitude as well so that areas of high $\mathrm{bEQD} - D_a$ difference are affected more. The impact was much lower for the CTV where the overall total dose differences were smaller. The impact of around $1\mathrm{Gy}$ to the total dose deviation in high deviation areas is generally lower than the difference between bEQD and D_a.

In summary, **late responding organs** that show a low α/β value are at risk for a stronger underestimation of the achieved biological effect and total biological dose. The uncertainties of the α/β values induce an uncertainty of bEQD. The impact on the bEQD uncertainty is higher for tissues with a lower α/β. The magnitude of uncertainty of bEQD can therefore be lowered by improving the accuracy in the α/β-estimation. This magnitude is lower than the difference between bEQD and D_a but need to be considered in the overall uncertainty of bEQD.

Thought experiment: The influence of dose variations on the prediction of the α/β-value Very broadly summarized, there are two ways to predict the α/β-value of a tissue. Either it is based on clonogenic assays, as already described before in section 2.2.2.1; or it is based on patient-data, which means it is an endpoint-related α/β-value based on dose-to-outcome data from patients. This would involve the following steps: Dose volume parameters of the dose distribution are calculated that are normally linked to the normal tissue complication probability (NTCP) calculated for this tissue and for a certain endpoint (for example the equivalent uniform dose (EUD)), see section 2.2.2.2. Then the NTCP curve is generated based on the chosen model and the treatment outcome data. This needs to be done for different fractionation schemes. For each fractionation scheme, the D50 is calculated from the obtained NTCP curve (D50: the dose that results in a 50% complication probability). These D50 values are then plotted against the dose they are linked to, which gives the so called *Douglas-Fowler plot*. The slope of the resulting curve gives the α/β-ratio.
It is beyond the scope of this thesis to predict any α/β-value from reported outcome due to the lack of data based on the only very recently acquired patient datasets. In the following, a thought experiment is presented on how the effect accumulation approach might impact the α/β prediction.
Assuming a certain biological effect E as an endpoint is observed clinically, this can either be assign to the dose- or the effect-accumulation-based survival fraction. Starting from the above equation setting $E_d/\alpha = E_{di}/\alpha$ in order to convert from varying to constant treatment doses, it follows that:

$$\frac{D^2/n}{\alpha/\beta_d} = \frac{\sum_i d_i^2}{\alpha/\beta_{di}} \tag{5.3}$$

For this, a constant α for both approaches was assumed. The α-term is not affected by dose variations since it only links to the total dose of the treatment (equation 3.10).
With this, the influence of dose variations on the prediction of the α/β-value can be estimated by:

$$\frac{D^2/n}{\sum_i d_i^2} = \frac{\alpha/\beta_d}{\alpha/\beta_{di}} = \beta_{di}/\beta_d \tag{5.4}$$

According to Jensen's inequality (equation 3.3), it follows that:

$$\frac{(\sum_i d_i)^2}{n} \leqslant \sum_i d_i^2 \tag{5.5}$$

which concludes to:

$$\alpha/\beta_d \leqslant \alpha/\beta_{di} \tag{5.6}$$

Thus, conventional dose accumulation would predict a systematically lower α/β-value than effect accumulation. This is shown in figure 5.2 for a fictional voxel analysis for different dose variation magnitudes given as the dose standard deviation. It can be seen that for standard deviations below 20% the impact remains

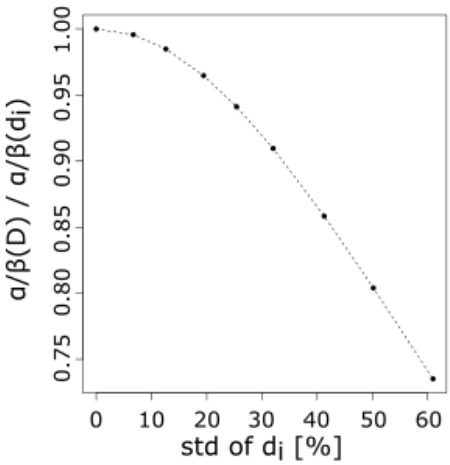


Figure 5.2. – The impact of dose variations on α/β prediction based on the same fictional voxel analysis as presented in section 4.1, according to eq. 5.4. Standard deviation (std) with respect of the mean dose.

below 5%. However, for higher dose variations, the impact quickly increases to 20% for 50% standard deviation of dose variation.

Assuming that the variation of the fractional dose can be described by a mean dose value with a dose variation factor given by a probability distribution, then it can be written as $\sum_i d_i = \overline{d} \sum_i g_i$ with $g \in \mathbb{R}_0^+$. This leads to:

$$\beta_{di}/\beta_d = \frac{D^2/n}{\sum_i d_i^2} = \frac{\overline{d}^2(\sum_i g_i)^2/n}{\overline{d}^2 \sum_i g_i^2} = \text{const.} \tag{5.7}$$

Therefore, the impact of the dose variations over therapy on the α/β-value is independent of the mean treatment dose and only depends on the magnitude of dose variations.

The impact of the fractionation scheme The difference between the two accumulation methods scales with the total delivered dose. It is therefore higher the more fractions are accumulated and the higher the dose per fraction. The latter leads to an increased risk of total effect underestimation in normal tissue for hypofractionation, based on the higher prescription dose and potentially steeper dose gradients. A higher dose per fraction increases the difference in the dose accumulation methods since the difference is given in the β-term that is quadratic to dose. Though the total treatment dose is higher for standard fractionation compared with hypofractionation, it was found that a higher dose per fraction given for the latter case is dominating the effect of deviation between bEQD and D_a, leading to higher difference amplitudes for hypofractionated treatment. In rescaling the dose to moderate hypofractionation with 3.0 Gy per fraction (21 fractions) and strong

hypofractionation with 6.8 Gy per fraction (5 fractions) it was found that the difference magnitude increased strongly in the majority of cases, again with highest difference in dose gradient regions in OARs. Going to moderate hypofractionation, highest differences increased to 6.8 Gy (compared to 4.9 Gy in 34 fractions). For strong hypofractionation, three cases already showed a total dose difference above 8 Gy which is already 24% of the tumour prescription dose but within the normal tissue. This is in agreement with the theoretical predictions.
One should note that the comparison of the total dose difference for different fractionation schemes is not only influenced by the choice of the treatment dose, but also by the "choice" of the fractions that are taken into account in the presented analysis. It might be that by reducing the number of fractions, this might exclude or include fractions of high motion amplitude present in the analysis with more fractions. This will impact the outcome of the observed deviations and can explain the lower dose difference found for hypofractionation observed for 3 out of the 9 cases. To best simulate the hypofractionated treatments, fractions 1-21 and 1-5 were chosen, respectively. However, hypofractionation might lead to an earlier decrease in tumour volume and thus a different pattern of organ deformation than for the presented analysis.
It should also be noted that scaling of the dose was chosen over setting up new hypofractionated treatment plans to ensure a more direct comparison of the total dose difference. Scaling was performed on the already deformed doses $\mathtt{d}_i$. Therefore, the observed difference did not depend on the treatment plan optimization or the DIR but solely on the fractionation scheme and analysed fractions. It was found that by creating new treatment plans, resulting differences can change a lot in hotspot magnitude and location based on different beam weights that might or might not traverse areas of higher or lower motion amplitude.
The LQM is only recommended for use in standard fractionation dose regimes. Therefore, for the high dose regions, its validity might be questioned. However, other models like the linear-quadratic-cubic (LQC) model imply even higher uncertainties in the additional tissue dependent parameter γ and with 6.8 Gy per fraction, the high dose regions are still within the applicable range from a number of suggestions [34]. Therefore, uncertainties might not be decreased with the use of these models. However, in case a cubic parameter is included, this would only enhance the observed trend by further increasing the weight of individual dose variations.

In summary, differences between bEQD and $\mathtt{D}_\alpha$ are higher, the higher the dose per fraction, since the effects from daily dose variations are enhanced. Therefore, **hypofractionated** treatment schemes are expected to show heavily increased differences between bEQD and $\mathtt{D}_\alpha$. Difference magnitudes above 8 Gy difference were observed for strong hypofractionation (about 25% of the tumour prescription dose) in normal tissue.

Local vs. volumetric analysis The difference magnitudes found locally are in a clinically significant range. However, the only small differences visible in the DVHs of bEQD and D_a for standard fractionation indicate that when clinical decision or dose-response modelling is solely linked to few distinct DVH-parameters, there will likely be no change in outcome when using one or the other dose accumulation strategy. This is in accordance with conclusions by de Xivry et al. [17]. Impact was however strongly increased in hypofractionated cases with generally higher dose deviations also visible in DV-parameters and here, the difference between the two accumulation strategies might have clinical impact.
With respect to the local differences that can be seen in the three-dimensional (3D)-data slices, this highlights the limitation of the use of dose-to-volume-based simplification of the 3D dose information and is a serious over-simplification of the spatial dose distribution in finding correlation of dose to toxicity. Especially based on this rather local than global occurrence of total dose difference, the impact in using bEQD rather than D_a has a higher importance for serial type organs like spinal cord or rectum. Generally, the computation of dose-volume parameters is more meaningful for parallel-type organs for which the volume that receives a certain dose is of interest. The here investigated organs of rectum and bladder show a more serial-type nature. Therefore, the found hot spots in deviation are more important than what is visible in the DVHs.
For the derived **dose-volume metrics**, the difference between the two accumulation methods was generally low on average, although systematic as expected, especially in the comparison with the difference to the planned dose, which was generally higher by at least one order of magnitude. The difference to the planned dose showed median values very close to 0 for all metrics. This is mostly due to the fact that the results differed in both positive and negative direction. Conclusions on dose variations should therefore only be taken when looking at the range of results, rather than a mean or median value.
The **Quantitative Analyses of Normal Tissue Effects in the Clinic (QUANTEC) recommendations** for dose constraints were evaluated for the DKFZ pelvis cohort in the three investigated fractionation schemes. It needs to be noted that the recommended constraints for rectum [51] and bladder [75] are only given for standard fractionation. The here used constraints were converted using the EQD_2 approach with the according α/β values for rectum and bladder, but are not to be considered clinically approved recommended constraints. As can be seen in the results, due to the fact that the α/β values for bladder and rectum were higher than that of the prostate, the highest dose regimes calculated from conversion for the OAR V_{xGy} are not even delivered to the patient in the CTV so that these values were not reached. Furthermore, especially for the volume of the bladder but as well for the rectum, the QUANTEC reports clearly state that the application of DV constraints is highly problematic since they are based on static planning CTs and highly susceptible to changes in volume, while volume changes in the bladder occur in every fraction. Similar caution must be taken for the constraints for the rectum.

The results showed for the majority of cases, that the constraints were still fulfilled for the accumulated doses (bEQD and D_a) but for some exception, violations of the constraints were found. This underlines the importance of using dose accumulation in general, especially for treatment adaptation in individual cases. However, the violations were already observed for D_a. bEQD only resulted in a small additional deviation but to no case were constraints were fulfilled by D_a but not bEQD.
The differences in the DV-metrics for bEQD and D_a were visibly increased for hypofractionation, in both median and range. In the case of strong hypofractionation, an outlier was observed with 7% difference in V_{21Gy} for the rectum and 4% in V_{23Gy} for the bladder. Violations of according appropriate constraints for hypofractionation might therefore occur. This indicates that the bEQD might probably not be relevant in a statistical manner but can have high impact for individual cases. This will be further discussed for the VoxTox cohort below.
For the DKFZ cohort, the dose metrics of **bladder and rectal wall** were compared to those of the entire organ. The results for the organ walls were generally shifted to smaller values (either closer to 0 or higher in negative numbers). Especially the bladder volumes that received certain dose values tend to shift towards negative values in contrast to the whole bladder. This indicates that the largest dose difference to the planned dose and between the accumulation methods is located within the organ volume. However, the differences between results of the respective investigated volumes remained small in all cases. The more labour intense contouring of the organ walls can however improve the accuracy of the investigated DV-parameters.
For the **VoxTox DSM cohort**, it was found that the higher the threshold for the dose width computation, the more negative the median difference to the planned dose. At 30 Gy threshold, the difference is close to 0 and shifts towards -3 (relative lateral extend) at 70 Gy. In comparison to the difference to the planned dose, the difference between both accumulation strategies remains marginal, but still showing the expected systematic difference. As discussed before, not using DIR to map the dose on a voxel-basis might have washed out local dose deviation hot spots so that observed dose variations magnitudes were decreased wrongly. The magnitude of difference from accumulated to planned dose metrics is generally very high for the dose width and the EUD. The EUD showed to be systematically lower for the accumulated than the planned dose in the vast majority of cases. As was described before, the mean dose was systematically higher for all cases. This is however no contradiction but is simply based on the volume-effect parameter in the EUD calculation (see section 2.2.2.2). With $a = 11.11$, a rather serial type nature is assumed for the rectum, thus, enhancing the impact of local dose hot spots. Based on this observation, apparently the dose distribution showed high dose gradients and hot spots for the planned dose, while this effect was washed out for D_a and bEQD since dose hot spot location varied from day to day. The wash out effect then impacts the result of the EUD calculation with the chosen a to a high extent leading to the observed lower than planned EUD. This underlines the previous assumption that D_a and bEQD were generally washed out either truthfully or to an increased extent

based on not using DIR. EUD for bEQD and D_a thus showed to be very similar. The DV-parameters for the spinal cord observed in the **VoxTox H&N cohort** showed no impact from the dose accumulation methods. Generally, higher deviation were observed for the brainstem than for other organs, though, as discussed above, the origin of the difference hot spots in the peripheral image regions cannot be linked to dose variations without doubt. This will be discussed below for the DIR accuracy.

