Master of Science Thesis
HT2021

Breathing Effects in Proton Scanning for Hodgkin's Lymphoma

John Nilsson

Supervisors

Marika Enmark, Anneli Edvardsson, Ingrid Kristensen, Lund and Karin Andersson, Christina Vallhagen Dahlgren, Uppsala

Medical Radiation Physics, Lund
Faculty of Science
Lund University
www.msf.lu.se

Andningseffekter för protonbehandling av Hodgkins lymfom

Idag behandlas nästan hälften av alla cancerpatienter i Sverige med någon form av strålbehandling, som komplement till annan behandling eller som självständig terapi. Strålningen är optimerad för att koncentrera effekten på tumörområdet och minimera biverkningar på intilliggande organ och vävnad. Dagens strålbehandling utnyttjar flera strålslag som tillför olika behandlingsformer och valbarheter. Majoriteten av alla cancerpatienter har traditionellt sätt behandlats med fotoner, vilket är strålning med tillräckligt hög energi för att ge upphov till skada på vävnad. Numera är även laddade partiklar i form protoner en vanlig behandlingsform. Protoner har fysikaliska egenskaper som medför att partiklarna avger majoriteten av sin energi på en distinkt räckvidd, jämfört med fotoner som deponerar sin energi över ett större behandlingsdjup. Den väldefinierade räckvidden hos protoner tillåter tumöreffekten att koncentreras naturligt, vilket resulterar i att behandlingsvolymen minskar vid klinisk protonbehandling jämfört med konventionell strålbehandling. Partikelräckvidden är däremot starkt påverkad av densitetsvariationer i kroppen, vilket vid större förändringar, orsakat av till exempel andningsrörelser, kan förskjuta protoner bort ifrån tumörområdet. Behandlingsområden såsom lunga har således tidigare undvikits eftersom andningsrörelser är starkt närvarande i lungområdet. Modern protonbehandlingen använder idag en skannad stråle som utgörs av flera monoenergetiska protoner, vilka levereras till olika delar av tumörvolymen. Målet är att bestråla tumörområdet homogent, vilket möjliggörs då den skannade strålen sveps över hela tumörvolymen. Under bestrålning justeras energin och positionen av protonerna som ska överlämna energin. Skanningen är dock inte momentan i tiden, vilket, i kombination med andningsrörelser, kan leda till att den skannade strålen inte sveps uniformt över tumörvolymen. En sådan effekt gör att bestrålning inte blir homogen och är ett exempel på en så kallad interplayeffekt.

I det här arbetet undersöktes och kvantifierades påverkan av andningsrörelser vid skannad protonbehandling av mediastinal Hodgkins lymfom, en lymfkärlsjukdom lokaliserad i lungområdet där andningsrörelser kan vara av betydelse. Idag är det endast rekommenderat att tillåta protonbehandling av Hodgkins lymfom för tumörrörelser mindre än 5 mm i någon riktning för att försäkra sig om att tumörområdet bestrålas homogent och att angränsade områden inte bestrålas. Syftet med arbetet var att jämföra andningseffekter på tumören för tre olika andningsstorlekar: $5, 10$ och $15\,\text{mm}$ och undersöka huruvida gränsen för 5 mm tumörrörelse kan höjas. Andningen simulerades m.h.a. en periodisk vågrörelse för tre patienter med olika behandlingsplaner och tumörgeometrier som tidigare fått protonbehandling. En retrospektiv analys genomfördes där andningsparametrar såsom vågrörelsens form, frekvens eller diverse behandlingsparametrar som i teorin kan minimera andningseffekten undersöktes och vilken inverkan parametrarna hade på patienterna.

Studien tillkännager att interplayeffekt är patientspecifik med stora variationer, men att behandlingsparametrar kan justeras för att mildra interplayeffekten. Arbetet har även belyst nödvändigheten av en protonbehandling med homogen bestrålning, då andningseffekten var större för inhomogena bestrålningar. Resultaten presenteras med optimism och ger indikationer på att tumörrörelser över 5 mm kan tillåtas, och viktigaste av allt, att protonbehandlingen är säker för de rörelserestriktioner som finns idag.

Abstract

Purpose/Background: Proton beam scanning (PBS) for mediastinal Hodgkin's Lymphoma (HL) is limited when treating in free-breathing (FB) due to breathing motion in the thoracic region. Tumour motion and the time structure of the beam delivery will cause interplay and dose blurring effects, resulting in discrepancies of dose in the planned dose distribution both inside and at the edges of tumour volume (TV). Current national guidelines only allow tumour motion less than 5 mm in any direction to ensure target dose coverage. This study assesses the effects of simulated breathing motions in PBS delivery for various mediastinal HL dose distributions using different tumour motion amplitudes and beam delivery parameters. Specifically, this study aims to quantify the dosimetric effects of large breathing amplitudes on PBS delivery in the cases included in the study and identify parameters that may mitigate breathing effects.

Material and methods: The impact of breathing was estimated by computing relative dose differences D_{rel} between dose distributions during motion and corresponding static conditions. Dose distributions were measured using a 2D ionisation chamber array detector on top of a modified motion platform at the Swedish proton facility in Uppsala. Planar dose distributions from three HL patients with clinical robust treatment plans were measured in solid water for different target depths by allocating various amounts of solid water plates on top of the detector. The respiratory tumour motion was simulated using an asymmetrical sine-curve with amplitudes of 5, 10 and 15 mm in the superior-inferior (SI) direction. Tumour motion parameters were assessed by changing period time and motion regularity. We also studied a range of beam delivery parameters by modifying the air gap, spot spacing and accumulating fractions. Breathing effects were quantified by computing a 98th percentile of the absolute value of D_{rel}.

Results: Breathing effects increased with larger amplitudes but could be mitigated by accumulating fractions, increasing the air gap and reducing the spot spacing. The interplay effect dominated the overall breathing effect and varied drastically between patients at 5 mm motion amplitude, with a single-fraction 98th percentile value ranging from 10.3% in the worst-case scenario, representing a heterogeneous dose distribution, to 2.2% in the best-case scenario, corresponding to a homogeneous dose distribution.

Conclusion: For the patients considered, interplay effects are highly patient-specific, highlighting that the impact of respiratory motion has a high dependency on the beam delivery and plan characteristics and not only on the tumour motion pattern. PBS may be allowed for tumour motion above 5 mm in the SI direction. However, to reduce the interplay effect, care should be given to avoid a heterogeneous dose distribution in the tumour volume, keeping the spot intensities low if achievable.

Abbreviations

HL - Hodgkin's Lymphoma
PBS - Pencil Beam Scanning
OAR - Organ At Risk
PT - Proton Therapy
DIBH - Deep Inspiration Breath-hold
FB - Free Breathing
PDD - Percentage Depth Dose
SOBP - Spread-Out Bragg Peak
TV - Tumour Volume
IBA - Ion Beam Applications
ESS - Energy Selection System
PSPT - Passive Scattering Proton Therapy
IMPT - Intensity-Modulated Proton Therapy
FWHM - Full Width Half Maximum
MU - Monitor Units
HU - Houndsfield Units
CT - Computed Tomography
CTV - Clinical Target Volume
AP - Anterior-Posterior
SI - Superior-Inferior
LR - Left-Right
τ_{ss} - Spot Settling Time
τ_{es} - Energy Switching Time
SFO - Single-Field Optimised
MFO - Multiple-Field Optimised
COM - Centre Of Mass
RBE - Relative Biological Effectiveness
WET - Water Equivalent Tissue

Contents

1 Introduction

Hodgkin's lymphoma (HL) is a rare malignancy that stems from Reed-Sternberg cells, an enlarged and cancerous form of white B-cells most often formed in the mediastinal lymph nodes inside the thoracic region of the body. In the worst cases, the Reed-Sternberg cells spread through the lymphatic system and sometimes beyond it. [1, 2] Patients with early staged HL most often have to endure a combination of chemotherapy and radiation therapy. [3, 4] In the past, photon therapy techniques were used as consolidation therapy following chemotherapy. Today, Proton Pencil Beam Scanning (PBS) has become a promising alternative to treating HL. The PBS technique has demonstrated high tumour control while maintaining a low toxicity profile to organs at risks (OARs). [5, 6] However, a major confounder for an accurate PBS delivery in the thoracic region is intra-fractional motion. In contrast to inter-fractional variations, which can greatly be reduced by daily imaging and accurate set-up protocols [7], intra-fractional motion is random and unpredictable. [8] In the thoracic region, intra-fractional variations are mainly caused by respiratory motion, hence very patient-specific and will affect the planned dose distribution differently depending on many parameters. [9, 10] The tumour motion correlated with the respiratory motion [11] and the temporal characteristics of the beam delivery in PBS proton therapy (PT) will cause interplay and dose blurring effects that induce discrepancies between the planned and delivered dose distribution. [5, 7, 12–14] Intra-fractional movements are significant concerns for mediastinal HL due to its location inside the thoracic region. The Swedish national treatment guidelines recommend HL patients to perform Deep Inspiration Breath-hold (DIBH). [3] In DIBH, the tumour is fixated and generally has many dosimetric advantages over treatment in free-breathing (FB). [5, 15] However, DIBH might not be suitable for all patients and treatment in FB may be the only option. Although PBS in FB is possible, it is only advisable for patients with tumour motion less than 5 mm along any translation axes to ensure target coverage. [3]

1.1 Aim

This study aimed to quantify the dosimetric influence of breathing effects in PBS PT for various mediastinal HL dose distributions in FB. We assessed the current recommendations on tumour motion by evaluating different tumour motion amplitudes of 5, 10 and 15 mm along the superior-inferior (SI) direction, motion regularity and different breathing period times. In theory, choosing specific field and plan characteristics, such as using appropriate spot spacing and air gap, can also affect the dosimetric impact of intra-fractional motion. [5, 10, 16, 17] Hence, the impact of plan and field configurations and motion characteristics on breathing effects were included in the study.

2 Theoretical background

2.1 Proton physics

Proton beams interact differently in matter compared to photon beams. In contrast to photons, proton beams are defined by a range parameter specific to the medium. The range approximately determines the proton beam's distance before losing all of its energy and stops. Thus, the range parameter becomes an indirect measure of the proton beam energy. The range parameter is not solely dependent on the initial energy of the proton beam but also on the angle of incidence and the density variation along the path of the beam. [6, 18] Plotting a percentage depth dose curve (PDD) of a proton beam demonstrates the dependency on the range (figure 1). In general, the PDD curve describes the dose deposition pattern of a particular particle inside a medium as a function of depth. Figure 1 visualises that protons have minimal exit dose compared to photons. Instead, most of the energy is deposited at a particular depth known as the Bragg peak, which is uniquely defined by the range parameter. [6, 7, 18] In modern PT, the Bragg peak characteristics of protons are utilised to form a Spread-Out Bragg Peak (SOBP) in-depth using proton beams of different energies. In this way, PT can attain high homogeneity and conformity of dose distribution over the entire tumour volume (TV). [7, 18, 19]

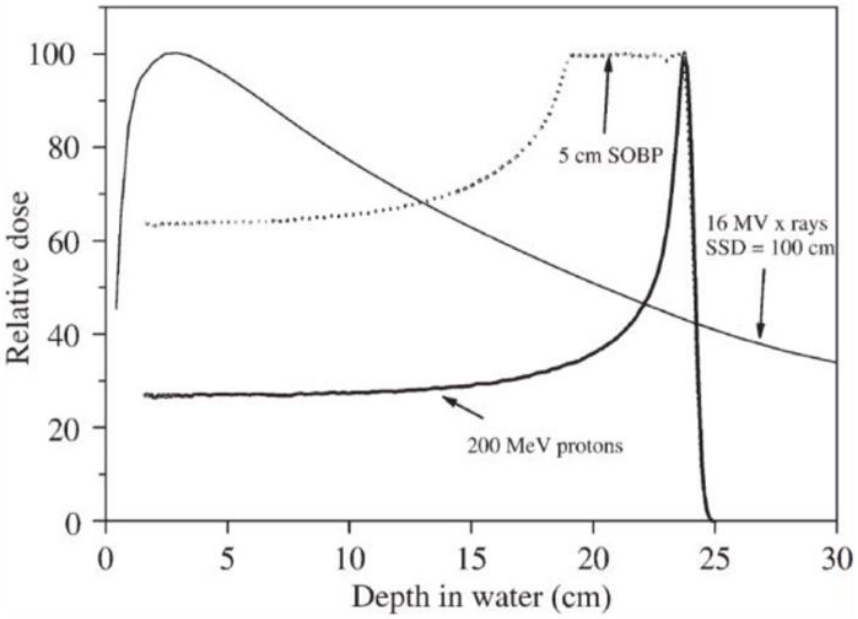


Figure 1: Normalised PDD curves at maximum dose for a proton and x-ray (photon) beam ($10x10$ cm^2 field size). A 5 cm SOBP is also visualised. [18]

2.2 The PT gantry

One of the major manufacturers for PT systems is Ion Beam Applications (IBA). IBA proton facilities utilise a cyclotron to produce protons with a fixed energy of 227 MeV. Subsequently, protons will travel towards the treatment room along a beamline. The proton beam's energy is reduced in the beamline by an Energy Selection System (ESS) to obtain therapeutic energies of e.g. $60 - 227$ MeV. Before irradiating the patient, the beam will enter and traverse through many important technical components located in the so-called nozzle. The nozzle and technical devices inside the nozzle are described and visualised in figure 2. Inside the nozzle, several ionisation chambers (IC1-3) monitor and verify the beam characteristics and measure the dose output before leaving the nozzle. [7] Other components include scanning magnets used to focus the beam position laterally inside the TV [7, 20] and a range shifter that can be inserted perpendicular to the beam path at the nozzle exit. The range shifter degrades the beam's energy, suitable for superficial tumours. [7, 20]

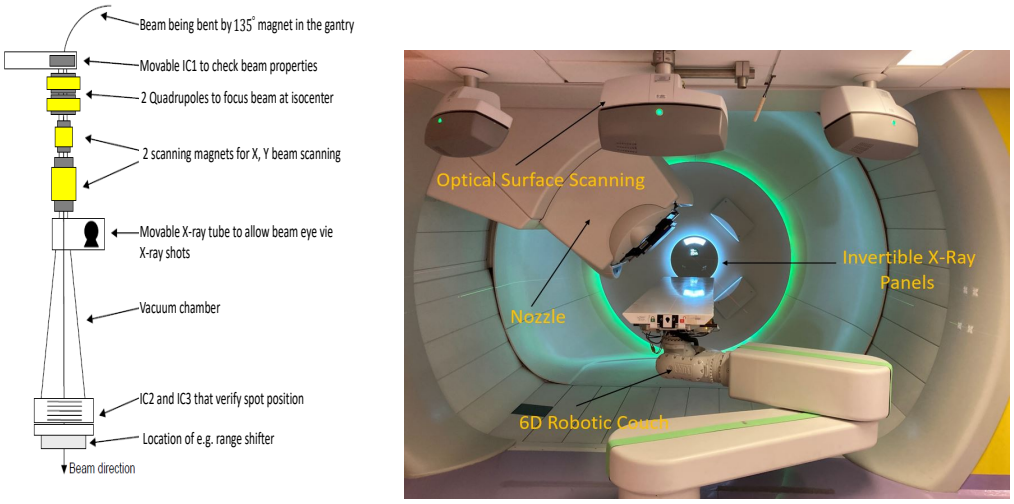


Figure 2: Left: Important components used inside the nozzle. Right: One of the treatment rooms at the Swedish proton facility in Uppsala. The treatment couch and nozzle can rotate in six dimensions for most gantries.

2.3 Pencil beam PT

The latest advancement in PT is the implementation of PBS. Contrary to its ancestor, Passive Scattering PT (PSPT), PBS PT can better conform the dose distribution to the TV located under challenging geometries within the body without inducing a higher risk of overdosing OARs. PBS offers a variety of optimisation tools. One of the most advanced PT techniques is PBS Intensity-Modulated PT (IMPT), which can modulate the beam intensity of the pencil beams based on objective functions and constraints on the target region. [21]

In PBS, the nozzle emits multiple beamlets (pencil-beams) that deliver the dose. A beamlet is a collection of mono-energetic proton beams. Each proton within the beamlet will deposit its

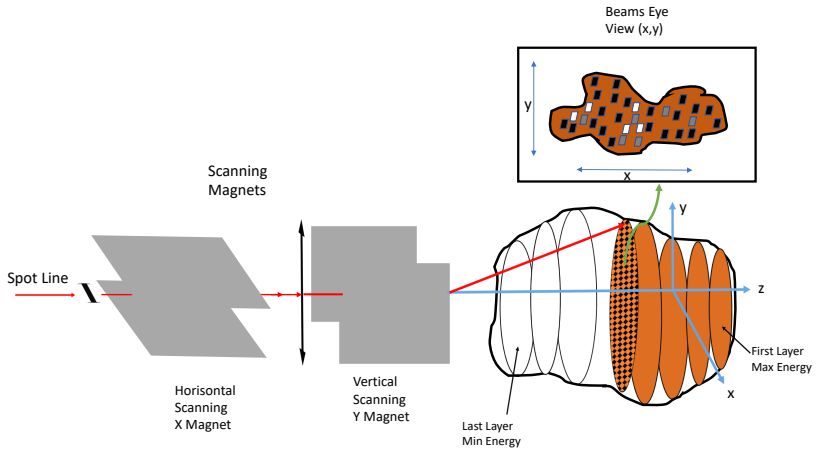
energy at a mostly unified position, called the "spot". Due to proton beam scattering inside the nozzle and patient, the spot will have a finite size [22], which can be estimated by the Full-Width Half Maximum (FWHM) of a beamlet at, e.g. iso-center in air for a specific beam energy. [20, 23]

Beamlets are paired in groups, forming several energy layers depending on their energy. Beamlets are spread out inside the TV laterally for a particular energy layer by scanning magnets located inside the nozzle, forming a spot pattern inside the energy layer. [22] The scanning magnets are illustrated in figure 2 and explained in figure 3a. In PSPT, mechanical devices such as apertures conform the dose distribution laterally. However, in PBS, the scanning magnets realises the spot patterns determined by a treatment planning system to form the dose distribution. The spot pattern is optimised to conform the dose distributions laterally (and proximally) by limiting spot positions near the edges of the TV, changing the number of spots in an energy layer, and altering the spot's intensity (Monitor Units [MU] per spot) for every spot inside an energy layer. [7, 18, 19, 21] The overall dose distribution of the TV is delivered by accumulating energy layers starting with the highest energy corresponding to the deepest depth, then descending to lower energies for shallower depths. Notably, the most distal layers will also contribute to the dose in the proximal layers. By delivering many energy layers of Bragg peaks, an SOBP is formed in-depth. This phenomena is showcased in figure 3b. [7, 18, 19, 21]

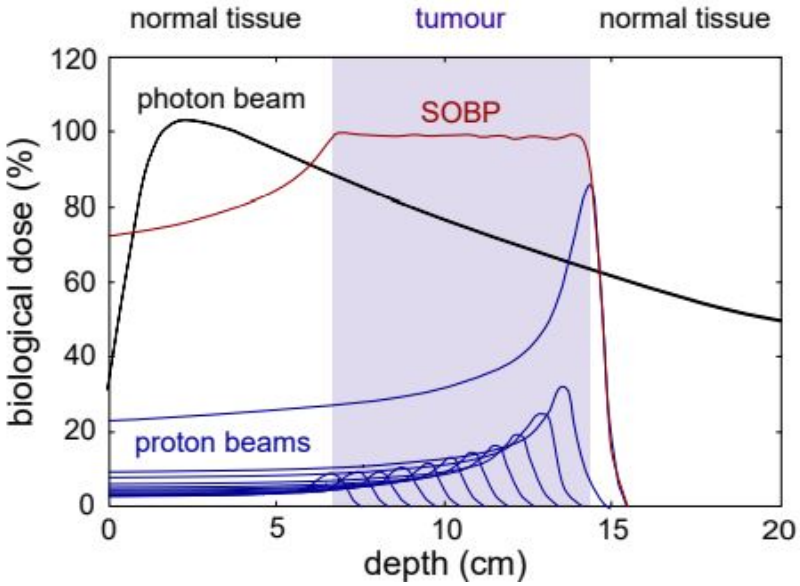
2.4 Uncertainties in PBS PT

The unique dose characteristics of protons with sharp dose gradients make them much more sensitive to density variations along the direction of the beam than photons. Even small shifts of range can severely impact the dose distribution and, in worst-case scenarios, result in little or no dose to parts of the TV. [8, 24, 25]

