Vrije Universiteit Brussel

Faculty of Engineering

A comparison of two clinical correlation models for dynamic tumor tracking with a focus on geometrical accuracy

Graduation thesis submitted in partial fulfilment of the
requirements for the degree of Master of Engineering: Biomedical Engineering

Jennifer Dhont

Promoter: Prof. Dr. Dirk Verellen
Advisors: MSc. Kenneth Poels
Dr. Ir. Tom Depuydt

2014

Vrije Universiteit Brussel

Faculteit Ingenieurswetenschappen

A comparison of two clinical correlation models for dynamic tumor tracking with a focus on geometrical accuracy

Eindwerk ingediend voor het behalen van de
graad van Master in de Ingenieurswetenschappen: Biomedische Ingenieurstechnieken

Jennifer Dhont

Promoter: Prof. Dr. Dirk Verellen
Begeleiders: MSc. Kenneth Poels
Dr. Ir. Tom Depuydt

2014

Acknowledgements

This thesis would not have been possible without the support of various people.

Foremost, I would like to express my sincere gratitude to Professor Dirk Verellen for allowing me to participate in the Vero project, for promoting this thesis and for granting me the opportunity to present this work abroad.

I would also like to express my appreciation to my advisers Kenneth Poels and Tom Depuydt, for answering my numerous questions and giving me insightful comments and suggestions. In particular I would like to thank Kenneth for the work he provided to support this thesis.

I would also like to thank the entire staff of the Radiotherapy Department of the UZ Brussels for making me feel welcome and for helping me find my way through the hospital. In addition, I would like to thank Thomas Lacornerie from the Centre Oscar Lambret for providing me with the necessary information about the Cyberknife system.

I would also like to offer a special thanks to Ruben, not only for his endless support and encouragement, but also for helping me through several ICT-related obstacles.

Lastly, I would also like to thank my parents and sister for their generous support and encouragement, not only throughout the duration of this thesis but during my entire education.

Abstract

A comparison of two clinical correlation models for dynamic tumor tracking with a focus on geometrical accuracy

Jennifer Dhont, Master in Biomedical Engineering, 2014

Keywords: Radiotherapy, Tracking

Objective In radiation therapy, respiration induced motion of tumors in or near the thoracic region compromises the treatment accuracy. For this reason, several techniques have been introduced over the past few decades to compensate or manage respiratory motion. The most recent technique proposed is real-time tumor tracking (RTTT), a technique in which the position of the radiation beam is adjusted in real-time to assure that the tumor remains aligned with the beam at all time. To allow for this kind of treatment, the position of the internal target has to be known at all time, while maintaining a low dose to the surrounding healthy tissue. This can be achieved using a surrogate breathing signal and a correlation model that correlates the surrogate signal to the motion of the target.

There are currently two systems clinically in use that are capable of performing RTTT, the SRTS Cyberknife system and the Vero SBRT system. In both systems a different correlation algorithm is implemented. The objective of this study is to compare these algorithms in terms of geometrical accuracy.

Methods Both correlation models were simulated off-line and log files from a patient population of 10 patients, previously treated on the Vero system with RTTT, were used. With each model, internal target positions were estimated and compared with the actual target positions to compute the correlation model errors. This was done for three different intra-fraction update scenarios of the correlation model and in each scenario the errors were compared.

Results In two of the three update scenarios, a significant difference was found between the accuracy of the Vero model and that of the Cyberknife model. It was observed that the difference in build and the difference in dimension of the input data to the model, as well as the difference between the three update scenarios, influenced the accuracy of the two models. It could also be observed that variations of the breathing cycle over time significantly decreased the accuracy of both models.

Conclusion Despite the higher complexity of the Cyberknife model, the Vero model performed slightly better for our patient population in two of the three update scenarios. The higher accuracy is most likely due to a more extensive data set to build the model and the higher dimension of the input data to the model during treatment. We can also conclude that for both models, some form of intra-fraction updating is required to compensate for variations in the correlation between the internal target motion and surrogate signal over time.

Samenvatting

A comparison of two clinical correlation models for dynamic tumor tracking with a focus on geometrical accuracy

Jennifer Dhont, Master in Biomedical Engineering, 2014

Keywords: Radiotherapy, Tracking

Objectief In de radiotherapie kan ademhaling-geïnduceerde beweging van tumoren in of nabij de thorax de accuraatheid van de behandeling in het gedrang brengen. Voor die reden zijn de voorbije jaren nieuwe technieken voorgesteld om ademhalingsbeweging te compenseren of onder controle te houden. De meest recente techniek is real-time tumor tracking (RTTT), een techniek waarbij de positie van het stralingsveld in real-time wordt aangepast zodat de tumor zich steeds in dit veld bevindt. Om deze techniek mogelijk te maken dient het systeem ten allen tijde de locatie van het target te kennen, terwijl de stralingsdosis in het gezond weefsel van de patiënt zo laag mogelijk blijft. Om dit te verwezenlijken wordt gebruik gemaakt van een surrogaat ademhalingssignaal en een correlatiemodel om het surrogaat signaal te correleren met de beweging van het target.

Momenteel zijn er klinisch twee toestellen in gebruik die RTTT kunnen toepassen, het Cyberknife SRTS systeem en het Vero SBRT systeem. Beiden beschikken over een verschillend correlatie algoritme. Het doel van deze studie is deze algoritmes te vergelijken op basis van geometrische accuraatheid.

Methode Beide correlatiemodellen werden offline gesimuleerd en logbestanden van een populatie van 10 patiënten die eerder met RTTT op het Vero systeem werden behandeld werden toegepast. Met elk model werden tumor posities bepaald die konden vergeleken worden met de werkelijke tumor posities om zo correlatiefouten te berekenen. Dit werd toegepast in drie verschillende update scenario's en in elk scenario werden de fouten met elkaar vergeleken.

Resultaten Bij twee van de drie update scenario's kon een significant verschil worden vastgesteld tussen de accuraatheid van het Cyberknife model en dat van het Vero model. Er kon ook worden vastgesteld dat het verschil in opbouw en het verschil in dimensie van de input data in het model tijdens de behandeling, als ook de verschillen tussen de drie update scenario's, invloed hadden op de accuraatheid van de modellen. Een laatste vaststelling was dat ook de variaties van de ademhalingscyclus een significante invloed hadden op de accuraatheid.

Conclusie Ondanks dat het Vero model eenvoudiger is dan het Cyberknife model, presteerde het Vero model toch beter in twee van de drie scenario's. De hogere accuraatheid is hoogst waarschijnlijk te wijten aan de grotere dataset om het model te bouwen en de hogere dimensie van de input data in het model tijdens de behandeling. Een laatste conclusie is dat voor beide modellen een update methode noodzakelijk is om te compenseren voor variaties van de ademhalingscyclus die optreden na verloop van tijd.

This work was selected through peer review process for oral presentations at the 26^{th} Annual meeting of the Belgian Hospital Physics Association (February 2014, Louvain-La-Neuve, Belgium) and the 33^{th} Annual ESTRO meeting (April 2014, Vienna, Austria).

Contents

List of Figures

List of Tables

Glossary

1D One-dimensional.

2D Two-dimensional.

4D Four-dimensional.

CBCT Cone-beam computed tomography.

CCD Charge-coupled device.

CK Cyberknife.

CM Correlation model.

CNS Central nervous system.

CTV Clinical target volume.

DMLC Dynamic multi-leaf collimator.

DNA Deoxyribonucleic acid.

DT Dynamic Tracking.

E95 95^{th} percentile.

EPID Electronic portal imaging device.

FF Full Fluoroscopy.

IGRT Image guided radiotherapy.

IMRT Intensity modulated radiotherapy.

ITV Internal target volume.

kV kilo-Volt.

LINAC Linear accelerator.

MLC Multi-leaf collimator.

MRI Magnetic resonance imaging.

MV Mega-Volt.

NSCLC Non-small cell lung cancer.

NTCP Normal tissue complication probability.

PTV Planned target volume.

RMS root mean square.

RTTT Real-time tumor tracking.

SBRT Stereotactic body radiation therapy.

SI Superior-inferior.

SRTS Synchrony Respiratory Tracking System.

US Ultra-sound.

VMAT Volumetric Modulated Arc Therapy.

Chapter 1

Introduction

1.1 Cancer and Radiation Therapy

Cancer comprises over one hundred different pathologies that are all characterized by abnormal and uncontrolled cell growth. Since the beginning of the 21st century, the number of cancer incidences has increased on both a national and international level. In 2011, 64 thousand Belgians and 14 million people worldwide were diagnosed with some form of cancer [1][2]. On a worldwide scale it is estimated that, if this trend of increasing incidences continues, by 2030 20 million new cancer diagnoses will be made annually [3]. Besides a high incidence rate, cancer is also the second most important cause of death in both Europe and America. In Belgium, lung-cancer in men and breast-cancer in women are the number one causes of death for the middle-aged population [4].

Fortunately, over the past few decades great progress has been made in treating and curing cancer [5]. Diagnostic and surgical techniques but also chemotherapeutic treatments and radiation therapy have advanced tremendously.

Radiation therapy or radiotherapy is a treatment technique in which ionizing radiation is used to control and/or kill cancerous cells in the body. Ionizing radiation comprises all types of radiation that have sufficient energy to create ions in matter by removing electrons from a stable atom. As such, damage can be inflicted on a sub-cellular level. The most significant damage and the damage that one wants to create in cancerous cells is that of the DNA (Deoxyribonucleic acid) [6][7]. DNA damage by ionizing radiation can either occur direct or, more frequently, indirect. Direct DNA damage occurs when the ionizing radiation impinges on a part of the DNA molecule and ionizes the DNA molecule at that location. Indirect damage

occurs when the ionizing radiation ionizes the water molecules around the DNA molecule and produces radicals which chemically damage the DNA. When the DNA of a cell is damaged, different scenarios can occur depending on the severity of the damage. A correct repair of the DNA is possible and most probable in the case of single stranded breaks as the complementary strand is still intact. A wrongful repair can result in either a non-viable mutation which will lead to apoptosis or in a viable mutation which can proliferate and lead to possible malignancies or birth deficiencies depending on the cell type.

The interaction mechanism between ionizing radiation and healthy or cancerous tissue is very similar as ionizing radiation is not selective for the type of exposed tissue [8]. However, the biological consequences after radiation damage differ among different types of healthy tissue and tumor cells. External factors such as heat and the environment of the cells can also influence the cells radio-sensitivity. These differences can be used during radiotherapy to spare healthy adjacent tissue while still sufficiently treating the target. A way to differentiate the biological response of healthy tissue and malignant cells is by fractionating the delivery of the radiation dose. Dividing the necessary therapeutic dose in small amounts has shown to induce fewer short- and long-term complications in the surrounding healthy tissue. This is believed to be the consequence of a better repair mechanism for sub-lethal damage in normal tissue compared to cancerous tissue. Fractionation can thus result in a greater cumulative damage in tumor cells than in the healthy nearby tissue, leaving the possibility to treat tumors with ionizing radiation.

Standard fractionation was long considered the only safe option to treat cancerous tissue with ionizing radiation as single large doses lead to the uncontrollable risk of acute and long-term complications of the healthy nearby tissue. However, mid-twentieth century, a new approach was successfully introduced called stereotactic radiosurgery (SRS) which broke with the conventional fractionation scheme. It was a technique that used large-dose single sessions of radiation, delivered to the central nervous system (CNS) [9]. The main factor that allowed this kind of administration was the technique with which the radiation was delivered. Techniques taken from stereotactic surgery using stereotactic frames fixed to the patient with an attached 3D coördinate system enabled a high spatial accuracy. The dose was delivered with a large number of non-overlapping beams, each carrying a relatively low weight and consequently causing little damage to the normal tissue they crossed to reach the deep seated tumor. The cumulation of these multiple beams created a hot-spot in the tumor and a steep dose fall-off was generated outside the target towards adjacent healthy tissue. Delivering these higher, more potent doses resulted in a lower re-population of the cancerous cells, resulting in a more effective treatment.

The technique of SRS and hypo-fractionation was later extended with the help of newly developed treatment devices targeting the tumor with a high spatial accuracy. The so-called stereotactic body radiotherapy (SBRT) is now being applied to many other treatment cites besides the CNS, where the four main requirements of hypo-fractionation are met; effective patient immobilization, precise target localization, accurate dose delivery and the ability to produce a steep dose fall-off towards the surrounding healthy tissue [10][11].

1.2 Respiratory-induced motion in radiotherapy

The average respiratory rate of a healthy human being at rest is around 12 to 16 breaths per minute. During respiration, the thoracic region expands to induce an expansion of the lungs so air is drawn in. The process of respiration not only causes the lung tissue to move, it also causes the displacement of several other organs situated near the thoracic region such as the liver and pancreas [12][13]. A tumor located in the lungs or in these nearby organs often exerts a similar motion as the tissue it is surrounded with. Non-small cell lung cancers (NSCLC) located closely near the diaphragm for example have shown to make displacements in the order of centimetres with each breath.

As a consequence of this intra-fraction motion, treating tumors in or near the thoracic region with static external beam radiotherapy can cause serious complications. Not only is there a higher probability that the tumor volume is insufficiently irradiated, there is also an increase in the probability of irradiating surrounding healthy tissue. The fact that these kind of pathologies, accounting for about a fifth of all new cancer diagnoses, are also often inoperable or unresponsive to other treatments gives them a particularly poor outcome with an overall 5-year survival rate of 6-16% [14][15]. However, radiation therapy has shown to improve the local control probability of static non-small cell lung cancers and so there has been a significant amount of research around accurately treating moving targets with ionizing radiation.

1.3 Current technologies in radiotherapy for treating moving targets

The variety of radiotherapy techniques to treat moving targets situated in the thoracic region only is quite extensive. The treatment option is mainly chosen based on the amount of movement, the location of the target and the options available in the medical center. An overview of all motion management and motion compensating techniques is given next. It should be noted that all techniques are classified in the field of stereotactic body radiation therapy (SBRT) and image-guided radiation therapy (IGRT) [16][17].