In summary, due to the local impact of the bEQD, deviations of bEQD to D_a showed only marginal impact in conventional DV-metrics in standard fractionation. The difference between the accumulation methods in **conventional DV-analysis only becomes clinically relevant for hypofracionation**. With regard to the observed local deviation hotspots of several Gy, this underlines the limitations of DV-analysis. bEQD should especially be investigated for **serial-type organs** which are affected most by local dose hot spots. Statistically, the difference between both accumulation methods is low compared to the difference in planned dose but can be high for individual cases. DV-metrics for organ walls showed to be lower than those for the whole organs to a small degree.

γ-analysis The result from the γ-analysis underline once more the importance of local dose analysis. The γ-value varied strongly, even within one organ. It was also found to be important to differentiate between regions that received lower than planned dose from those that received higher doses in the γ-analysis since otherwise the γ-passing rates would be much lower than reasonable. However, it should be noted that this is only true for the normal tissue regions since in tumour tissue, underdosage instead of overdosage is critical. Within the CTV, areas of negative γ value are thus of interest, nevertheless, these regions generally showed values close to 0 and are therefore within the limits of acceptable γ-criterion.
It was found that the difference hot spots between bEQD and D_a often correlated to regions of already failed γ-criterion, thus, further increasing the γ-value but not necessarily increasing the γ-failing rate. The equivalent is given for $\gamma \leqslant -1$. This explains the only low impact of the accumulation strategy on the overall passing rates. It is therefore of higher importance to consider the actual magnitude of γ in the failed regions, rather than the passing-rate alone. When only passing- and failing-rates are considered, while the magnitude of the γ-value is neglected, a region of $\gamma \geqslant 1$ will be treated in the same manner independent of the fact that it shows values close to 1 or far beyond. Here, it was found that the γ-value can be strongly increased in failure regions and that based on this analysis, clinical decisions on D_a and bEQD, respectively, might have a different outcome. Therefore, for γ-analysis, the bEQD approach should be considered.
For the analysis a criterion of 3% / 3 mm was chosen. For stricter criteria such as 2% / 2 mm or 1% / 1 mm, a similar trend is expected with even stronger influence

of bEQD, since bEQD impacts locally, which is an aspect that could be investigated in future work.

In summary, analysis of the absolute magnitude of the γ-value is proposed rather than conventionally reported γ-passing rates. **bEQD increased the local γ-value in regions of already failed γ-criterion further** and might therefore lead to different clinical conclusions than D_a. γ-passing rates were lowered in all cases by 1% on average, with a maximum of 3%.

DIR accuracy Propagation of DIR uncertainties to the conventionally accumulated dose has been a popular field of research due to its increasing importance for new emerging methods but there is not yet a standardized quality assurance procedure available (section 2.2.1.3). **Investigations of dose uncertainties from DIR in literature** are for example based on deformable phantoms using 3D dosimetry gel [83] [32], numerical phantoms [56], different approaches of statistical error sampling procedures [54][73] or by comparison of registration strategies [74]. All these approaches require the implementation of specially designed soft- or hardware solutions that are not publicly available at this point. Literature rarely reports on the quantitative magnitude of dose errors propagated from DIR but only certain dose criteria or algorithm performance, mostly due to a missing ground truth to compare to. Nassef et al. [56] investigated conventional dose accumulation in 24 prostate cancer patients and found mean differences of D_a to the planned dose of 6.9 Gy for the bladder and 2.0 Gy for the rectal wall. They investigated the dose uncertainty from DIR by using a numerical phantom and the same DIR as for the patients. They found a mean DIR uncertainty of 2.7 Gy in the bladder and 1.2 Gy in the rectal wall. The statistical error sampling procedures investigated by Murphy et al. [54] showed dose error hot spots around 1 - 1.5 Gy, especially in dose gradient regions. The use of biomechanical models is proposed by a range of authors to improve the anatomical plausibility of the registration results, especially in regions of homogeneous imaging values. This was for example tested recently in depth by Zachiu et al. [85] who concluded that especially in regions of homogeneous imaging values, biomechnanical models and additional quality assurance criteria can improve the registration outcome.
It is beyond the scope of this work to perform an in depth analysis of the DIR accuracy. As can be seen from the reports in literature and the results within this thesis discussed below, uncertainties will strongly depend on the choice of the underlying DIR algorithm and the used parametrizations and constraints. Within this work, commercially approved DIR software implemented in the RayStation was used. Organ contours were additionally used to guide the DIR to improve registration quality in terms of anatomically realistic deformations. Nevertheless, image registration is arguably the biggest source of uncertainty and has been a controversial topic for dose accumulation [68]. This, however, holds true for both D_a and bEQD.

The impact of the DIR on D_α and bEQD was investigated in two ways in this work: by a worst-case estimation and by comparing the outcome of different DIR algorithms.

The **worst case dose mapping error analysis** was performed to solely focus on the difference in the impact of dose mapping errors between both accumulation strategies, independent of the choice of the DIR algorithm. This is, however, a worst case estimation and should be regarded as such. Especially at organ boundaries, often located in the gradient regions around the target, smaller errors in the mapped doses than estimated here are expected when using the contour information in the DIR process. The observed worst case magnitudes of error in total dose were generally higher or equal to those found for the difference between the accumulation strategies and reached up to 5 Gy for D_α and up to 9 Gy for bEQD. Thus, the magnitude of error induced by using conventional dose accumulation instead of effect-based accumulation strategies, is in the same range as the highly investigated impact of DIR uncertainties. Quality assurance of the DIR for dose accumulation (both bEQD and D_α) is urgently needed to reduce uncertainties in both workflows. Higher error magnitudes were found for $\Delta bEQD$ than for ΔD_α. Highest uncertainties were located in the dose gradient regions in both cases, due to the nature of the error estimation based on neighbouring voxel doses. The increased uncertainty for bEQD is due to the fact that the individual errors of the fractions are taken into account, in the same way that bEQD is systematically higher than D_α by consider the individual fraction doses in the accumulation. This interesting aspect will be further discussed in section 5.3 for the overall uncertainty budgets of bEQD compared with D_α.

For the second part of the analysis, a **comparison of DIR algorithms** was performed: the Raystation-based registration was compared in the total dose outcome with results from a simple demons and b-splines algorithm within the plastimatch software (without landmark or contour guidance). Substantial difference in the results were observed. It was also found that anatomically plausible deformations could not be guaranteed, especially for the b-splines registration. The demons algorithm seemed to be more robust in the investigated cases. Large deformations of the bladder, that induced very strong deformations in b-splines resulted in the fact that especially in the bladder regions, high dose differences occurred. While this might not be a problem for registration of CT images, since the area of interest is not affected and might be represented in a qualitatively satisfying manner, it will cause large deviations in the calculation of the accumulated dose (D_α and especially bEQD) and might therefore wrongly indicate high bladder doses after accumulation. For the demons algorithm, the dose difference in the accumulation methods were generally the lowest, also in comparison with the RayStation registration. This suggests low degrees of deformation and without quality control, it cannot be assured that the organ deformations were correctly backprojected in this case. The b-spline and demons algorithms seem to present two extreme cases of very large and very low deformation results, while the results from the RayStation are in be-

tween both. For the presented cases, it was assumed that the RayStation algorithm gives the most accurate representation of deformation in this comparison since it was additionally guided by organ contours. The results underline that DIR without geometry guidance is not sufficient in the pelvic region. Nevertheless, the quality of the RayStation DIR results are not known. In contrast to the b-splines and demons results, the DVFs could not be investigated. The former show high degrees of grid folding and anatomically unrealistic deformations for both plastimatch implementations which suggest that the results are not valid for dose accumulation and that deformation grids need to be investigated generally. The comparison of the b-splines- and the demons-based results for the mapped dose showed high differences even for fractions that seemed similarly acceptable for the CT deformation in both cases. This underlines that a simple qualitative inspection of the deformed CTs is not a valid quality assurance measure for the mapped doses. For the RayStation registration, if possible, quality assurance on the DVF should be performed (but exporting of these was not possible), for example by using the Jacobian determinant matrix that indicates grid folding and the degree of volume change of voxels (see section 2.2.1.3). It should be noted that measures of the spatial overlap of the resulting structures (e.g. Dice-coefficient) should not be used for the RayStation registration, since this was part of the DIR optimization process within the algorithm and is therefore an inappropriate choice as a quality measure.
DIR was also used for the **registration of the VoxTox H&N cases**. High dose differences were found especially in the peripheral regions of the registered images. This might have been caused by the varying positions of the daily MVCT scans, that did not overlap entirely in all cases causing ill-posed problems for the DIR algorithm by mapping different anatomies. Cropping of the mapped doses in the areas that did not overlap in all fractions was considered. However, it was concluded to rather report on the dose deviations in the whole volume. This was because for clinical applications, a conservative approach would be preferred when dealing with potential overdosage of critical structures. Thus, if it cannot be ruled out for sure that the observed differences were caused by erroneous DIR, but might have been the result of actual biological overdosage due to fractional dose variations, the results of the biologically accumulated dose should still be considered.

In summary, **DIR induces uncertainties in D_a and bEQD that are in the same range as the systematic deviation between the two**. bEQD inherits higher magnitudes of uncertainty than D_a. Quantitative quality assurance of DIR for dose accumulation is, thus, urgently needed but as of yet not available. For the pelvic region, DIR guidance with anatomical features is necessary to obtain anatomically plausible deformations or new advances from biomechanical modelling should be investigated.

Interpolation As discussed and concluded in detail in section 3.1.1, it is the dose to a volume of homogeneous dose (the smallest of which is a voxel due to discretization) that needs to be conserved in the mapping process in order to be **biologically consistent** and with this conserve the obtained biological effect of the fraction in the mapping process. Interpolation in this work is based on standard trilinear interpolation techniques. It can be shown mathematically that trilinear interpolation is equivalent to a volume-weighted average of the contributing voxel fractions. Interpolated dose is, thus, weighted by the number of contributing cells. As outlined before, interpolation induces deviations of the dose assigned to a cell in case different dose levels are summarized over a certain volume (discussed in section 3.1.1). This is especially the case in regions of strong dose gradients. This is however a general problem induced by discretization. Already on the level of a voxel, cells might have received different dose levels but are being assigned the spatial dose average of the volume of the voxel. In the mapping process, even if dose-conserved mapping is applied, the dose a cell is assigned with will be different after mapping due to the fact that interpolation will take into account the dose of the neighbour voxels as well. Therefore, the change in assigned dose in the mapping process due to interpolation does not necessarily induce a greater deviation to the true delivered dose to the cell than is already given by discretization alone. This is also given for cells that might have been shifted to another neighbouring voxel in the mapping process. For an accurate assessment of this problematic, it would be necessary to model the dose distribution on the level of the individual cells and quantify the induced deviation from discretization and interpolation. As concluded before, within the limits of small volume changes and high grid resolution (or low gradients), the uncertainties induced by interpolation and discretization are assumed to be small.
Discussed before was the difference between a dose- or energy-*conserving* interpolation. This must not be confused with a **dose- or energy-*based* interpolation**. The interpolation method in this work (trilinear interpolation) is based on dose, a form of *direct dose mapping (DDM)*. A problem in using this approach arises as well for the ill-posed problem of volume changing structures. For volume growth, mapping results in volume reduction and the doses from several voxels might be overlapped into one voxel. This opens the question on how to implement the dose mapping in case of overlap. Instead of DDM, another approach maps the energies of the voxels. A good overview of both approaches is given by Li et al [44] and Milz et al. [52]. The energy-based mapping of dose (called *energy-mass transfer (EMT)*) requires the estimation of the voxel mass from imaging gray values to derive the deposited energy. Transferred energy and mass are being considered. In essence, this approach is a mass-weighted average of the dose instead of the volume-weighted average of DDM described above for trilinear interpolation. For the presented work, EMT was not applied or investigated. This was due to two factors: in case physical density is constant, as is the case for incompressible soft tissue, the change in both mass and volume is directly proportional and the difference between both methods becomes negligible. This is different in case of lung tissue, where density changes

and EMT becomes more accurate than DDM. Another reason for not utilizing EMT is the location of occurring volume change. In the pelvic region, volume change is mainly driven by the filling states of both bladder and rectum. However, the dose to the bladder and rectum content that changes in volume, is not of interest (investigated by comparing whole organ contours and organ walls). EMT might also play a role on tissue surfaces to other structures with significantly different mass density (e.g. bone or air). However, the structures of interest show very similar mass densities, also at organ boundaries. Tumour shrinkage was not investigated in this work, however, in case of strong volume changes of the tumour, it might be worth considering EMT.