Different elements during treatment planning and beam delivery will result in range uncertainties. One of the main contributors to range uncertainties stems from estimated patient-specific stopping power values used to predict the range inside the patient. The stopping power values are calculated and related through Houndsfield Units (HU) in the Computed Tomography (CT) image. CT images with good image quality (high image resolution, no artefacts) are estimated to have a range uncertainty of $\sim$ 3%. After that, additional uncertainties will add to this value due to anatomical changes inside the patient throughout the fractionated treatment, resulting in variations in the beam path or to an "in-vivo range". [8] This include set-up errors (daily misalignment of patient positions relative to the planned position), inter-fractional changes (weight gain, tumour shrinkage) or intra-fractional changes (mainly due to respiratory motion), amongst others. [5, 7, 8, 12, 24] Daily imaging can reduce set-up errors; however, intra-fractional or inter-fractional variations are far less controllable and result in an uncertain in-vivo range. Photon therapy accounts for most uncertainties by adding a margin around the tumour and its sub-clinical microscopic spread (Clinical Target Volume, CTV). For instance, set-up errors or internal motion are considered in this way. [7] These concepts are still used in proton treatment planning; however, only adding a spatial volume expansion to the CTV is not sufficient. [22, 26] For protons, the CTV structure is expanded based on range uncertainties, whose margin is created in the optimisation process of the treatment plan. This technique is known as robust treatment planning since this treatment



((a)) The beamlet traverses through the nozzle, where scanning magnets focus the beamlet to a particular position inside the TV, perpendicular to the beam axis (the z-axis in this case). The layered brown segments inside the TV represent layers that have been delivered. The beamlet is delivered to a particular layer by changing the beam energy. One layer has black dots, which corresponds to a spot pattern with different spot intensities, shown in the beams eye view illustration as different colours. Re-illustrated from [7]



((b)) PDD curves for a photon beam and several proton beams forming the SOBP in depth. [7]

Figure 3: Illustration of the SOBP and scanning magnets.

plan accounts for e.g. set-up variations during treatment. [26] A robust plan will ultimately be less conformal than an un-robust plan, but beamlets with a high risk of degrading the plan will be penalised iteratively in the optimisation process, reducing their intensity. [7, 8] One consequence

of robust treatment planning is that high-intensity spots may be created in the border of high-density variations, such as bone-lung segments in the body. To account for range uncertainties in such regions, high MU/spot are formed since beamlets delivered through softer tissues, as the lung, would travel further than beamlets traversing through bone for the same spot's intensity. [7]

2.5 Intra-fractional uncertainties

The main cause of intra-fractional variations in the thoracic region is respiratory motion, which extorts a force on the TV and other tissues, resulting in an asymmetrical motion. Respiratory organ motion affects both the spatial and density information of the CT data, resulting in degradation of the desired dose inside the target region. In a worst-case scenario, the dose will mainly be delivered within the regions of OARs. [5, 7] The dose distribution is affected by two dosimetric effects. Firstly, the intra-fractional organ motion will smear and average the dose, even for a uniform delivery. [5, 16, 27–29] This effect is known as *dose blurring*, and the geometrical shifts will widen the beam penumbra and reduce conformity near the edges. In summation, dose blurring might shift the dose away from the TV and into neighbouring organs. [5, 14, 27]

The second dosimetric effect is due to *interplay*, caused by the simultaneous movement between the temporal characteristics of PBS and organ motion. In effect, the planned range might be greater or less than the delivered range, causing overshoot or undershoot, i.e. under-and over-dosing inside the TV. In other words, the interplay effect will enhance the variability of dose, even for regions that would receive a uniform dose under a static scenario. Thus, interplay reduces the dose homogeneity of the TV, and in the worst-case scenario, the undershooting of the range will translate to over-dosing of OARs, especially at the distal edge of the TV. [5, 7, 8, 13, 14, 16]

The effect of dose blurring can be reduced by increasing the internal margin centred around the CTV, but increasing margins will not mitigate the impact of interplay. [22, 28] In fact, the study [29] found that interplay had the largest impact on the target dose distribution for all motion amplitudes in lung cancer patients treated with IMPT. [29] Therefore, techniques mitigating interplay will be described more closely in the following sections.

2.6 Mitigating the interplay effect

Many studies in radiotherapy have investigated different methods to mitigate and limit the effect of intra-fractional movements and how to control the interplay effect better. [7, 12, 20] The interplay effect is highly patient-specific and depends on several parameters. [9, 10] For instance, in one study, the effect of respiratory motion on a static dose distribution was estimated for liver patients using IMPT. The average homogeneity index based on $D_5 - D_{95}$ increased from 8.8% with no motion to 23.4% including motion. [28] In another example with pediatric HL patients and IMPT, the $D_5 - D_{95}$ value reached 22.8% for one patient in a single fraction. [30] Overall, fractionation generally reduces the interplay effect [12, 16, 17, 29, 30], however not observed completely for all cases. Instead, studies have found converging limits where fractionation stops mitigating the interplay effect after 6 fractions. [12, 30] To further describe the root of the interplay effect,

Meijers et al. [9] have distinguished interplay into three categories: Patient characteristics, plan characteristics and delivery characteristics. These three categories will be discussed thoroughly in the following subsections.

2.6.1 Patient characteristics

Patient characteristics include individual parameters such as tumour motion amplitude, period times and motion regularity. [9] Many studies states that larger motion is highly correlated with interplay. [7, 10, 12, 17, 29–31] In detail, the study [29] found that interplay effects followed a linear-quadratically behaviour as a function of motion amplitude for IMPT and lung cancer patients. Several other studies found that irregular breathing patterns cause a greater impact of interplay on the TV than symmetric breathing patterns. [12, 16, 17] In addition, it was observed in [10] that a longer breathing period is associated with a reduction of homogeneity, with the largest effect observed for longer periods and short treatment times. Lastly, studies have also observed that the size of the TV is associated with interplay, with a greater impact on smaller TV than large ones. [12, 29]

Today, many techniques exist to hamper the effect of patient-specific parameters. Motion management techniques exist, such as DIBH, where patients hold their breath for around 20 seconds. [13] The Swedish national treatment guidelines for HL recommend that patients perform DIBH. [3] DIBH has the potential to reduce the magnitude of motion [5] and generally reduce doses to OARs compared to FB. [15] However, DIBH is not suitable for all patients. According to current guidelines, PT is still possible in FB as long as target motion is less than 5 mm along any of the Anterior-Posterior (AP), Superior-Inferior (SI) or Left-Right (LR) axes. [3]

2.6.2 Delivery characteristics

Delivery characteristics include e.g. spot sizes, spot spacing and rescanning techniques. [9] The initial energy, the distance between the patient and the range shifter or nozzle (air gap) and the possible use of a range shifter all influence the size of each spot. Large spot sizes reduce the impact of interplay since the risk of under-or over-shooting becomes less probable. Larger air gaps combined with a range shifter will increase the spot size and are often recommended to mitigate interplay. In addition, the treatment plan can be more resistant to interplay effects with smaller spot spacing. A smaller spot spacing increases the overlap between adjacent spots, which smears the dose deviations inside the TV. [5, 10, 16, 17] Moreover, rescanning or repainting techniques are other tools used for smoothing dose irregularities caused by interplay. [5, 31] One of the more straightforward but still effective repainting technique is volumetric rescanning. Here the beam is delivered across the entire TV twice or more for each fraction. The prescribed dose per fraction is the same; however, for N repaintings, the dose delivered during one painting will be scaled by 1/N. [20] Zeng et al. [20] showcased that volumetric repainting greatly minimised the influence of interplay for seven mediastinal HL patients. A similar trend was observed when rescanning techniques were applied to liver tumours affected by abdominal respiratory motion [28]

The potential risk with volumetric repainting is that spots might miss some phases during a

breath cycle due to coherence between each repainting and the breathing pattern. [5, 20] The treatment time during one fraction is not instant; instead, the delivery for a single fraction takes several minutes and individual paintings tenfold of seconds. Hence, a tumour might be positioned in many locations during a single fraction. [10, 20, 31] When all protons have been delivered in one particular spot position, the lateral magnets need time to transition to the subsequent spot position, commonly referred to as the spot settling time τ_{ss}. [10, 20] τ_{ss} is longer than the delivery time of individual spots in a position. A sufficient τ_{ss} is needed to ensure adequate dose delivery for each position. τ_{ss} is system dependent controlled by how the magnets function and spot position's dose accuracy. [10, 20] Overall, the total treatment time during a single fraction are dependent on the number of energy layers and the delivery time of beamlets for each energy layer, which in turn is dependent on the time properties of a particular spot position and the number of positions in the layer. As been mentioned, the interplay reduces for longer treatment times, which is correlated with a longer τ_{ss}. Notably, reducing the spot spacing implies more spots per energy layer, which increases the accumulated sum of τ_{ss} (total) per layer. [10] Furthermore, the time needed to switch between different energy layers, namely the energy switching time, τ_{es}, is of the order $\sim$ 1 s. [31] Thus, τ_{es} is more time comparable to a person's breath cycle (typically $3-5$ s). [10, 20, 31] The magnitude of τ_{es} makes it theoretically possible for the beam scanning to be synchronised with the respiratory motion. However, since breathing periods varies from cycle to cycle and the total τ_{ss} varies over time per iso-energy layer, τ_{es} will not affect the dose homogeneity inside the TV, regardless of the magnitude of τ_{es}. [20] Lastly, the temporal characteristics of PBS are one explanation to why smaller TV are more affected by interplay. Larger targets need more spots to be delivered in each layer, thus increasing the delivery time, which populates spots more evenly over the entire breathing period. [29]

2.6.3 Plan characteristics

Plan characteristics involve parameters such as field orientations or field optimisation. [9] In Single-Field Optimised (SFO) plans, each field is optimised to deliver the same dose to the target region. [32] In contrast, in Multiple-Field Optimised (MFO) plans, each field is optimised simultaneously to achieve a more conform dose distribution. Thus, MFO plans can generally achieve better OARs sparing than SFO plans; however, as a compromise, MFO techniques are less robust and more susceptible to uncertainties such as the interplay effect. [5, 16]

2.7 Motion simulation

Simulating tumour movement has been a great focus over several decades. Lujan et al. [33] and Seppenwolde et al. [34] proposed a simplified model that characterises tumour motion by a periodic sine or cosine wave. The periodic function follows the motion of a single point, which is usually the Centre Of Mass (COM) of the tumour. [7, 12, 33–36] The underlying assumption is that the tumour or parts of the tumour move due to the respiratory motion and no other external force. Eq. 1 describes the most important parts of periodic function:

$$S(t) = A \sin\left(\frac{\pi \cdot t}{T} + \phi\right)^{2n} - A/2 \quad (1)$$

Here, A represents the peak-to-peak target motion amplitude for the tumour's motion in the LR, AP and SI direction, respectively. T is the respiratory cycle period, and n is a fitting parameter to model asymmetrical breathing patterns. n is a positive integer, and a larger value corresponds to more time spent during expiration than inspiration. [31] White et al. [37] observed from a patient cohort of 50 people that most patients (68%) spent more time in exhalation than inhalation, suggesting a fitting parameter larger than 1 for most patients. Lastly, ϕ is the starting phase of the motion curve, i.e. ϕ will determine the starting point of the motion when beam delivery begins. [31] The phase parameter will greatly influence the magnitude of interplay. [10, 13, 31] Generally, distal energy layers contribute to most of the spots during delivery; thus, changing the initial phase may deliver distal layered spots when the TV is out of position. [10] Despite this, due to the fluctuation of starting phase, this parameter should have a negligible dosimetric impact on a fractionated treatment. [33] For instance, without gating techniques such as DIBH, each measurement will be performed during random time points, and each phase during one breath-hold is equally probable. [20, 33], which will average the homogeneity loss due to phase differences throughout treatment.

3 Material and Methods

3.1 The proton beam therapy system

The measurements were performed at the Swedish proton facility, the Skandion clinic in Uppsala. The proton centre (Ion Beam Application S.A., Louvainla-Neuve, Walloon Brabant, Belgium; IBA) has two gantries and can deliver pencil beams with energies between $60 - 226$ MeV. Each gantry has a 3.5 cm range shifter, which can be removed from the beam path when needed. The snout position can be adjusted by changing the position of the range shifter closer or further away from the patient, which enables different air gaps between the patient and the snout. This study was performed at gantry 1 with the range shifter equipped. The pencil beams are delivered in a square spot grid pattern for each energy layer, where each scan plane is perpendicular to the beam axis.

3.2 Patient selection

Three mediastinal HL patients were selected. The CTV tumour volumes were 152, 232 and 336 cubic centimetre (cc) for patient 1, 2 and 3, respectively.

3.3 Treatment plans

The clinical treatment plans were acquired from Eclipse's treatment planning system (Varian Medical Systems, Palo Alto, CA, 15.6). All three patients' treatment plans needed to be re-optimised in a FB setting since the original treatment plans were based on DIBH. Organ and tumour structures

had already been delineated in a FB CT-set for patient 1, but structures in patient 2 and 3 needed to be re-drawn to represent a FB anatomical setting. The re-delineation was accomplished by applying Eclipse's rigid and deformable image registration tools between the CT-set in DIBH and the CT-set in FB, thus rearranging the anatomy from DIBH to FB. If the structure changes were deemed unrealistic, structures were corrected using Eclipse's contouring tools.

Secondly, the treatment plans were re-optimised according to the clinical practice set by Skåne university hospital. The treatment plans had a prescribed dose of 29.75 Gy (in Relative Biological Effectiveness, RBE dose) divided in 17 fractions and a constant RBE of 1.1 for protons was used. Hence the prescribed physical dose per fraction was approximately 1.591 Gy. SFO, a 3.5 cm range shifter and a 5 mm spot spacing were all employed in the treatment plans. In addition, two anterior fields with a small angle in between and an equal weight of 0.5 were used to accommodate an almost volumetric repainting effect. In the end, the treatment plans were clinically acceptable and robust since they met the hospital's clinical objectives and robustness criteria.

Lastly, the treatment plans were re-calculated in Water Equivalent Tissue (WET) values according to clinical procedures at the hospital. We analysed the treatment plans based on the dose distribution in water by investigating the lateral dose profiles at different water depths. The dose distribution was measured at two depths of the TV for each patient, which is referred to as *target depths*. In table 1 some of the treatment plan characteristic's are summarised, including the target depths. There is also an appendix presenting the lateral dose profiles for each target depth, shown in figures 25, 26 and 27.

Table 1: Plan characteristics of the dose distribution based on WET values for the patients included in the study. The range parameter describes the maximum distance to the 95% dose level. The SOBP parameter is the distance between the min and max point of SOBP in the target region. Energy layers and the number of spots present field characteristics (F1/F2). Lastly, the target depth parameter refers to which depth the dose distribution was measured.

Patient	Range [cm]	SOBP [cm]	Energy layers F1/F2	# Spots F1/F2	Target depths [cm]
1	7.8	6.6	29 / 28	4564 / 4234	4.5 / 6.0
2	7.7	7.0	34 / 39	5756 / 5506	2.6 / 4.2
3	10.9	10.9	42 / 41	7135 / 4233	4.7 / 7.0

3.4 The experimental set-up

Breathing motion was simulated with eq. 1 using a modified version of the Hexamotion motion platform (ScandiDos Inc., Uppsala, Sweden). The Hexamotion device has six degrees of freedom; however, only translation along the SI axis was utilised. The motion platform was modified by mounting a MatriXX 2D ionisation chamber array detector (IBA Dosimetry, Schwarzenbruck, Germany) on top of the motion platform. The MatriXX detector consists of a $32x32$ array of small ionisation chambers with 1020 detectors (coroner detectors not present), separated by 7.6 mm. [38] The dose measured in each ion chamber contribute to the overall dose distribution in a horizontal

plane. The target depths presented in table 1 were simulated by allocating different amounts of solid water plates on top of the MatriXX detector. Each patient was delivered with a gantry angle of 0°, 90° relative to the MatriXX detector due to the angular dependency of the MatriXX detector. A snout position of 30 cm was employed in all treatment plans. This detector set-up corresponded to an air gap of approximately 21 cm. Lastly, since the motion platform needed to be modified to fit our experimental set-up, a heavy cylinder was put on the half of the MatriXX where no ion chambers were located to stabilise the set-up. A 3D version of the set-up is described in figure 4, illustrating the Hexamotion platform and MatriXX detector. In figure 28 in the appendix, a photo of the set-up is also presented.

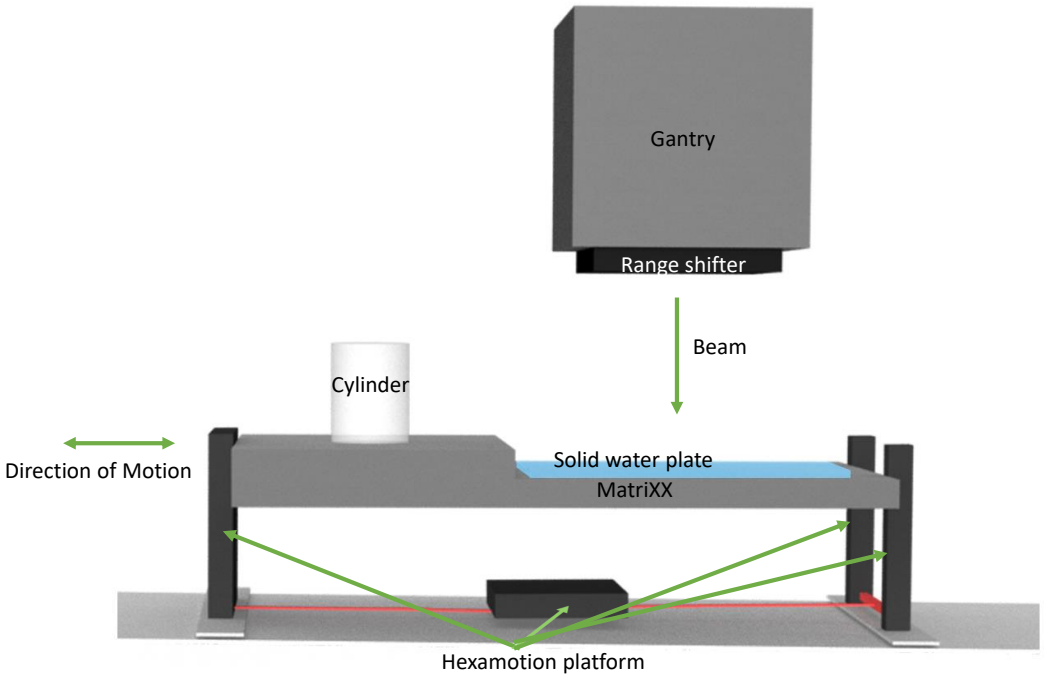


Figure 4: The experimental set-up with the MatriXX detector and the Hexamotion platform. The solid water plates simulated target depths (table 1). At all times, the Matrixx detector was aligned 90° to the gantry, located at 0°. The air gap, i.e., the distance between the detector and the 3.5 cm range shifter was 21 cm for most measurements. The MatriXX detector was restricted to motion in the SI-direction only. A cylinder was placed on the MatriXX to stabilise the motion path.