Motion management using margins - the ITV approach

In general, a safety or treatment margin is applied to the clinical target volume (CTV) to ensure that the complete CTV receives the prescribed dose. This margin is applied to account for inaccuracies during imaging, treatment planning and delivery. It does not only take into account the uncertainties in position, size and shape of the target, it also includes variations in patient and beam positions. The CTV plus safety margins result in the planned target volume or PTV [18]. Taking into account substantial movement of the target can be done using an additional margin in the PTV definition, resulting in the internal target volume or ITV. This kind of motion management can be applied when the respiratory-induced motion of the target is in the order of a few millimetres. It is not advised to be used for more extensive motions as the amount of healthy tissue irradiated can become significant, increasing the normal tissue complication probability (NTCP) [19][20].

Abdominal compression techniques

Abdominal compression is a technique that is used to reduce diaphragmatic motion during treatment delivery. A stereotactic body frame with a pressure plate that is pressed against the abdomen of the patient is applied to immobilize and position the patient and to induce forced shallow breathing [21][22]. As such the thoracic motion is limited, while still allowing the patient limited normal breathing. Markers and other indications on the stereotactic body frame can also be used to reproduce the patient position at each fraction, however, imaging is still essential before each treatment fraction to ensure correct tumor positioning. Studies have shown that this approach is capable of reducing the motion of certain tumors to

under 5 mm [23]. However, depending on the location of the tumor and the principal directions of movement, certain tumor motions can not be compressed under acceptable levels. The reproducibility of the compression system and the level of comfort of the patient during treatment are two other adverse characteristics of this motion management approach.

Breath-hold and free-breathing gating techniques

Gated radiotherapy is an image-guided radiotherapy technique in which there is only an administration of radiation at a selected moment in the breathing cycle. Meaning that only at a certain phase in the breathing cycle the radiation beam is on and the target is irradiated, while at all other breathing phases the beam is off. The position and width of the treatment window or gate are determined based on the periodic movement of the target and is a trade-off between time efficacy and irradiating nearby healthy tissue. Localisation of the target with respect to the treatment window is performed using external or internal respiratory signals, with or without fiducial markers implanted in or near the malignant tissue [24][25]. Although the additional dose of radiation in image-guided gating is not that high, it is still a concern to some clinicians. The less invasive gating option using an external surrogate signal is therefore preferred.

As the beam is not always on while the patient is treated, the treatment time of one fraction can be prolonged remarkably compared to non-gated treatments. This makes it a sub-optimal treatment technique in terms of economical factors and patient-comfort. A combination of gating with voluntary breath-hold or shallow-breathing techniques can be applied to optimize the duty cycle and decrease treatment time [26][27][28].

Motion compensation through tracking

The most recent technique for respiratory-induced motion compensation is real-time tumor tracking (RTTT) [29]. During RTTT the position of the radiation beam is continuously adjusted in real-time so that the target remains in the same relative position with respect to the beam. As such, highly conformal treatment plans can be applied to a moving target. Compared to gating, the beam remains on at all time and the patient is allowed to breath freely during each entire treatment fraction.

Real-time tumor tracking can be performed in several ways. Assuring that the target and the beam remain aligned at all time can be achieved by either mo-

ving the patient support system with respect to a static beam or by moving the beam with respect to a stable patient position. The first has shown to be feasible, however, comfort of the patient in such case is questionable and it is not certain that the target position is not influenced by this kind of constant patient movement. The latter has shown to be feasible with real-time adjusted moving linear accelerators and is clinically in use right now in two different systems. Another approach is with the use of a dynamic multi-leaf collimator (DMLC). However, DMLC tracking has shown to be very dependent of the orientation of the tumor motion with respect to the leaf motion direction. The maximum leaf speed is also a limiting factor in DMLC tracking and the combination of intensity modulated radiation therapy (IMRT) with DMLC might be an unrealistic demand on the MLC [30].

A main requirement in applying tracking to treat moving tumors is to know at all time where the tumor is located. Although respiratory motion can be assumed to be approximately periodic, realistically it is not. Continuously imaging the target and a part of its surrounding to determine its location would be the most straightforward approach. However, as treatment times per fraction can go up to half an hour and more, the imaging dose to the patient, which would mainly consist of dose to the skin and other healthy tissue, would reach unacceptable levels [31][32]. Especially because a high temporal resolution of the target position is required as input to a prediction algorithm implemented to compensate for unavoidable system latency. Therefore, instead of imaging, a surrogate breathing signal that shows a strong relation with the motion of the target has to be used to determine the target positions. This indirect way of tumor localization entails an increase in uncertainty of the dose delivery during tracking.

1.4 Purpose of this work and thesis outline

The use of a surrogate breathing signal in real-time tumor tracking requires the application of a motion model to correlate the surrogate signal to the motion of the target. Many different motion models, more often called correlation models, have been proposed and sometimes tested on clinical data.

Two of these correlation models are those implemented in the only two current systems clinically performing RTTT, the Vero SBRT System and the Cyberknife (CK) Robotic Radiosurgery system. Both systems are equipped with a correlation model (CM) to estimate the internal target positions based on an external surrogate signal. Both CMs have been sufficiently described and evaluated in the past [33][34][35][36][37], but to our knowledge they have never been compared in the same clinical setting. The purpose of this work is to simulate, validate and compare both correlation models in terms of geometrical accuracy, for a population of patients previously treated with RTTT on the Vero SBRT system.

In Chapter 2, the use of correlation models is further explained, together with a general introduction to motion models that could be used in real-time tumor tracking. Linear and polynomial correlation models will be discussed in more detail as the Cyberknife and Vero CMs are based on this type of models. After this general overview, both the Vero system and the Cyberknife system are explained, together with an extensive description of their correlation model. Finally, an overview of how the comparison was carried out is described.

All results of the comparison together with the conclusions that could be drawn from them can be found in chapter 3 and 4, respectively. In Chapter 5, a discussion is started concerning the outcome of this comparison, its limitations and some future work that could be carried out to continue this study on motion models for real-time tumour tracking.

Chapter 2

Correlation Models

2.1 General overview

Correlation models are used in both gating and real-time tumor tracking. They are applied to define the relation between a surrogate breathing signal and the motion of the internal target [38]. It is a mathematical algorithm that is built before the start of each treatment by using both internal target and surrogate data. During the treatment, the surrogate signal is used as input to the correlation model (CM) and tumor positions are calculated as output. Due to the strong interaction, a correlation model might be considered as the combination of 3 equally important components; a surrogate signal, a correlation algorithm and a motion signal of the target [39].

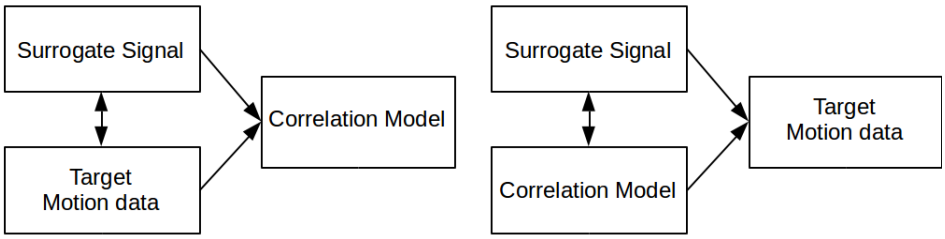

Figure 2.1: The correlation model is defined before treatment using both the surrogate breathing signal and the target motion data. During treatment, the surrogate breathing data is used as input to the correlation model to determine the target motion.

The surrogate signals proposed in literature to be used in tumor tracking come in a wide variety. From the use of spirometry [40][41][42], to markers indicating the

displacement of the chest or abdomen, to the use of ultrasound, surface monitoring or magnetic resonance imaging (MRI). The main requirement is that the surrogate signal shows a strong correlation with the motion of the target, together with the ability to be easily acquired and processed with a high temporal resolution, causing no harm to the patient.

The motion data of the internal target can also be acquired in a number of ways. However, these are most often simply different imaging modalities. Only one exception was proposed in which the use of an electromagnetic signal was described [43]. The most common imaging modalities are X-ray imaging, CT and cone-beam CT (CBCT) and MRI. Due to poor image resolution compared to CT, ultra-sound (US) has not generally been used to obtain motion data, however, it has been proposed as a surrogate signal due to its high temporal resolution and the use of sound waves instead of ionizing radiation [44][45].

The choice of correlation model strongly depends on the surrogate signal, the motion data and how accurate the user wants to model the relationship between both and over which period of time. It is also a question whether the model should be patient-specific or used cross-population. Although the latter seems fairly impossible due to strong variations in breathing motion among patients, not to mention the differences in pathology, they have been proposed by Ehrhardt et al., Fayad et al. and others [46][47][48][49]. Also, as mentioned before, breathing motion might be assumed to be periodical, but realistically it is not. There are both intra-cycle and inter-cycle variations that might influence the accuracy of the correlation model. Intra-cycle variations or variations occurring during one breathing cycle are for example the different trajectories during inhalation and exhalation, also called hysteresis. Inter-cycle variations or differences between several breathing cycles are for example a more extensive maximum exhalation or inhalation compared to the previous cycles. Both phenomena can be observed in Figure 2.2.

In general, any direct correlation model (f) can be represented as follows

$$\mathbf{T} = f(\mathbf{S}) \tag{2.1}$$

In which $\mathbf{T}$ is the 3D target position estimation and $\mathbf{S}$ is the surrogate signal. The capabilities of the correlation model strongly depend on the surrogate signal and the motion data. Intra-cycle variations for example can only be modelled if the surrogate signal is able to distinguish between in- and exhalation, if the motion data samples both breathing phases and if the model can make different estimations for in- and exhalation. If for example, f is one linear function (f :

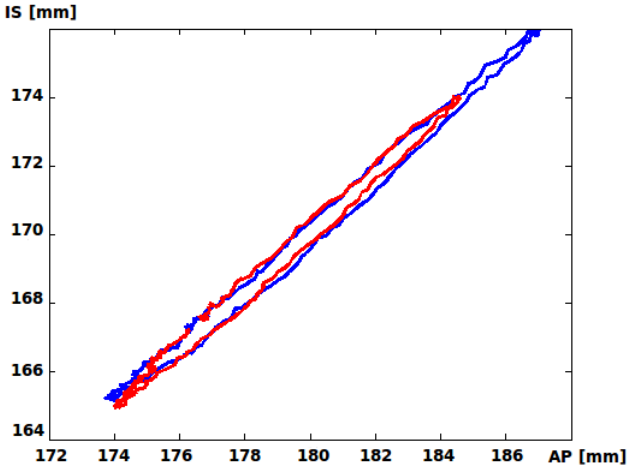

Figure 2.2: Two breathing cycles represented by the surrogate breathing signal in the vertical and medial direction of patient 1. The presence of hysteresis can be observed, as well as a difference in maximum inhalation between the two subsequent cycles.

$\mathbb{R} \mapsto \mathbb{R}^3$), it is not capable to make different estimations for in- and exhalation as each complete breathing cycle is modelled as a straight line. In this case the relation between the target position and the surrogate signal is represented as follows

$$\mathbf{T}_i = A_{1i}\mathbf{S} + A_{2i} \quad i = 1, 2, 3 \tag{2.2}$$

A_{1i} and A_{2i} are coëfficients determined when the model is fitted to the target motion and surrogate data. The linear correlation model is robust and easy and relatively fast to build. Although it doesn't represent a very realistic breathing signal, for some patients it models the relation between the motion of the surrogate signal and the target motion well and provides accurate tracking results. However, the majority of patients requires a more complex model.

This model can be extended to higher order polynomials and/or a combination of multiple polynomials of similar or different degree. Depending on the degree of the polynomials, the number of polynomials used and the type and dimension of the surrogate data, a breathing loop can be modelled and a distinction can be made between in- and exhalation. These models offer more flexibility and therefore model the internal target motion of most patients with a higher accuracy compared to linear models.

It was already briefly mentioned that the correlation model is built before the treatment. Clinically, this is done before each treatment fraction. The motion data and surrogate data are obtained in a synchronized way, and the correlation model is fitted to this data. In most cases, a least square optimization process is applied. The amount of data obtained for the build of the correlation model is another factor that influences the accuracy of the model. If the data is only obtained during one breathing cycle, inter-cycle variations are difficult to be included in the model, depending on the model. The same goes for intra-cycle variations if the data is only obtained during inhalation or exhalation. A dataset obtained over several breathing cycles at a sampling rate high enough to sample intra-cycle variations is therefore most optimal. However, it should also be noted that models requiring a rather large amount of training data, such as models based on neural networks or single-vector regression, are clinically unfavourable or even inapplicable, although they might model the breathing motion with a higher accuracy [50][51][52].

A final point that should be addressed concerning correlation models is the implementation of updates. Most models correlate the surrogate motion to the target motion with a so-called black-box approach in which only the input and output data are considered, but not the actual underlying physical system and its behaviour and characteristics. As such, not all variations of the correlation between the surrogate signal and the target motion can be handled by the model, meaning that they can cause an increase in modelling error. Therefore, most models have to be updated regularly during one treatment fraction, if not completely rebuild. The downside to updating and rebuilding is that these can increase the treatment time significantly, influencing patient comfort which in turn can influence the patient's breathing. Coaching methods based on audio-visual signals, or one of the two, have shown to decrease these inter-cycle variations, but including them in the correlation model itself would be a more straight-forward and less patient-dependent approach.