Besides trilinear interpolation, a sub-voxelization-based interpolation methods was tested in this work, referred to as the **Rosu-technique** based on the underlying publication by Rosu et al. [64]. The magnitude of difference between both interpolation strategies for the total doses of bEQD and D_a reached several Gy, both in negative and positive direction. The differences found for the investigated cases were higher than what was reported by Rosu et al., which is surprising since higher changes would have been expected for lung tissue due to the higher magnitude of volume changes. Regions of high differences were mostly found within the bladder which is were the strongest volume changes occurred. Furthermore, the differences were higher when using the demons based registration which in turn showed lower differences between the two accumulation strategies for the investigated case. This might be a result of more washed out dose gradients in the b-splines registration that are indicated by the larger area of total dose difference. Larger differences would be expected in strong dose gradients and in the investigated case, the demons algorithm seems to have produced a mapped dose with sharp gradients in contrast to the b-splines registration result. Overall, it needs to be noted that the Rosu implementation within the AVID workflow is as of yet not being used by other users and the strong differences might indicate that another proof of the implementation validity is necessary. In case the observed deviations are indeed a result of high uncertainties induced by trilinear interpolation, then the impact is similarly high as the impact of different dose accumulation strategies and should be taken into account and investigated further. This would mean that the advanced method by Rosu et al. should also be considered for use in the pelvic region and not only in lung tissue.

In summary, **dose-conserving dose mapping** needs to be implemented in order to facilitate a biologically consistent workflow and basis for accumulation.
The choice of the interpolation strategy shows a strong impact on the outcome of bEQD and D_a and especially **the sub-voxelization method by Rosu et al. should be investigated further**.

Dose-toxicity correlation For the preformed receiver operator characteristics (ROC) analysis, only weak correlations of dose to toxicity were observed in all cases except for rectal bleeding (RB) grade 2. This was found for all investigated total doses (planned dose, D_α and bEQD). It can be said for this specific dose-toxicity correlation, that neither conventional nor biological dose accumulation showed to be better predictors of toxicity than the planned dose. In clinical practice, dose constraints are mostly given for the rectum in parameters that link a certain dose to a volume. The fact that the found correlations are very weak in almost all cases underlines that the simple use of these constraints might not be appropriate for the rectum and more sophisticated dose-toxicity models including volume effects and organ architecture should be investigated, especially for the locally affecting 4bEQD.
The performed dose-toxicity correlation was based on DSM, a method to investigate the dose distribution on a local scale. The use of dose-width fits was used in order to find summarizing descriptive dose metrics for the given distributions without loosing all the spatial information of the DSM. Nevertheless, this method only takes the central dose cluster into account and hot spots in the peripheral region might be overlooked. It was found that especially in these peripheral regions, the difference to the planned dose as well as between the accumulation strategies was the highest, which therefore might have not been entirely included in the final correlation process. EUD summarized the DSM as a weighted mean. Although, with $a = 11.11$ a serial-type organ is assumed, which emphasizes the impact of local dose hot spots, this metric reduced the investigated dose distribution to a single value. This value should be interpreted carefully, also since a itself is an estimated empirical quantity only. As was discussed above, accumulation was not based on a voxel-wise dose mapping technique. This might have washed out the accumulated doses (D_α and bEQD), resulting in lower magnitudes of EUD as well as of $bEQD - D_\alpha$, since high local dose variation amplitudes might not have been observed correctly.
The differences found between D_α and bEQD to the planned dose, respectively, were much higher than the difference between the accumulation strategies. Furthermore, the impact on the single dose metrics was low. Therefore, only minor changes for the dose-toxicity correlation is expected since this was only based on single metrics and without an underlying model, as in NTCP modelling. This was accordingly observed within the ROC analysis for the result of the area under the curve (AUC). In many cases, the dose metrics obtained from the planned dose were higher than those from D_α. Since bEQD is systematically higher than D_α, results from bEQD were closer to that of the planned dose in many cases.
For most cases, surprisingly, the planned dose showed higher correlation than the results from bEQD and D_α. The difference was small in all cases between the three investigated doses, though high differences to the planned dose were observed. Generally the outcome of the ROC analysis showed only very low correlation to outcome. This indicates a low sensitivity of the correlation method to describe the relation between dose and toxicity. It was however hypothesized that the accumulated dose would be a better predictor than the planned dose since in principle, it should be

closer to what was actually delivered to the patients. This was also published by Shelley et al. [70] on a prior test cohort to the here used cases. There are two reasons why this might not be the case: i) there might be a systematic deviation of the accumulated dose (present for both D_α and bEQD) to the true delivered dose. The mapping process used in this workflow that does not utilize voxel-wise mapping is a limitation that was discussed before and might have a systematic impact. Also the automatic contouring of the rectum might have altered the outcome. ii) the correlation of dose to toxicity might be more complex for the rectum, so that a simple correlation of single dose metrics is not appropriately describing the dose-toxicity relationship. For most endpoints, except for RB grade 2, the AUC was close to 0.6 which is set as the limit for a correlation assessment, which underlines hypothesis ii). Thus, only for RB grade 2, and for dose widths fits between thresholds of 50-65 Gy, a correlation above an AUC of 0.7 was found, which indicates more strongly a relationship of dose to toxicity in this regime.
Within the investigated cohort, patients were partially treated with standard fractionation and partially with moderate hypofrationation. EQD_2-based conversion was performed to use both cohorts in the same analysis. Regarding the above discussed results on hypofractionation, it should be investigated if the found correlations differ when investigating both cohorts separately. Since the impact of bEQD is strongly increased for hypofractionation, this leads to the hypothesis that difference in AUC might be higher (comparing bEQD and D_α) for the sub-cohort treated with moderate hypofractionation.

In summary, the performed dose-toxicity correlation is only **marginally impacted by the use of bEQD**. Generally, the performed **ROC-analysis showed to have only low descriptive power to describe dose-toxicity correlations of single dose metrics**: i) correlation results were generally low, only for RB grade 2, the mid to high dose thresholds of the dose widths showed significant correlations above an AUC of 0.6. ii) results were similar for the use of the planned dose, bEQD and D_α, despite high differences between bEQD and D_α to the planned dose, respectively. iii) the planned dose unexpectedly showed higher correlation than both D_α and bEQD in the majority of toxicity endpoints and dose metrics. Within this toxicity-correlation analysis, it could not be shown, that D_α or bEQD are better predictors of toxicity than the planned dose. Based on the local impact of bEQD, **more advanced toxicity models should be investigated** that take organ architecture into account, especially for hypofractionation. This also questions the use of single dose-volume constraints for rectal doses in clinics.

5.3 CONCLUSION: SHOULD WE ACCUMULATE EFFECT INSTEAD OF DOSE?

This section tries to answer two questions with regard to the above discussed aspects and results: i) Is the use of the conventionally accumulated or the biological accumulated dose still advantageous compared with the planned dose when considering the uncertainty budget? ii) For which applications is biological dose accumulation clinically favourable?

The uncertainty budget of bEQD and D_α

The Linear Quadratic Model (LQM) to estimate the biological effect from a radiotherapy (RT) treatment is used in various clinical applications (section 2.2.2.2). It describes a non-linear relationship of the biological effect to dose. Nevertheless, in current practice, the total delivered dose of the treatment, obtained from simply using the planned dose or conventional dose accumulation (D_α), is used for clinical decisions on treatment adaptation and correlation to the biological treatment outcome. To obtain the mathematically correct biological effect as defined by the LQM as a (cell) survival fraction (SF), the derived effect accumulation approach $SF(d_i)$ has to be used. The result of this calculation can also be obtained by the use of the total biological dose (bEQD) within the LQM. bEQD is therefore referred to as the total biological dose, as it provides the mathematically correct survival fraction. It was shown, that bEQD is systematically larger than D_α. The computation of both D_α and bEQD in patient data is prone to a number of uncertainties that were described individually in section 5.2. This opens the question, whether the difference between both approaches is clinically relevant when considering the overall uncertainty budget.

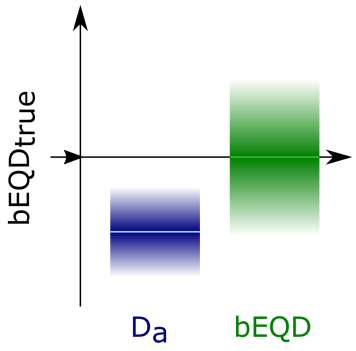


Figure 5.3. – Qualitative comparison of the mathematical difference between D_α and bEQD: D_α shows a systematic underestimation of the total dose with respect to the total biological effect, thus, lower *accuracy*. In comparison, bEQD shows lower *precision* due to higher impact of uncertainties inherited from DIR.

Due to uncertainties in the computation of bEQD and D_α, which cannot be quantified (specifically deformable image registration (DIR) and α/β, as discussed in section 5.2), their *true value* is not known in patient data. Within this section, the accuracy and precision of D_α and bEQD are therefore discussed in a qualitative comparison with respect to an assumed true biological dose ($bEQD_{true}$), to assess the overall uncertainty budgets and relations between the approaches.

The mathematical difference between the bEQD and D_α approaches is a systematic deviation: the SF as calculated from D_α is systematically larger than the mathe-

matically correct value $SF(d_i)$. Thus, D_α is systematically too low in comparison with $bEQD_{true}$. Thus, bEQD increases the *accuracy* of biological effect prediction by avoiding this systematic deviation, i.e. underestimation of the obtained total biological effect. Figure 5.3 illustrates this principle.
Both bEQD and D_α are as well prone to uncertainties of their own. Above all, the uncertainties from DIR that propagate through to the mapped voxel doses represent the bottleneck in the achievable *precision* of both accumulation approaches. The presented results show the high impact of the registration uncertainties on both dose accumulation strategies, though direct quantification is not possible (section 5.2). As discussed above, it was however found that bEQD as well as $SF(d_i)$ are prone to higher uncertainties inherited from DIR than D_α. This is depicted in figure 5.3 which shows a wider spread for bEQD than for D_α. Although the underlying assumed dose error was equal (Δd_i), its propagation depends on the accumulation strategy. As a result, the estimated uncertainties for bEQD were found to be up to two times higher than those of D_α. Therefore, in a first conclusion, the precision of D_α is higher than that of bEQD, while bEQD has a higher accuracy (figure 5.3).

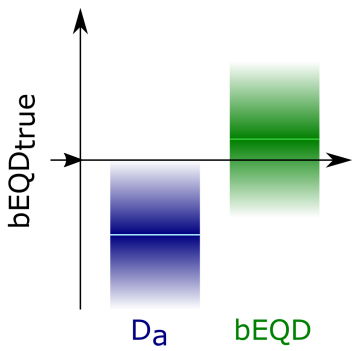


Figure 5.4. – Qualitative comparison of the mathematical difference between D_α and bEQD: ΔD_α underestimates the uncertainties from Δd_i in the same manner the total dose is underestimated with respect to the biological effect, a higher uncertainty should be assumed. bEQD shows an additional systematic uncertainty due to the assignment of α/β, which is, however, lower than the systematic difference to D_α.

The systematic deviation between D_α and bEQD is based on the non-linear relation of the SF to dose. This is also inherited by the uncertainties of D_α in error propagation. Therefore, in consequence, it needs to be concluded that ΔD_α also underestimates the influence of the individual dose errors Δd_i in the same manner as D_α underestimates the total biological effect from the individual d_i. This should be considered in the assessment of D_α accordingly. Thus, both the precision and the accuracy of D_α are lowered, which is illustrated in figure 5.4.
The assignment of the α/β-value to compute bEQD represents a limitation to its accuracy that is not present for D_α. It was found that this induced uncertainty in bEQD is smaller than the potential registration uncertainty or the difference to D_α. Nevertheless, due to its potential magnitude of more than 1 Gy, it needs to be considered in the overall uncertainty budget. The magnitude depends on the uncertainty in α/β and is therefore organ specific and can be improved by better knowledge of the α/β value. As for the uncertainty emerging from DIR, the uncertainty from α/β can only be estimated since the true deviation is not known (section 5.2). In figure 5.4 this uncertainty is depicted as a systematic deviation since bEQD will for example be systematically too low, in case the assigned α/β is higher than the true α/β of

the tissue.
The uncertainty from interpolation is another source of uncertainty that impacts both bEQD and D_a. Due to the potentially high impact observed in the patient data analysis using the Rosu implementation, this aspect should be investigated further. As of now, it cannot be said whether the impact is systematic or random.
From the presented results, it can be seen that the magnitudes of the estimated uncertainties of both D_a and bEQD are in a similar range as the difference between the two approaches. This underlines two aspects: first, the use of dose accumulation is generally prone to high uncertainties from several sources that need to be quantified and taken into account for both D_a and bEQD; second, the biological inaccuracy by the use of conventional dose accumulation has a similarly strong impact on the overall uncertainty as other highly investigated sources such as DIR. In contrast to the DIR uncertainty, it can, however, be easily avoided by use of bEQD.