Breathing motion was simulated with the Hexamotion platform and by choosing appropriate values for the target motion amplitude A, the period time T and the fitting parameter n in eq. 1. In the majority of all measurements, a breathing period of 5 s was used, based on respiratory traces presented in several other studies. [10, 20, 33, 34] In some measurements a 2.5 s breathing period was also used to compare results obtained for a 5s period. Furthermore, we chose the fitting parameter $n = 2$ based on previous findings in the literature. [31, 34, 37] The motion regularity of any motion trace is summarised on the form *Asym* or *Sym*. Asym corresponds to an asymmetrical sine curve in the form of eq. 1 with $n = 2$, while Sym means a pure symmetrical sine wave. We disregarded the influence of the starting phase ϕ parameter, which means that each

measurement will begin at a random point somewhere along the sine curve. Lastly, to assess and verify typical peak-to-peak motion amplitudes A, the CTV motion of three additional HL patients were evaluated in Eclipse using the 4D-CT data of the patients. The tumour motion was assumed to correlate with the CTV's COM variations in AP, SI and LR-direction. Evaluation of the three 4D CT data sets showed that changes of COM had a mean peak-to-peak motion amplitude of 2 mm, 4 mm and 1 mm for the AP, SI and LR-direction, respectively, which are comparable to seven other HL patients motion amplitudes presented in [20]. From both examinations, motion along the SI direction is the greatest, and many authors have also confirmed this for several other tumour types. [31, 34, 39, 40] Current guidelines for HL restricts PT for motion amplitudes larger than 5 mm in any direction. Although non of the ten patients analysed in this section had an SI-motion greater than 5 mm, breathing effects should be quantified for higher amplitudes. Based on this, we investigated 5, 10 and 15 mm target motion amplitudes in SI-direction only. One of the motion traces to mimic the tumour motion is illustrated in figure 5.

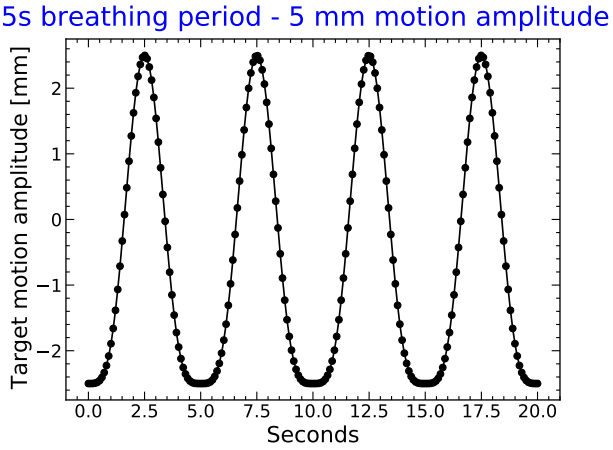


Figure 5: One of the motion traces that was used to evaluate breathing effects. This motion trace corresponds to an asymmetrical sine curve with a 5 mm motion amplitude (peak-to-peak) and a 5 s breathing period (eq. 1).

3.5 Series of measurements

In total, six different types of motions patterns were simulated, i.e. 6 different motion traces with a unique set of parameters described by eq. 1. The impact of breathing was estimated by comparing measured dose distributions under motion D_{motion} and corresponding static conditions D_{static}. 2D gamma evaluations were computed between static dose distributions measured for a particular target depth to validate and verify the consistency of the experimental set-up.

Depending on the patient, different series of measurements were performed. To simplify, flow diagrams are presented for the various patients, shown in figure 6, 7 and 8. Patient 1 was the only patient that had an organ and CTV structures delineated in a FB CT-set. Therefore these structures were deemed to be more reliable than the structures of the other two patients. Moreover, out of the two depths, the regional dose distribution at 4.5 cm was spread over a wider region in WET

than the target region at 6 cm depth. Thus, we assumed that the dose contribution at 4.5 cm depth would better represent the CTV region than the contribution at 6 cm depth. Therefore, several more measurements were performed at 4.5 cm target depth than at 6 cm. For instance, the snout position and spot spacing were adjusted for some measurements at target depth 4.5 cm (figure 7). The snout position was lowered from 30 cm to 18 cm, corresponding to a decrease in air gap from 21 cm to 9 cm, while the spot spacing was changed from 5 mm to 3 mm. N.b. the treatment plans had to be re-optimised when one of the two beam delivery parameters were changed.

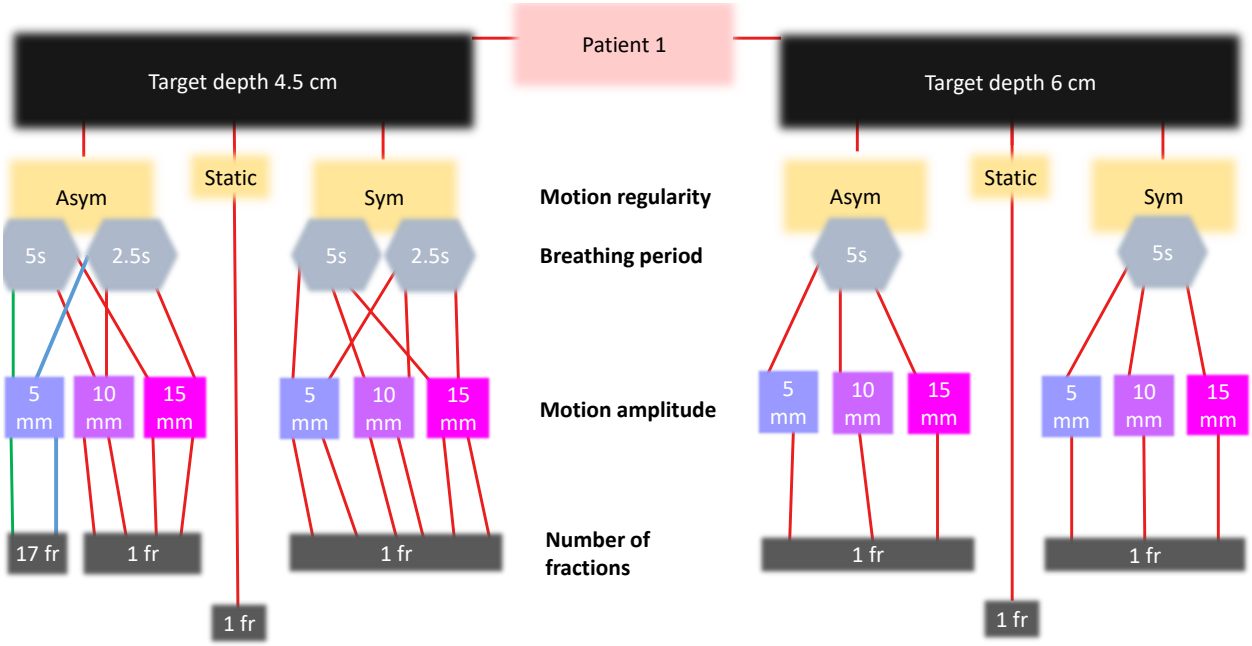


Figure 6: Flow diagram of measurement for patient 1. For instance, one motion trace has motion regularity "Asym", breathing period "5s" and target motion amplitude "5mm" since this describes one possible breathing pattern of eq. 1. The asymmetrical motion trace with 5 mm motion amplitude was measured in 17 fractions for two different breathing periods. All other measurements were only measured once. N.b, the "static" rectangle corresponds to one of the static dose distributions that were measured for a particular target depth.

3.5.1 Dose blurring

The dose blurring effect was estimated at 4.5 cm target depth for all motion amplitudes. Dose blurring was simulated by moving the patient couch in small (mm) steps along the SI-direction following the motion curve. A static dose distribution was measured for every step but with a different offset from the starting position. For instance, with the 5 mm target motion amplitude, the following steps along the SI-direction were used: $-2.5, -2, -1, 0, 1, 2, 2.5, 2, 1, 0, -1, -2$ for a motion trace starting at minimum peak amplitude (-2.5). These 12 offsets correspond to a unique dose distribution. The mean value of all dose distributions was computed to estimate a static dose

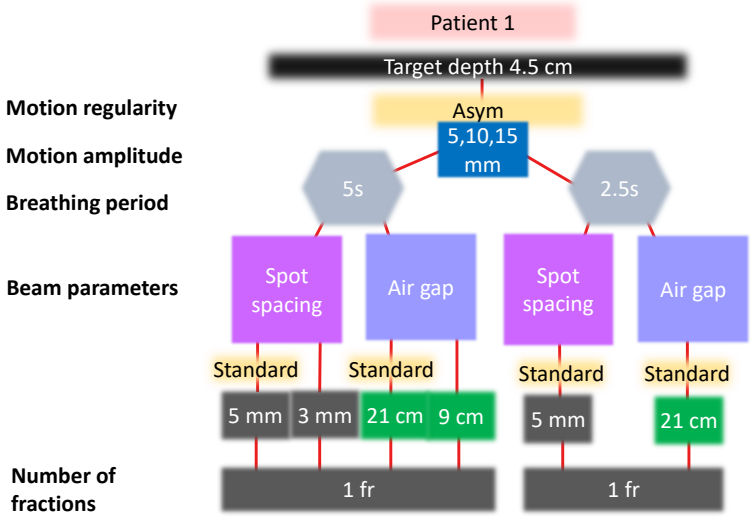


Figure 7: An extension of the flow diagram in figure 6 for target depth 4.5 cm and the asymmetrical motion trace. Additional measurements were performed by changing the spot spacing or air gap for the 5*s* breathing period. For a constant amplitude of either 5, 10 or 15 mm, the snout position was lowered from 30 cm to 18 cm, corresponding to a decrease in air gap from 21 cm to 9 cm, while the spot spacing was changed from 5 mm to 3 mm. The "Standard" rectangle notes what is currently used in all clinical treatment plans. Only one parameter was changed at a time, e.g. when 3 mm spot spacing was investigated, the air gap remained at 21 cm according to clinical settings.

distribution, including dose blurring. In this way, the sole impact of the interplay effect on the target dose could be determined since the static measurements had incorporated dose blurring.

3.6 Evaluation metrics and dosimetric analysis

The measured dose distributions, 32*x*32 pixel arrays, were evaluated individually per field, but for most cases, as the sum of dose distribution of both fields, corresponding to a dose distribution of a single fraction.

3.6.1 Dose mask

Before any analysis could be performed on the measured dose distributions, the target region needed to be re-defined. During the treatment planning process, the tumour location is lost when converting the CT values to WET, and the only information left is dose values. Therefore, "to find and isolate" the actual region of the CTV, a dose mask was constructed by only considering the high-dose area of the measured dose distribution. The mask was created such that only values above the cut-off dose 1511 mGy (95% of the actual prescribed dose) were considered, while

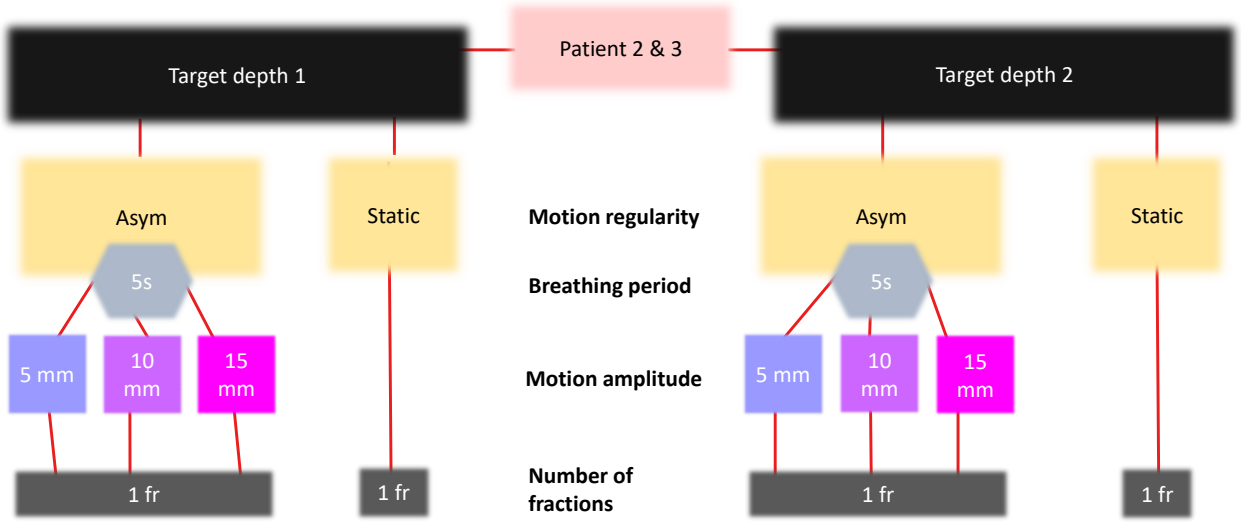


Figure 8: Flow diagram of measurements performed on patient 2 and 3. The target depths are found in table 1. Only asymmetrical motion traces were measured using a single breathing period of 5 s and three different motion amplitudes. Motion dose distributions were measured once and compared to a static dose distribution measurement.

dose values outside this limit were set to zero. This computation is visualised in figure 9, where the measured dose distribution (figure 9 left) is reduced, and the dose mask (figure 9 middle) is formed. However, no analysis was based on the dose mask. Instead, all the results depend on an eroded dose mask (figure 9 right). The eroded dose mask is formed by removing one pixel (7.6 mm) in every direction around the dose mask through erosion, which is an approximation to isolate the target area from its margins.

The eroded dose mask was created for every target depth's static dose distribution. The same eroded dose mask is thereafter assigned to the corresponding dose distribution under motion for each target depth.

3.6.2 Relative dose differences

Relative dose differences between the eroded static and motion dose distribution was computed according eq. 2:

$$D_{rel} = \left(\frac{D_{motion} - D_{static}}{D_{static}} \right) \cdot 100 \tag{2}$$

where D_{motion} and D_{static} are the eroded motion and static dose distributions, respectively. D_{static} and D_{motion} will be $32x32$ pixel arrays, hence D_{rel} will also be a $32x32$ array whose pix-

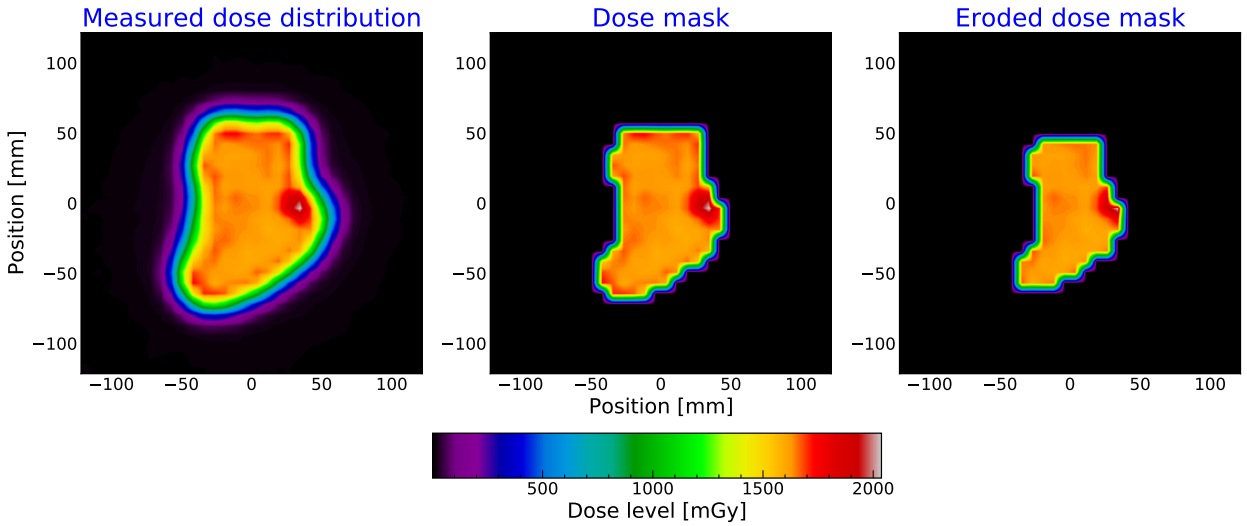


Figure 9: This figure describes how the eroded dose mask was constructed and the intermediate steps with an example. The figure visualises a static dose distribution of patient 1 at target depth 4.5 cm. The data matrices in each of the sub-figures above have been interpolated using the bilinear interpolation parameter before the image is displayed with the imshow function using the Matplotlib library (version 3.3.4) in Python. [41] At this target depth, the dose distribution has a high dose region, which is also highlighted in the lateral dose profile in figure 25.

els corresponds to the relative dose differences. When we moved the patient couch to estimate the extent of dose blurring, the static term in eq. 2 is replaced with a corresponding static term including dose blurring $D_{Static,blur}$. In the study, D_{rel} was used to assess breathing effects qualitatively. For instance, an eroded static and motion dose distribution for patient 1 and 4.5 cm target depth are compared in figure 10.

For the case of multiple fractions, D_{rel} are obtained through computation of accumulated static and motion dose distributions. E.g for fraction 3 ("fr3"), the motion dose distribution D_{motion} is the sum: $D_{motion} = D_{motion,fr1} + D_{motion,fr2} + D_{motion,fr3}$. Meanwhile, the static dose distribution D_{static} is $D_{static,fr3} = 3 \cdot D_{static}$, i.e. a multiple of the current fraction number, since static measurements was deemed to be similar for each target depth. Only after e.g. $D_{motion,fr3}$ and $D_{static,fr3}$ have been computed the eroded dose mask was created and D_{rel} was computed.

3.6.3 98th percentile value

The absolute value of D_{rel} was considered to quantify breathing effects, only accounting for relative dose differences corresponding to the eroded dose mask region. A 98th percentile was computed of the absolute relative dose differences, i.e. a 98th percentile of $|D_{rel}|$. The 98th percentile ($\delta D98$) gives an upper bound estimate of the breathing effects and was used to compare different measurements with each other.