For that reason the physico-mechanical model proposed by Wilson and Meyer [53] is definitely worth mentioning as an example that is not based on polynomial functions. This model approaches the problem of unknown target location by modelling lung tumor motion, or more precisely the spatio-temporal relationship between chest motion and tumor motion, as a contact problem of elasticity theory by describing the physiology using an arrangement of springs and dash-pots. Solving the model is done by solving three coupled non-dimensional differential equations through an optimization routine. While the mathematical model is general for any tumor in the lung, the parameters are obtained specifically for every patient and every tumor-location with a set of 2 min training data during the optimization process. Adjustments were done by Ackerley et al. [54] who simplified

the three-dimensional model to a one-dimensional model, shown possible again by Wilson and Meyer in case the tumor motion is predominantly in endangeringthe superior-inferior (SI) direction. The most remarkable feature of these physico-mechanical models was described by Ackerley et al. when they tested their model on clinical data of 10 patients previously treated on a Mitsubishi Real-Time Radiation Therapy System with up to ten fractions and 4 beams per fraction. It was shown in Stage 2 of their study that not only was the physico-mechanical model able to model the correlation well, it was also possible to use the parameters, obtained at the first fraction, throughout the entire treatment session with only a small increase of the tracking error. Meaning that this model is not only capable of taking into account inter-cycle variations, but also inter-fraction variations.

What follows in the next two sections is a more detailed description of the two clinical correlation models that were compared in this study, together with a description of the systems in which they are implemented.

2.2 Dynamic tumor tracking on the Vero SBRT System

One of the correlation models that was evaluated in this study is that of the Dynamic Tracking (DT) on the Vero SBRT System. The next few sections contain a short description of the Vero system and the workflow during RTTT, together with an extensive description of its implemented correlation model, the surrogate signal and the acquisition of the target motion data.

The Vero SBRT System

The Vero SBRT System is a joint product of BrainLAB (BrainLAB AG, Feldkirchen, Germany) and MHI (Mitsubishi Heavy Industries, Tokyo, Japan). It is a four-dimensional IGRT system designed with a gimbaled X-ray head [56][57]. It consists of a small and light 6 MV C-band linear accelerator (LINAC) with a fast MLC mounted on an O-ring gantry, permitting a full 360° rotation. The MLC allows for a maximum field size of 150 by 150 mm, with a maximum leaf speed of 50 mm/s. Two orthogonal gimbals support the linac-MLC assembly which can be driven to rotate in a perpendicular and parallel direction with respect to the couch. This mechanism enables the possibility to perform RTTT of moving targets, decoupled from the DMLC intensity modulation. The maximum excursion of the

beam axis is 4.4 cm (2.5°) at the isocenter plane in both perpendicular and parallel direction. Besides an electronic portal imaging device (EPID) for MV imaging, the Vero system is also equipped with two on-board orthogonal kV radiograph and fluoroscopy devices that are attached to the O-ring gantry at 45° from the MV beam axis. This imaging system allows for cone-beam CT and simultaneous acquisition of orthogonal X-ray images at 15 frames/second. An ExacTrac (BrainLAB AG, Feldkirchen, Germany) automated infra-red (IR) reflective marker-based optical tracking system with a 5 degrees of freedom robotic treatment couch is also integrated into the Vero system. This optical tracking system is used for both pre-treatment patient positioning and for the acquisition of the surrogate breathing signal of the patient.

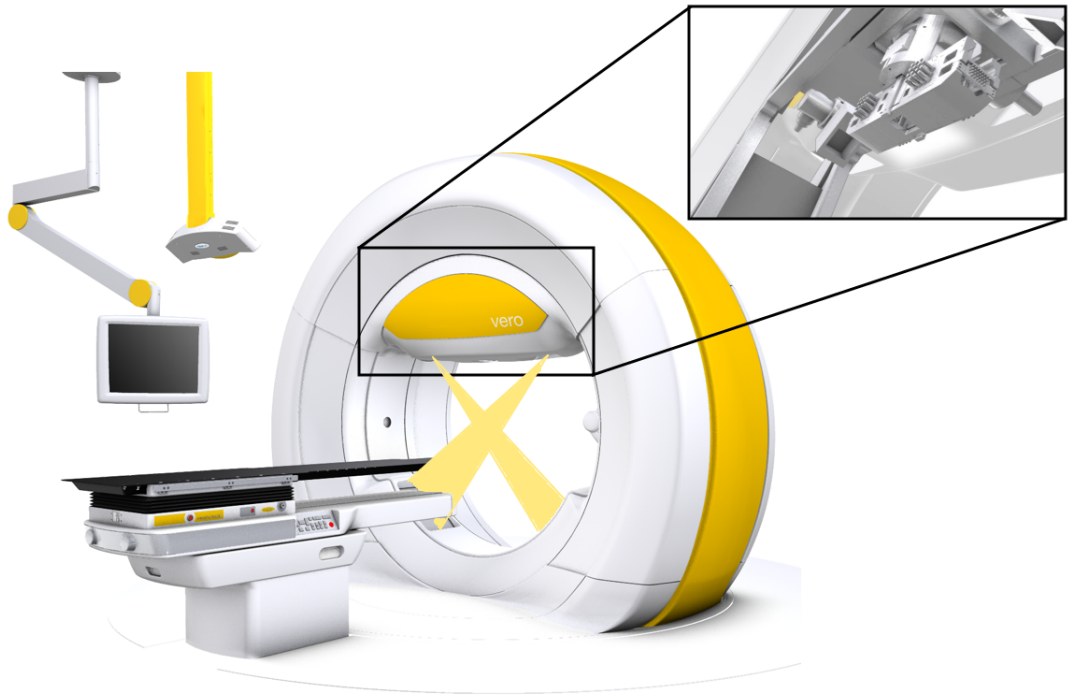

Figure 2.3: The Vero SBRT System with robotic couch and ExacTrac infra-red camera with a detailed representation of the gimbaled X-ray head and two orthogonal kV imaging systems. *Image courtesy of Brainlab AG (BrainLAB AG, Feldkirchen, Germany).*

A treatment fraction with real-time tumor tracking starts with pre-positioning the patient laying on the robotic treatment couch. This is done by placing reflective IR markers on several stable parts of the skin, e.g. sternum and rib-cage, which were also placed during the acquisition of the planning CT. After initial pre-positioning, a cone-beam CT (CBCT) or orthogonal kV X-ray set is acquired for

image guided positioning based on bony structures. A Visicoil (IBA, Louvain-la-Neuve, Belgium) gold marker, see Figure 2.4, is used as a fiducial and is implanted in or near the tumor several days prior to the planning CT to ensure a stable position during the course of treatment. Once the patient is set up close to the isocenter, the 4D modelling is initiated. First, the surrogate breathing signal is obtained from an additional set of IR markers positioned on the higher abdomen or chest, moving with the breathing cycles of the patient. Once the breathing signal is picked up, a synchronized acquisition of orthogonal X-ray images begins. This can take between 20 to 40 s, covering at least 4 breathing cycles. A region-growing based detection algorithm is used to localize the internal fiducial marker with an accuracy of 0.5 mm or lower. The surrogate breathing signal and internal target positions are then used to fit a correlation model that is used during treatment to estimate the internal target position based on the surrogate breathing signal. A linear forward prediction algorithm is also implemented to compensate for a known 50 ms latency. The operator allows for a small level of deviation between the estimated and actual target position, defined by a tolerance circle or error threshold. The motion of the gimbals is recorded at a rate of 50 Hz in a log file. Combining this data with the detected internal target positions from the monitoring X-rays (0.5 frames/s) enables tracking error calculations [37][55].

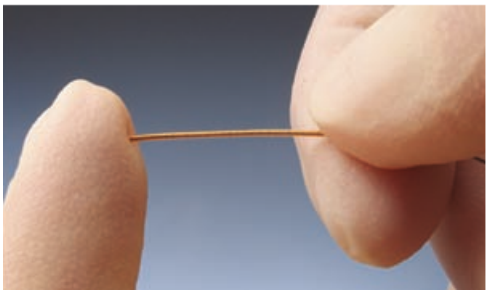

Figure 2.4: Magnified image of a visicoil^{TM} implantable marker for real-time tumor tracking. *Image courtesy of IBA (IBA, Louvain-la-Neuve, Belgium).*

Important to note, with respect to the work performed in this study, is that before and during each treatment fraction log files are generated. These log files store, with the appropriate time stamps, the internal target positions and external marker positions, as well as the internal target estimations by the correlation model implemented in the Vero SBRT system. The estimated target positions and those obtained through imaging are both three-dimensional while only the vertical positions of the external markers are stored.

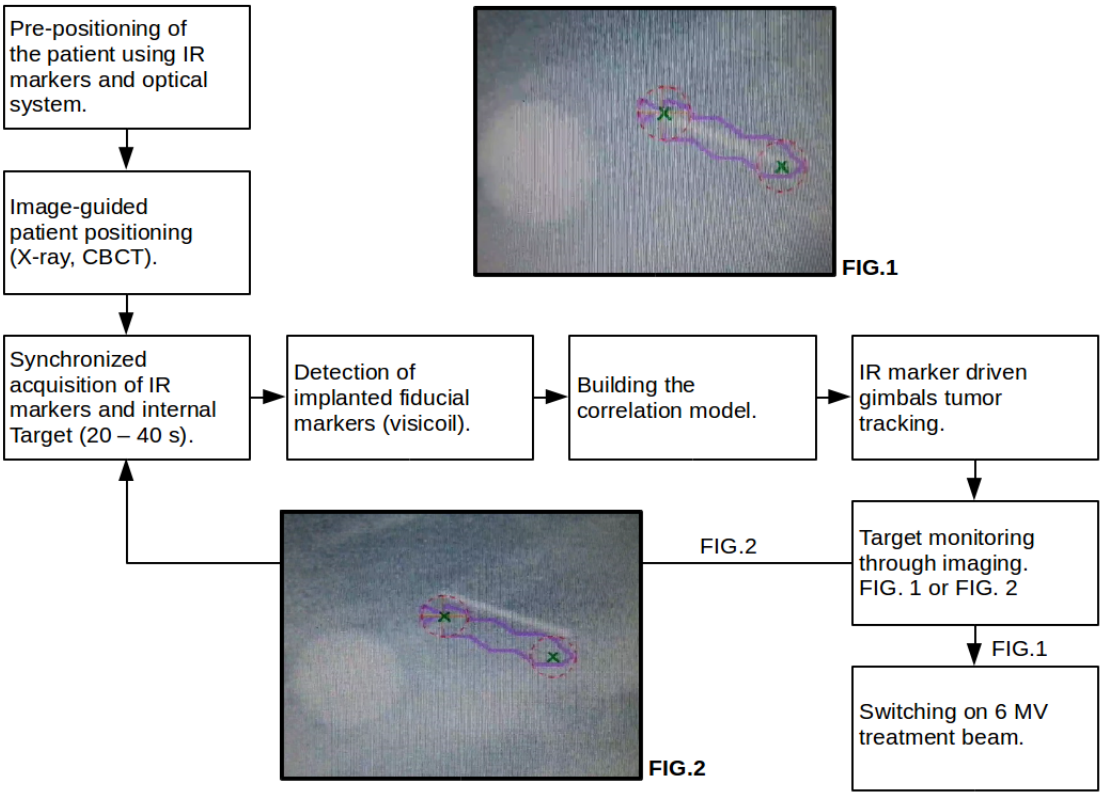

Figure 2.5: Schematic representation of the Dynamic Tracking workflow on the Vero SBRT System. The treatment beam is switched on only if the tracking error is smaller than 3 mm in any direction, represented by a tolerance circle at both ends of the implanted visicoil marker. *Image courtesy of Depuydt T., et al. [32]*

The correlation model

As mentioned in the description and workflow of the Vero SBRT system, an optical tracking device is present in the treatment room that is capable of tracking IR markers positioned on the chest or higher abdomen of the patient. This signal serves as the surrogate breathing signal that is applied as input to the correlation model. The number of markers that is used is preferably five or six, however, possible ambiguities resulting from the 3D reconstruction based on unsorted points match in combination with intermittent loss of signal due to reflection issues or lack of motion can decrease this number to four in the clinic.

The motion data of the internal target consists of the registration of the position of an implanted visicoil marker in or near the target. The implantation

is rather invasive, but markerless tracking is not yet implemented in the system. The visicoil positions are acquired through fluoroscopy using the orthogonal X-ray imaging devices at a frequency of 11 Hz to build the model and at a frequency of approximately 2 Hz for validation during the treatment.

The correlation model clinically implemented in the Vero SBRT system to conduct dynamic tracking is a second-degree polynomial function (F) with the 1-dimensional external marker positions (x) and external marker derivatives of the positions ($\dot{x}$) as variables [58].

$$F(x, \dot{x}) = ax^2 + bx + c + d\dot{x}^2 + e\dot{x} \tag{2.3}$$

The velocities are calculated by a weighted [1,...,1,3,5] fit of a first-degree polynomial function through the last 15 IR data-points. Making direct use of the velocity in the correlation model enables modelling possible hysteresis of the breathing signal as the velocity is opposite in sign for the different breathing phases. The parameters a, b, c, d and e are computed before each treatment fraction through a least-square gradient descent optimization process using a dataset containing both internal target positions and external marker positions obtained through at least 20 s full fluorosopy (FF) at 11 Hz. A model is fitted for every external marker and every direction of motion of the target. The use of 5 external markers for example leads to 15 different correlation functions. The final target position estimation in each direction is obtained by averaging over all markers.

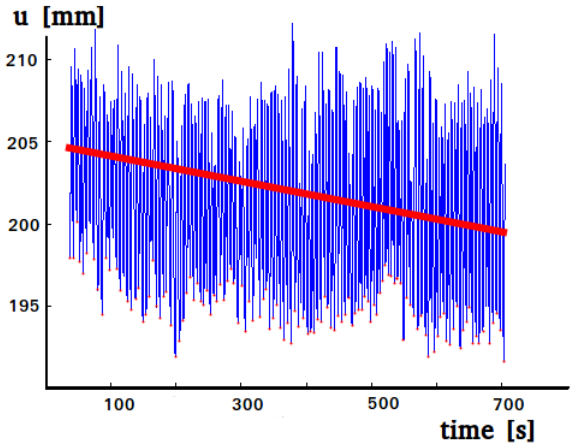

Figure 2.6: Baseline drift of the respiratory signal of patient 4 during fraction 1, over a time-range of 11 min.