For both investigated pelvis cohorts, the difference of the planned dose to D_a and bEQD, respectively, was higher than the found uncertainties. In the absence of other not investigated sources of uncertainty within the accumulation process, it can therefore still be assumed, that D_a and bEQD are closer to $bEQD_{true}$ than the planned dose. An illustration of this conclusion is depicted in figure 5.5. As shown, D_a is systematically below $bEQD_{true}$, while the planned dose can deviate in either direction. Depending on the direction of the difference to the planned dose, bEQD is either closer to the planned dose (figure 5.5, cases 1) or shows a higher difference (figure 5.5, cases 4-7) to the planned dose than D_a.

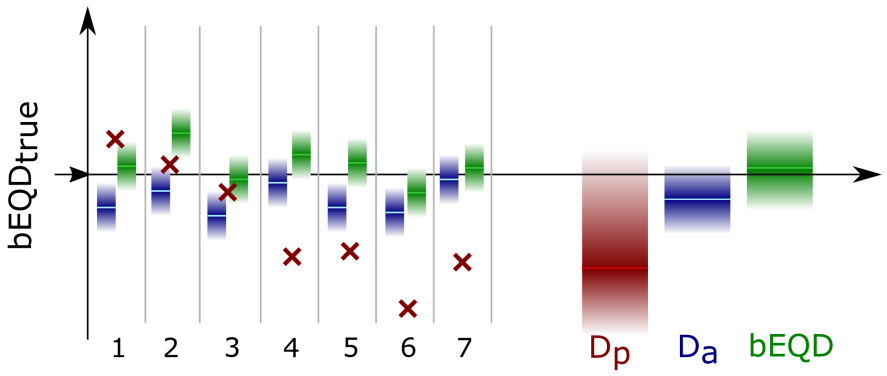


Figure 5.5. – Qualitative comparison of the planned dose (D_p) with D_a and bEQD for a schematic number of examples. bEQD is always larger or equal to D_a. On average, for a large cohort, D_a will be close to the assumed $bEQD_{true}$, though systematically lower. The difference between the accumulation strategies (bEQD and D_a) is smaller in comparison with the difference to the planned dose. However, the difference can be both high to the planned dose and/or between the accumulation methods in individual cases.

For the investigated dose-volume (DV)-parameters as well as for reports from literature, the deviations of bEQD and D_a from the planned dose often counterbalance

when looking at entire patient cohorts. The systematic offset of bEQD compared with D_a resulted in a shift in the overall values that was small compared with the range of difference to the planned dose within the entire cohort. A shift in median values was, however, visible for hypofractionation. Moreover, individual outliers occur for both the difference to the planned dose as well as for the difference between bEQD and D_a. Figure 5.5 shows a qualitative example of this situation which can be directly compared to the DV results in figure 4.11 and 4.21. This shows that the use bEQD instead of the planned dose is important in individual cases while statistically, its impact is lower.

Possible applications of bEQD To begin with, three potential areas where bEQD and effect accumulation might have an impact on clinical decisions and outcome were introduced: treatment adaptation, re-treatment and dose-response modelling.
It was found that statistically, D_a as well as bEQD are close to the planned dose on average for large cohorts, especially when looking at dose-volume metrics. Individual cases, however, showed high deviations. This is similar when looking at three-dimensional (3D)-data: while hotspots occur, differences are small on average. As was shown above, the magnitude of difference between bEQD and D_a is low in comparison to the difference to the planned dose, but again, high in individual cases. The here performed dose-toxicity correlation did not show significant impact of neither D_a nor bEQD compared with the planned dose and generally low correlation for the majority of investigated endpoints and dose metrics. In combination, these findings indicate that: i) The accumulated dose and especially bEQD should be considered for treatment adaptation and re-treatment. Although on average, the impact is low, individual cases might benefit significantly. However, for standard fractionation, this is only the case when local doses are investigated. For hyofractionation, the impact is much larger and therefore also visible in DV-analysis. ii) The impact on dose-response modelling is low when using simple dose metric-based approaches as the one presented in this thesis. Since the impact of bEQD is low on average, no clinical impact is to be expected. Thus, the use of bEQD is only relevant when investigating more sophisticated approaches in normal tissue complication probability (NTCP) modelling that consider volume effects, organ architecture and interplay of different dose-volume relations. It should for example be investigated if bEQD has an impact when looking at functional sub-units (FSUs) of organs, especially those with a serial-type nature. Moreover, dose-response modelling for hypofractionated cohorts should be investigated.
bEQD should further be investigated for the accumulation of intrafractional doses, since for example the interplay effect in gated treatments can cause high local dose variations. This can occur both intrafractionally as well as between fractions. However, for this case, the energy-mass transfer (EMT)-based dose interpolation approach should be considered (see section 5.2).
Zavgorodni et al. [87] hypothesise that the similar approach of the equivalent con-

stant dose (ECD) (section 5.1) might impact the choice of treatment margins. Indeed, when using bEQD (or $\mathsf{d}_\mathfrak{b}$ which is equivalent to ECD), isodose lines will be shifted due to the systematic offset to $\mathsf{D}_\mathfrak{a}$. Thus, when considering the bEQD concept in for example probabilistic treatment optimization which considers possible daily dose variation, the result and the choice of margins might be different.

6

SUMMARY

In this thesis, a biologically consistent way to perform dose accumulation in image-guided radiotherapy (RT) was developed. It was shown that both the accuracy and precision of the estimated total delivered dose are affected by biological mechanisms. The first aim of this thesis therefore was the development of a workflow for dose accumulation that is in accordance with the commonly used Linear Quadratic Model (LQM). It was shown how biological consistency needs to be addressed in two aspects of the dose accumulation workflow (section 3.1): the dose mapping and dose accumulation itself. The implementation of biological dose mapping was investigated and discussed, showing that the recent claim for an energy conserving workflow implies biological inconsistency which needs to be overcome by using a dose-conserving approach. The methodology for biological dose accumulation was established. On the basis of the daily deformed doses (on the background of biologically consistent mapping), the described effect accumulation approach yields the mathematically correct (cell) survival fraction (SF) according to the LQM. The conversion of this approach to a total dose which yields the correct SF, can be performed by using the derived total biological dose (bEQD) approach or the biological fraction dose d_b. This enables comparison to the planned dose and commonly used dose constraints. In comparison, it was shown that the conventional dose accumulation workflow implies a systematic underestimation of the obtained biological effect.
The second aim of this thesis was the data-based conclusion on clinical relevance and applicability of the derived biological dose accumulation approach with respect to the current standard. In an in-depth analysis, the impact of all stages of the dose accumulation workflow of bEQD in comparison with the conventional dose accumulation was investigated, concerning magnitude of difference and the occurring uncertainties. This was addressed theoretically and by analysis of the data of three patient cohorts. It was shown that using bEQD increases the accuracy of the dose accumulation workflow with respect to the biological effect prediction by avoiding the systematic inconsistency in the application of the LQM.
The quantitative results showed local hotspots with differences between bEQD and D_a with magnitudes around 4 Gy for standard and 8 Gy for hypofractionation. This difference is largest for tissues with a low α/β value, in regions of large motion amplitudes and in hypofractionated treatments. Highest deviations were observed in normal tissue in steep dose gradient regions. However, dose-volume based analysis showed only very little dependency on the used dose accumulation strategy. This means that clinical decision of dose-summarizing dose metrics and dose-toxicity correlations based on these, will most likely not be affected. Local dose analysis is

recommended to supplement dose-volume-based analysis.
The precision of dose accumulation needs to be improved by reducing the investigated underlying uncertainties from α/β assignment, deformable image registration (DIR) and interpolation. bEQD showed to be more sensitive to DIR uncertainties than D_a which underlines that these uncertainties are underestimated in conventional dose accumulation in the same way that the total dose is underestimated. It was shown that the difference $bEQD - D_a$ (thus the "biological inaccuracy of D_a) is in the same range as the overall uncertainty budget from α/β assignment, DIR and interpolation. This underlines the importance of biological dose accumulation (comprehensive conclusion in section 5.3).
The analysis in the data from the three cohorts showed little impact in cohort-averaged values for bEQD but individual cases of strong deviations to D_a. Therefore, the presented approaches should be considered especially for treatment adaptation, to avoid overdosage of organ(s) at risk (OAR) in individual cases. No clinically relevant impact on the performed dose-toxicity correlation for the rectum using single dose metrics was observed of either bEQD nor D_a, compared with the planned dose. Based on the observed high difference to the planned dose and the systematic difference between bEQD and D_a, dose-toxicity correlation should consider advanced volume-effect models and organ architecture, to investigate the impact of biological dose accumulation further, especially for hypofractionation (comprehensive conclusion in section 5.3).
Dose accumulation is entering clinics and its use will further increase in the future. This fact underlines the need for biological precision and accuracy. The presented investigations provide a biologically consistent workflow and the results show the necessity for its use in specific cases.

Physics ~~vs.~~ *and* biology: reasons that motivated the need for this thesis

APPENDICES

A

APPENDIX A: LISTS

LIST OF PUBLICATIONS

Manuscripts submitted for publication

N. I. Niebuhr, C. Karger, K. Harrison, D. J. Noble, R. Jena, R. O. Floca, J. Seco, M. Wilson, A. Pfaffenberger (2019). On the biological rationale of dose mapping in the presence of volume changes. *Phys Med Biol*, (under review).

N. I. Niebuhr, M. Splinter, T. Bostel, J. Seco, C. M. Hentschke, J. Hörner-Rieber, M. Alber, P. E. Huber, N. H. Nicolay and A. Pfaffenberger (2019). Biologically consistent dose accumulation using daily patient imaging. *Phys Med Biol*, (under review).

Conferences (own contributions)

N. I. Niebuhr, L. E. A. Shelley, M. Splinter, K. Harrison, N. H. Nicolay, J. Hörner-Rieber, R. Jena, A. Pfaffenberger (2019). Biologically consistent dose accumulation and its uncertainties and implications in pelvic patients. *BiGART2019, Aarhus, Denmark*

N. I. Niebuhr, T. Bostel, K. Harrison, R. Jena, N. H. Nicolay, D. J. Noble, L. E A. Shelley, M. Splinter, A. Pfaffenberger (2019). Comparing biological and conventional dose accumulation using daily imaging of H&N and pelvis cases. *ESTRO 38, Milano, Italy*

N. I. Niebuhr, M. Splinter, T. Bostel, N. H. Nicolay, A. Pfaffenberger (2018). Uncertainties in accumulation of local physical and biological dose from daily patient imaging. *6th MR in RT, Utrecht, The Netherlands*

N. I. Niebuhr, A. Pfaffenberger (2018). On the impact of dose variations in contrast to mean dose on EQDx and α/β estimation. *ESTRO 37, Barcelona, Spain*

N. I. Niebuhr, C. M. Hentschke, R. O. Floca, M. Alber, K. M. Kraus, A. Pfaffenberger (2017). Should we accumulate the biological effect instead of dose? *AAPM 59th annual meeting and exhibition, Denver, USA*

N. I. Niebuhr, C. M. Hentschke, R. O. Floca, M. Alber, K. M. Kraus, T. Bostel, F. Sterzing, A. Pfaffenberger (2017). Application of Local Effect Accumulation in Contrast to Dose Accumulation *BiGART2017, Aarhus, Denmark*

Paper contributions during the time but outside the scope of this thesis

N. I. Niebuhr, W. Johnen, G. Echner, A. Runz, M. Bach, M. Stoll, K. Giske, S. Greilich and A. Pfaffenberger (2019). The ADAM-pelvis phantom: and anthropomorphic, deformable and multimodal phantom for MRgRT. *Phys Med Biol 64(4): 04NT05.*

O. Bohoudi, F. J. Lagerwaard, A. M. E. Bruynzeel, N. I. Niebuhr, W. Johnen, S. Senan, B. J. Slotman, A. Pfaffenberger, M. A. Palacios (2019). Evaluation of dose accumulation in MRI-guided adaptive radiotherapy for prostate cancer using an anthropomorphic deformable pelvis phantom. *Radiotherapy & Oncology*, (accepted for publication, September 2019).

D. Hoffmans, N. I. Niebuhr, O. Bohoudi, A. Pfaffenberger, M. A. Palacios (2019). An end to end test for MR-guided online adaptive radiotherapy. *Phys Med Biol*, (under review).

L. Volz, L. Kellerer, S. Brons, C. Graeff, N. I. Niebuhr, R. Radogna, C. Schömers, S. Jolly and J. Seco (2019). Experimental exploration o a mixed helium/carbon beam for online treatment monitoring in carbon beam therapy. *Phys Med Biol*, (under review).