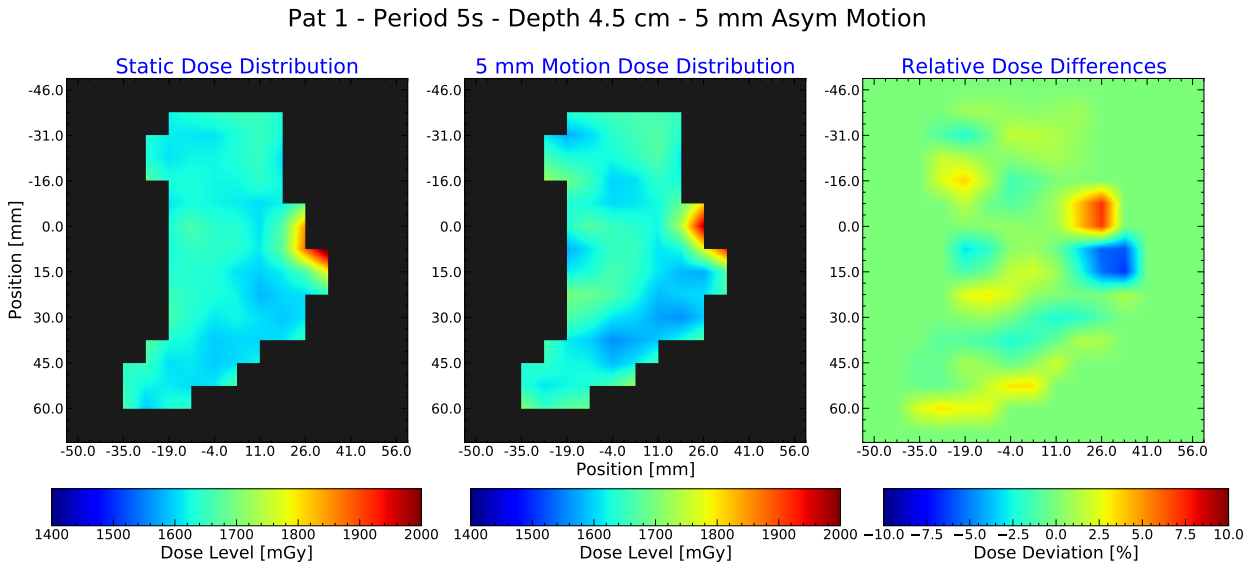


Figure 10: The three sub-figures show eroded dose distributions and relative dose differences D_{rel} for patient 1 and target depth 4.5 cm. Left: The static dose distribution. Middle: Motion dose distribution. The motion trace has an asymmetrical pattern (Asym) with a 5 s breathing period and 5 mm target motion amplitude. Right: Relative dose differences computed according to eq. 2. A positive relative dose difference (red/yellow) corresponds to a larger value in the motion dose distribution than in the static dose distribution and vice versa. The data matrices in each of the sub-figures above have been interpolated using the bilinear interpolation parameter before the image is displayed with the imshow function using the Matplotlib library (version 3.3.4) in Python. [41]

3.6.4 Other evaluation tools

Python (version 3.6.3) was used to form the eroded dose mask region and compute relative dose differences and $\delta D98$ values. The python script also performed 2D gamma evaluations between the eroded static and motion dose distributions. The gamma evaluation reveals large relative dose differences (high maximum and minimum values) within the eroded dose mask region. Mainly the *PyMedPhys* library (version 0.36.1) was used to perform the gamma evaluation. [42] PyMedPhys was validated by comparing pass rates computed for several static dose measurements with the *myQA software* (IBA Dosimetry). [43] For all gamma evaluations, a local dose criteria with a low dose cut-off of 10% was used. Lastly, statistical testing was performed to find a significant difference between two data sets for some cases. The non-parametric Mann–Whitney U-test (U-test) and Spearman's correlation test were used. The U-test's criteria are satisfied since all data are independent and have the same but unknown underlying distribution, i.e. the data is independent and identically distributed (idd). The samples are relatively small (less than 20), which validates the use of a non-parametric test. [44] For such data, Spearman's correlation can also be utilised. Spearman's correlation test indicates whether there is a monotonic relationship between two data sets or whether the data sets are independent. Spearman coefficient of $\rho = \pm 1$ highlights a perfect monotonic behaviour between the data, while $\rho = 0$ describes no association. [45]

3.7 Time structure analysis

Treatment log files of patients 1, 2 and 3 were analysed to reveal the beam delivery's time structure and spot characteristics. The treatment log file contains information about each energy layer's start and stop time, including the total number of spots and the total MU at each energy layer. By analysing the start and stop time for each layer, τ_{es} and τ_{ss} values could be deduced and plotted in combination with the spot characteristics for every energy layer.

4 Results

4.1 Consistency of the experimental set-up

The experimental set-up was deemed stable through gamma evaluation of measured static dose distributions for different target depths. The overall MatriXX position uncertainty was determined to be $0.5 - 1$ mm by assessing pass rate values for a 0.5%/0.5 mm and a 1%/1 mm criteria. For the 0.5%/0.5 mm criteria, the dose agreement dropped below 90% for some gamma evaluations done for patient 1. The disagreement was limited to the borders of the dose distribution; thus, the eroded dose mask region would have had higher pass rates.

4.2 Relative dose differences

In figure 11 and 12 relative dose differences are shown for two patients and one of their target depths. For figure 11, a static measurement at target depth 4.5 cm is compared to motion dose distributions for three different target motion amplitudes with an asymmetrical motion pattern and a 5 s breathing period. Similarly is done for patient 3 at 7 cm target depth (figure 12). Figure 11 visualise the worst-case scenario and relative difference values up to $\pm$20% are observed at the 15 mm motion amplitude. Both dose blurring and interplay effects contribute to the relative dose differences, but the interplay effect manifests itself as dense red (hot) and blue (cold) spots in the figure. In comparison, breathing effects in figure 12 are less prominent. Figure 12 is one of the best-case scenarios and relative dose values are less than $\pm$10% for the 15 mm motion amplitude. Overall, it is evident in both figures that the breathing effects become more pronounced with higher amplitudes.

4.3 Multiple fractions

This section presents the $\delta D98$ for the 17 fractions measured for patient 1 at 4.5 cm target depth. Two motion traces were simulated. Both breathing patterns were asymmetrical with a 5 mm motion target amplitude and two different period times 2.5 and 5 s, respectively. For both breathing periods, breathing effects (interplay and dose blurring) and only interplay were investigated.

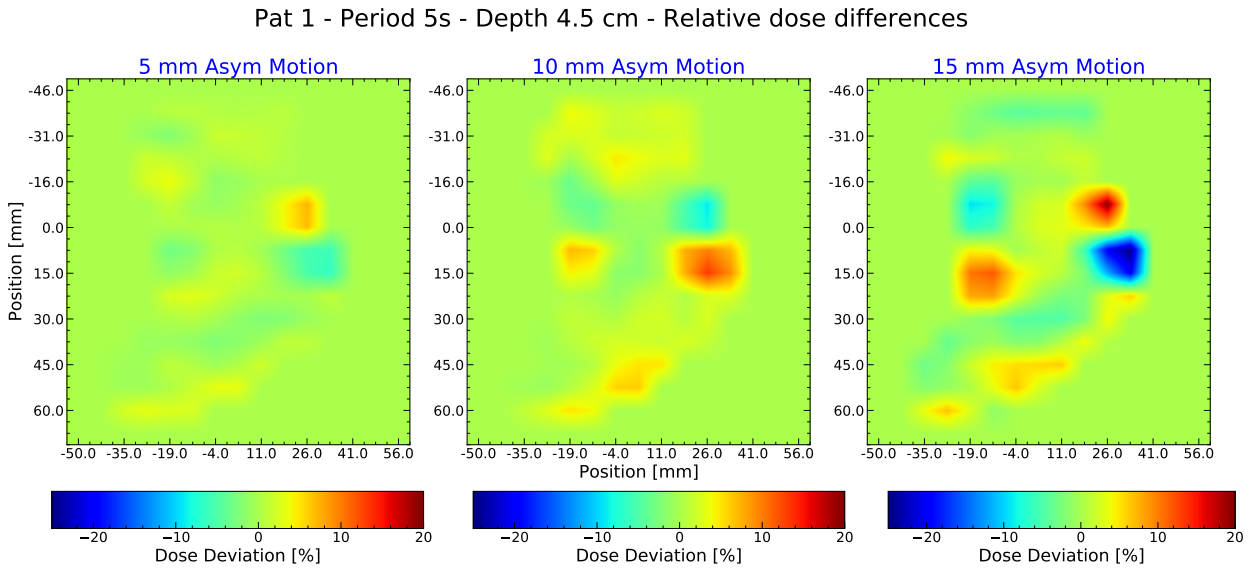


Figure 11: One example of relative dose differences for patient 1 and target depth 4.5 cm. This figure compares a static dose distribution to three motion dose distributions influenced by asymmetrical motion traces (Asym) with varying amplitudes and a 5 s breathing period. Note that a positive relative dose value (yellow/red) corresponds to a larger value in the motion dose distribution than in the static dose distribution and vice versa. The colour bar encloses the largest dose difference observed for any of the measurements in this study. Both dose blurring and interplay effects are considered. Note that the data matrices in each subfigure are interpolated using the bilinear interpolation parameter before being displayed with the imshow function using the Matplotlib library (version 3.3.4) in Python.

4.3.1 Individual fractions

The inter-fractional variation of $\delta D98$ are presented in figure 13. In this figure, both dose blurring and interplay are considered. No general trend exists among the $\delta D98$ values, suggesting that the impact of breathing is random between fractions.

It is evident that the $\delta D98$ values plotted in figure 13 are influenced by the random starting phase for each measurement. This is further illustrated in figure 14 where different relative dose difference plots are shown similarly as in figures 11 and 12, but in this case for the same patient, target depth, motion amplitude and regularity (patient 1, target depth 4.5 cm, with a 5 mm asymmetrical target motion amplitude). The only difference between subplot 14a and 14b is the breathing period (2.5 and 5 s). Both dose blurring and interplay are considered.

4.3.2 Accumulating fractions

The $\delta D98$ value decreases for a higher number of accumulated fractions (figure 15), where breathing effects and interplay effects are separated for two period times. Despite the large breathing effects observed for single-fraction measurements, breathing effects cancel out throughout a fractionated treatment. Although there is no significant difference between measurements of different period times (U-test, p=1.0), they do not show the same behaviour since the $\delta D98$ value increases

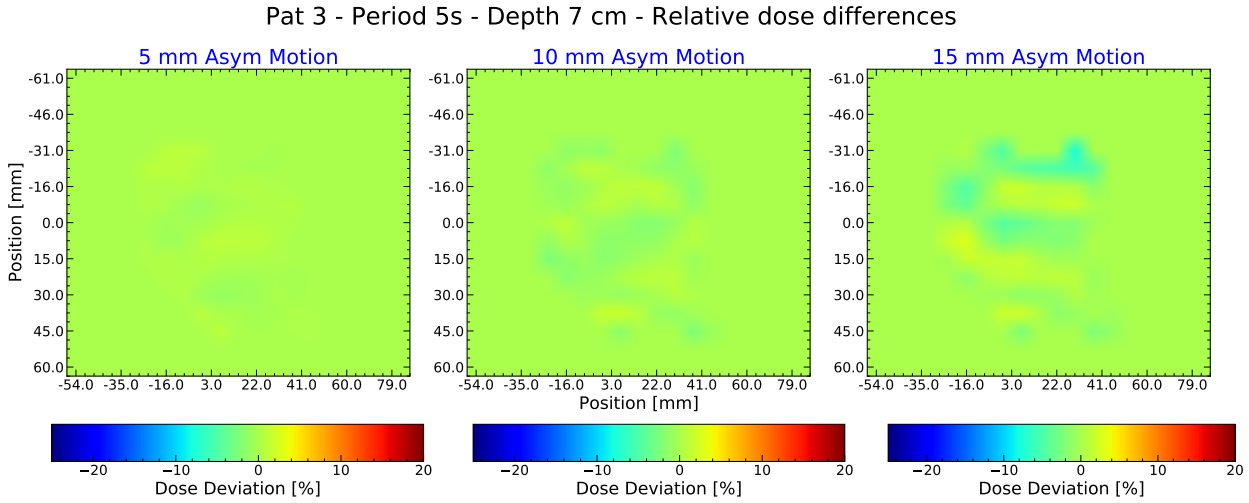


Figure 12: Similar plots to figure 11, but this figure illustrates relative dose differences of patient 3 at target depth 7 cm. The static dose distribution is compared to corresponding dose distributions influenced by motion with varying amplitude, a 5 s breathing period and asymmetrical motion regularity (Asym). The colour bar encloses the largest dose difference observed for any of the measurements in this study. Positive relative dose values correspond to larger values in the motion dose distribution than in the static dose distribution and vice versa. Both dose blurring and interplay effects are considered. The data matrices in each sub-figure are interpolated using the bilinear interpolation parameter before being displayed with the imshow function using the Matplotlib library (version 3.3.4) in Python.

after fraction 5 for measurements corresponding to a 5s breathing period. Furthermore, there is no difference when removing the contribution of dose blurring, which indicates that dose blurring has a non-significant (U-test, p=1.0) impact on target dose distribution for this particular motion trace, patient and field configuration.

As observed in figure 15, $\delta D98$ values does decrease when accumulating fractions. Lower $\delta D98$ values represent a smaller spread of relative dose differences inside the dose region (less hot and cold spots). The spread of relative dose differences can be visualised by plotting a histogram demonstrating the frequency of pixels corresponding to a particular relative dose difference. Sub-figures 16a and 16b represent individual histograms for different accumulated fractions corresponding to measurements with period times 2.5 and 5 s, respectively, including both dose blurring and interplay effects. As the number of fractions increases, the spread converges inside the two red horizontal lines, i.e. into the $\pm 3\%$ dose limit. The *ratio* value presented for each histogram in figure 16 describes the number of pixels with relative dose differences below $\pm 3\%$ relative to the total number of pixels inside the eroded dose mask. For an increasing number of fractions, the ratio converges to 1, equivalent to a smaller spread of relative dose differences. A tabular summary of the ratio values for all cumulative measurements shown in figure 15 are presented in table 2. Similar to $\delta D98$ values, the ratio values showcase that dose blurring has a negligible effect on the dose distribution since ratio values for interplay and dose blurring are similar to ratio values when interplay is only accounted for.

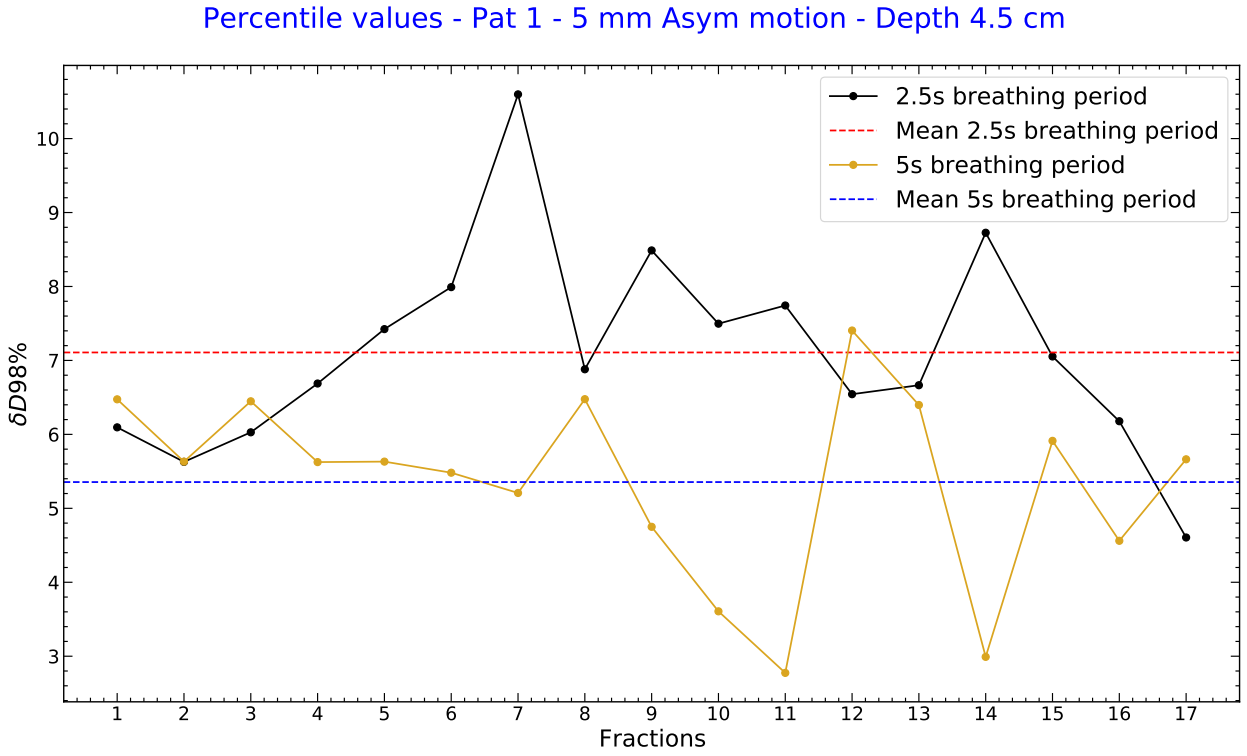


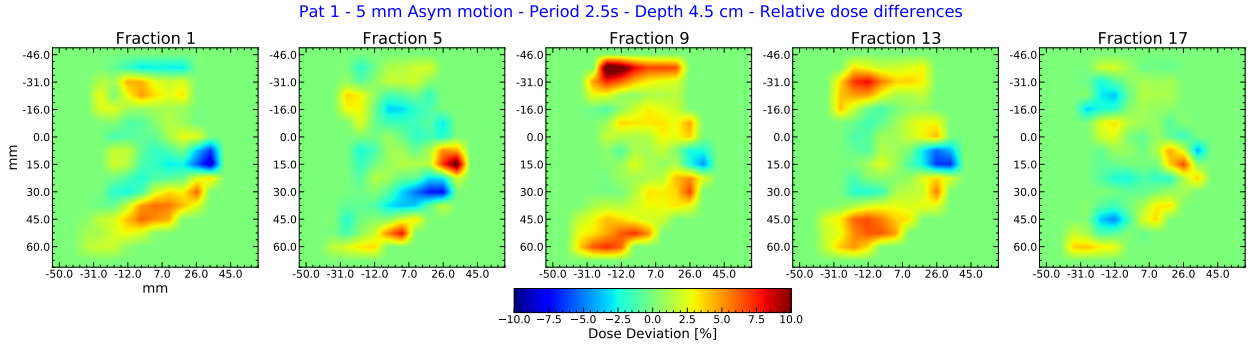
Figure 13: Variation of $\delta D98\%$ value over different single-fractions measurement for patient 1 at target depth 4.5 cm. Motion traces are similar with an asymmetrical breathing pattern and a 5 mm motion amplitude but have different breathing periods. Both dose blurring and interplay are considered.

Table 2: Ratio values for patient 1 at target depth 4.5 cm. The ratio value represents the number of pixels with relative dose differences below $\pm 3\%$ relative to the total number of pixels inside the eroded dose mask. Ratio values are computed for measurements presented in figure 15, i.e., using motion traces with an asymmetrical breathing pattern, a 5 mm motion amplitude and with either 2.5 or 5 s breathing period. Values are separated between interplay and dose blurring (total impact of breathing) and interplay only.

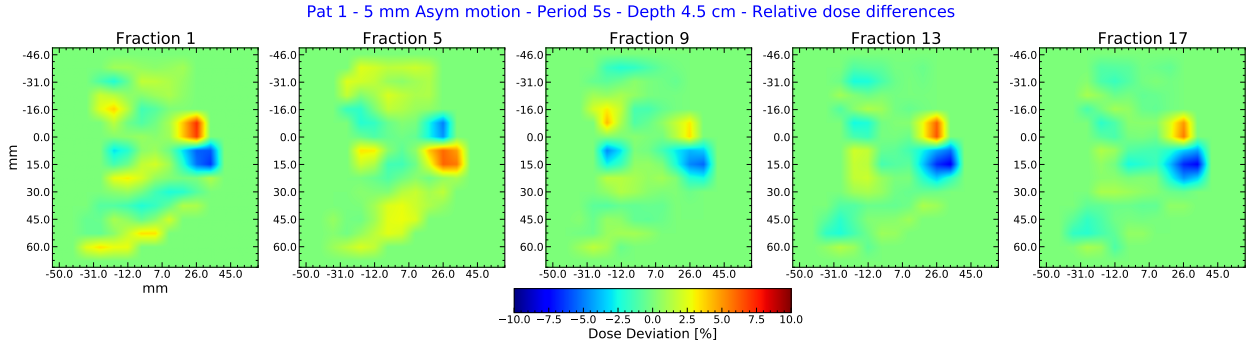
Fraction	**Ratio**			
	Interplay & dose blurring		Interplay Only	
	2.5s breathing period	5s breathing period	2.5s breathing period	5s breathing period
1	0.79	0.86	0.78	0.86
5	0.90	1.00	0.91	1.00
9	0.98	0.97	0.98	0.97
13	0.97	0.97	0.95	0.96
17	1.00	0.96	1.00	0.95

4.3.3 Individual fractions and fields

Field 1 and 2 are analysed individually in figure 17 corresponding to measurements with period times 2.5 and 5 s, respectively. Larger dose differences are present than in figure 13 where the sum of both fields was analysed. Thus, per single fraction, the volumetric repainting effect does



((a)) Relative dose difference plots for a 2.5s breathing period.



((b)) Relative dose difference plots for a 5s breathing period.

Figure 14: Figures of relative dose differences of patient 1 and target depth 4.5 cm for five different fractions (measurements). Motion traces have an asymmetrical breathing pattern with a 5 mm motion amplitude and different breathing periods. Both dose blurring and interplay are considered. The data matrices in each subfigure are interpolated using the bilinear interpolation parameter before being displayed with the imshow function using the Matplotlib library (version 3.3.4) in Python.

mitigate breathing effects to some degree. Notably, field 1 has significantly greater $\delta D98$ values than field 2 for both breathing periods (U-test with p-value 0.001 with period times 2.5 and 5 s, respectively).