The correlation model is used throughout the entire treatment fraction unless the tracking error exceeds a threshold of 3 mm in any direction. At that moment the treatment is manually interrupted and a completely new data set is obtained to rebuild the correlation model. An important contribution to the increase in tracking error over time is drift of the surrogate breathing signal and/or internal target, the first shown in Figure 2.6. Baseline drift is the frequently occurring phenomenon where the average position of the external marker lowers with time, while the range of the target motion remains the same. This baseline drift of the external markers is partially due to the patient who often relaxes after a few minutes on the treatment table and relaxes his/her muscles, and partially due to the internal anatomy that changes a little during that same time period after laying down in a supine position. As a result, the target estimations by the correlation model are also shifted compared to the actual target positions and a larger tracking error is established. Other changes in the breathing cycle, such as variations in the amount of hysteresis or an exceptionally deep in- or exhalation can also cause the random error to exceed the predefined threshold. Important to note is that it is always up to the system control operator whether or not to rebuild the correlation model.

2.3 Tumor tracking on the SRTS Cyberknife System

The second correlation model that was evaluated in this study is that of the Synchrony Respiration Tracking system integrated in the Cyberknife Robotic Radiosurgery System. What follows is a short overview of the Cyberknife system, together with an extensive description of its clinically implemented correlation model.

The Cyberknife Robotic Radiosurgery System

The Cyberknife Robotic Radiosurgery System by Accuray Incorporated (Accuray Incorporated, Sunnyvale, CA, USA) was initially designed to perform frameless stereotactic radiosurgery. It was only later that the Synchrony Respiratory Tracking System for motion compensation was included. The Cyberknife system consists of a miniature lightweight 6 MV X-band linac mounted to an industrial multi-jointed robotic arm. The robotic arm moves freely and is capable of aiming the radiation beam with six degrees of freedom, enabling the delivery of many

independently targeted (non-isocentric) and non-coplanar treatment beams. To ensure optimal dynamic tracking capabilities, the weight of the beam-line was reduced to below 200 kg. In the compact design the bending magnet and the flattening filter was removed and the collimation devices were limited to fixed cones of various diameters. However, a first update allowed the implementation of an Iris diaphragm-like collimator while a later update allowed the implementation of an MLC. Two rigidly fixed off-board X-ray imaging systems are orthogonally configured in the treatment room at 45° and 315° from the vertical axis. This imaging system is used to acquire the internal target positions, either with or without the help of implanted fiducial markers. The surrogate breathing signal is acquired from a set of multiple light emitting diode (LED) markers typically fixed on a tightly fitting vest the patient wears during treatment, see Figure 2.8. The marker motion is obtained at approximately 30 Hz, with a continuously measuring stereo-camera containing 3 CCD camera's [59][60].

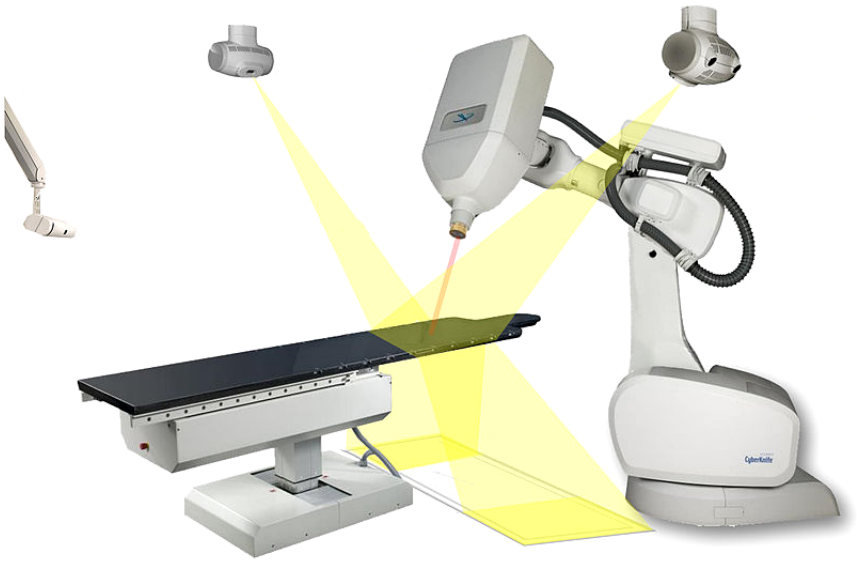

Figure 2.7: The Cyberknife Robotic Radiosurgery System with off-board orthogonal kV imaging systems, robotic couch and stereo-camera. *Image courtesy of Accuray Incorporated (Accuray Incorporated, Sunnyvale, CA, USA).*

A first pre-positioning of the patient is done using the X-ray image guidance system and an adjustable treatment table. A final fine alignment of the target relative to each treatment beam is done by moving the robotic arm relative to the patient. A set of both internal target data and external marker data is used to build the correlation model before each treatment fraction. The system latency of the Cyberknife/Synchrony system is 200 or 115 ms, depending on the version,

which is compensated with an adaptive prediction algorithm [61]. Periodically acquired X-ray images are used to calculate the tracking error and update the target position estimation algorithms if necessary.

Figure 2.8: LED markers fixed on a tightly fitting vest the patient wears to obtain a surrogate breathing signal during a real-time tumor tracking treatment on the SRTS Cyberknife system. *Image courtesy of Accuray Incorporated (Accuray Incorporated, Sunnyvale, CA, USA).*

The correlation model

As mentioned in the description and workflow of the SRTS Cyberknife system, an optical tracking device is present in the treatment room that is capable of tracking LED markers fixed to a tightly fitting vest the patient wears during treatment, similar to the previously described Vero approach except here the LED markers can be individually identified (each having its own frequency) avoiding the problems related to unsorted point matching. It is this signal that serves as the surrogate breathing signal that is applied as input to the correlation model.

The motion data of the internal target consists of the registration of the position of multiple markers implanted near the target. The marker positions are acquired through X-ray imaging using the orthogonal X-ray imaging devices.

The correlation model implemented in the Cyberknife/synchrony system is built from two second-degree and/or one first-degree polynomial function, with the 1D external marker positions (x) as the only variable [51].

$$f(x) = \begin{cases} f_{in}(x) \\ f_{ex}(x) \\ f_{lin}(x) \end{cases}$$

The functions of the correlation model are fitted before each treatment fraction to a dataset containing maximum 15 data pairs of internal target data and external marker data, obtained over several breathing cycles [62]. One quadratic function is fitted to all data points obtained during inhalation (f_{in}) and one is fitted to all data points obtained during exhalation (f_{ex}). A distinction between both breathing phases is made based on the velocity (v) of the external markers which is opposite in sign for the different breathing phases. So for all data points n:

$$v_n := \frac{x_{n-1} - x_{n+1}}{2\Delta t} \tag{2.4}$$
$$\Rightarrow \begin{cases} n \in inhalation & if \quad v_n \geq -v_t \\ n \in exhalation & if \quad v_n \leq v_t \end{cases}$$

With Δt the sampling period of the surrogate signal (s) and v_t a threshold speed. Not taking v_t equal to zero allows data points at slow speed to be assigned to both in- and exhalation, the reason for which will become clear later on.

The linear function (f_{lin}) is fitted to both in- and exhalation points. The eventual combination of functions to model the breathing loop (bi-quadratic, quadratic-linear or only linear) is chosen depending on the modelling error (e_m), slightly favouring the more simpler models.

$$e_m = \sqrt{\frac{\sum e_i^2}{n-m}} \tag{2.5}$$

With e_i being the difference between the modelled point and the actual data point, n being the number of points in the model dataset and m being the number of parameters in the model, i.e. 2 for linear, 3 for quadratic. Important to note is that not all markers should have the same correlation model. The breathing motion of marker one can for example be better modelled with a fully linear function while the breathing motion of marker 2 is more accurately modelled with the bi-quadratic model [34].

In case the bi-quadratic model is most favourable, a breathing loop is modelled using the quadratic inhalation and exhalation function. However, the linear model is also used as a kind of fall-back mechanism. Over time, it is possible that the amplitude of the breathing signal changes, reaching positions that were not registered during the acquisition of the data that was used to fit the functions of the model. The linear function is then applied during treatment to all this external marker data that is located outside the range of external marker data that was used

to fit the three functions ($\mathcal{R}$), plus a predefined margin resulting in $\mathcal{R}^+$, see Figure 2.9. As such, large extrapolation errors are reduced which can be present when using higher order polynomials. A blending mechanism is also implemented to make sure that there is a smooth transition between the three main functions, and a continuous breathing loop is modelled. This is also where the earlier mentioned non-zero threshold speed comes in, allowing a smooth transition between the two breathing phases.

The final bi-quadratic model can be represented as in Figure 2.10 and is made up as follows, with $\mathcal{R}$ being the initial collection of external marker data used to fit the model and σ a parameter determining the amount of overlap between the higher and lower order polynomials in the blending mechanism:

$$r_1 = \min\mathcal{R}, \quad r_2 = \max\mathcal{R}, \quad r_3 = \sigma(r_2 - r_1)$$
$$\mathcal{R}^- = [r_1 + 2r_3, r_2 - 2r_3]$$
$$\mathcal{R}_1^{\leftarrow} = [r_1 + r_3, r_1 + 2r_3], \quad \mathcal{R}_1^{\rightarrow} = [r_2 - 2r_3, r_2 - r_3]$$
$$\mathcal{R}_2^{\leftarrow} = [r_1 - r_3, r_1 + r_3], \quad \mathcal{R}_2^{\rightarrow} = [r_2 - r_3, r_2 + r_3]$$
$$\mathcal{R}^+ = \mathcal{R}_2^{\leftarrow} \bigcup \mathcal{R}_1^{\leftarrow} \bigcup \mathcal{R}^- \bigcup \mathcal{R}_1^{\rightarrow} \bigcup \mathcal{R}_2^{\rightarrow}$$

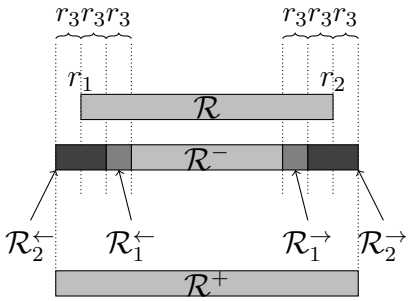

Figure 2.9: Representation of the ranges in the blending mechanism of the Cyberknife correlation model. *Image courtesy of Ernst F., et al. [63].*

$$d_{x,1}^{\leftarrow} = \frac{x - r_1 - 2r_3}{2r_3}, \quad d_{x,1}^{\rightarrow} = \frac{x - r_2 + 2r_3}{2r_3}$$
$$d_{x,2}^{\leftarrow} = \frac{x - r_1 - r_3}{r_3}, \quad d_{x,2}^{\rightarrow} = \frac{x - r_2 + r_3}{r_3}$$

During treatment, the velocity of the external breathing data is calculated to determine if the data points originate from the exhalation (ϵ_{ex}) or inhalation (ϵ_{in}) breathing phase. Based on this division, they are used as input into the associated correlation model to calculate the internal target position.

$$\mathbf{T} = \begin{cases} f_{in}(x) & \text{if } x \in \mathcal{R}^- \bigcap \epsilon_{in} \\ f_{ex}(x) & \text{if } x \in \mathcal{R}^- \bigcap \epsilon_{ex} \\ \left(1 - d^{\rightarrow}_{x,1}\right) f_{in}(x) + d^{\rightarrow}_{x,1} \frac{1}{2}\left(f_{in}(x) + f_{ex}(x)\right) & \text{if } x \in \mathcal{R}^{\rightarrow}_1 \bigcap \epsilon_{in} \\ \left(1 - d^{\rightarrow}_{x,1}\right) f_{ex}(x) + d^{\rightarrow}_{x,1} \frac{1}{2}\left(f_{in}(x) + f_{ex}(x)\right) & \text{if } x \in \mathcal{R}^{\rightarrow}_1 \bigcap \epsilon_{ex} \\ \left(1 - d^{\leftarrow}_{x,1}\right) f_{in}(x) + d^{\leftarrow}_{x,1} \frac{1}{2}\left(f_{in}(x) + f_{ex}(x)\right) & \text{if } x \in \mathcal{R}^{\leftarrow}_1 \bigcap \epsilon_{in} \\ \left(1 - d^{\leftarrow}_{x,1}\right) f_{ex}(x) + d^{\leftarrow}_{x,1} \frac{1}{2}\left(f_{in}(x) + f_{ex}(x)\right) & \text{if } x \in \mathcal{R}^{\rightarrow}_1 \bigcap \epsilon_{ex} \\ \left(1 - d^{\rightarrow}_{x,2}\right) \frac{1}{2}\left(f_{in}(x) + f_{ex}(x)\right) + d^{\rightarrow}_{x,2} f_{ex}(x) & \text{if } x \in \mathcal{R}^{\rightarrow}_2 \\ \left(1 - d^{\leftarrow}_{x,2}\right) \frac{1}{2}\left(f_{in}(x) + f_{ex}(x)\right) + d^{\leftarrow}_{x,2} f_{ex}(x) & \text{if } x \in \mathcal{R}^{\leftarrow}_2 \\ f_{lin}(x) & \text{if } x \notin \mathcal{R}^+ \end{cases}$$

The correlation functions are gradually updated during the treatment fraction to reduce the tracking error. Every 30 seconds to five minutes, one of the 15 data pairs used to fit the model is substituted with a new pair, based on a first-in first-out principle. A complete rebuild can also be executed in case the tracking error exceeds a predefined threshold of 5 mm two times in a row. During a Cyberknife treatment, an increase in tracking error can originate from baseline drift of the external markers or from other variations in the breathing cycle of the patient [34].