ACRONYMS

2D two-dimensional

3D three-dimensional

4D four-dimensional

AUC area under the curve

BED biologically effective dose

bEQD total biological dose

CBCT cone beam computed tomography

CT computed tomography

CTV Clinical Target Volume

DDM direct dose mapping

DIR deformable image registration

DKFZ German Cancer Research Center

DSM dose surface map

DV dose-volume

DVF deformation vector field

DVH dose-volume histogram

ECD equivalent constant dose

EMT energy-mass transfer

EQD_2 equivalent dose in 2 Gy fractions

EUD equivalent uniform dose

FSU functional sub-unit

GTV Gross Tumour Volume

HLUT Hounsfield look-up table

HU Hounsfield Units

H&N head and neck

ICRU International Commission on Radiation Units & Measurements

IMRT intensity-modulated radiotherapy

linac linear accelerator

LKB Lyman-Kutcher-Burman

LQM Linear Quadratic Model

MLC Multi Leaf Collimator

MRI magnetic resonance imaging

MVCT mega voltage computed tomography

NTCP normal tissue complication probability

NTD normalized total dose

OAR organ(s) at risk

PE photoelectric

PTV Planning Target Volume

QA quality assurance

QUANTEC Quantitative Analyses of Normal Tissue Effects in the Clinic

RB rectal bleeding

ROC receiver operator characteristics

ROI region of interest

RT radiotherapy

SF (cell) survival fraction

TCP tumour control probability

TERMA Total Energy Released per unit Mass

TPS treatment planning system

LIST OF FIGURES

LIST OF TABLES

B

APPENDIX B: ADDITIONAL DATA AND RESULTS

PARAMETER FILS FOR DIR

Parameter files for deformable image registrations (DIR) using **plastimatch** in this thesis to register the daily CT images to the planning CT images for the DKFZ pelvis cohort (in comparison to the RayStation based DIR.

Bspline-based DIR

```
[STAGE]
xform=align_center

[STAGE]
xform=bspline
optim=lbfgsb
impl=plastimatch
threading=openmp
max_its=50
convergence_tol=5
grad_tol=1.5
regularization_lambda=0.002
grid_spac=100 100 100
res=4 4 2

[STAGE]
xform=bspline
optim=lbfgsb
impl=plastimatch
threading=openmp
convergence_tol=5
grad_tol=1.5
max_its=50
regularization_lambda=0.0003
grid_spac=50 50 50
res=2 2 2
```

```
[STAGE]
xform=bspline
optim=lbfgsb
impl=plastimatch
threading=openmp
convergence_tol=5
grad_tol=1.5
max_its=50
regularization_lambda=0.0005
grid_spac=20 20 20
res=1 1 1
```

Demons-based DIR

```
[STAGE]
xform=align_center

[STAGE]
xform=vf
impl=itk
optim=demons
optim_subtype=fsf
#demons_gradient_type=symmetric
demons_smooth_update_field=0
demons_std_update_field=1
demons_smooth_deformation_field=1
demons_std_deformation_field=1
demons_step_length=10;
histo_equ=1
num_hist_levels=1024
num_matching_points=7
thresh_mean_intensity=1
max_its=100
res=10 10 10

[STAGE]
xform=vf
impl=itk
optim=demons
optim_subtype=fsf
#demons_gradient_type=symmetric
demons_smooth_update_field=0
demons_std_update_field=1
```

```
demons_smooth_deformation_field=1
demons_std_deformation_field=1
demons_step_length=3;
histo_equ=1
num_hist_levels=1024
num_matching_points=7
thresh_mean_intensity=1
max_its=100
res=5 5 5

[STAGE]
xform=vf
impl=itk
optim=demons
optim_subtype=fsf
#demons_gradient_type=symmetric
demons_smooth_update_field=0
demons_std_update_field=1
demons_smooth_deformation_field=1
demons_std_deformation_field=1
demons_step_length=1;
histo_equ=1
num_hist_levels=1024
num_matching_points=7
thresh_mean_intensity=1
max_its=100
res=2 2 2

[STAGE]
xform=vf
impl=itk
optim=demons
optim_subtype=fsf
#demons_gradient_type=symmetric
demons_smooth_update_field=0
demons_std_update_field=1
demons_smooth_deformation_field=1
demons_std_deformation_field=1
demons_step_length=0.5
histo_equ=1
num_hist_levels=1024
num_matching_points=7
thresh_mean_intensity=1
```

```
max_its=20
res=1 1 1
```

For the **Elastix** DIR used for the VoxTox H&N data, the following settings were used:

- Similarity metric: mutual information using stochastic gradient descent
- Four resolution stages: (8,8,4); (4,4,2); (2,2,1); (1,1,1)
- affine transform followed by a cubic-bspline transform with a grid-point spacing of 30 mm.

ADDITIONAL RESULTS

DV-metrics results complementing sections 4.2.3 and 4.2.4.

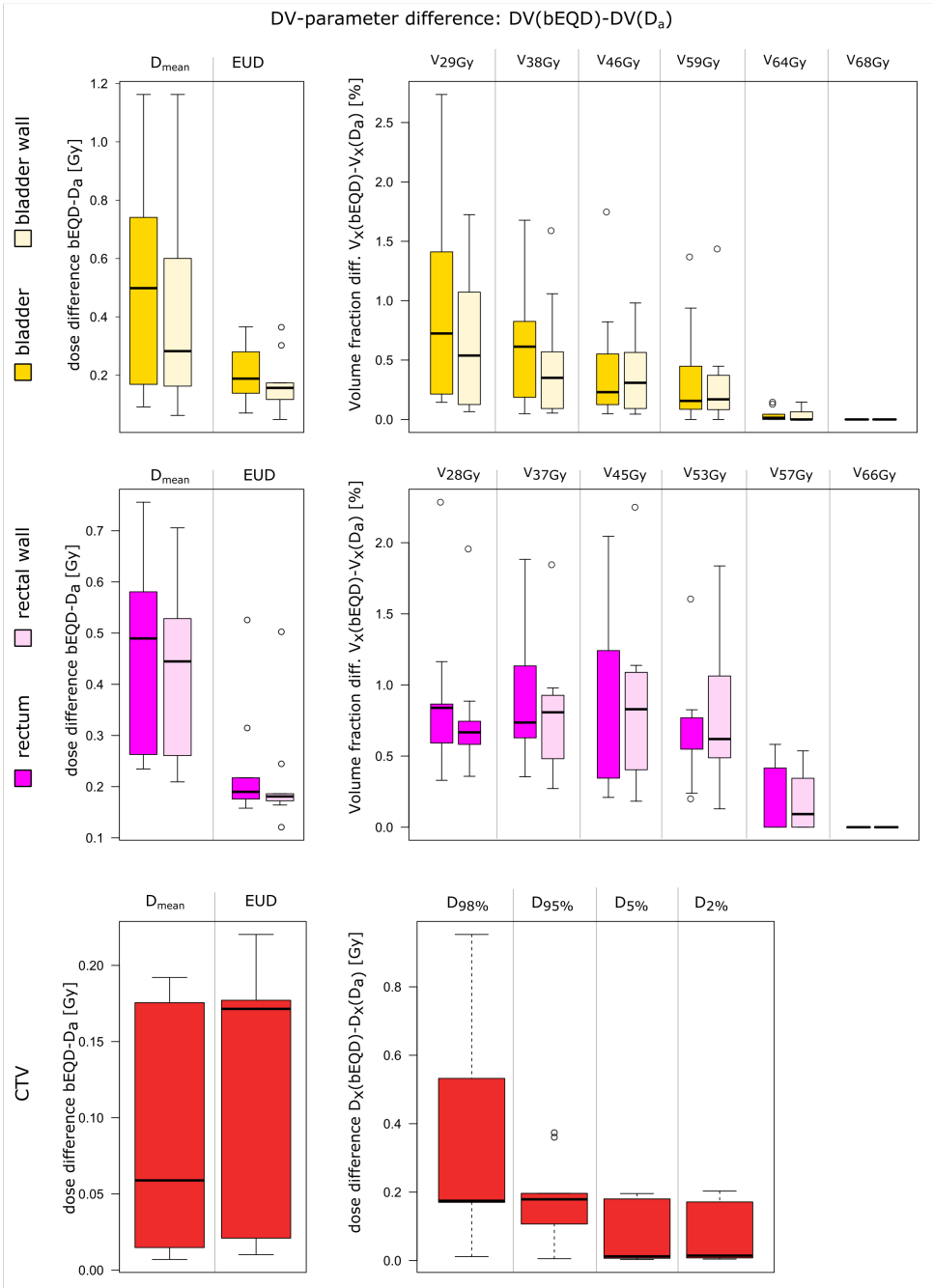


Figure B.1. – Comparison of DV parameters for moderate hypofractionation as the difference between DV(bEQD) and DV(D_a) of the structures of interest of bladder(B), bladder wall (BW), rectum (R), rectal wall (RW) and tumour (Clinical Target Volume (CTV)) for various dose metrics. The respective V_{xGy} were converted for hypofractionation with n=21 fractions according to the equivalent dose in 2 Gy fractions (EQD_2) approach. with $\alpha/\beta = 5Gy$ for the bladder and $\alpha/\beta = 3Gy$ for the rectum.

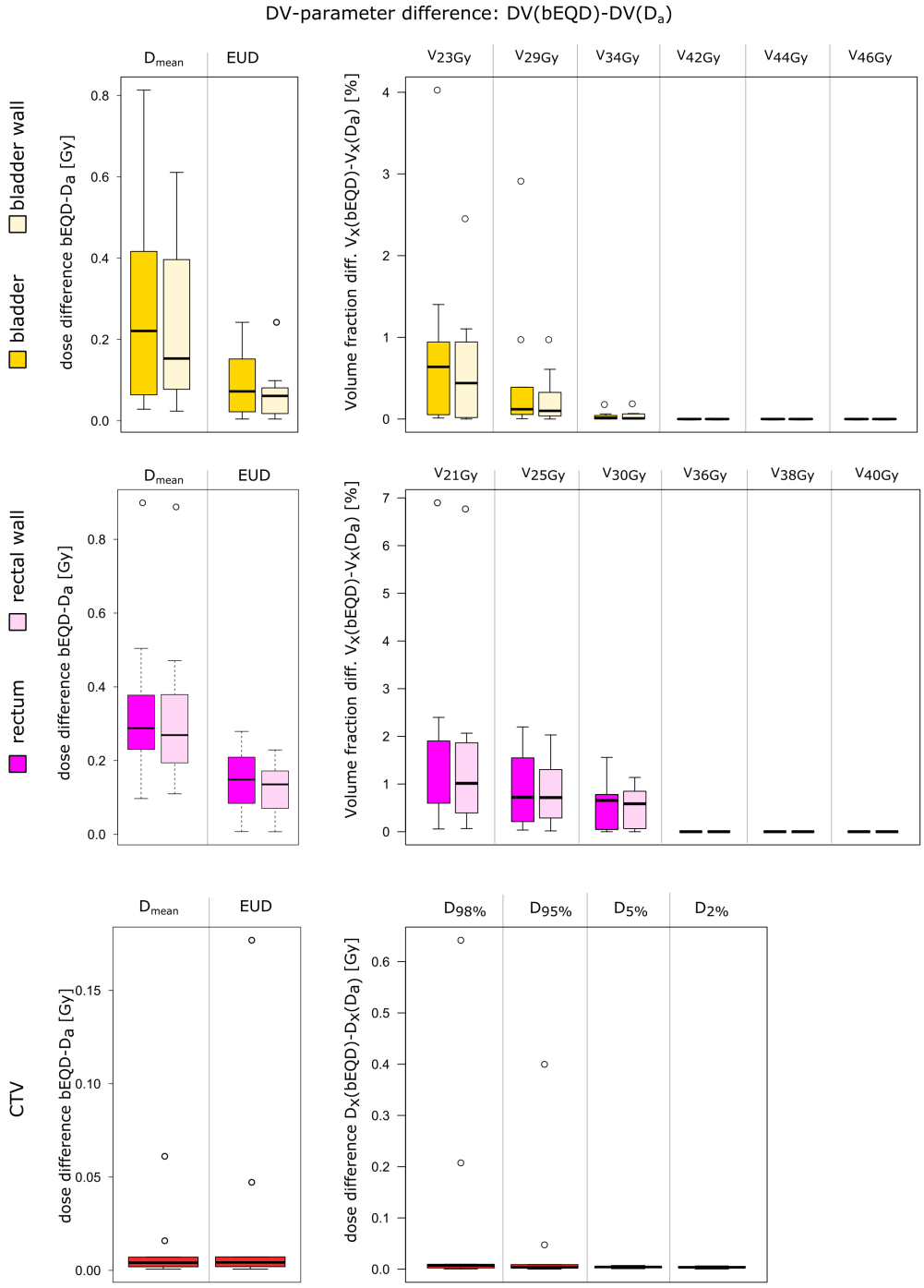


Figure B.2. – Comparison of DV parameters for strong hypofractionation as the difference between DV(bEQD) and DV(D_a) of the structures of interest of bladder(B), bladder wall (BW), rectum (R), rectal wall (RW) and tumour (CTV) for various dose metrics. The respective V_{xGy} were converted for hypofractionation with n=5 fractions according to the EQD_2 approach. with $\alpha/\beta = 5Gy$ for the bladder and $\alpha/\beta = 3Gy$ for the rectum.