4.3.4 Accumulating fractions for different fields

The cumulative effect for each field is illustrated in figure 18. All measurements but measurements associated with period time 5 s and field 1 converge to a minimum value after 10+ fractions. Instead, the $\delta D98$ value associated with the 5 s breathing period and field 1 increases after fraction 5.

The behaviour depicted in figure 18 were analysed further with gamma evaluations. Gamma distributions were computed based on static and motion dose distributions for each field for different accumulated fractions using a passing criteria of 3%/0.01% mm. The computed pass rates were deemed reliable since the mean deviation of pass rates between the myQA software and the PyMedPhys library was 0.2%. In figure 19, gamma evaluations on static and motion dose distri-

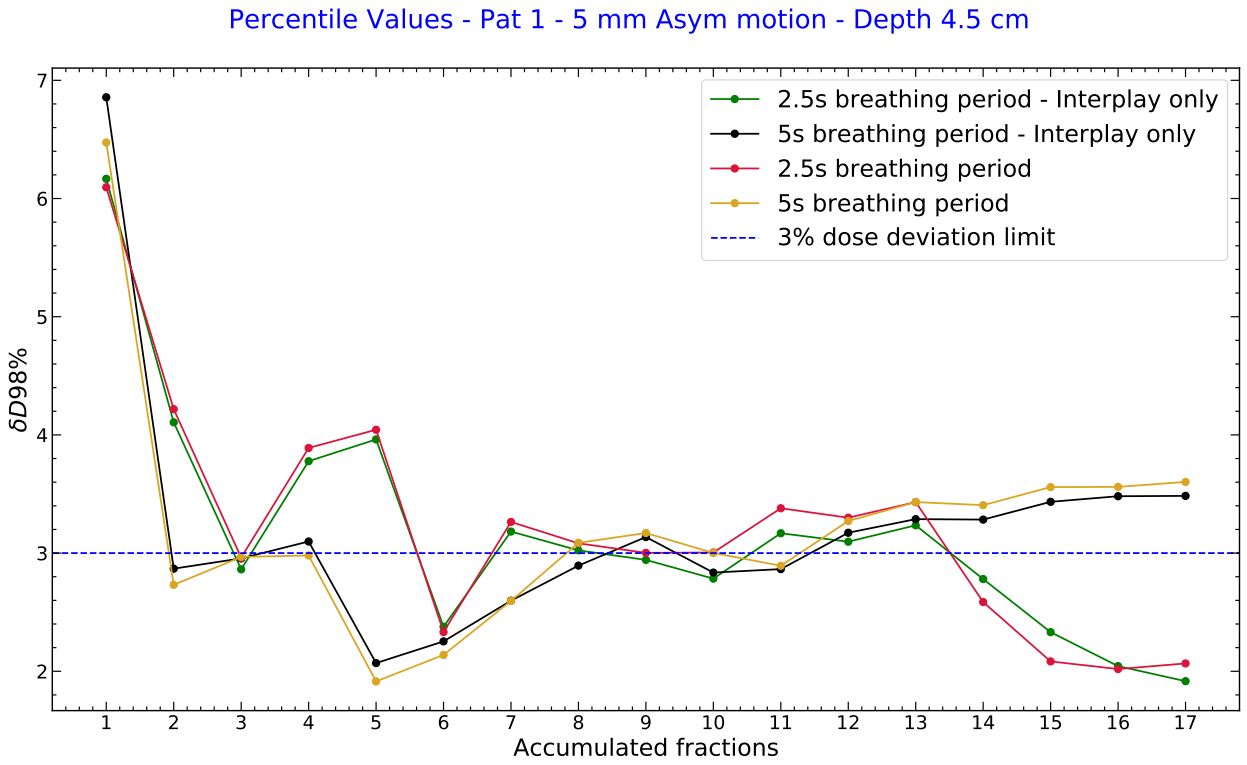
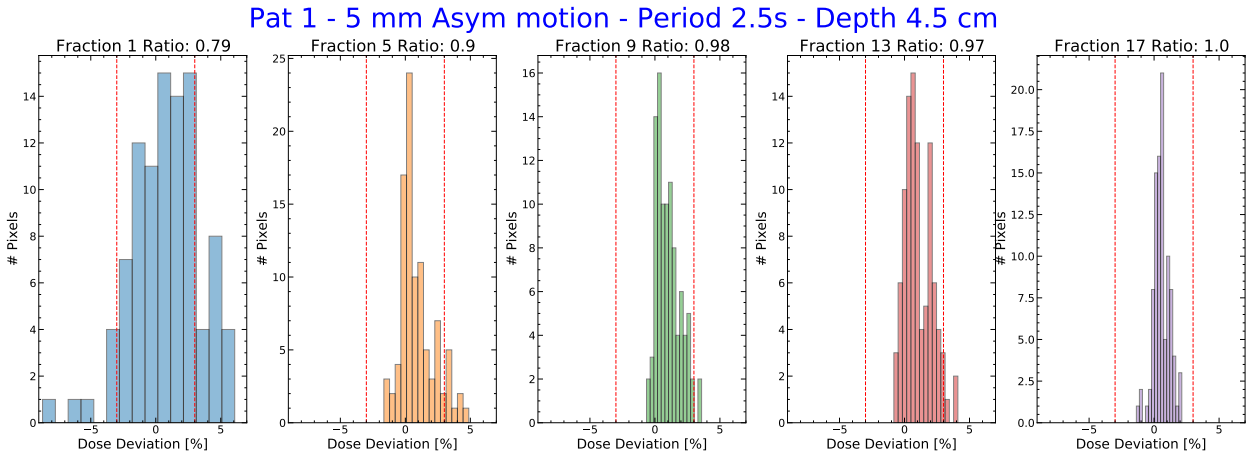


Figure 15: $\delta D98$ value as a function of accumulated fractions for patient 1 at target depth 4.5 cm. Asymmetrical breathing patterns with a 5 mm motion amplitude were utilised with either a 2.5 or a 5 s breathing period. Measurements "Interplay only" solely depend on interplay, while the other two cases include both dose blurring and interplay.

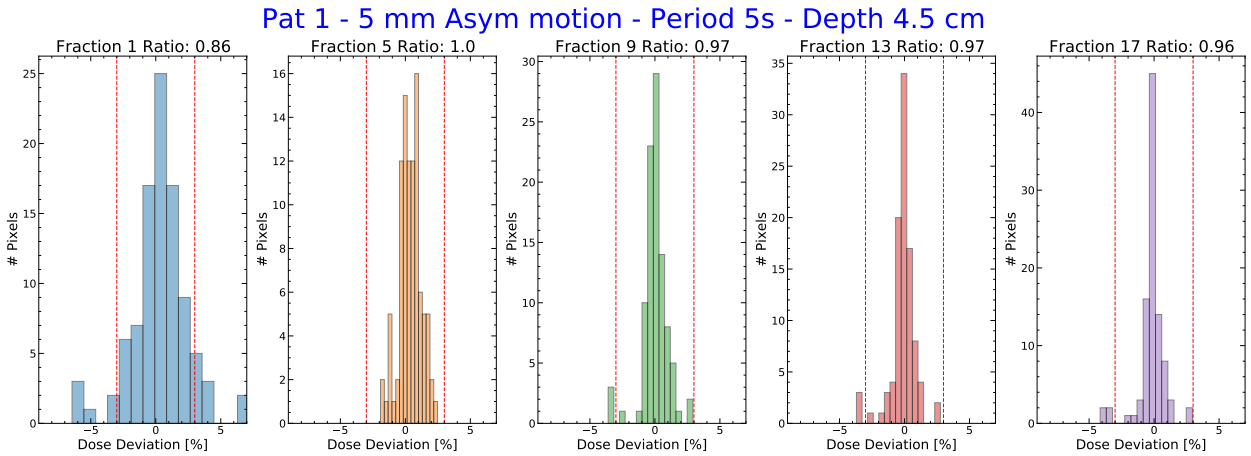
butions are illustrated for field 1 and 2 for different accumulated fractions with the period time 2.5 s. Similarly, the same gamma evaluations are computed for the 5 s breathing period (figure 20). It becomes clear that the greatest dose differences are present for field 1, which agrees with the result illustrated in figure 17. Notably, for the measurement associated with a 5 s period time and field 1, the greatest dose differences are restricted to a small region in the dose distribution for later fractions. This region is a high dose region within the eroded dose mask (figure 9), which is also illustrated in figure 25 demonstrating the lateral dose profile for patient 1 at target depth 4.5 cm. However, the impact of the high dose area is less severe with a period time of 2.5 s.

4.4 Treatment characteristics

In this section, additional $\delta D98$ values for patient 1 at target depth 4.5 cm are presented. Three different asymmetrical breathing patterns were used with varying target motion amplitudes and breathing periods. Moreover, two beam delivery parameters were adjusted, using a 9 cm air gap instead of 21 cm and the spot spacing was changed from 5 to 3 mm. The result of the measurements is presented in figure 21. Figure 21 indicates that the $\delta D98$ value increases for higher motion amplitudes, with a large increase (roughly $+10$ percentage units) when changing the motion am-



((a)) Histogram distributions of different accumulated fractions with 2.5 s period time.



((b)) Histogram distributions of different accumulated fractions with 5 s period time.

Figure 16: Histogram distributions of different accumulated fractions of relative dose differences for patient 1 at target depth 4.5 cm. Motion traces are similar with an asymmetrical breathing pattern and 5 mm motion amplitude, but with two different breathing periods. The red horizontal lines correspond to relative dose differences equated to $\pm 3\%$. Both dose blurring and interplay are considered. The ratio value is the quotient of the number of counts inside the red lines and the total number of counts.

plitude from 5 to 15 mm for all cases except the measurement with reduced spot spacing. Figure 21 also illustrates that a smaller spot spacing results in lower $\delta D98$ values for all amplitudes. Likewise, using a larger air gap, the $\delta D98$ value increased for the 5 mm and 10 mm motion amplitude.

4.5 Patient and target depths

Figure 22 displays comparisons of different patients and target depths using an asymmetrical motion trace with three different motion amplitudes but with a constant 5 s breathing period. Similarly to figure 21, the $\delta D98$ values increase with higher motion amplitudes. For all $\delta D98$

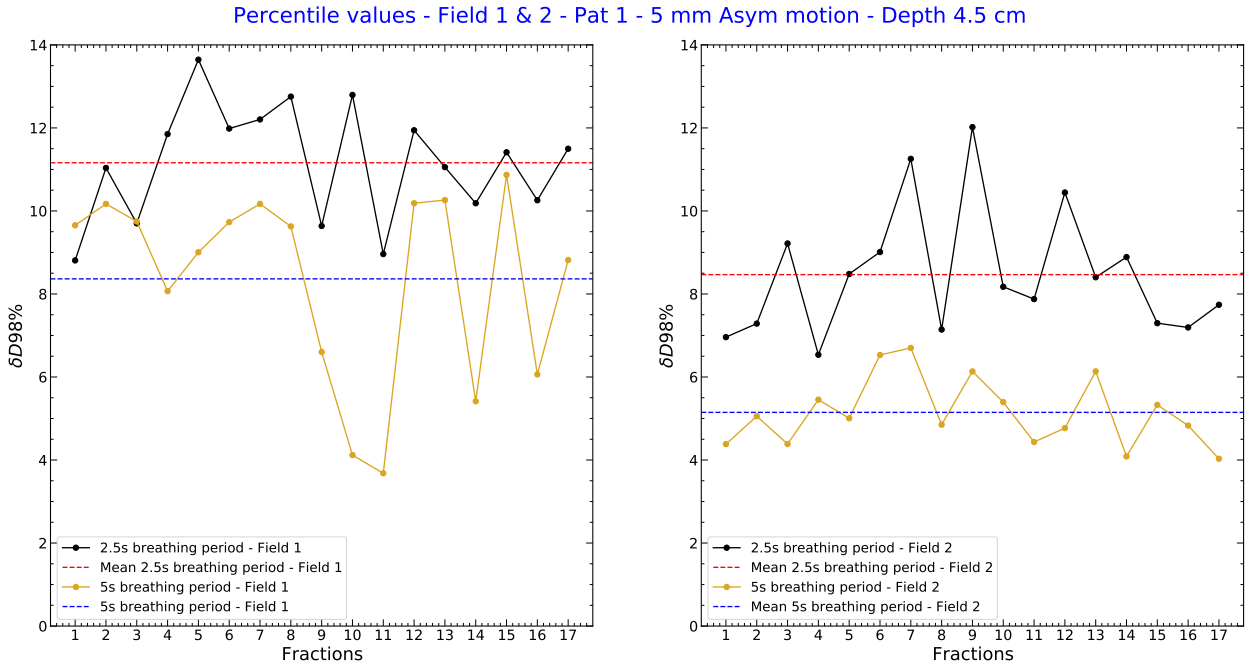


Figure 17: Variation of $\delta D98$ values over different single-fractions measurements for patient 1 at target depth 4.5 cm, similar to figure 13, but here for both fields. Motion traces simulate an asymmetrical breathing pattern with a 5 mm motion amplitude and 2.5 or 5 s period times. Both dose blurring and interplay are considered.

values except one, the $\delta D98$ values associated with patient's 1 target depths lies distinctly above other patients $\delta D98$ values.

4.6 Effect of motion regularity

The regularity of motion was investigated for patient 1 at two target depths, with different amplitudes and breathing periods (figure 23). For most measurements, symmetrical measurements have slightly less $\delta D98$ values than the corresponding asymmetrical measurements. However, in general, no real trend is observed.

4.7 Time structure of the beam delivery

Figure 24 concludes different parameters obtained from the treatment log files of patient 1 using the spot scanning system at the Skandion clinic. Similar figures for time structure of the beam delivery is illustrated for patient 2 and 3 in the appendix (see figure 29 and 30). Note that "layer 0" corresponds to the energy layer with the highest energy and layers with higher numbers corresponds to shallower depths with lower energies. For instance, target depth 4.5 cm is a relatively shallow depth positioned around the 17th energy layer. Importantly, the dose distribution measured at target depth 4.5 cm accumulates all energy layers up to and including energy layer 17. The subplot "time delivering dose per layer" in figure 24 describes the approximately total τ_{ss} for

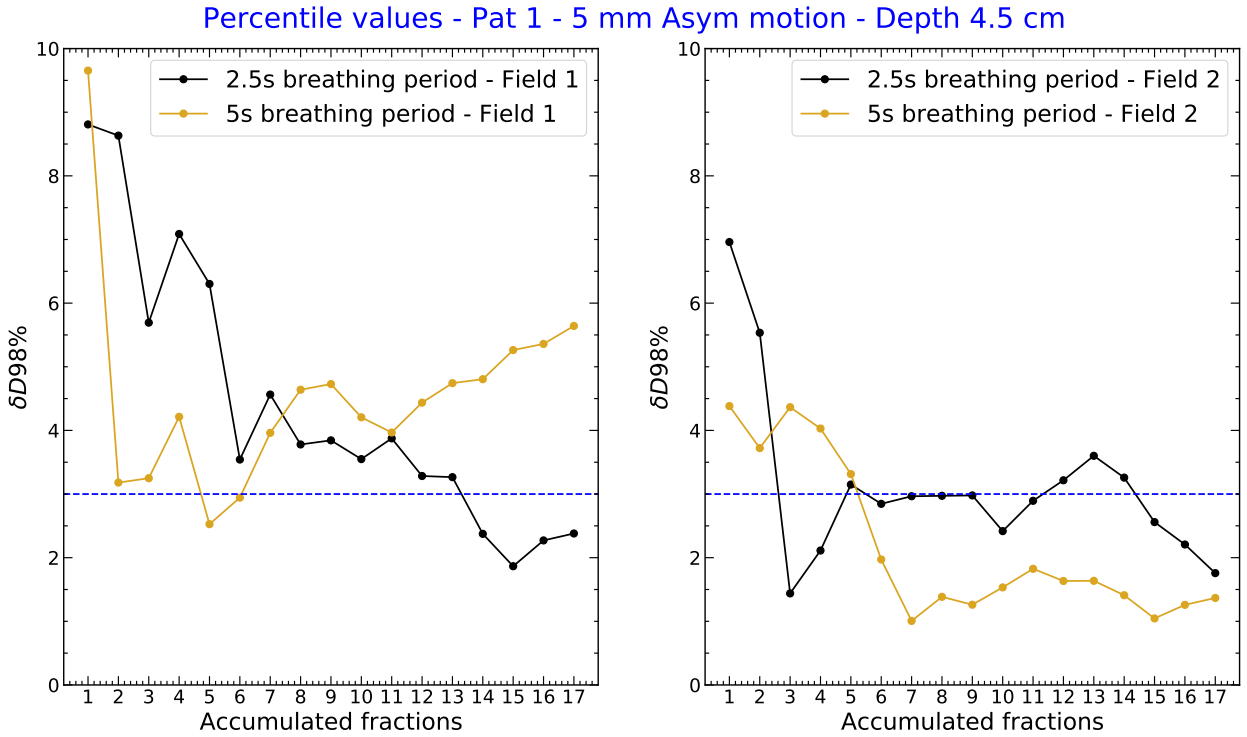


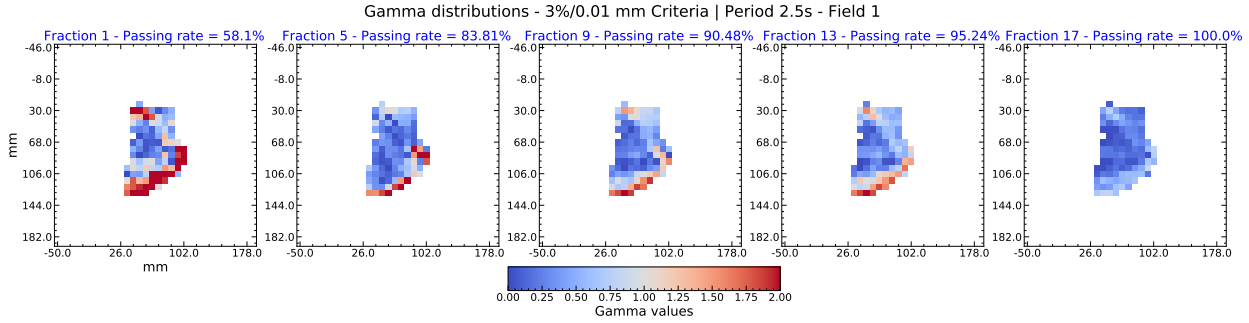
Figure 18: Variation of the $\delta D98$ value for different accumulated fractions for patient 1 at target depth 4.5 cm, separated by field. Motion traces have an asymmetrical motion regularity, a 5 mm amplitude and either a 2.5 s or a 5 s period time. Both dose blurring and interplay are considered.

each energy layer, while the "total time per layer" depicts the total τ_{ss} plus τ_{es}. Larger markers correspond to more spots in a layer, which increases the time to deliver the specific energy layer. Spearman's correlation test presents that larger markers roughly follow a positive monotonic relationship with higher MU; (r=.99 and p<.001) and (r=.97 and p<.001) for field 1 and field 2, respectively. The value of τ_{es} is random since no correlation is observed between the size or number of MU and τ_{es}. Due to the random behaviour of τ_{es}, the "total time per layer" is just shifted by a constant relative to the plot "time delivering dose per layer".