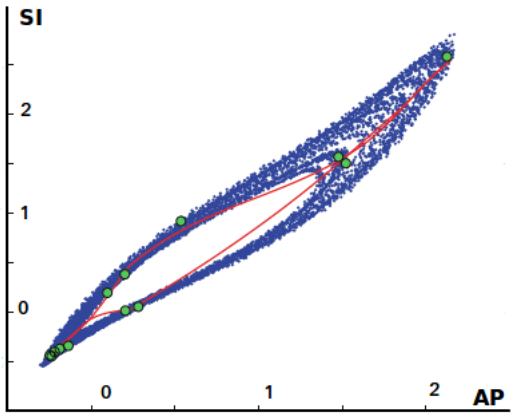

Figure 2.10: Representation of the bi-quadratic CK model (red) built with maximum 15 data points (green) and patient breathing data (blue).

2.4 Methodology of the comparison

In the first stage of the comparison, both the Cyberknife correlation model and the Vero correlation model were simulated offline in MatLab (Natick, MA, USA). For the Cyberknife model, this was performed based on descriptions of the model found in literature [33][34][63][64]. The Vero model was simulated based on literature and private communication with Brainlab (BrainLAB AG, Feldkirchen, Germany)[37][58]. Both correlation models were also validated in some way to ensure no errors had occurred during simulation. The Cyberknife model was validated, to a certain extent, by comparing the outcomes of two models simulated by two different people independently. If both were simulated as intended, there should be no significant difference between the outcomes. The Vero model was validated by comparing the outcomes of the simulated model, i.e. the target positions, to the outcomes of the clinically implemented model which are stored in log files generated before treatment. It should however be noted that those target positions were estimated after a forward prediction algorithm was applied to the surrogate breathing signal, while the estimations with the simulated models were made using a surrogate signal that was not forward predicted. Small discrepancies between the outcomes of the simulated model and the outcomes stored in the log files are therefore expected, mainly around the maximum inhale and maximum exhale positions. However, despite this difference, it is expected that they agree with sub-millimetric accuracy and a validation should be possible.

Log files, all listed below, from a patient population of 10 lung and liver patients were used not only to validate the Vero model, but also to train both models and to use the surrogate signals as input to the models so target positions could be estimated. An overview of those log files is listed below. These 10 patients were previously treated on the Vero SBRT system with real-time tumor tracking for a maximum of 10 fractions per patient and maximum 9 beams per fraction.

- The correlation_model.log file contains the 20 - 40 s 3D internal target positions and 1D vertical external marker positions with which the correlation model is built before the treatment fraction, together with corresponding timestamps. It also contains 3D target position predictions, with timestamps, computed by the correlation model implemented in the Vero SBRT system. These predictions were computed from external marker positions to which the forward prediction algorithm was applied. All positions are stored with a 0.01 mm precision and timestamps with a 1 ms precision.

- The RespiratoryMotionTracking.csv file contains the 3D external marker positions recorded during treatment, stored with a 0.01 mm precision. Corresponding time stamps are stored with a 1 ms precision.
- The TumorPositionTracking.csv file contains the 3D internal target positions predicted by the clinically implemented correlation model, again stored with a 0.01 mm precision. Corresponding time stamps are stored with a 1 ms precision.
- The XrayTimeTracking.csv file contains the time stamps at which X-ray images of the target are taken, with a precision of 1 ms.

The Cyberknife model was simulated so that only the bi-quadratic model with fall-back and blending mechanism was applied. The complete linear model and linear-quadratic models were omitted. As such, both correlation models represented a continuous breathing loop throughout the entire treatment to model the relation between the surrogate breathing signal and the internal target motion. While the linear and linear-quadratic models might be sufficient for some patients or certain discrete parts of a treatment fraction, their accuracy will always be inferior to that of a bi-quadratic model. Therefore, to ensure a fair comparison and to be able to obtain unambiguous results about the spatial accuracy of both models, the simulated Cyberknife model was not completely equal to the model implemented in the clinical SRTS Cyberknife system.

Both correlation models were also simulated with three different update scenarios. In all three scenarios, the standard procedure of building the model before each fraction was implemented. Then in the first scenario, no intra-fraction retraining was allowed so that the initial models were used throughout the entire fraction. As such, they were subjected to a substantial baseline drift of over 1 mm. In a second scenario, a gradual update similar to that of the clinical Cyberknife algorithm was implemented. Every minute, a selection of the data pairs to build the model was substituted with new data from the same phase in the breathing cycle, and the models were re-trained. For the Cyberknife model, one out of 12 data points were substituted while for the Vero model around 15 data points were renewed. In a last scenario, a complete rebuild, the update implemented in the clinical Vero algorithm, was implemented. This rebuild was performed every time the tracking error acceded the predefined threshold of 3 mm in any direction. For our population, this was after about 6 min.

Finally, the Cyberknife correlation model was also simulated in two different ways. Once with the clinical Cyberknife build with 12 data points and once with a dataset obtained through 20 - 40 s full fluoroscopy (FF), the clinical build of the

Vero model. As such, it could be evaluated whether or not the difference in build between the Vero and the Cyberknife model has any significant influence on the geometrical accuracy of the model. Both Cyberknife models were simulated with the three update scenarios described above.

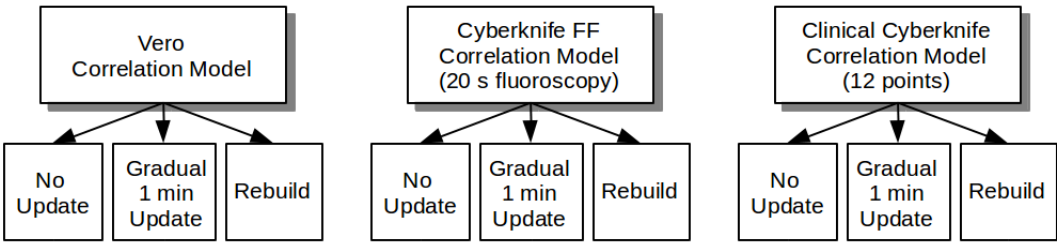

Figure 2.11: Three correlation models (CMs) were simulated; the Vero CM, the Cyberknife FF CM built from 20 s fluoroscopy and the clinical Cyberknife CM built with 12 data points. All three models were simulated in three different update scenarios; no update throughout the treatment fraction, a gradual update every minute and a complete rebuild when the tracking error exceeds a threshold of 3 mm in any direction.

In the second stage of the comparison, the data in the log files was used to simulate the behaviour of the correlation models during each treatment fraction that was clinically delivered on the Vero system. This was carried out by retrospectively applying the data that was used to train the correlation model during the actual treatment, to train the models we had simulated in Matlab. Afterwards, the surrogate breathing signal, which was also stored in the log files, was applied to all three models, in the three different update scenarios, to estimate the internal target positions. This resulted in nine sets of target position estimations.

In the third and final stage, these target position estimations were compared with the actual target positions obtained from orthogonal X-ray images (every 2 s) and in-house developed marker detection software. As such, root-mean-square errors (RMS) and 95th percentiles (E95), meaning the value below which 95% of the errors are located, could be calculated. This was done for all three models in each update scenario, per patient and averaged over the whole population. Comparing the geometrical accuracy of all three models, with respect to the update scenarios, is now a question of comparing the RMS and E95. T-tests were carried out to see if the differences in RMS between models were statistically significant.

Chapter 3

Results

3.1 Validation

In the first stage of the comparison both models were simulated offline and validated. An independent simulation of the Cyberknife model was carried out and results were compared. The internal target position estimations of the simulated Vero model were compared with the internal target position predictions of the clinically implemented Vero model which were stored in log files (correlation_model.log).

Vero correlation model

The root-mean-square errors between the target position estimations of the simulated model and the target position predictions of the clinically implemented model, and the 95th percentile (E95), averaged over all ten patients, can be found in Table 3.1. The three directions of motion of the target; lateral, longitudinal and vertical, show an error of 0.38, 0.79 and 0.44 mm, respectively. Paired t-tests were performed to evaluate possible statistical significances.

For 7 out of 10 patients, there was no statistically significant difference ($p>0.1$), in any direction, between the target positions of the simulated Vero model and those of the clinically implemented model. Patient 2 showed a significant difference in the lateral direction ($p<0.02$), patient 3 a significant difference in the vertical direction ($p<0.02$) and patient 10 in both the lateral ($p<0.02$) and longitudinal direction ($p<0.02$).

	lateral	**longitudinal**	**vertical**
RMS [mm]	0.38	0.79	0.44
E95 [mm]	0.70	1.50	0.79

Table 3.1: Validation of the Vero model; calculating the root-mean-square (RMS) errors between the outcomes of the simulated model with those of the clinically implemented model stored in log files before treatment.

Figures 3.1 and 3.2 show the estimated target motion (computed by the simulated Vero model) and the predicted target motion (computed by the clinically implemented Vero model) in the lateral, longitudinal and vertical direction during fraction 2 in the treatment of patient 2. A difference between the computed target positions of both models in the lateral direction can indeed be observed. However, when comparing the lateral direction to the two other directions of motion, it can be observed that the lateral direction is not the main component of motion. Although the difference between both motion signals is statistically significant, the root-mean square error (0.30 mm) remains sub-millimetric. During the two other fractions, fraction 1 and 3, the correspondence between the two motion signals is similar to those represented in Figures 3.1 and 3.2, with the largest discrepancy in the lateral direction observed during fraction 2 and represented in Figure 3.1.

Figure 3.3 shows the target motion computed by the simulated Vero model and computed by the clinically implemented model in the vertical direction during fraction 1 in the treatment of patient 3, while Figure 3.4 shows the same but during fractions 2, 3 and 4. Patient 3 was treated with 4 fractions in total. It can be observed that a discrepancy between the results is present only during fraction 1, where the predicted target motion by the clinically implemented Vero model shows strong random variations and jumps. It can also be observed that the vertical direction is not the main component of motion. While the difference between the predictions of both models might be statistically significant, the root-mean square error (0.32 mm) remains sub-millimetric.

For patient 10, there is a statistically significant difference between the results from the simulated Vero model and those from the clinically implemented model in both the lateral and longitudinal direction. These differences however are only present during the first part of the first of 10 fractions. The first part is defined as the part before the first complete rebuild of the correlation model. So far it is unclear why the discrepancy, which is both a mismatch in amplitude and phase, is present.

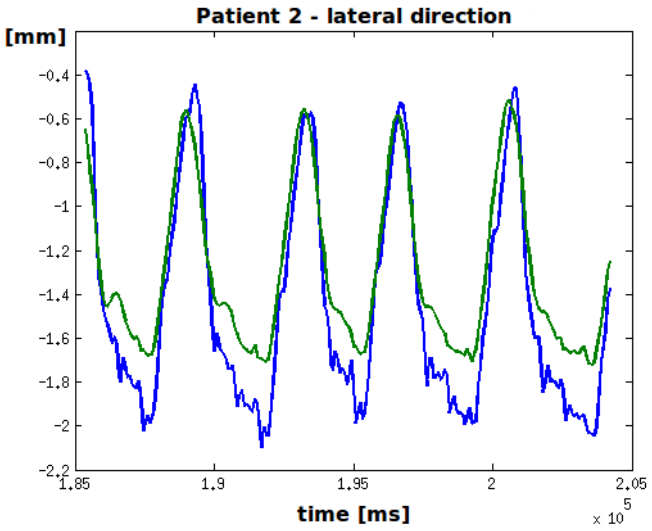

Figure 3.1: The estimated target motion (green) and the predicted target motion (blue) in the lateral direction during fraction 1 in the treatment of patient 2.

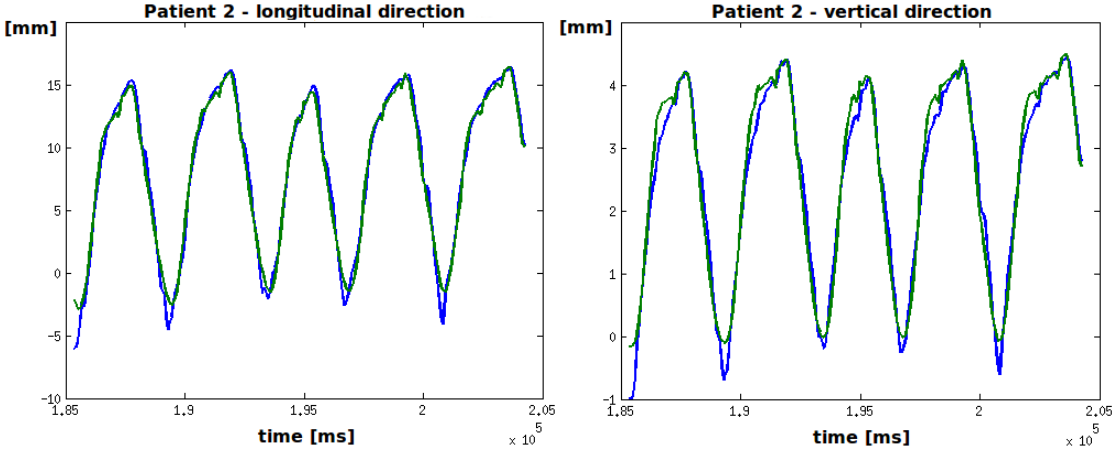

Figure 3.2: The estimated target motion (green) and the predicted target motion (blue) in the longitudinal (left) and vertical (right) direction during fraction 2 in the treatment of patient 2

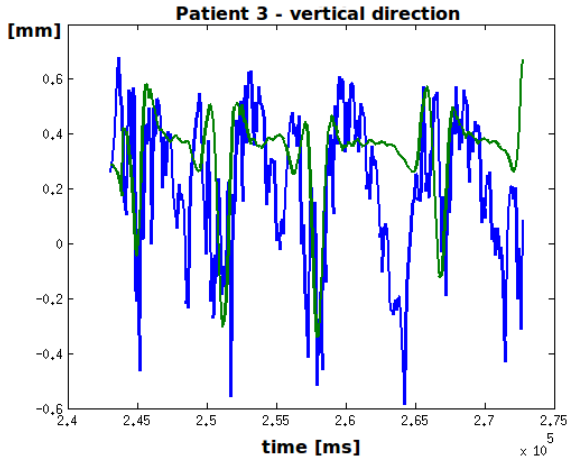

Figure 3.3: The estimated target motion (green) and the predicted target motion (blue) in the vertical direction during fraction 1 in the treatment of patient 3.