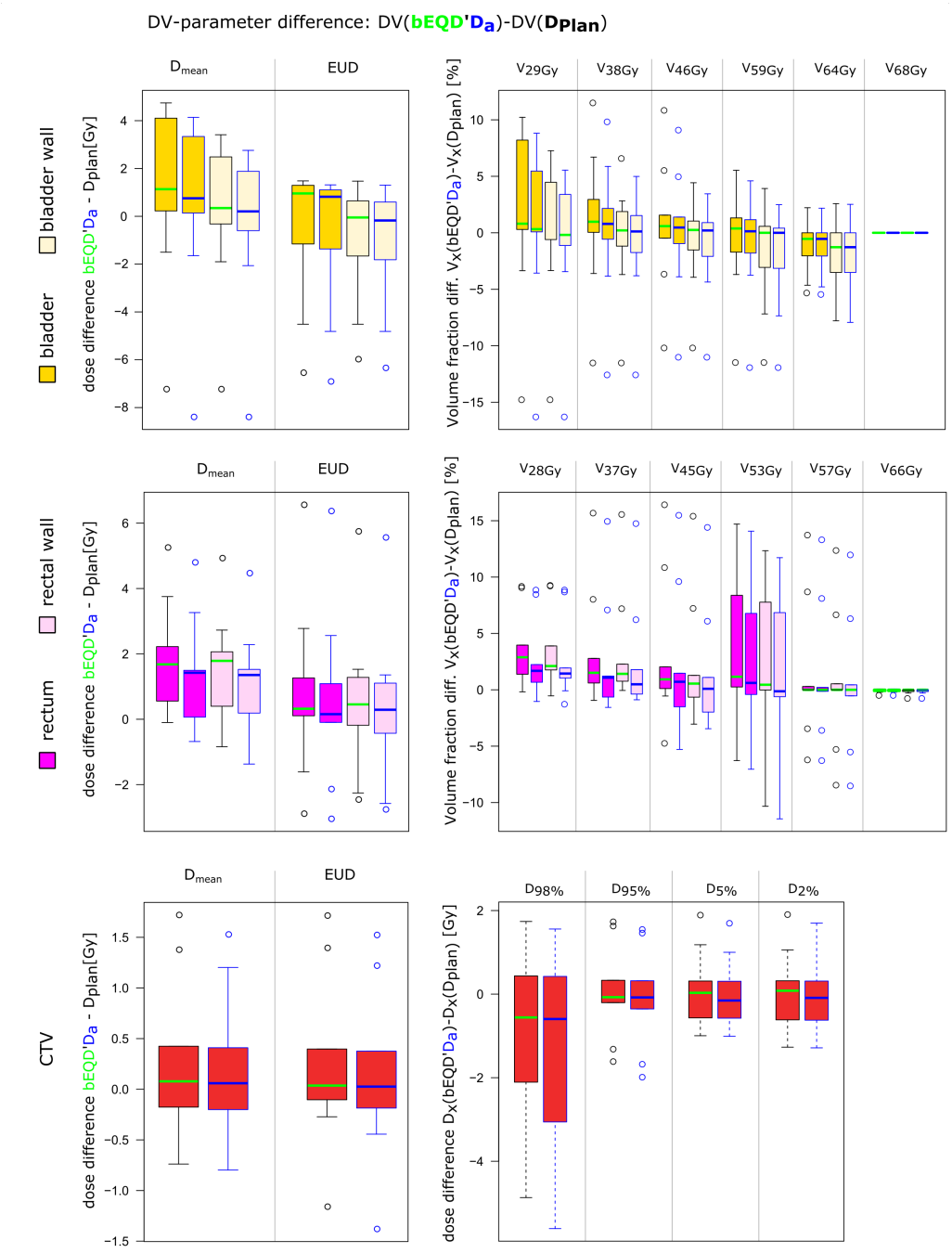


Figure B.3. – Comparison of DV parameters for moderate hypofractionation (n=21) as the difference between either DV(bEQD) or DV(D_a) to DV(D_{plan}), respectively. Shown for the structures of interest of bladder(B), bladder wall (BW), rectum (R), rectal wall (RW) and tumour (CTV) for various dose metrics. The respective V_{xGy} were converted for hypofractionation with n=21 fractions according to the EQD_2 approach. with $\alpha/\beta = 5Gy$ for the bladder and $\alpha/\beta = 3Gy$ for the rectum.

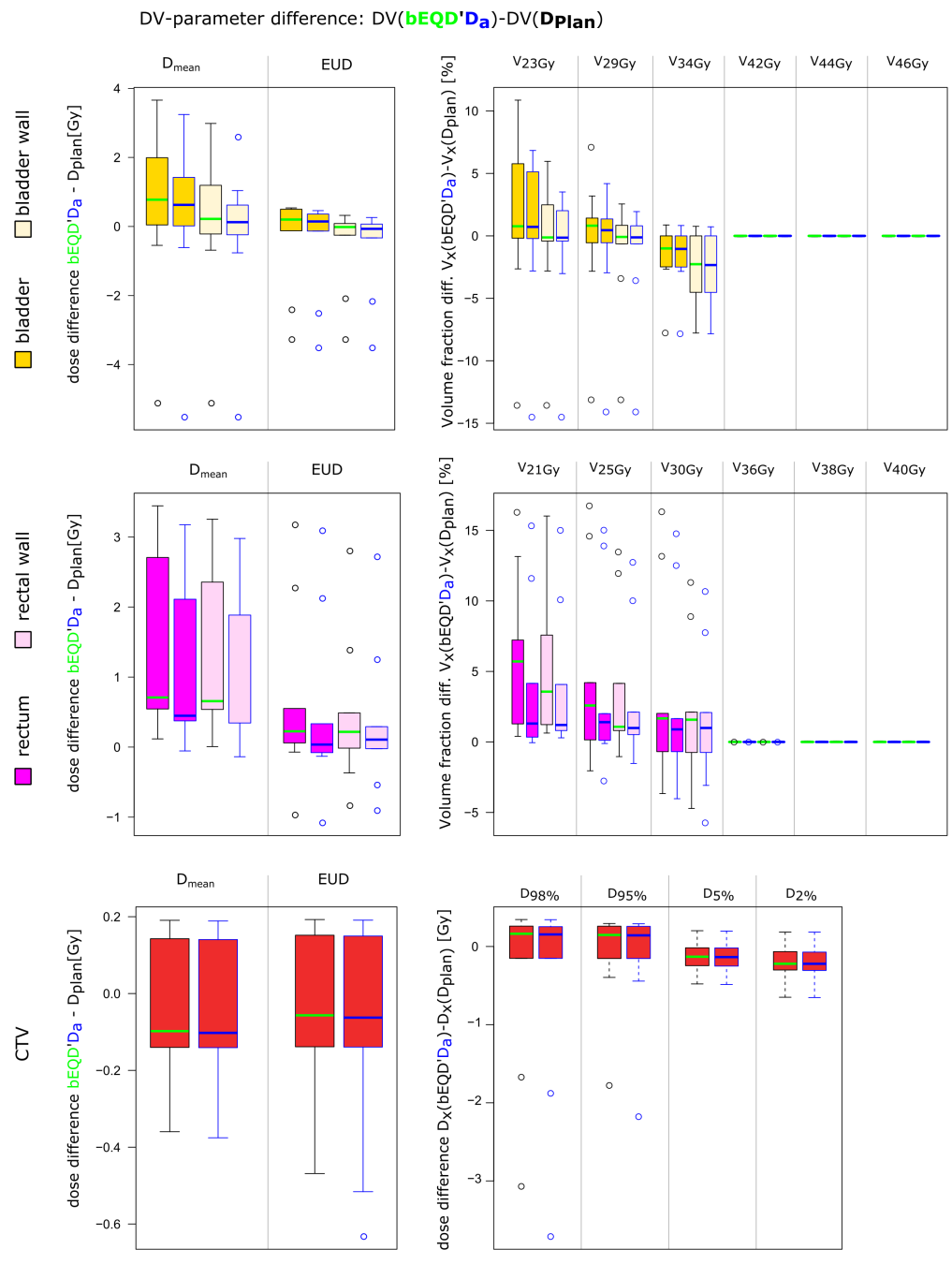


Figure B.4. – Comparison of DV parameters for strong hypofractionation (n=5) as the difference between either DV(bEQD) or DV(D_a) to DV(D_{plan}), respectively. Shown for the structures of interest of bladder(B), bladder wall (BW), rectum (R), rectal wall (RW) and tumour (CTV) for various dose metrics. The respective V_{xGy} were converted for hypofractionation with n=5 fractions according to the EQD_2 approach. with $\alpha/\beta = 5Gy$ for the bladder and $\alpha/\beta = 3Gy$ for the rectum.

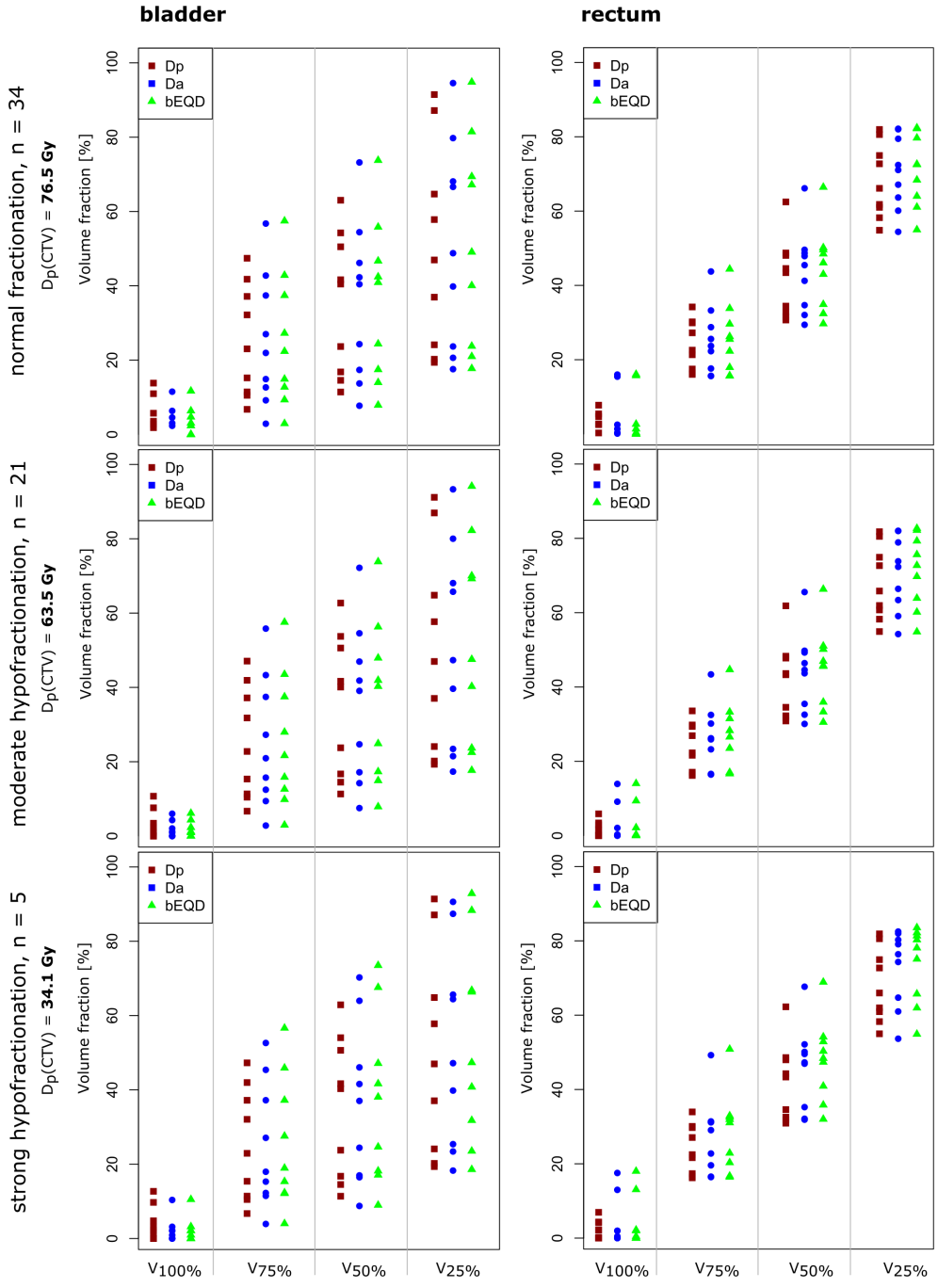


Figure B.5. – $V_{x\%}$ (volume receiving at least x% of the planned dose) comparison between planned (D_p), accumulated (D_a) and biological accumulation dose (bEQD) for rectum and bladder.