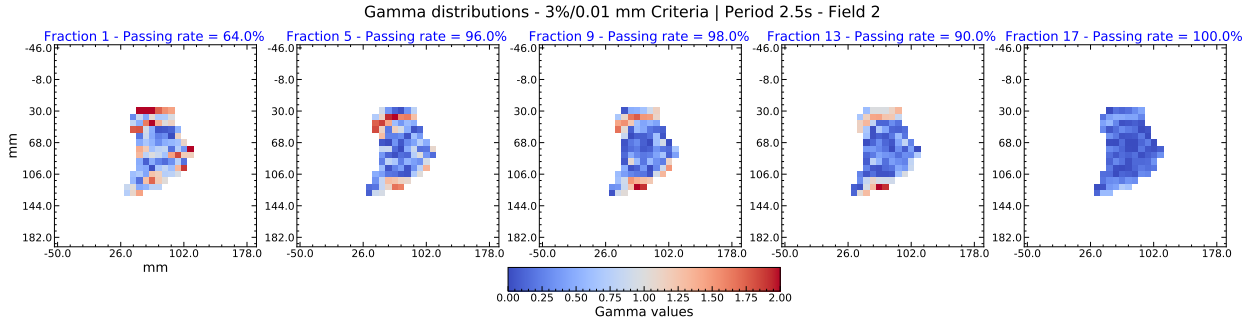
5 Discussion

5.1 The influence of target motion amplitude

Figure 11 and 12 show qualitatively that higher amplitudes result in larger breathing effects. Overall, this is true for all patients and target depths (figure 22) or different treatment characteristics (figure 21), which agrees with current literature. [7, 10, 12, 17, 30, 31] The 17 fractions measured for patient 1 at 4.5 cm target depth with a 5 mm asymmetrical motion pattern showed, on average, a single fraction $\delta D98$ variation about 7.1% and 5.5% with period times 2.5 and 5 s, respectively. Moreover, a large variation of $\delta D98$ values were observed between patients, which is shown in



((a)) Gamma distributions for field 1



((b)) Gamma distributions for field 2.

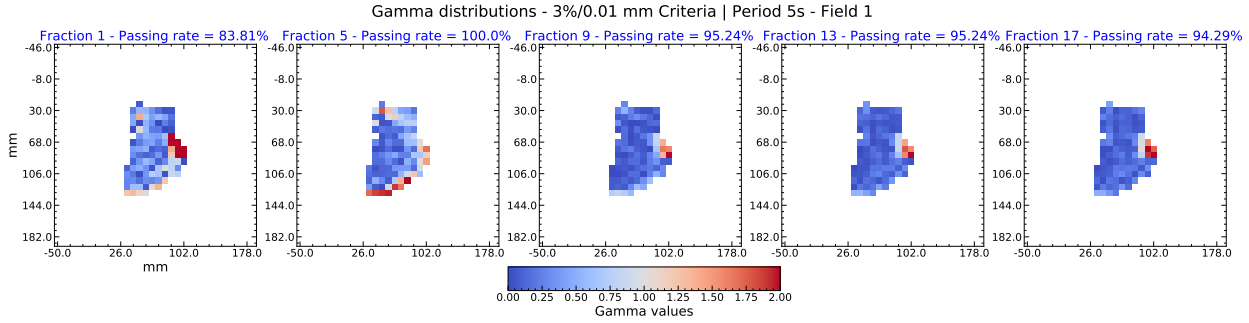
Figure 19: Field-wise gamma evaluation for different accumulated fractions for patient 1 at target depth 4.5 cm using an asymmetrical breathing curve with a target motion amplitude of 5 mm and a 2.5 s breathing period. The gamma criteria was set to 3%/0.01mm. Both dose blurring and interplay are considered.

for instance figure 22. For all patients considered in this study, single-fraction $\delta D98$ values varied between $\sim$ 10.3% in the worst-case scenario to $\sim$ 1.0% in the best-case scenario.

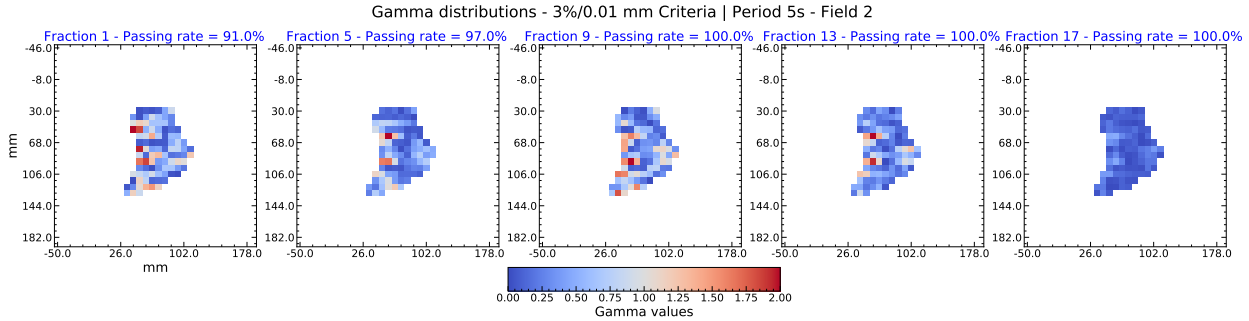
5.2 Multiple fractions

Despite being limited to one patient and target depth, some conclusions can be drawn from figure 15. Firstly, breathing effects are reduced when accumulating fractions, in line with other studies [12, 16, 17, 29, 30] The fractionation effect averages the relatively large variations of $\delta D98$ values present for individual fractions (figure 13) and individual fields (figure 17). The large inter-fractional variability of the $\delta D98$ values suggests that the starting phase has a large influence on the interplay effect. However, due to the fractionation effect, the impact of the starting phase have a negligible effect over an entire treatment course. Furthermore, figure 14 illustrates the stochastic nature of interplay since the hot and cold spots inside the dose distribution are located differently depending on the starting phase (fraction number). The interchangeableness between hot and cold spots inside the high dose area is caused by the randomly synchronised or non-synchronised movement between the tumour (detector) and the temporal beam delivery structure.

Some studies [12, 30] have reported that accumulating fractions only reduces dose variations



((a)) Gamma distributions for field 1.



((b)) Gamma distributions for field 2

Figure 20: Field-wise gamma evaluation for different accumulated fractions for patient 1 at target depth 4.5 cm using an asymmetrical breathing curve with a 5 mm target motion amplitude and 2.5 s period time. The gamma criteria was set to 3%/0.01mm. Both dose blurring and interplay are considered.

inside tumour regions to some extent. Figure 15 illustrates this phenomenon partially. Measurement associated with a 2.5s breathing period reaches a minimum for the last accumulated fraction, suggesting that at least 15+ fractions are needed to sufficiently reduce the influence of breathing. However, for measurements associated with a 5 s breathing period and field 1, the minimum $\delta D98$ value occurs at fraction 5 before increasing above the ± 3 dose deviation limit for later fractions. The trend associated with the 5 s breathing period measurement has the opposite tendency compared to what is depicted in the literature. [12, 16, 17, 29, 30] One possible explanation is the apparent peak and degree of inhomogeneity of the lateral dose profile for field 1 at the 4.5 cm target depth (figure 25). In figure 18, all measurements except 5 s breathing period and field 1 show a fractionation effect, i.e. converging to a minimum value for 10+ fractions. Instead, the $\delta D98$ value increase after fraction 5 when accumulating fractions, which is similar to the behaviour illustrated in figure 15. Gamma evaluations presented in figure 20 depict that the inhomogeneous lateral dose profile of field 1 is highly influential for the increase of the $\delta D98$ value for later fractions. The gamma evaluation shows that the large relative dose differences are located at or near the high dose region (figure 9) for measurements associated with a 5 s breathing period and field 1. However, the impact is much less pronounced for measurements associated with a 2.5 s period time, which might be deduced from the greater averaging effect for shorter periods. For instance, by ob-

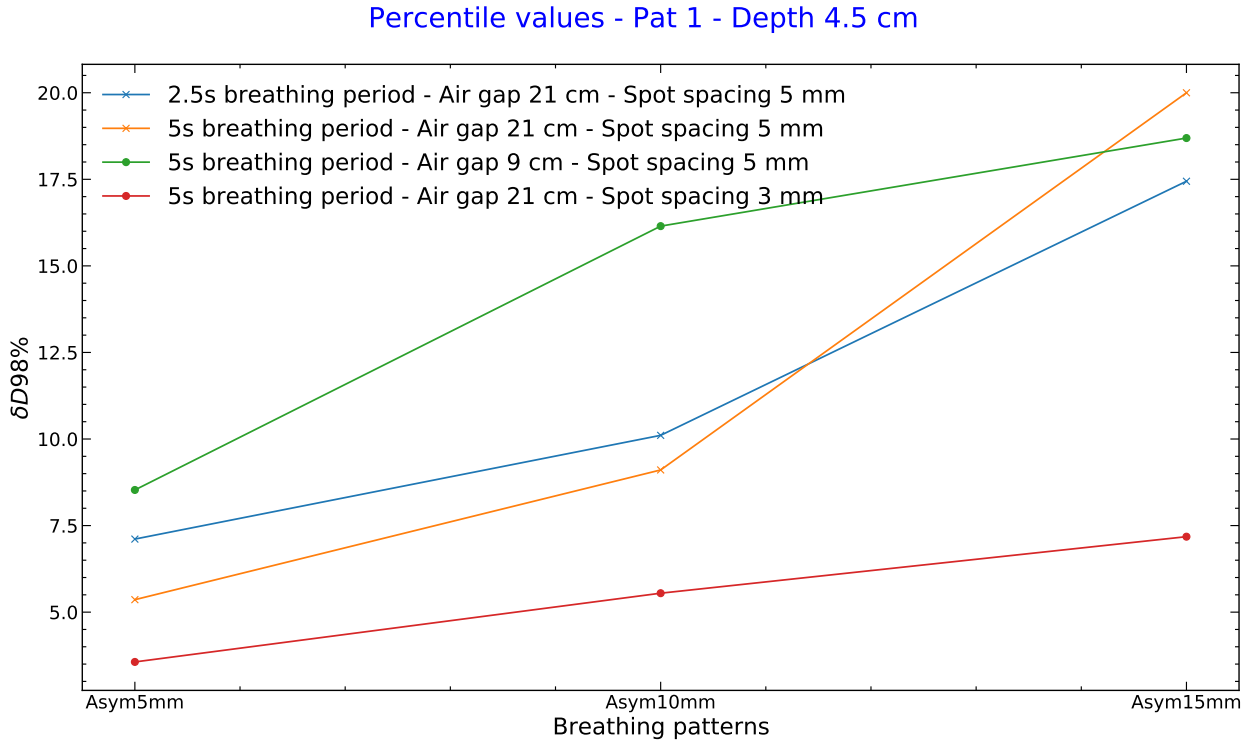


Figure 21: Variation of $\delta D98$ values for different treatment characteristics for patient 1 at 4.5 cm target depth, using an asymmetrical breathing pattern, three different motion amplitudes and two different breathing periods. Either the period, air gap or spot spacing were altered. If the air gap or spot spacing was changed, the breathing period was kept at 5s. N.b, the percentile value at 5 mm motion amplitude is the mean value out of the 17 fractions measured for period times 2.5 and 5 s, respectively (the red and blue line in figure 13, respectively). Both dose blurring and interplay are considered.

serving the time structure of the spot scanning delivery for patient 1 (figure 24), the target depth at 4.5 cm lies directly after the delivery of energy layers that are densely populated with spots. The target dose distribution measured at target depth 4.5 cm accumulates all prior energy layers, including the 17th energy layer. A shorter period may be less affected by layer shifts caused by breathing, as the spots will be more evenly distributed over the entire breath cycle. This is supported in [10], where the greatest homogeneity loss occurred for the longest breathing period. However, it should be noted that the difference between with period times 2.5 and 5 s, respectively (figure 15), was not found statically significant (U-test, p=1.0), and in patient anatomy, the breathing period is not absolute. Instead, the breathing period can likely change drastically both inter-fractionally and intra-fractionally. The change of breathing period will further add to the stochastic nature of interplay and average the dose distribution even more, reducing the interplay effect in the process.

Lastly, figure 15 suggest that there is minimal effect of dose blurring inside the high dose region since there is a none significant effect on the $\delta D98$ values when removing the uncertainties caused by dose blurring. Although this disagrees with the theory that dose blurring and interplay are two additive dose uncertainties, the impact of dose blurring might be less significant inside the

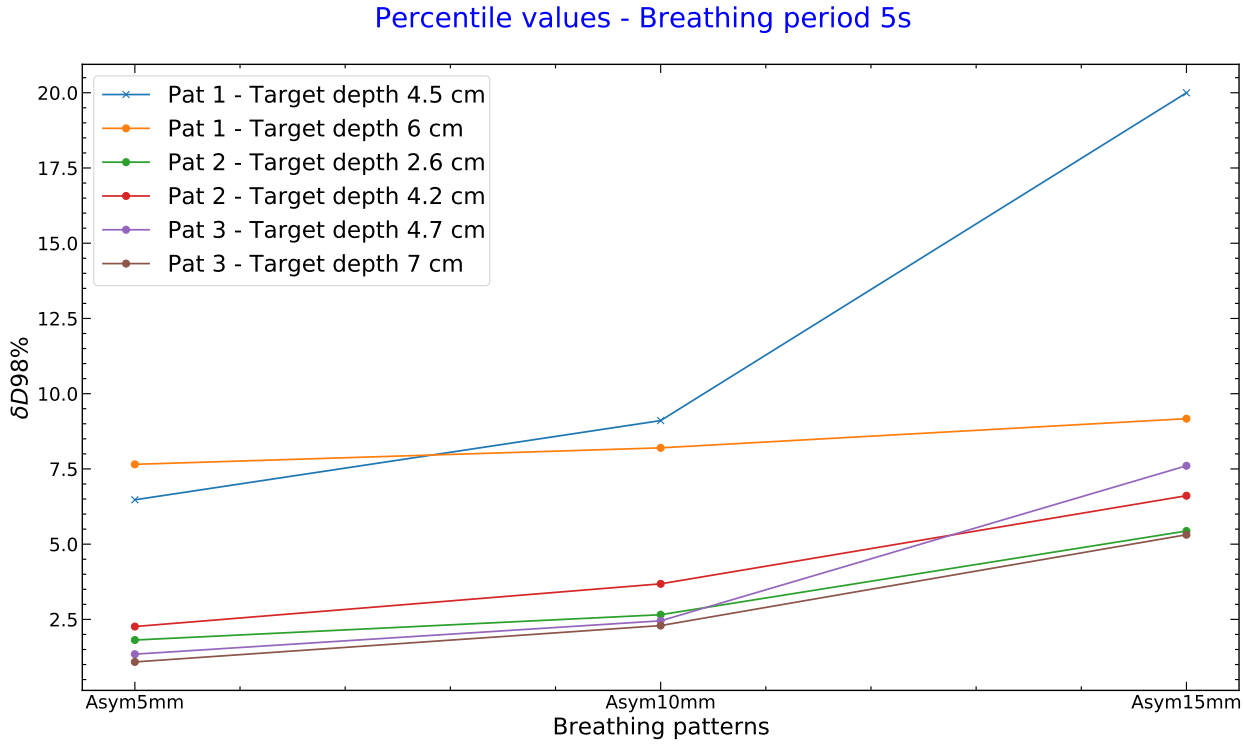


Figure 22: Variation of $D98$ percentile values for different patients and target depths. All cases were simulated with an asymmetrical motion trace with a constant 5 s breathing period but with different target motion amplitudes. The percentile value at 5 mm motion amplitude is the mean value out of the 17 fractions (blue line in figure 13). Both dose blurring and interplay are considered.

isolated high dose region. When creating the eroded dose mask region, the edges of the measured dose distribution are removed, which might reduce the influence of dose blurring. Additionally, [29] also showed that dose blurring had a smaller impact on the dose distribution than interplay. For instance, for a 5 mm motion amplitude, it was found that the effect of dose blurring was less than 1% of the total breathing effects.

5.3 Patients and target depths

Among the patients in figure 22, the magnitude of the target motion amplitude is the only variable parameter apart from target geometry. Despite this, there is considerable variation of $\delta D98$ values between patient 2 for target depth 4.5 cm relative to other patients and target depths. The only considerable differences between the patients are TVs and the lateral dose profiles (25, 26 and 27). In this study, patient 1, 2 and 3 had TVs of 152, 232 and 336 cc, respectively. A larger TV will result in a longer beam delivery time since more spots need to be delivered for each field (table 1). This can directly be observed by comparing the time needed to deliver all energy layers for both fields for the different patients ("total cumulative time" in figure 24, 29 and 30, respectively). Two studies [12, 29] have reported that the interplay effect is dependent on the TV, where smaller TVs are associated with more pronounced interplay effects since the beam delivery time is shorter.

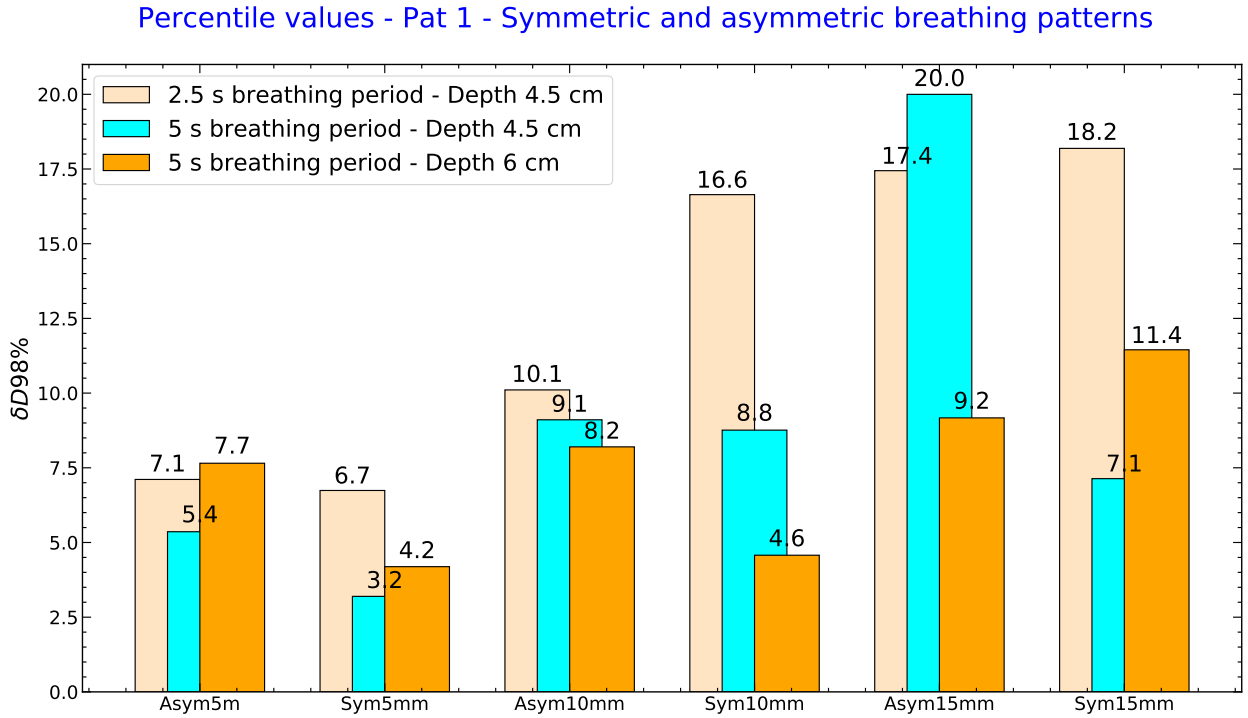


Figure 23: Bar plot showcasing the variation of $\delta D98$ values for patient 1 for both target depths and different target motion amplitudes, breathing periods and motion regularity. Both dose blurring and interplay are considered.

Hence, spots will be distributed more evenly in time and over the entire breath cycle, smearing the dose distribution and reducing the interplay effect.

The inhomogeneous lateral dose profile at the 4.5 cm target depth (25) for patient 1 is another possible reason for the larger $\delta D98$ values at 4.5 cm target depth compared to other target depths (22). The lateral dose profiles for the other two patients are relatively flat (26 and 27), which may reduce the impact of interplay more. The dose distribution at target depth 4.5 cm is affected by high spot weights (high dose region in figure 9) and a shift of these spots can cause large discrepancies to the static dose distribution (e.g. figure 20).