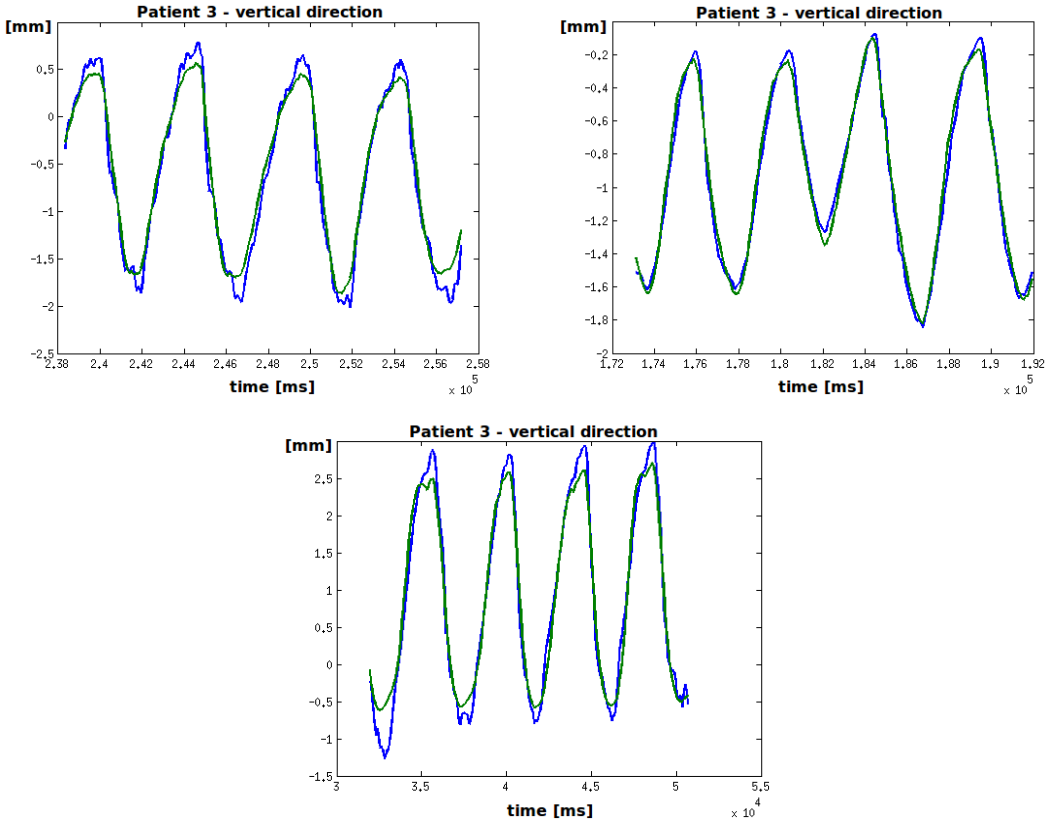

Figure 3.4: The estimated target motion (green) and the predicted target motion (blue) in the vertical direction during fraction 2 (top left), 3 (top right) and 4 (bottom) in the treatment of patient 3.

Cyberknife correlation model

The validation of the Cyberknife model had to be carried out based on two independent simulations of the model by two different people. The main reason why no validation was carried out based on log files is because there was no available access to a Cyberknife system. Secondly, the Cyberknife correlation model that was compared in this study was based on that implemented in the SRTS Cyberknife system, but certain parts of the clinical model were deliberately omitted. Meaning that a difference between target positions estimations of the simulated model and those from the implemented model are to be expected, making a validation inconvenient. Therefore, to ensure that the correct Cyberknife model was compared with the Vero model, the CK model was simulated twice, independently, and the outcomes were compared. Possible errors in the simulation were not likely to be simulated twice and would be noticed by a significant difference in outcomes.

	CK FF Model 1	**CK FF Model 2**
RMS [mm]	1.94	1.79

Table 3.2: Validation of the Cyberknife model; root-mean-square (RMS) tracking errors averaged over the patient population of the simulated FF Cyberknife model 1 and the simulated FF Cyberknife model 2, both in the rebuild scenario.

The independent simulation and comparison was carried out with the Cyberknife model built with 20 s full-fluoroscopy in the complete rebuild scenario. The RMS tracking errors were calculated, averaged over the patient population and compared. Results can be found in Table 3.2. The RMS tracking error of the FF Cyberknife model 1 averaged over the population equalled 1.94 mm, the error of FF Cyberknife model 2 equalled 1.79 mm.

The difference of 0.15 mm was considered low enough to assume that the intended Cyberknife model was simulated and a comparison with the Vero model could be carried out.

3.2 Comparison

The tracking errors, defined as the difference between the estimated target positions and the actual target positions, and 95^{th} percentiles (E95) of the Vero correlation model, the Cyberknife correlation model built with 20 s full fluoroscopy (CK FF) and the Cyberknife correlation model built with 12 data points (CK clin) can be found in Table 3.3 for all three update scenarios. The 95^{th} percentiles per patient can be found in Figures 3.5 to 3.10. Figures 3.5 to 3.7 show the 95th percentiles per update scenario for all three models per patient, to compare the geometrical accuracy amongst the 3 models at the same scenario. Figures 3.8 to 3.10 show the 95th percentile per model for all three scenarios per patient, to compare the influence of the different update scenarios on each model.

In the next two sections, the models will be compared with each other per update scenario and each model will be evaluated taking into account the influence of each update scenario. More specifically, the Vero model will be compared with the clinical Cyberknife model to evaluate their difference in response to baseline drift, and the Cyberknife FF model will be compared with the clinical Cyberknife model to evaluate the influence of the number of data used to build the model.

	Vero	**CK FF**	**CK clin**	**t-test**
RMS [mm]				
no update	2.53	2.70	2.86	$p > 0.1, p > 0.1$
1-min update	1.95	2.30	2.87	$p < 0.02, p < 0.005$
rebuild	1.73	1.94	2.21	$p < 0.02, p < 0.02$
E95 [mm]				
no update	4.77	5.04	5.17	
1-min update	3.64	4.32	5.43	
rebuild	3.17	3.52	3.79	

Table 3.3: Geometrical accuracy of the Vero model, the clinical (clin) 12-point Cyberknife (CK) model and the CK model built with full fluoroscopy (FF) data. t-tests were performed between the Vero model and respectively the FF CK model and the clinical 12-point CK model.

3.2.1 Evaluation of the three correlation models

In the no update scenario, see Table 3.3, the tracking error of the Vero model (2.53 mm) and the Cyberknife FF model (2.70 mm) and the Vero model (2.53 mm) and the clinical Cyberknife model (2.86 mm) show no significant difference ($p>0.1$, $p>0.1$). The geometrical accuracy of the models in this scenario, for our population, is similar for all three models. Figure 3.5 illustrates the results for all 10 patients individually.

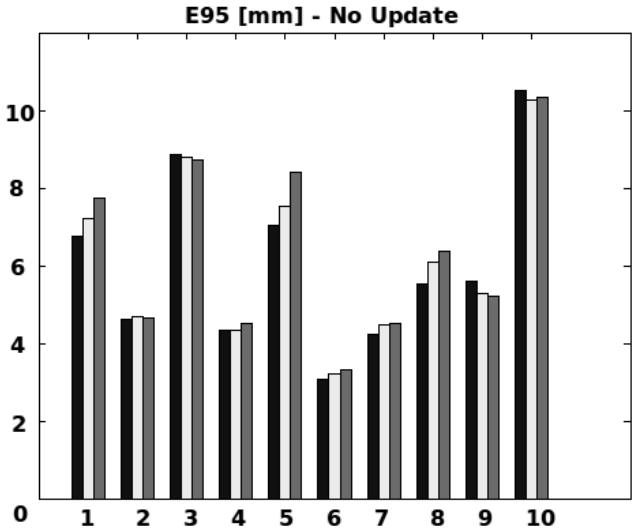

Figure 3.5: 95th percentiles of the Vero model (black), the Cyberknife FF model (white) and the clinical Cyberknife model (grey) per patient (1→10) in the no update scenario. No significant difference is observed between the Vero model and both Cyberknife models ($p>0.1$, $p>0.1$).

In this scenario all three models are subjected to a significant baseline drift (> 1 mm up to 6 mm). Meaning that after a certain time period the input data to the correlation model, which is the external marker data, is partially located outside the range of data that was used to build the model. This should lead to extrapolation errors due to the use of higher order polynomial functions. These errors however should be reduced with the linear fall-back mechanism which is implemented in the Cyberknife model only. Next to extrapolation errors, baseline drift also leads to a shift in the estimated target positions that is not presented in the actual positions, which largely contributes to the tracking error. The hypothesis was that the Vero model would be more sensitive to baseline drifts, due to the lack of a fall-back mechanism, and would potentially show larger tracking errors due to extrapolation errors, this however is not the case. A shift between the estimated

target positions and the actual target positions however was observed for some patients.

When comparing in this no update scenario the accuracy of the models built with 20s FF data (Vero model and CK FF) with the accuracy of the clinical Cyberknife model which used only 12 data points to build, no significant difference can be observed. Building the model with more data for our population did not ensure a better geometrical accuracy in the presence of substantial baseline drift.

In the gradual 1-min update scenario, the tracking error of the Vero model (1.95 mm) and the Cyberknife FF model (2.30 mm) and the Vero model (1.95 mm) and the clinical Cyberknife model (2.87 mm) show a significant difference ($p<0.02$, $p<0.005$). The geometrical accuracy of the Vero model in this update scenario is, for our population, superior compared to both Cyberknife models. These findings are illustrated per patient in Figure 3.6.

Similar results can be found for the complete rebuild scenario where the tracking errors equal 1.73 mm (Vero model), 1.94 mm (Cyberknife FF model) and 2.21 mm (clinical Cyberknife model) with a significant difference between the Vero model and both Cyberknife models ($p<0.02$, $p<0.02$). These findings can also be observed per patient, represented in Figure 3.7. It can be observed that the difference in accuracy between the Vero model and Cyberknife models is not as extensive as in the gradual update scenario.

In the latter two scenarios all three models are gradually updated or rebuild and thus less subjected to baseline drift, making the tracking error more dependent on how good the correlation models are capable of modelling the breathing cycle and the possible presence of hysteresis. The characteristics of the Vero model such as directly using both the external marker positions and velocities seem to ensure that this model is better at modelling the breathing cycles of our patient population.

Comparing the models built with data obtained through 20 s FF at 11 Hz (Vero model and CK FF) with the clinical Cyberknife model which used only 12 data points, it is seen that the difference in build has a significant influence in both the gradual update scenario and the rebuild scenario. Building the model with more data did ensure a better modelling of the breathing cycle of our patients, which was to be expected.

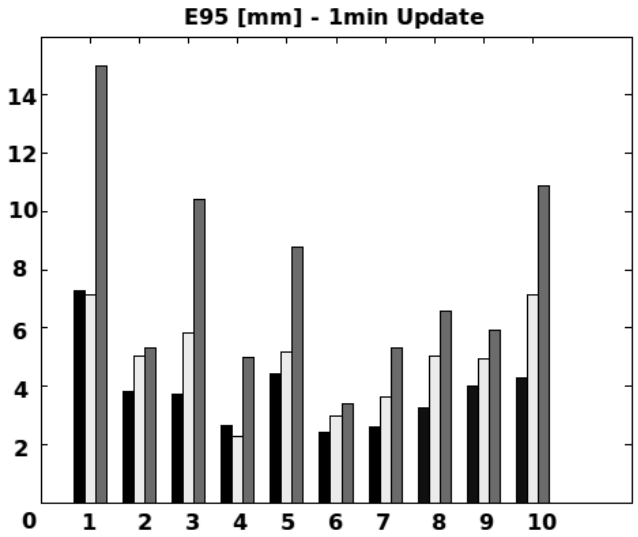

Figure 3.6: 95th percentiles of the Vero model (black), the Cyberknife FF model (white) and the clinical Cyberknife model (grey) per patient (1→10) in the 1-min update scenario. A significant difference is observed between the Vero model and both Cyberknife models ($p<0.02$, $p<0.005$).

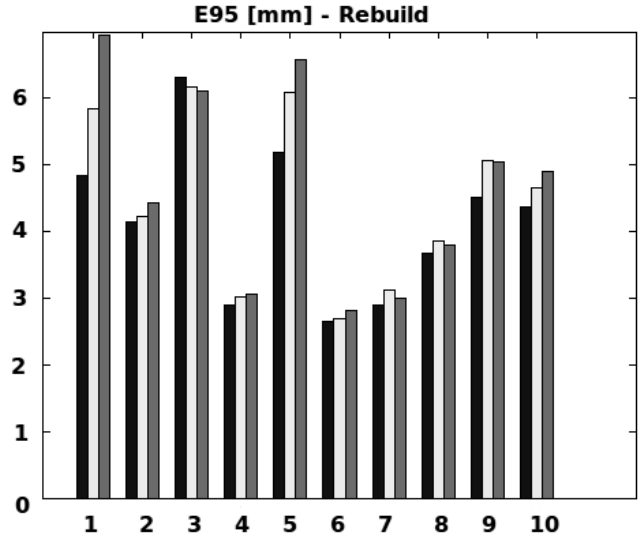

Figure 3.7: 95th percentiles of the Vero model (black), the Cyberknife FF model (white) and the clinical Cyberknife model (grey) per patient (1→10) in the rebuild scenario. A significant difference is observed between the Vero model and both Cyberknife models ($p<0.02$, $p<0.02$).