BIBLIOGRAPHY

[1] H. Afsharpour, S. Walsh, C.-A. Collins-Fekete, E. Vigneault, F. Verhaegen, and L. Beaulieu. On the sensitivity of alpha/beta prediction to dose calculation methodolog in prostate brachytherapy. Int J Radiation Oncol Biol Phys,, 88(2):345–350, 2014. 32, 114

[2] A. Ahnesjö, P. Andreo, and A. Brahme. Calculation and application of point spread functions for treatment planning with high energy photon beams. Acta Oncol, 26(1):49–56, 1987. 11

[3] E. S. Andersen, L. P. Muren, T. S. Sorensen, K. O. Now, M. Thor, J. B. Petersen, M. Hoyer, L. Bentzen, and K. Tanderup. Bladder dose accumulation based on a biomechanical deformable image registration algorithm in volumetric arc therapy for prostate cancer. Physics in Medicine and Biology, 57:7089–7100, 2012. 64

[4] L. Bodgi, A. Canet, L. Pujo-Menjouet, A. Lesne, J.-M. Victor, and N. Foray. Mathematical models of radiation action on living cells: From the target theory to the modern approaches. a historical and critical review. Journal of Theoretical Biology, 394:93–101, 2016. 28, 29

[5] E. Boman, M. Kapanen, P. L., and S.-L. Lahtela. Importance of deformable image registration and bioloigcal dose summation in planning of radiotherapy retreatments. Medical Dosimetry, 42(4):296–303, 2017. 2, 18

[6] T. Bortfeld, J. Burkelbach, R. Boesecke, and W. Schlegel. Methods of image reconstruction from projections applied. Physics in Medicine and Biology, 35:1423–1434, 1990. 1

[7] T. Bortfeld and H. Paganetti. The biologic relevance of daily dose variations in adaptive treatment planning. Int J Radiation Oncology Biol Phys, 65(3):899–906, 2006. 107, 108

[8] T. Bostel, I. Sachpazidis, M. Splinter, N. Bougatf, T. Fechter, C. Zamboglou, O. Jäkel, P. E. Huber, D. Baltas, J. Debus, and N. H. Nicolay. Dosimetric effects of interfractional variations in prostate cancer radiotherapy – implications for imaging frequency. submitted, 2019. 58, 59, 81

[9] A. Brahme, J. E. Roos, and I. Lax. Solution of an integral equation encountered in rotation therapy. Physics in Medicine and Biology, 27:1221–1229, 1982. 1, 9

[10] D. J. Brenner, A. A. Martinez, G. K. Edmundson, C. Mitchell, H. D. Thames, and E. P. Armour. Direct evidence that prostate tumors show high sensitivity to fractionation (low alpha/beta ratio), similar to late-responding normal tissue. Int J Radiation Oncology Biol Phys, 52(1):6–13, 2002. 32

[11] 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. Medical Physics, 44(7):e43–e76, 2017. 21

[12] F. Buettner, S. L. Gulliford, S. Webb, M. R. Sydes, D. P. Dearnaley, and M. Partidge. Assessing correlations between the spatial distribution of the dose to the rectal wall and late rectal toxicity after prostate radiotherapy: an analysis of data from the mrc rt01 trial (isrctn 47772397). Physics in Medicine and Biology, 54(21), 2009. 66, 68

[13] N. G. Burnet, J. E. Scaife, M. Romanchikova, S. J. Thomas, A. M. Bates, E. Wong, D. J. Noble, L. E. A. Shelley, S. J. Bond, J. R. Forman, A. C. F. Hoole, G. Barnett, F. Brochu, M. P. D. Simmons, R. Jena, K. Harrison, P. L. Yeap, A. Drew, E. Silvester, P. Elwood, H. Pullen, A. Sultana, S. Y. K. Seah, M. Z. Wilson, S. G. Russell, R. J. Benson, Y. L. Rimmer, S. J. Jefferies, N. Taku, M. Gurnell, A. S. Powslon, C.-B. Schönlieb, X. Cai, M. P. F. Sutcliffe, and M. A. Parker. Applying physical science techniques and cern technology to an unsolved problem in radiation treatment for cancer: the multidisciplinary 'voxtox' research programme. CERN Ideasq J Exp Innov, 1(1):3–12, 2017. 3, 65

[14] I. Chetty and M. Rosu-Bubulac. Deformable image registration for dose accumulation. Seminars in Radiation Oncology, 29(3):198–208, 2019. 42

[15] I. J. Chetty and J. Fontenot. Adaptive radiation therapy: Off-line, on-line, and in-line? Int. J. Rad. Onc. Biol. Phys., 99(3):689–691, 2017. 42

[16] M. Curie. Sur l'étude des courbes de probabilité relatives à l'action des rayons x sur les bacilles. Comptes-Rendus de l'Acdémie des Sciences, 188:202–204, 1929. 28

[17] J. O. de Xivry, P. Castadot, G. Janssens, J. A. Lee, X. Geets, V. Grégoire, and B. Macq. Evaluation of the radiobiological impact of anatomic modifications during radiation therapy for head and neck cancer: Can we simply summate the dose? Radiotherapy and Oncology, 96:131–138, 2010. 107, 110, 118

[18] A. Dhawan, M. Kohandel, R. Hill, and S. Sivaloganathan. Tumour control probability in cancer stem cells hypothesis. PLoS ONE, 9(5):e96093, 2014. 35

[19] K. Dolde, P. Naumann, C. Dávid, M. Kachelriess, A. J. Lomax, D. C. Weber, N. Saito, L. N. Burigo, A. Paffenberger, and Y. Zhang. Comparing the

effectiveness and efficiency of various gating approaches for pbs proton therapy of pancreatic cancer using 4d-mri datasets. Physics in Medicine and Biology, 64(8):085011 (11pp), 2019. 18

[20] M. Elkind and G. Whitmore. The radiobiology of cultured mammalian cells. Gordon and Breach, 1967. 29

[21] B. Emami, J. Lyman, A. Brown, L. Coia, M. Goitein, J. E. Munzenrider, S. B., L. J. Solin, and W. M. Tolerance of normal tissue to therapeutic irradiation. Int J Radiat Oncol Biol Phys, 21:109–122, 1991. 36

[22] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J.-C. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka, J. Buatti, S. Aylward, J. Miller, S. Pieper, and R. Kikinis. 3d slicer as an image computing platform for the quantitative imaging network. Magn Reson Imaging, 30(9):1323–1341, 2012. 61

[23] J. F. Fowler. The linear-quadratic formula and progress in fractionated radiotherapy. Br J Radiol, 62:679–694, 1989. 33

[24] Y. C. Fung. Biomechanics: Mechanical Properties of Living Tissues. Springer Science and Business Media, 2013. 41

[25] M. Glitzner, S. P. M. Crijns, B. D. de Senneville, C. Kontaxis, F. M. Prins, J. J. W. Lagendijk, and B. W. Raaymakers. On-line mr imaging for dose validation of abdominal radiotherapy. Physics in Medicine and Biology, 60(2):8869–8883, 2015. 26

[26] A. Godley, E. Ahunbay, C. Peng, and X. A. Li. Accumulating daily-varied dose distributions of prostate radiation therapy with soft-tissue–based kv ct guidance. Journal of applied Clinical Medical Physics, 13(3):98–107, 2012. 1, 17

[27] E. J. Hall and A. J. Giaccia. Radiobiology for Radiologists. Lippincott Williams and Wilkins, 2006. 30

[28] J. A. Hatton, P. B. Greer, C. Tang, P. Wright, A. Capp, S. Gupta, J. Parker, W. C., and J. W. Denham. Does the planning dose–volume histogram represent treatment doses in image-guided prostate radiation therapy? assessment with cone-beam computerised tomography scans. Radiotherapy and Oncology, 98(2):162–168, 2011. 1, 16

[29] J. H. Hubbell and S. M. Seltzer. Radiation physics division, pml, nist. © 1989, 1990, 1996 copyright by the U.S. Secretary of Commerce on behalf of the United States of America. All rights reserved. NIST reserves the right to charge for these data in the future. Last Update to Data Content: July 2004; NISTIR 5632. https://dx.doi.org/10.18434/T4D01F. 6

[30] G. D. Hugo, C. Dial, and J. V. Siebers. Dose accumulation for tumors undergoing mass changes during fractionated radiation therapy: In regard to zhong and chetty. Int. J. Rad. Onc. Biol. Phys., 99(5):1308–1310, 2017. 44

[31] 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. Radiation Oncology Biol. Phys., 76(3):S135–S139, 2010. 18, 40

[32] G. Janssens, O. de Xivry, S. Fekkes, A. Dekker, B. Macq, P. Lambin, and W. van Elmpt. Evaluation of nonrigid registration models for interfraction dose accumulation in radiotherapy. Medical Physics, 36(9):4268–4276, 2009. 121

[33] J. L. W. V. Jensen. Sur les fonctions convexes et les inégaltés entre les valeurs moyennes. Acta Math, 30:175–193, 1906. 49

[34] M. Joiner and A. van der Kogel. Basic Clinical Radiobiology. Hodder Arnold, 2009, 4th Edition. 33, 69, 117

[35] D. Jones. Icru report 50 — prescribing, recording and reporting photon beam therapy. Medical Phyics, 21(6), 1993. 12

[36] A. M. Kellerer and H. H. Rossi. The theory of dual radiation action. Curr. Top. Radiat. Res., 8:85–158, 1972. 2, 29

[37] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. Pluim. elastix: A toolbox for intensity-based medical image registration. IEEE Transactions on Medical Imaging, 29(1):196–205, 2010. 112

[38] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. W. Pluim. elastix: a toolbox for intensity based medical image registration. IEEE Transactions on Medical Imaging, 29(1):196–205, 2010. 69

[39] P. Källman, A. Agren, and B. A. Tumour and normal tissue responses to fractionated non-uniform dose delivery. Int J Radiat Oncol Biol Phys, 62:249–262, 1992. 37

[40] K. M. Kraus, O. Jäkel, N. I. Niebuhr, and A. Pfaffenberger. Generation of synthetic ct data using patient specific daily mr image data and image registration. Phys. Med. Biol., 62:1358–1377, 2017. 24

[41] P. Kupelian, K. M. Langen, O. A. Zeidan, S. L. Meeks, T. R. Willoughby, T. H. Wagner, S. J. Jeswani, K. J. Ruchala, H. J., and G. H. Olivera. Daily variation in delivered doses in patients treated with radiotherapy for localized prostate cancer. Int J Radiation Oncology Biol Phys, 66(3):876–882, 2006. 1, 16

[42] G. J. Kutcher and C. Burman. Calculation of complication probability factors for non-uniform normal tissue irradiation: the effective volume method. Int J Radiat Oncol Biol Phys, 16:1623–1630, 1989. 35

[43] T. Landberg, J. Chavaudra, J. Dobbs, J.-P. Gerard, G. Hanks, J.-C. Horiot, K.-A. Johansson, T. Möller, J. Purdy, N. Suntharalingam, and H. Svensson. Report 62 - prescribing, recording and reporting photon beam therapy. Journal of the International Commission on Radiation Units and Measurements, 1(os32), 1999. 12

[44] H. S. Li, H. Zhong, J. Kim, C. Gilde-Hurst, M. Gulam, T. S. Nurushev, and I. J. Chetty. Direct dose mapping versus energy/mass transfer mapping for 4d dose accumulation: fundamental differences and dosimetric consequences. Phys. Med. Biol., 59:173–188, 2014. 124

[45] X. Liu, E. Huang, Y. Wang, Y. He, H. Luo, M. Zhong, D. Qiu, C. Li, H. Yang, G. He, J. Zhou, and F. Jin. Dosimetric comparison of helical tomotherapy, vmat, fixed-field imrt and 3d-conformal radiotherapy for stage i-ii nasal natural killer t-cell lymphoma. Radiat Oncol, 12(76), 2017. 10

[46] J. T. Lyman. Complication probability as assessed from dose-volume histograms. Radiat Res Suppl, 8:S13–S19, 1985. 35

[47] L. Marks, E. D. Yorke, A. Jackson, R. K. Ten Haken, L. S. Constine, A. Eisbruch, S. M. Bentzen, J. Nam, and J. O. Deasy. Use of normal tissue complication probability models in the clinic. Int. J. Radiation Oncology Biol. Phys., 76(3):S10–S19, 2010. 18, 36

[48] B. Marples and M. Joiner. The repsonse of chinese hamster v79 cells to low radiation doses: evidence of enhanced sensitivity of the whole cell population. Rad. Res., 133:41–51, 1993. 33

[49] S. J. McMahon. The linear quadratic model: usage, interpretation and challenges. Physics in Medicine and Biology, 64(1):24pp, 2019. 2, 28, 30

[50] H.-G. Menzel. The international commission on radiation units and measurements - report 83. Journal of the ICRU, 10(1), 2015. 37

[51] J. M. Michalski, H. Gay, A. Jackson, S. Tucker, and J. O. Deasy. Radiation dose-volume effects in radiation-induced rectal injury. Int. J. Radiation Oncology Biol. Phys., 76(3):S123–S129, 2010. 60, 63, 118

[52] S. Milz, J. J. Wilkens, and W. Ullrich. A dose error evaluation study for 4d dose calculations. Phys. Med. Biol., 59:6401–6415, 2014. 124

[53] R. Miralbell, S. A. Roberts, E. Zubizaretta, and J. H. Hendry. Dose-fraction sensitivity of prostate cancer deduced from raiotherapy outcomes of 5969 pateints

in seven international institutional datasets: alpha/beta = 1.4 (0.9-2.2) gy. Int J Radiation Oncology Biol Phys, 82(1):e17–e24, 2012. 32

[54] 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. Medical Physics, 39(2):573–580, 2012. 121