Lastly, it should be noted that the $\delta D98$ values described in figure 22 are based on single-fractions, hence strongly affected by the starting phase. As observed in figure 13, the $\delta D98$ value changes randomly from fraction to fraction. Thus, the $\delta D98$ values for target depths other than 4.5 cm might be lower by chance. However, since all $\delta D98$ values for the other target depths lies distinctly below $\delta D98$ values for target depth 4.5 cm, it is not reasonable to believe that all target depths other than 4.5 cm are an under-estimation of the "true" $\delta D98$ value.

5.4 Treatment characteristics

Figure 21 indicates that adding a larger number of spots per energy layer reduces interplay, which is in agreement with the current literature. [5, 10, 16, 17] With a smaller spot spacing, the total

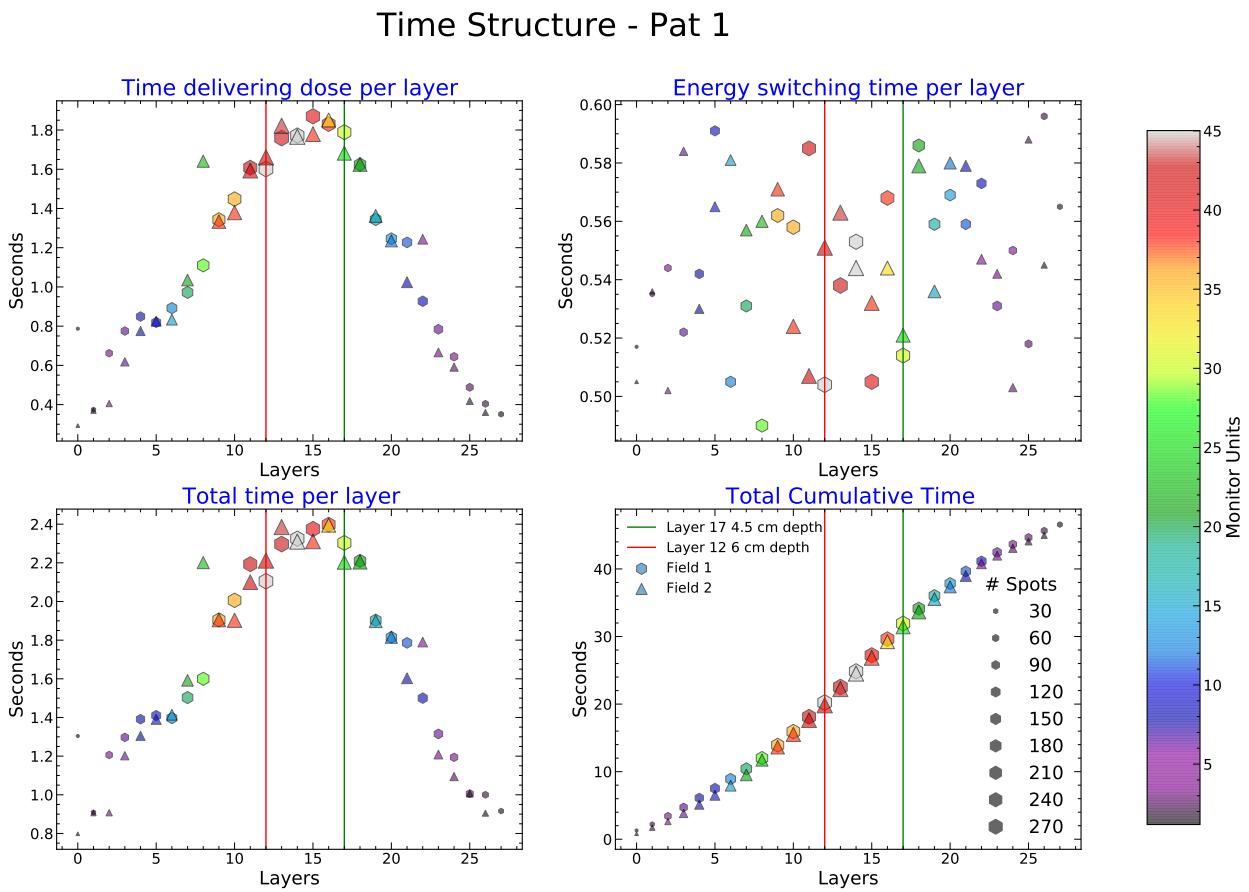


Figure 24: This figure concludes information about the time structure and beam characteristics during the delivery of each field for patient 1 (table 1), with layer 0 corresponding to the energy layer with the highest energy. For instance, the 4.5 cm target depth corresponds to 17th energy layer. The "time delivering dose per layer" refers to when the dose is delivered for each energy layer. The "energy switching time per layer" describes the time structure of τ_{es}. The sub-plot "total time per layer" explains the time needed to deliver each energy layer, including τ_{es}. Lastly, "total cumulative time" is the accumulated total time for different layers, starting with layer 0. The colour bar describes the amount of MU delivered for a specific energy layer and the size of each marker reflects the number of spots delivered for each layer. Larger markers correspond to more spots and vice versa.

duration of τ_{ss} is increased since more spots are delivered per energy layer. Thus, spots will be spread more uniformly over the entire breath cycle. In addition, there is a larger overlap of spots, resulting in a larger smoothing effect, reducing the interplay effect. The result suggests that the impact of dose inhomogeneities can be mitigated since the $\delta D98$ values observed with a 3 mm spot spacing (figure 21) are in line with $\delta D98$ values noted for other target depths (figure 22) with 5 mm spot spacing. However, these results need to be validated with more measurements. Similarly to a reduced spot spacing, a larger air gap reduces interplay. An increased air gap leads to larger spot sizes, resulting in a greater average effect, as [10, 20] also demonstrate.

5.5 Motion regularity

Current literature depicts that irregular motion traces affect the planned dose distribution more than regular motion. [12, 16, 17] However, the result illustrated in figure 23 is ambiguous with no distinct trends.

6 Conclusion

This study aimed to analyse simulated breathing motions in PBS for different mediastinal HL dose distributions. Breathing effects were most prominent for larger amplitudes but could be reduced by accumulating fractions, increasing spot sizes and using smaller spot spacing. Although the data is sparse, no clear dependency on motion regularity was observed. Breathing effects were predominantly caused by interplay, with dose blurring contributing in a non-significant way (U-test, p=1.0). The interplay effect was patient-specific, suggesting that treatment plan characteristics and beam parameters play a crucial part in breathing effects. Single-fractions $\delta D98$ values varied between $\sim$ 10.3% in the worst-case scenario to $\sim$ 1.0% in the best-case scenario. Overall, PBS may be allowed for tumour motion above 5 mm in the SI direction. However, more care should be given to avoid heterogeneous dose distributions in the tumour volume, keeping the spot intensities low if achievable.

7 Future prospects and limitations

There are some drawbacks to this thesis. The clinical SFO treatment plans are inherently homogeneous, however, when re-calculated in WET values, the dose distributions becomes more inhomogenous. Adding that both fields were measured perpendicular to the gantry, not considering the angle between the two fields, further distorted the dose distribution. Accordingly, breathing effects observed clinically may be lower than those found in this thesis. Another problem with converting treatment plans to WET values is losing the TV position. In this thesis, the target area was approximated by the 95% iso-dose level of the prescribed dose, and the dose mask excluded all points outside the 95% iso-dose level. In addition, an eroded dose mask was created for every static dose distribution to serve as an approximation to isolate the target area from its margins and to analyse the effects of the "true" target region. Creating a dose mask is questionable since all breathing effects outside the eroded dose mask are neglected, and this approach would not be relevant for MFO plans due to their inhomogeneous spot patterns. The methodology could be improved by implementing artificial CTV in WET values based on typical HL CTV geometries found in the clinic rather than clinical treatment plans. In this case, the dose mask could be avoided since the TV location is not lost.

The spatial resolution of the MatriXX detector could have also been another disadvantage ($7.6 \cdot 7.6$ mm^2). The large pixel dimensions could reduce the influence of breathing since motion with a 5 mm amplitude is small compared to the pixel size. Nonetheless, the impact of the spatial resolution was minimal, with a gamma value of $\sim$ 98% for a 3%/0.01 mm passing criteria when

comparing a static dose measurement for patient 1 at target depth 4.5 cm with similar static dose distribution, but with a 4 mm off-set in the SI-direction. Furthermore, the experimental set-up was not entirely stable since the starting position of the MatriXX detector was altered between motion traces, despite the Hexamotion platform being reset to its starting point after motion. The set-up was deemed stable enough for motion in one direction but not in two or three. In the future, the set-up might be polished further to allow motion in more than one dimension. Otherwise, the existing set-up could be used to analyse new parameters or treatment plans. For instance, it would be interesting to compare with different spot spacing to clarify if the same reduction of breathing effects observed in this study can be seen in other treatment plans. This study is also limited by the number of target depths and patients selected and the number of fractions measured. In order to enhance the significance of the study, at least 17 fractions should be measured for all targets depths. In this way, the measurements will be less dependent on the starting phase. The analysis could have also considered over- and under-dosage separately. Instead of computing the absolute value of D_{rel}, 98th and 2nd percentiles of D_{rel} could be computed to estimate over- and under-dosage, respectively. However, we considered that $\delta D98$ based on the absolute value of D_{rel} to be the most suitable parameter, and the histogram analysis in figure 16 consider over- and under-dosage specifically for one patient.

Simulation studies may be the most effective method for future research since larger combinations of conditions and parameters can be manipulated. Simulations are faster and they might be able to create a more realistic patient-like set-up. In this thesis, the experimental set-up, solid water plates and simulated breathing patterns are relatively simplistic. For instance, under patient-like conditions, the variation of tissue densities will affect the range of each spot differently compared to the case of a homogeneous water phantom. The presence of different densities during breathing will result in range uncertainties, causing beamlets within the same energy layer to be distorted, adding further uncertainty to the total breathing effects.

8 Acknowledgements

I am grateful for all help I have gained while working on my master thesis. It has been a tone of fun and it has been a pleasure meeting a lot of new friendly faces [at a distance :)] during the past semester.

Specials thanks to:

- Marika Enmark, thank you for your critical thinking and outstanding enthusiasm.
- Anneli Edvardsson, thank you for your expertise and for always being available during the semester.
- Ingrid Kristensen, thank you for helping me get started in Eclipse and for your expertise in treatment planning.
- Karin Andersson & Christina Vallhagen Dahlgren, thank you for making me feel welcome at the Skandion clinic and for your insights in PBS, despite being at a distance.

- All personnel at the radiotherapy department in Lund and the Skandion clinic, thank you for your hospitality and for making my master thesis much more enjoyable.
- Thomas Olausson & Markus Strickert, thank you for your excellent programming knowledge and helping me be calm under hectic periods.

References

1. Molin, D. *Hodgkins lymfom* https://www.internetmedicin.se/behandlingsoversikter/onkologi/hodgkins-lymfom/. Accessed October 28, 2021.
2. *Hodgkin's lymphoma (Hodgkin's disease)* https://www.mayoclinic.org/diseases-conditions/hodgkins-lymphoma/symptoms-causes/syc-20352646. Accessed September 10, 2021.
3. *Hodgkins lymfom. Nationellt vårdprogram.* https://kunskapsbanken.cancercentrum.se/diagnoser/hodgkins-lymfom/vardprogram/. Version 4.0. Februari 08, 2022.
4. *Key Statistics for Hodgkin Lymphoma* https://www.cancer.org/cancer/hodgkin-lymphoma/about/key-statistics.html. Accessed January 21, 2021.
5. Chang, J. Y. *et al.* Consensus guidelines for implementing pencil-beam scanning proton therapy for thoracic malignancies on behalf of the PTCOG thoracic and lymphoma subcommittee. *International Journal of Radiation Oncology* Biology* Physics* **99,** 41–50 (2017).
6. Loap, P. *et al.* Current Situation of Proton Therapy for Hodgkin Lymphoma: From Expectations to Evidence. *Cancers* **13,** 3746 (2021).
7. Pfeiler, T. *Dynamic evaluation of 4D robust optimisation for motion management in scanned proton therapy of hepatocellular carcinoma.* PhD thesis (Technical University Dortmund., 2018).
8. Lomax, A. J. Myths and realities of range uncertainty. *The British journal of radiology* **93,** 20190582 (2020).
9. Meijers, A. *et al.* Evaluation of interplay and organ motion effects by means of 4D dose reconstruction and accumulation. *Radiotherapy and Oncology* **150,** 268–274 (2020).
10. Dowdell, S., Grassberger, C., Sharp, G. & Paganetti, H. Interplay effects in proton scanning for lung: a 4D Monte Carlo study assessing the impact of tumor and beam delivery parameters. *Physics in Medicine & Biology* **58,** 4137 (2013).
11. Van Herk, M., Remeijer, P., Rasch, C. & Lebesque, J. V. The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy. *International Journal of Radiation Oncology* Biology* Physics* **47,** 1121–1135 (2000).
12. Lee, E. *et al.* Measurement-based study on characterizing symmetric and asymmetric respiratory motion interplay effect on target dose distribution in the proton pencil beam scanning. *Journal of applied clinical medical physics* **21,** 59–67 (2020).
13. Edvardsson, A. *Dosimetric effects of breathing motion in radiotherapy* PhD thesis (Lund University, 2018).

14. Bortfeld, T., Jiang, S. B. & Rietzel, E. *Effects of motion on the total dose distribution* in *Seminars in radiation oncology* **14** (2004), 41–51.

15. Edvardsson, A. *et al.* Comparative treatment planning study for mediastinal Hodgkin's lymphoma: impact on normal tissue dose using deep inspiration breath hold proton and photon therapy. *Acta Oncologica* **58,** 95–104 (2019).

16. Han, Y. Current status of proton therapy techniques for lung cancer. *Radiation oncology journal* **37,** 232 (2019).

17. Grassberger, C. *et al.* Motion interplay as a function of patient parameters and spot size in spot scanning proton therapy for lung cancer. *International Journal of Radiation Oncology* Biology* Physics* **86,** 380–386 (2013).

18. Mohan, R. & Grosshans, D. Proton therapy–present and future. *Advanced drug delivery reviews* **109,** 26–44 (2017).

19. Mohan, R., Das, I. J. & Ling, C. C. Empowering intensity modulated proton therapy through physics and technology: an overview. *International Journal of Radiation Oncology* Biology* Physics* **99,** 304–316 (2017).

20. Zeng, C. *et al.* Proton pencil beam scanning for mediastinal lymphoma: the impact of interplay between target motion and beam scanning. *Physics in Medicine & Biology* **60,** 3013 (2015).

21. Tommasino, F. *et al.* Clinical implementation in proton therapy of multi-field optimization by a hybrid method combining conventional PTV with robust optimization. *Physics in Medicine & Biology* **65,** 045002 (2020).

22. Lambert, J., Suchowerska, N., McKenzie, D. & Jackson, M. Intrafractional motion during proton beam scanning. *Physics in Medicine & Biology* **50,** 4853 (2005).

23. Kraan, A., Depauw, N., Clasie, B., Madden, T. & Kooy, H. Impact of spot size variations on dose in scanned proton beam therapy. *Physica Medica* **57,** 58–64 (2019).

24. Paganetti, H. Range uncertainties in proton therapy and the role of Monte Carlo simulations. *Physics in Medicine & Biology* **57,** R99 (2012).

25. Fattori, G. *et al.* The dependence of interplay effects on the field scan direction in PBS proton therapy. *Physics in Medicine & Biology* **64,** 095005 (2019).

26. Liu, W., Zhang, X., Li, Y. & Mohan, R. Robust optimization of intensity modulated proton therapy. *Medical physics* **39,** 1079–1091 (2012).

27. Ceberg, S., Ceberg, C., Falk, M., af Rosenschöld, P. M. & Bäck, S. Å. J. *Evaluation of breathing interplay effects during VMAT by using 3D gel measurements* in *Journal of Physics: Conference Series* **444** (2013), 012098.

28. Zhang, Y., Boye, D., Tanner, C., Lomax, A. J. & Knopf, A. Respiratory liver motion estimation and its effect on scanned proton beam therapy. *Physics in Medicine & Biology* **57,** 1779 (2012).

29. Protik, A., van Herk, M., Witte, M. & Sonke, J.-J. The impact of breathing amplitude on dose homogeneity in intensity modulated proton therapy. *Physics and Imaging in Radiation Oncology* **3,** 11–16 (2017).

30. Boria, A. J. *et al.* Interplay effect of target motion and pencil-beam scanning in proton therapy for pediatric patients. *International journal of particle therapy* **5,** 1–10 (2018).

31. Seco, J., Robertson, D., Trofimov, A. & Paganetti, H. Breathing interplay effects during proton beam scanning: simulation and statistical analysis. *Physics in Medicine & Biology* **54,** N283 (2009).

32. Cubillos-Mesıas, M. *et al.* Impact of robust treatment planning on single-and multi-field optimized plans for proton beam therapy of unilateral head and neck target volumes. *Radiation Oncology* **12,** 1–10 (2017).

33. Lujan, A. E., Larsen, E. W., Balter, J. M. & Ten Haken, R. K. A method for incorporating organ motion due to breathing into 3D dose calculations. *Medical physics* **26,** 715–720 (1999).

34. Seppenwoolde, Y. *et al.* Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy. *International Journal of Radiation Oncology* Biology* Physics* **53,** 822–834 (2002).

35. Ferris, W. S., Kissick, M. W., Bayouth, J. E., Culberson, W. S. & Smilowitz, J. B. Evaluation of radixact motion synchrony for 3D respiratory motion: Modeling accuracy and dosimetric fidelity. *Journal of Applied Clinical Medical Physics* **21,** 96–106 (2020).

36. King, R. *et al.* Time-resolved dosimetric verification of respiratory-gated radiotherapy exposures using a high-resolution 2D ionisation chamber array. *Physics in Medicine & Biology* **61,** 5529 (2016).

37. White, B. M., Zhao, T., Lamb, J. M., Bradley, J. D. & Low, D. A. Physiologically guided approach to characterizing respiratory motion. *Medical physics* **40,** 121723 (2013).

38. *MatriXX PT/ONE* https://www.iba-dosimetry.com/product/matrixx-one-pt/. Accessed: 2021-12-23.

39. Langen, K. M. & Jones, D. T. Organ motion and its management. *International Journal of Radiation Oncology* Biology* Physics* **50,** 265–278 (2001).

40. Schaake, E. E. *et al.* Differential motion between mediastinal lymph nodes and primary tumor in radically irradiated lung cancer patients. *International Journal of Radiation Oncology* Biology* Physics* **90,** 959–966 (2014).

41. *Matplotlib. matplotlib.pyplot.imshow* https://matplotlib.org/3.5.1/api/_as_gen/matplotlib.pyplot.imshow.html. Accessed: 2022-03-15.

42. *PyMedPhys* https://docs.pymedphys.com/. Accessed: 2021-12-02.

43. *myQA* https://www.iba-dosimetry.com/myqa-software/. Accessed: 2021-11-05.

44. Hazra, A. & Gogtay, N. Biostatistics series module 3: comparing groups: numerical variables. *Indian journal of dermatology* **61,** 251 (2016).

45. Schober, P., Boer, C. & Schwarte, L. A. Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia* **126,** 1763–1768 (2018).

9 Appendix

9.1 Lateral dose profiles

The lateral dose profile in water for different patient's target depths (table 1) from Eclipse's treatment planning system are presented in figure 25, 26 and 27. The lateral dose profiles describe the planned dose distribution but are deemed to represent the static dose distribution measured for different target depths using the experimental set-up described in figure 4.