3.2.2 Evaluation of the update scenarios

The Vero model root-mean-square errors (RMS) and 95th percentiles (E95) averaged over all patients (Table 3.3) indicate that both a gradual update and a complete rebuild of the model are superior to not updating the model. These results were expected as both updating and rebuilding compensates for baseline drift and as such decreases the tracking error. Looking at the patient specific representation in Figure 3.8, it can be observed that for 9 out of 10 patients a gradual update is more optimal compared to a complete rebuild, though with minor difference. This might be an indication that for some patients treated on the Vero system an automatic gradual update might have been more accurate than the implemented rebuild methodology.

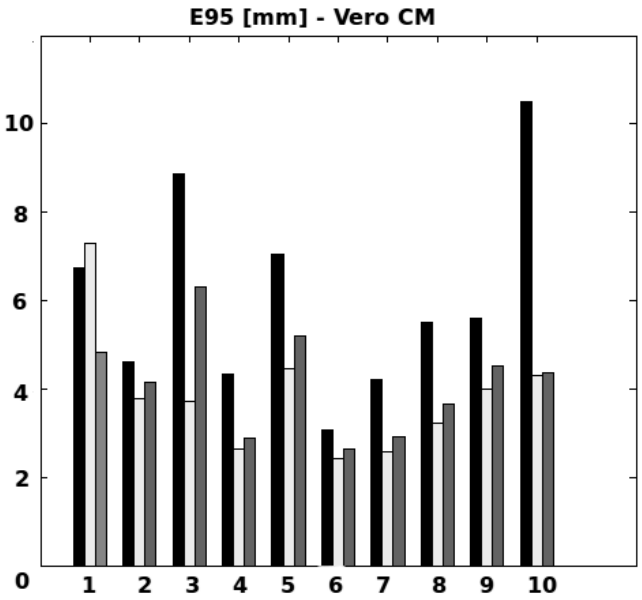

Figure 3.8: 95th percentiles of the Vero model per patient (1→10) for 3 update scenarios: no update (black), 1 min update (white), rebuild (grey).

The Cyberknife FF model RMS errors and E95 averaged over all patients (Table 3.3) indicates that both a gradual update and a complete rebuild of the model is better than to not update the model, similar as with the Vero model. Both results were again expected as updating and rebuilding compensates for baseline drift and as such decreases the tracking error. Looking at the patient specific representation in Figure 3.9, it can be observed that in the majority of cases (6 out of 10 patients) a complete rebuild is better in terms of accuracy, compared to a gradual update.

As the differences are minor, it can not be stated whether one update scenario is superior over another.

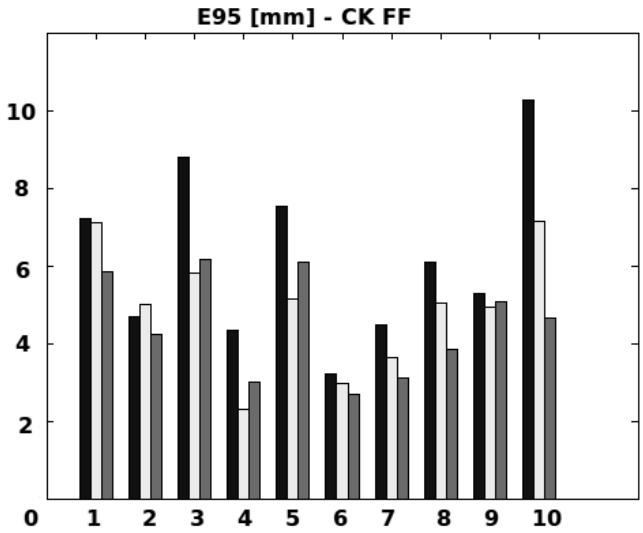

Figure 3.9: 95th percentiles of the Cyberknife FF model per patient (1→10) for 3 update scenarios: no update (black), 1 min update (white), rebuild (grey).

The clinical Cyberknife model RMS errors and E95 averaged over all patients (Table 3.3) indicate that only a complete rebuild of the model can improve its performance. For our population, the geometrical accuracy did not improve when applying a gradual 1-min update compared to a no update scenario. The patient specific representation (Figure 3.9) illustrates these findings. This is quite remarkable as the gradual update is the one implemented in the actual Cyberknife system. Most likely it is due to the fact that only 1 different data point can not contain enough information to represent the changes in the breathing cycle.

Overall, it can be observed that the models built with 20 s FF data benefit from both a gradual update and a complete rebuild, while the clinical Cyberknife model only benefits from a complete rebuild.

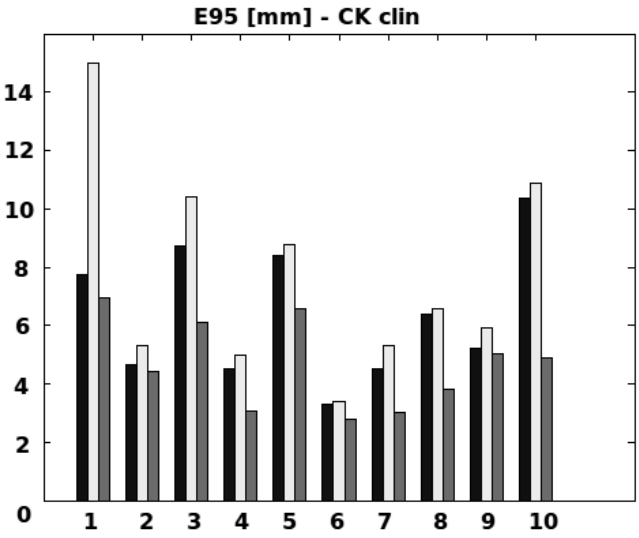

Figure 3.10: 95th percentiles of the clinical Cyberknife model per patient (1→10) for 3 update scenarios: no update (black), 1 min update (white), rebuild (grey).

Chapter 4

Conclusion

The first intention of this study was to simulate and validate the correlation models implemented in the Vero SBRT system for Dynamic Tracking and the Synchrony Respiration Tracking Cyberknife system. Simulating the Vero model was shown feasible. The validation based on log files confirmed the correspondence between the simulated correlation model and the one implemented in the Vero SBRT system. Validation of the Cyberknife model was based on the comparison of an independent simulation of the FF model in the rebuild scenario. A correspondence between the tracking errors averaged over the patient population was introduced to quantify the validation.

After simulating and validating both correlation models, the comparison based on geometrical accuracy and response to baseline drift and hysteresis of the respiratory signal could be performed.

In the presence of significant baseline drift, the correlation model of the Vero system without linear fall-back mechanism responded similar to the more complex Cyberknife model with a linear fall-back mechanism. In the case of intra-fraction re-training and updating and thus compensating for baseline drift, the geometrical accuracy of the Vero model was shown to perform with a higher accuracy compared to both the Cyberknife model build with 12 data points (clin CK) and that build with data from 20 s full fluoroscopy (FF). From the complete rebuild scenario, it could be observed that the models built with a relatively large set of data points (20s FF - 11 Hz) are capable of modelling the breathing motion with a higher accuracy than the Cyberknife model which is built with on average 12 data points. More so, the Vero model which, besides their positions, also directly incorporates the marker velocities, could estimate the target positions with a higher accuracy than both Cyberknife models. It was also shown that the gradual update did not

have any major influence on the clinical Cyberknife model, while an improvement in geometrical accuracy could be observed for the models built with data from 20s full-fluoroscopy. A complete rebuild positively influenced the accuracy of all models. However, based on the individual patient results, for the Vero model there might be an indication that an automatic gradual update is superior over a complete rebuild methodology, this was confirmed by Poels et al. [71].

Several conclusions can be drawn from these results. In general, it can be concluded that baseline drift and other variations of the breathing motion over time, such as amplitude and frequency changes, significantly decrease the accuracy of all correlation models. As a decrease of the accuracy equals larger tracking errors on the target position, this increase leads to the irradiation of adjacent healthy tissue and, due to tight geometrical margins, a geographical miss and under-dose of the target. This both increases the normal tissue complication probability and decreases the tumor control probability. Compensating for the above mentioned variations is thus absolutely necessary.

Two ways to do this is by introducing gradual updates or a complete rebuild of the model once the error exceeds a certain threshold. However, from the results it can be concluded that the impact of these intra-fraction retraining options strongly depends on the correlation model itself. While a gradual update can decrease the tracking error of models built with a relatively large data set, substituting only one data point in models built with a limited amount of data does not seem to influence the accuracy. A complete rebuild on the other hand can improve the accuracy of all models. However, initiating a complete rebuild requires a temporary interruption of the dose delivery, increasing the over-all treatment time. During the RTTT treatments, it was seen that increasing the treatment time can significantly decrease the comfort of patients, this and fatigue induced variations in their breathing pattern which again decreased the tracking error.

When variations of the breathing signal over time are excluded, the conclusion is that the models built with more data points are capable of modelling the breathing motion with a significantly higher accuracy. Moreover, from the rebuild scenario it could be concluded that for shorter treatment fractions, i.e. high dose rate delivery, models built with 20 s full-fluoroscopy are clinically favourable. If real-time tumor tracking should be introduced in treatments such as VMAT (Volumetric Modulated Arc Therapy), a correlation model built with a relatively large set of data would be beneficial over models built with only a limited amount of data. It should however be remarked that more data directly entails a larger imaging dose to the skin of the patient, rendering the use of an external surrogate signal and correlation model to be useless. A balance between this skin dose and the amount of data to build the correlation model therefore has to be established.

Also, among the models built with more data points, the model using both the external marker positions and velocities, instead of just the positions, showed to be capable of modelling the breathing motion with even yet a higher accuracy.

In general, the conclusion is that polynomial models built with sufficient data obtained over multiple breathing cycles, and if possible multiple surrogate signal parameters, are capable of modelling the relation between external marker motion and the internal target motion with a high accuracy. However, an update and/or rebuild scenario is still required to overcome variations over time. The implementation of an automatic gradual update might be more optimal than a rebuild methodology both in terms of accuracy, patient comfort, and economical factors as this does not require a temporary interruption of the dose delivery during the treatment fraction.

Chapter 5

Discussion

5.1 Discussion of the results

Input data to the correlation model

In both the Vero and Cyberknife tracking systems, optical markers positioned closely to the skin of the patient at a location moving strongly with respiration, are used to generate a surrogate breathing signal. The input to the correlation model however is different between the two. While the Vero model uses both the vertical marker positions and velocities, the Cyberknife model uses only their vertical positions.

It is believed that this difference can have a large impact on the correlation model accuracy. Ernst et al. stated that modelling respiratory motion should not be done by regarding only spatial displacement, but that including additional data will improve the accuracy [51]. Other options besides marker velocity or vectors indicating the direction of motion are for example parameters such as airflow and tidal volume, obtained through spirometry. It has been reported that surrogate data from spirometry shows a more linear relation with the internal target motion than optical markers [65][66]. A disadvantage of spirometry is that there can be considerable drift in the spirometry signals due to instrumentation errors or escaping air. These could however be corrected with the help of signals obtained with optical markers, meaning that an interplay of both spirometry and an optical marker tracking system would be ideal, but also rather cumbersome in clinical practice.

Tracking a fiducial marker

While the goal of real-time tumor tracking is to follow and treat a patient-specific moving volume of cancerous cells, in reality it is the movement of an implanted fiducial marker with a fixed volume on which the whole treatment is based. The internal target positions mentioned in this study are those of an implanted marker that can be automatically detected by the treatment system, not of the target itself. It was assumed that the relative position of the marker with respect to the PTV shape and location remained stable throughout the entire treatment period. Studies by Imura et al. and Brown et al. reported that this assumption is justified and that the fiducial position is stable and robust, but only over a limited period of time [67][68].

Despite the justification of our assumption, relying on a fiducial markers gives you motion information only at that location and not over the whole PTV, increasing treatment uncertainties. Also, these justifications only apply to certain markers and certain ways of implantation. A worst-case scenario would be if the marker starts to serve as a pivot point around which the tumor would start to rotate. In that case the tracked marker would be registered as stable while the target that has to be treated moves around. Also, implanting the fiducial marker is considered to be an invasive procedure. For those reasons, marker-less tracking is currently a much studied topic [69][70].

Inter-patient variability

An aspect that was not considered during the research performed in this study is the inter-patient variability. It was not the intention to evaluate these possible variations, unless inconsistencies with the results of the comparison were present among the patient population. The latter was not the case. It can however be noted, when evaluating the results, that certain inter-patient variations should be mentioned as they seem to be influenced by the different update scenarios evaluated in this study.

Going back to Figures 3.5 to 3.7, though not statistically evaluated, it can be observed that the accuracy of all models differs quite significantly among the patient population. In the no update scenario for example, the difference in accuracy of all models between patient 6 and patient 10 is a factor 3. This immediately raises the question of applying patient-specific margins, which are applied to overcome position uncertainties still present during RTTT. A similar conclusion was made by

Pepin et al. when studying the Cyberknife Synchrony respiratory tracking system [33].

More so, when comparing Figure 3.5 with Figure 3.7, it can be observed that the inter-patient variability decreases when the rebuild scenario is applied. The influence of rebuilding on the accuracy of the models is thus different among the patient population. For patients having an already relatively high accuracy during the no update scenario, the increase in accuracy due to rebuild is not as significant as for those patients showing a low accuracy during the no update scenario. Patient 6 for example shows almost no improvement from no update to rebuild, while the accuracy of patient 10 has doubled.

Overall, the accuracy of the models can vary quite significantly among the patient population when no updates are applied, while the accuracy of the different models per patient is consistent. However, the inter-patient variability diminishes when a complete rebuild is applied to improve the accuracy of the models. This is due to a smaller improvement of the accuracy of the models for those patient who already showed a relatively high accuracy when the models were not updated. Thus, patients showing a low accuracy when the models are not updated benefit significantly more from a rebuild scenario than those already showing a high accuracy with no update.

5.2 Limitations and future work

Validation of the Cyberknife model

Validating the simulated correlation models was part of the initial research performed in this thesis. While the Vero model could be validated based on log files, giving a decisive answer whether or not the simulated model corresponded to the model implemented in the system, the validation of the Cyberknife model was less conclusive.