[55] J. Murray, A. McQuaid, D. ad Dunlop, F. Buettner, S. Nill, E. Hall, D. Dearnaley, and S. Gulliford. Su-e-j-14: A novel approach to evaluate the dosimetric effect of rectal variation during image guided prostate radiotherapy. Medical Physics, 41(6), 2014. 66

[56] M. Nassef, A. Simon, G. Cazoulat, A. Duménil, C. Blay, C. Lafond, O. Acosta, J. Balosso, H. P., and R. de Crevoisier. Quantification of dose uncertainties in cumulated dose estimation compared to planned dose in prostate imrt. Radiotherapy and Oncology, 119(3):129–136, 2016. 1, 17, 40, 121

[57] N. I. Niebuhr, C. P. Karger, K. Harrison, D. J. Noble, R. Jena, R. O. Floca, J. Seco, M. Wilson, and A. Pfaffenberger. On the biological rationale of dose mapping in the presence of volume changes. Physics in Medicine and Biology, submitted August 2019. 3, 39, 104

[58] N. I. Niebuhr, M. Splinter, T. Bostel, J. Seco, C. M. Hentschke, R. O. Floca, J. Hörner-Rieber, M. Alber, P. Huber, J. Debus, N. H. Nicolay, and A. Pfaffenberger. Biologically consistent dose accumulation using daily patient imaging. Physics in Medicine and Biology, submitted September 2019. 3, 48, 56, 71, 74, 104

[59] A. Niemierko. A generalized concept of equivalent uniform dose (eud). Medical Phyics, 26(1101):Abstract WE–C2–9, 1999. 37

[60] A. Nimierko and M. Goitein. Modeling of normal tissue response to radiation: The critical volume model. Int. J. Rad. Onc. Biol. Phys., 25:135–145, 1993. 40, 41

[61] W. Nobnop, I. Chitapanarux, H. Neamin, S. Wanwilairat, V. Lorvidhaya, and T. Sanghangthum. Evaluation of deformable image registration (dir) methods for dose accumulation in nasopharyngeal cancer patients during radiotherapy. Radiotherapy and Oncology, 51(4):438–446, 2017. 17

[62] M. Nolden, S. Zelzer, A. Seitel, D. Wald, M. Müller, A. Franz, D. Maleike, M. Fangerau, M. Baumhauer, L. Maier-Hein, K. Maier-Hein, and I. Wolf. The medical imaging interaction toolkit: challenges and advances. Int J Comput Assist Radiol Surg, 8(4):607–620, 2013. 63

[63] F. Palorini, C. Cozzarini, S. Gianolini, A. Botti, V. Carillo, C. Iotti, T. Rancati, R. Valdagni, and C. Fiorino. First application of a pixel-wise analysis on bladder

dose–surface maps in prostate cancer radiotherapy. Radiotherapy and Oncology, 119:123–128, 2016. 1, 18

[64] M. Rosu, J. Chetty, M. Balter, M. L. Kessler, L. McShan, and R. K. Ten Haken. Dose reconstruction in deforming lung anatomy: Dose grid size effects and clinical implications. Medical Physics, 32(8):2487–2495, 2005. 27, 64, 92, 125

[65] J. E. Scaife, G. C. Barnett, D. J. Noble, R. Jena, S. J. Thomas, C. M. L. West, and N. G. Burnet. Exploiting biological and physical determinant of radiotherapy toxicity to individualize treatment. Br J Radiol, 88:150–172, 2015. 17

[66] W. Schlegel, O. Jäkel, and C. P. Karger. Medizinische Physik. Springer Spektrum, 2019. 5, 7, 8, 10

[67] T. E. Schultheiss, C. G. Orton, and R. A. Peck. Models in radiotherapy: Volume effects. Medical Physics, 10:410–415, 1983. 40, 41

[68] T. E. Schultheiss, W. A. Tomé, and C. G. Orton. Point/counterpoint: it is not appropriate to "deform" dose along with deformable image registration in adaptive radiotherapy. Medical Physics, 39(11):6531–6533, 2012. 2, 40, 112, 121

[69] D. P. Shamonin, E. E. Bron, B. Lelieveldt, M. Smits, K. S., and M. Staring. Fast parallel image registration on cpu and gpu for diagnostic classification of alzheimer's disease. Frontiers in Neuroinformatics, 7(50):1–15, 2014. 69

[70] L. E. A. Shelley, J. E. Scaife, M. Romanchikova, K. Harrison, J. R. Forman, A. M. Bates, D. J. Noble, R. Jena, M. A. Parker, M. P. F. Sutcliffe, S. J. Thomas, and N. G. Burnet. Delivered dose can be a better predictor of rectal toxicity than planned. Radiotherapy and Oncology, 123(3):466–471, 2017. 1, 17, 18, 65, 66, 68, 95, 99, 127

[71] J.-J. Sonke, M. Azanr, and C. Rasch. Adaptive radiotherapy for anatomical changes. Seminars in Radiation Oncology, 29(3):245–257, 2019. 1, 40

[72] S. J. Thomas, M. Romanchikova, K. Harrison, M. A. Parker, A. M. Bates, J. E. Scaife, M. P. F. Sutcliffe, and N. G. Burnet. Recalculation of dose for each fraction of treatment on tomotherapy. Br J Radiol, 89:20150770, 2016. 66, 69

[73] D. Tilly, N. Tilly, and A. Ahnesjö. Dose mapping sensitivity to deformable registration uncertainties in fractionated radiotherapy – applied to prostate proton treatments. BMC Medical Physics, 13(2), 2013. 121

[74] C. Veiga, M. A. Lourenco, S. Mouinuddin, M. van Herk, M. Modat, S. OUrselin, G. Royle, and J. McClelland. Toward adaptive radiotherapy for head and neck patients: Uncertainties in dose warping due to the choice of deformable registration algorithm. Medical Physics, 42(2):760–769, 2015. 40, 112, 121

[75] A. N. Viswanathan, E. D. Yorke, L. B. Marks, P. Eifel, and W. U. Shipley. Radiation dose-volume effects of the urinary bladder. Int J Radiation Oncology Biol Phys, 76(3):S116–S122, 2010. 63, 118

[76] E. M. Vásquez Osorio, M. S. Hoogeman, A. Al-Mamgani, D. N. Teguh, P. C. Levendag, and B. J. M. Heijmen. Local anatomic changes in parotid and submandibular glands during radiotherapy for oropharynx cancer and correlation with dose, studied in detail with nonrigid registration. Int. J. Rad. Onc. Biol. Phys., 70(3):875–882, 2008. 112

[77] O. Weistrand and S. Svensson. The anaconda algorithm for deformable image registration in radiotherapy. Med Phys, 42(1):40–53, 2015. 23, 60

[78] S. G. Williams, J. M. Taylor, Y. Tra, G. M. Duchesne, L. L. Kestin, A. Martinez, G. R. Pratt, and H. Sandler. Use of individual fraction size data from 3756 patients to directly determine the alpha/beta ratio of prostate cancer. Int J Radiat Oncol Biol Phys, 68(1):24–33, 2007. 31, 32

[79] J. Winiecki, T. Morgas, K. Majewskab, and B. Drzewiecka. The gamma evaluation method as a routine qa procedure of imrt. Reports of Practical Oncology and Radiotherapy, 14(5):162–168, 2009. 63

[80] H. R. Withers and J. M. Taylor. Treatment volume and tissue tolerance. Int J Radiat Oncol Biol Phys, 14:751–759, 1988. 36

[81] R. C. Wortel, M. G. Witte, U. A. van der Heide, F. J. Pos, J. V. Lebesque, M. van Herk, L. Incrocci, and W. D. Heemsbergen. Dose–surface maps identifying local dose–effects for acute gastrointestinal toxicity after radiotherapy for prostate cancer. Radiotherapy and Oncology, 117:515–520, 2015. 18

[82] R. J. Yaes, P. Patel, and Y. Maruyama. On using the linear-quadratic model in daily clinical practice. Int. J. Radiation Oncology Biol. Phys., 20(6):1353–1362, 1991. 52

[83] U. J. Yeo, M. L. Taylor, J. R. Supple, R. L. Smith, L. Dunn, T. Kron, and R. D. Franich. Is it sensible to "deform" dose? 3d experimental validation of dose-warping. Medical Physics, 39(8):5065–5072, 2012. 121

[84] R. Yeung, Y. McConnell, H. Warkentin, D. Graham, B. Warkentin, K. Joseph, and C. M. Doll. Intensity-modulated radiotherapy (imrt) vs helical tomotherapy (ht) in concurrent chemoradiotherapy (crt) for patients with anal canal carcinoma (acc): an analysis of dose distribution and toxicities. Radiat Oncol, 10(92), 2015. 10

[85] C. Zachiu, B. D. de Senneville, C. T. W. Moonen, B. W. Raaymakers, and M. Ries. Anatomically plausible models and quality assurance criteria for online mono- and multi-modal medical image registration. Physics in Medicine and Biology, 63(15), 2018. 121

[86] M. Zaider and L. Hanin. Tumor control probability in radiation treatment. Medical Physics, 38(2):574–583, 2011. 35

[87] S. Zavgorodni. The impact of inter-fraction dose variations on biological equivalent dose (bed): the concept of equivalent constant dose. Physics in Medicine and Biology, 49:5333–5345, 2004. 107, 131

[88] L. Zhang, M. Hub, S. Mang, C. Thieke, O. Nix, C. P. Karger, and R. O. Floca. Software for quantitative analysis of radiotherapy: Overview, requirement analysis and design solution. Computer Methods and Programs in Biomedicine, 110(3):528–537, 2013. 61

[89] H. Zhong and I. J. Chetty. Adaptive radi otherapy for nsclc patients: utilizing the principle of energy conservation to evaluate dose mapping operations. Physics in Medicine and Biology, 62(11):4333–4345, 2017. 42

[90] H. Zhong and J. Chetty. Caution must be exercised when performing deformable dose accumulation for tumors undergoing mass changes during fractionated radiation therapy. Int. J. Rad. Onc. Biol. Phys., 97(1):182–183, 2017. 2, 40, 42, 44, 103

[91] H. Zhong, M. Siddiqui, M. B., and I. J. Chetty. Evaluation of adaptive treatment planning for patients with non-small cell lung cancer. Physics in Medicine and Biology, 62(11):4346–4360, 2017. 42

ACKNOWLEDGEMENTS

I would like to thank Prof. Joao Seco, for being my first supervisor and for his support and guidance. I am very grateful for the possibility to carry out this thesis in the *Department of Medical Physics in Radiation Oncology* of Prof. Oliver Jäkel and for his role as second referee. I am especially thankful for all the opportunities this involved, including numerous conferences and a research visit. I would like to thank Prof. Norbert Herrmann and Prof. Joerg Jaeckel, for serving as examiners in the disputation of this thesis.
Special thanks go to Dr. Asja Pfaffenberger, for her supervision and ideas, giving me both scientific support and freedom, and the warm atmosphere in her group. Thank you to Prof. Markus Alber, for being a member in my thesis advisor committee and the feedback. I am particularly grateful for the computational assistance given by Dr. Ralf Floca and Dr. Clemens Hentschke. I have greatly benefited from the fruitful discussions with Prof. Christian Karger, for which I am very thankful. In clinical matters, I am grateful for the support from Dr. Nils Nicolay and Dr. Juliane Hörner-Rieber. Thank you to Mona Splinter, for her assistance with the data.
I'd like to express my deepest gratitude to Prof. Raj Jena, for his support and collaboration and the opportunity to pursue my research within the VoxTox group in Cambridge. Moreover, I'd especially like to thank Dr. Karl Harrison, Dr. David Noble and Leila Shelley, for the warm welcoming and outstanding support and guidance throughout my time in Cambridge. Without this exceptionally kind and fruitful collaboration, this thesis would not have possible.
The years of my PhD have been a journey for which I am very grateful I could travel in the company of numerous friends, in and outside of work - I am glad you were there! I'd like to appreciate all group-, lunch- and office-mates (current and former) and colleagues, for the great times we've spent, the kind support that was given, and the warm and fun working atmosphere. Especially, I'd like to say thank you to Kai, for his support, creativity and company along the "stony path"; to Lenny, for the fruitful discussions and encouragements; to Patrick, for recovering walks and talks; and to Josefine, for the successful joint journey to the finish line. Moreover, I want to thank Kathrin, for our time to find both strength and Zen together and, together with Oli, for the great times in the group, on numerous conferences and outside of work. Thank you to Jasper, for friendship, honesty and support, despite the distance. Thank you Markus, for your warm encouragements, from poster sessions to jumps in cold water, and all the joy along that way. Leonie, I am so glad to have you with me still outside the DKFZ-U-boot. Alina, thank you for every word and laugh, and all your warmth.
Very special thanks go to my family who have made all this possible, supporting and encouraging me, and letting me find my way. Felix, I want to deeply thank you for inspiring me with your courage and curiosity, and for all your love and support.

Erklärung

Ich erkläre hiermit, dass ich die vorgelegte Dissertation selbst verfasst und mich dabei keiner anderen, als der von mir ausdrücklich bezeichneten Quellen und Hilfen bedient habe.

Heidelberg, den 07.10.2019

.....