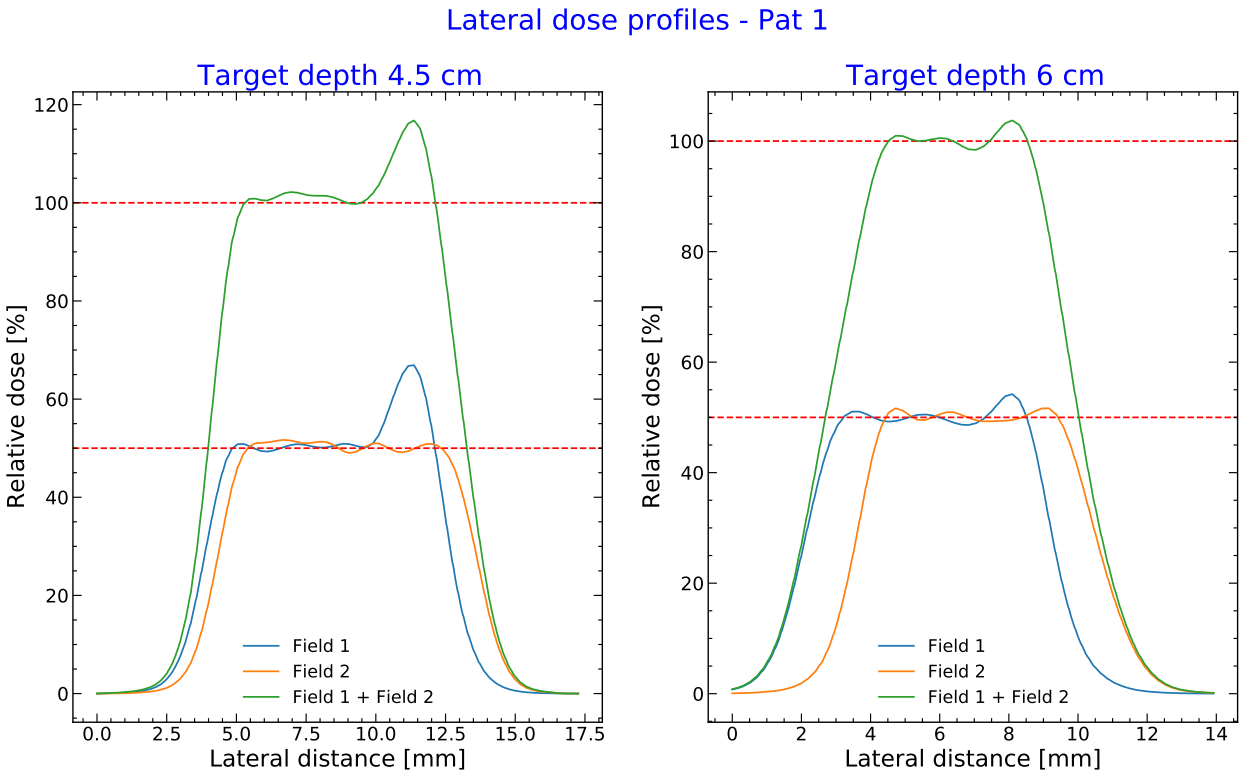


Figure 25: Lateral dose profiles in water at two target depths for patient 1.

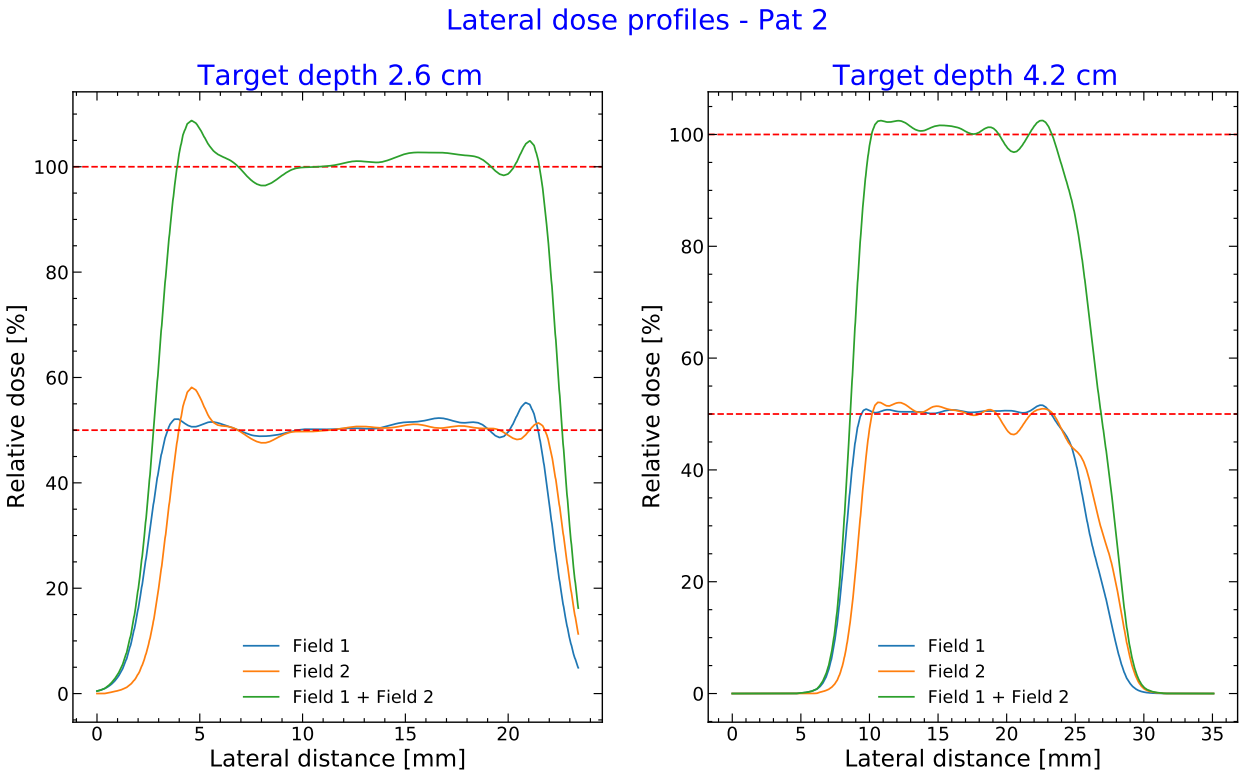


Figure 26: Lateral dose profiles in water at two target depths for patient 2.

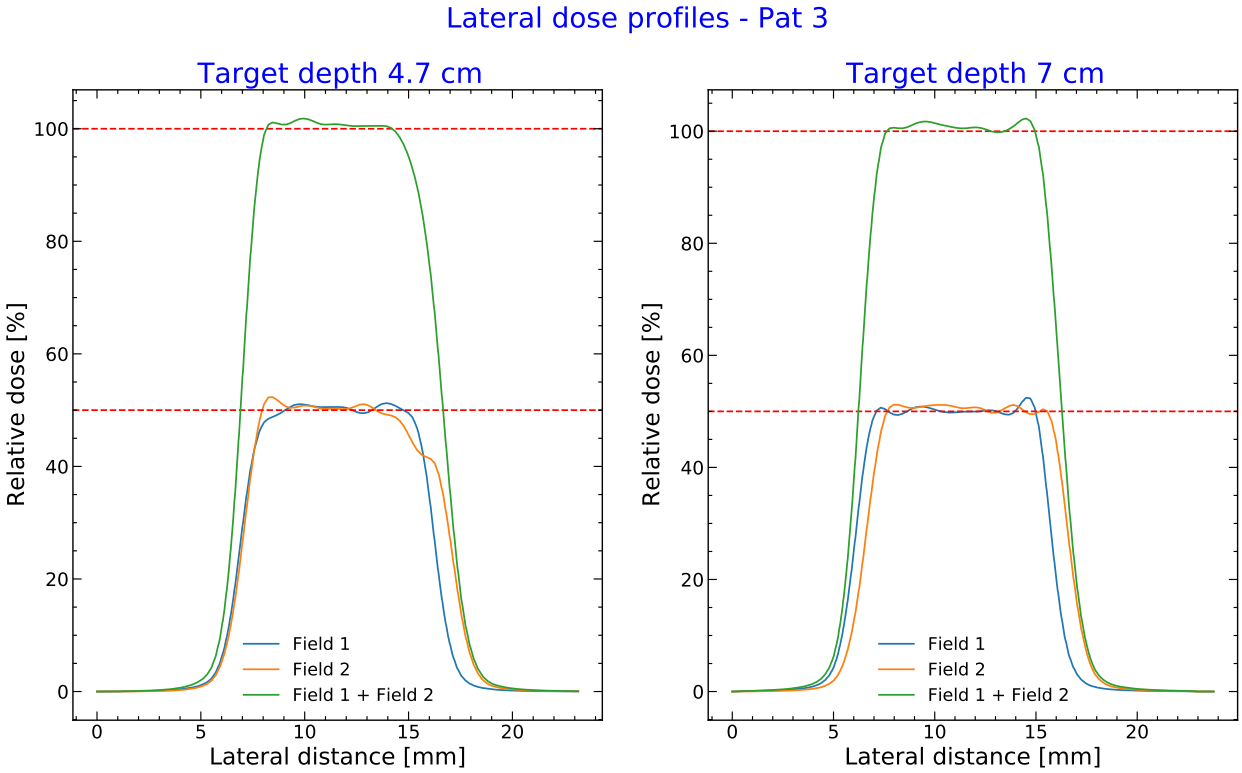


Figure 27: Lateral dose profiles in water at two target depths for patient 3.

9.2 Experimental set-up

The experimental set-up is illustrated in figure 28, highlighting the different components in figure 4 from another view.

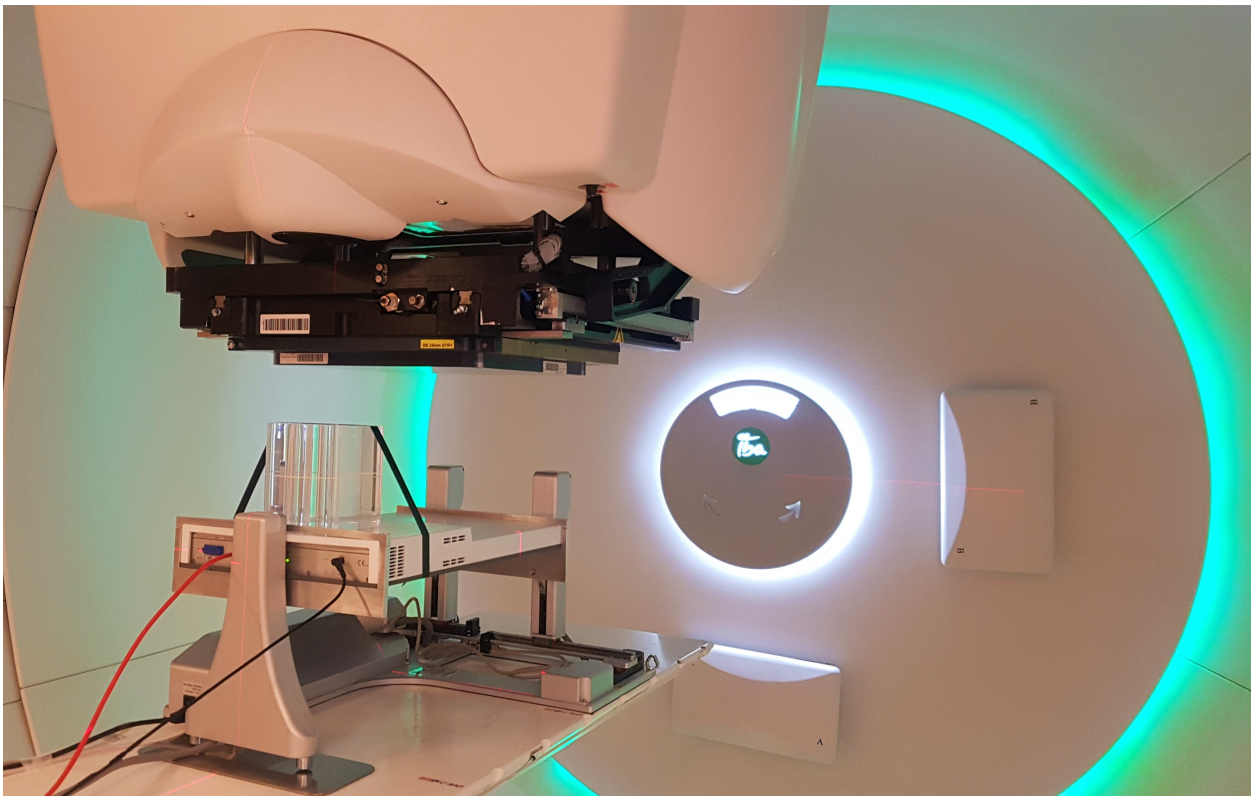


Figure 28: The experimental set-up with the MatriXX detector and the Hexamotion platform. Full description is presented in figure 4

.

9.3 Time Structure

Figure 29 and 30 (and figure 24) reveal information about the time structure and beam characteristics during the delivery of each field for the patients included in the study (table 1). Layer 0 corresponds to the energy layer with the highest energy. The dose distribution for one target depth at a certain energy layer accumulates all deeper energy layers (higher energies). The colorbar describes the amount of MU delivered for a specific energy layer and the size of each marker reflects the number of spots delivered for each layer. Larger markers correspond to more spots and vice versa.

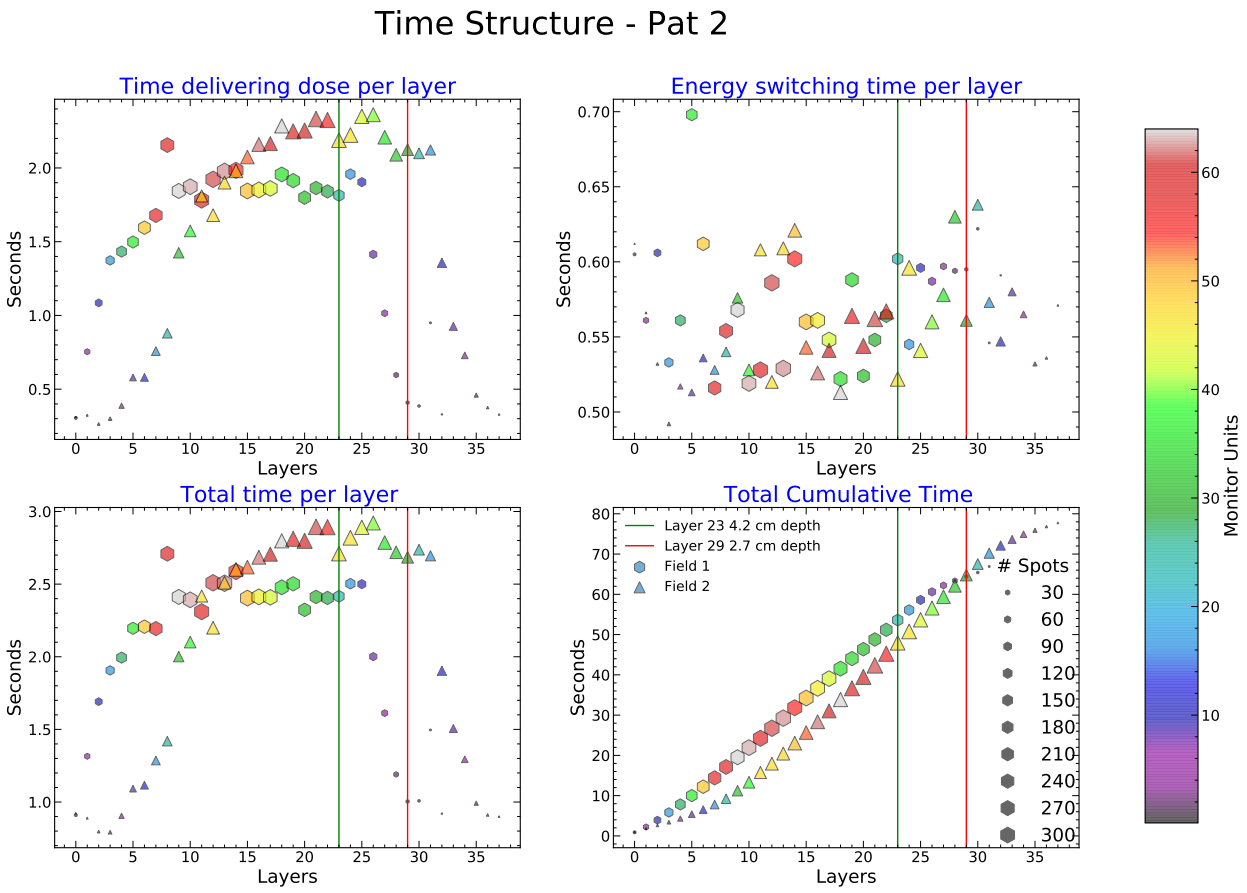


Figure 29: Time structure and beam characteristics of patient 2 during the delivery of each field.

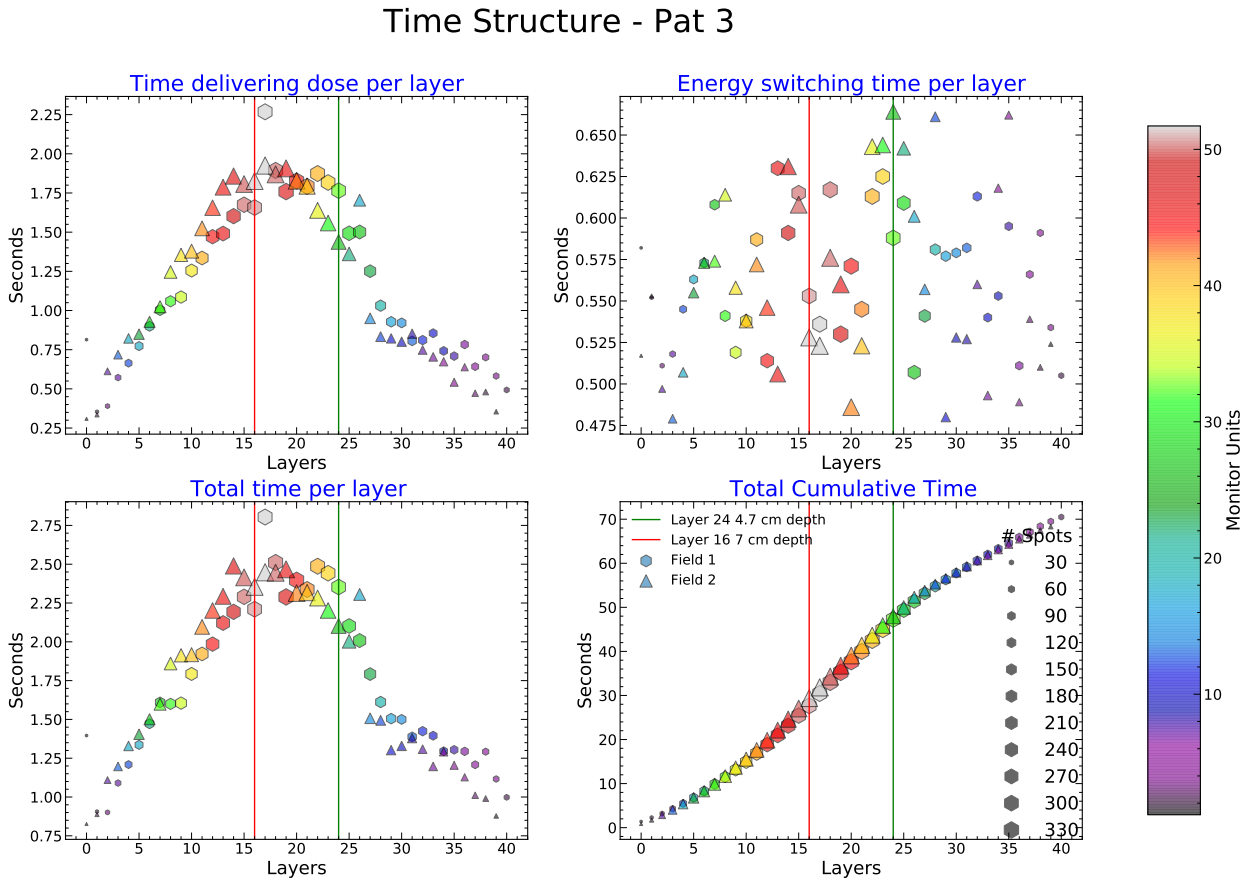


Figure 30: Time structure and beam characteristics of patient 3 during the delivery of each field.