Having log files at our disposal which were generated during a real-time tumor tracking treatment on the SRTS Cyberknife system would have greatly reinforced the validation process. This in turn would strengthen the conclusions which were drawn from the comparison of the simulated correlation models. However, it was never our intention to use the exact same model as the one implemented in the SRTS Cyberknife system in our comparison, making a validation based on Cyberknife log files less critical for this study.

Difference between the implemented and simulated Cyberknife model

While it was already mentioned a few times, it should be noted again that the Cyberknife correlation model which was simulated and used in the comparison is not completely equal to the model implemented in the Cyberknife SRTS system. Only the bi-quadratic part of the correlation model was simulated, while the linear model and linear-quadratic model were left out. The reason being that we wanted to compare only those models that represent a continuous breathing loop, making the comparison more meaningful. While a linear or linear-quadratic model might be sufficient for some patients during certain limited periods in the treatment, the accuracy of a more complex model representing a continuous breathing loop shows more continuity over a patient population and over the entire duration of a treatment. It is therefore mostly this component of the model that determines the overall accuracy. As such, comparing the Vero model with the bi-quadratic part of the Cyberknife model instead of the complete Cyberknife model makes the results less ambiguous.

Although this comparison is more straightforward, it is important to note that as such it is not correct to state that the model implemented in the Vero SBRT system is superior or inferior compared to the model implemented in the SRTS Cyberknife system, as the latter was not the model we simulated and used in the comparison.

Prediction and latency

Both correlation models, that of the Cyberknife and that of the Vero system, were simulated without their respective prediction algorithm. This prediction is present in both systems to compensate for system latency. Not including the prediction algorithms gave the advantage that we were able to conclude which simulated correlation model is capable of modelling breathing motion with a higher accuracy, without having to consider the influence of the prediction algorithms and the difference in latency (50 ms - 200 ms). In general, limiting the amount of variations between both models made it easier to evaluate which differences among the two models caused the difference in geometrical accuracy.

However, in a next stage it would be interesting to see the influence on the geometrical accuracy when the prediction algorithms are added to the correlation models. Moreover, when the goal is to make a comparison of the clinical performance of both systems, the implementation of the prediction algorithms would be required. The difference in latency can have such an impact on the accuracy

that the difference in performance of the correlation models might be negligible. For the Vero system, this could be tested by benchmarking the results of these simulations against the experimental verifications performed earlier by Depuydt T. et al [58].

If the prediction algorithm is included in the simulation, it might also be interesting to evaluate the influence of including the prediction in the build of the model. In the current workflow, the correlation models are built using external marker data to which no prediction is applied. In contrary, during the treatment, the external marker data is first forward predicted, which introduces an error, before it is used as input to the correlation model. Implementing the prediction to the external marker data that is used to build the model might result in more accurate target predictions, an approach which is used in correlation models based on neural networks and single vector regression [50][51].

Sensitivity tests

Using real patient data instead of artificial breathing data gave the opportunity to evaluate the models in real clinical circumstances. A downside to real patient data is that the breathing signals contain both hysteresis, baseline drift and other random variations over time. Making quantitative conclusions about the tracking error solely in terms of baseline drift or hysteresis is therefore not possible without excessive data manipulation.

An alternative would be to generate artificial breathing data with the help of a phantom, and introduce different levels of hysteresis or baseline drift, while keeping other variations to a minimum or completely exclude them. Using this data on the simulated models could yield more conclusive results about the response of the models to the different breathing phenomena. More so, possible differences in the response of both models could be observed.

Clinical implications

After comparing two types of polynomial correlation models in clinical circumstances, it has become clear that certain characteristics of the models, such as the type of build and the kind of input data or the type of update scenario, have a significant influence on the geometrical accuracy of the target tracking. This comparison has given an insight to which characteristics might be beneficial and which are absolutely required in order to have sufficient spatial accuracy. These

results are particularly interesting when considering treatment methods such as hypo-fractionation, where accurate target localization and precise dose delivery are crucial.

The work performed in this study was only a start to more extensive evaluations of the framework behind real-time tumor tracking systems and other treatment modalities that make use of a surrogate signal and correlation model, e.g. gating. Further evaluation of correlation algorithms together with the prediction algorithms and other input data and target position acquisition options should yield more information on how to perform more accurate target localization and precise treatment delivery. This eventually might lead to the possibility to extend real-time tumor tracking to moving targets located near other critical organs such as the heart.

Conflicts of interest

The research performed in this study was supported by the Flemish government through the Hercules foundation and corporate funding from Brainlab AG (BrainLAB AG, Feldkirchen, Germany).

Bibliography

[1] Belgium Cancer Registry. Belgium: Females/Males, number of invasive tumours by primary site and age group in 2011. http://kankerregister.org/

[2] World Health Organization, International Agency for Research on Cancer. Estimated Cancer incidence, mortality and prevalence worldwide in 2012: GLOBOCAN 2012. http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx, 2014.

[3] Ferlay J., et al. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. International Journal of Cancer 2010;127:2893-917.

[4] Belgian Federal Government, Algemene Directie Statistiek en Economische Informatie. Bevolking - Doodsoorzaken 1998-2011. http://statbel.fgov.be/nl/statistieken/cijfers/bevolking/sterfte_leven/oorzaken/, 2013.

[5] Edwards B.K., Brown M.L., Wingo P.A., et al. Annual Report to the Nation on the Status of Cancer, 1975 2002, Featuring Population-Based Trends in Cancer Treatment. Journal of the National Cancer Institute 2005;97:1407-27

[6] Lefemme C., Kelly W., Reynolds R., et al. Radiation, Cell Cycle and Cancer. Los Alamos Science 1995;23:51-89.

[7] Joiner M. and van der Kogel A. Basic Clinical Radiobiology. by Hodder Arnold 2009 ISBN 978-0-3409-2966-7.

[8] Slotman B.J., Solberg T.D., Verellen D. Extracranial Stereotactic Radiotherapy and Radiosurgery. by Informa Healthcare 2005 ISBN 978-0-8247-2697-3.

[9] Leksell L. The stereotaxic method and radiosurgery of the brain. Acta Chir Scand. 1951;102:316-9.

[10] Lax I., Blomgren H., Naslund I., et al. Stereotactic radiotherapy of malignancies in the abdomen. Acta Oncologica 1994;33:677-83.

[11] Blomgren H., Lax I., Naslung I., et al. Stereotactic high dose fraction radiation therapy of extracranial tumors using an accelerator. Acta Oncologica 1995;34:861-70.

[12] Seppenwoolde Y., Shirato H., Kitamura K., et al. Precise and real-time measurement of 3d tumor motion in lung due to beathing and heartbeat, measured during radiotherapy. International Journal of Radiation Oncology*Biology*Physics 2002;53:822-834.

[13] Keall P., Mageras G., Balter J., et al. The management of respiratory motion in radiation oncology. Report of AAPM Task Group 76. Medical Physics 33;10:3874-3900.

[14] US NAtional Institutes of Health, National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch. Surveillance, Epidemiology and End Results (SEER) Program Research Data (1975-2011). http://www.seer.cancer.gov, 2014.

[15] Liu H.H., Balter P., Tutt T., et al. Assessing respiration-induced tumor motion and internal target volume using four-dimensional computed tomography for radiotherapy of lung cancer. International Journal of Radiation Oncology*Biology*Physics 2007;68:531-40.

[16] Schlegel W., Bortfeld T., Grosu A.-L. New technologies in Radiation Oncology. by Springer 2006 ISBN 978-3-5400-0321-2.

[17] Verellen D., Ridder M.D., Linthout N., et al. Innovations in image-guided radiotherapy. Nature Review Cancer 2007;7:949-60.

[18] Stroom J.C. and Heijmen B.J.M. The effectiveness of an immobilization device in conformal radiotherapy for lung tumor: Reduction of respiratory tumor movement and evaluation of the daily setup accuracy. Radiotherapy and Oncology 2002;64:75-83.

[19] International Commission on Radiation Units and Measurements. ICRU report 62. Prescribing, recording and reporting photon beam therapy. Bethesda, MD:ICRU;1999.

[20] International Commission on Radiation Units and Measurements. ICRU report 83. Prescribing, recording and reporting intensity-modulated photon-beam therapy (IMRT). Bethesda, MD:ICRU;2010.

[21] Lax I., Blomgren H., Naslund I., et al. Stereotactic radiotherapy of malignancies in the abdomen. Methodological aspects. Acta Oncologica 1994;33:677-83.

[22] Blomgren H., Lax I., Naslund I., et al. Stereotactic high dose fraction radiation therapy of extracranial tumors using an accelerator. Clinical experience of the first thirty-one patients. Acta Oncologica 1995;34:861-70.

[23] Negoro Y., Nagata T., Aoki T., et al. The effectiveness of an immobilization device in conformal radiotherapy for lung tumor: Reduction of respiratory tumor movement and evaluation of the daily setup accuracy. International Journal of Radiation Oncology*Biology*Physics 2001;50:889-98.

[24] Wu H., Zhao Q., Berbeco R.I., et al. Gating based on internal/external signals with dynamic correlation updates. Physics in Medicine and Biology 2008;53:7137-50.

[25] Willoughby T.R., Forbes A.R., Buchholz D., et al. Evaluation of an infrared camera and X-ray system using implanted fiducials in patients with lung tumors for gated radiation therapy. International Journal of Radiation Oncology*Biology*Physics 2006;66:568-575.

[26] Guckenberger M., Kavanagh A., Webb S., et al. A novel respiratory motion compensation strategy combining gated beam delivery and mean target position concept - a compromise between small safety margins and long duty cycles. Radiotherapy and Oncology 2011;98:317-22.

[27] Linthout N., Bral S., Van de Vondel I., et al. Treatment delivery time optimization of respiratory gated radiation therapy by application of audio-visual feedback. Radiotherapy and Oncology 2009;91:330-5.

[28] Pollock S., Lee D., Keall P., et al. Audiovisual biofeedback improves motion prediction accuracy. Medical Physics 2013;40:041705.

[29] Verellen D., Soete G., Linthout N., et al. Quality assurance of a system for improved target localization and patient set-up that combines real-time infrared tracking and stereoscopic X-ray imaging. Radiotherapy and Oncology 2003;67:129-41.

[30] Keall P.J., Cattell H., Pokhrel D., et al. Geometric accuracy of a real-time target tracking system with dynamic multileaf collimator tracking system. International Journal of Radiation Oncology*Biology*Physics 2008;74:297-303.

[31] Hoogeman M., Prvost J.-B., Nuyttens J., et al. Clinical accuracy of the respiratory tumor tracking system of the cyberknife: assessment by analysis of log files. International Journal of Radiation Oncology*Biology*Physics 2006;65:1579-84.

[32] Depuydt T., Poels K., Verellen D., et al. Initial assessment of tumor tracking with a gimbaled linac system in clinical circumstances: A patient simulation study. Radiotherapy and Oncology 2013;106:236-40.

[33] Pepin E.W., Wu H., Zhang Y., et al. Correlation and prediction uncertainties in the Cyberknife Synchrony respiratory tracking system. Medical Physics 2011;38:4036-44.

[34] Seppenwoolde Y., Berbeco R.I., Nishioka S., et al. Accuracy of tumor motion compensation algorithm from a robotic respiratory tracking system; a simulation study. Medical Physics 2007;34:2774-84.

[35] Malinowski K.T., McAvoy T.J., George R., et al. Mitigating errors in external respiratory surrogate-based models of tumor position. International Journal of Radiation Oncology*Biology*Physics 2012;82:709-16.

[36] Hoogeman M., Prvost J-B., Nuyttens J., et al. Clinical accuracy of the respiratory tumor tracking system of the cyberknife: assessment by analysis of log files. International Journal of Radiation Oncology*Biology*Physics 2009;74:297-303.

[37] Depuydt T., Verellen D., Haas O., et al. Geometric accuracy of a novel gimbals based radiation therapy tumor tracking system. Radiotherapy and Oncology 2011;98:365-72.

[38] Hoisak J.D.P., Sixel K.E., Tirona R., et al. Correlation of lung tumor motion with external surrogate indicators of respiration. International Journal of Radiation Oncology*Biology*Physics 2004;60:1298-1306.

[39] McClelland J.R., Hawkesa D.J., Schaeffterb T., et al. Respiratory Motion Models: A Review. Medical Image Analysis 2012;17:19-42.

[40] Lu W., Low D.A., Parikh P.J., et al. Comparison of spirometry and abdominal height as four-dimensional computed tomography metrics in lung. Medical Physics 2005;32:2351-57.

[41] Low D.A., Parikh P.J., Lu W., et al. Novel breathing motion model for radiotherapy. International Journal of Radiation Oncology*Biology*Physics 2005;63:921-29.

[42] Lim S., et al. Guiding curve based on the normal breathing as monitored by thermocouple for regular breathing. Medical Physics 2007;34:4514-18.

[43] Timinger H., Krueger S., Borgert J., et al. Motion compensation for interventional navigation on 3D static roadmaps based on an affine model and gating. Physics in Medicine and Biology 2004;49:719-32.

[44] Atkinson D., Burcher M., DeClerck J., et al. Respiratory motion compensation for 3-D freehand echocardiography. Ultrasound in Medicine and Biology 2001;27;161520.

[45] Blackall J.M., Penney G.P., King A.P., et al. Alignment of sparse freehand 3-D ultrasound with preoperative images of the liver using models of respiratory motion. IEEE Transactions on Medical Imaging 2005;24:140516.

[46] Ehrhardt J., Werner R., Schmidt-Richberg A., et al. Statistical modeling of 4D respiratory lung motion using diffeomorphic image registration. IEEE Transactions on Medical Imaging 2011;30:25165.

[47] Fayad H., Clement J.F., Pan T., et al. Towards a generic respiratory motion model for 4D CT imaging of the thorax. IEEE Nuclear Science Symposium/Medical Imaging Conference 2009

[48] He T., Xue Z., Xie W., et al. Online 4-D CT estimation for patient-specific respiratory motion based on real-time breathing signals. Proceedings Medical Image Computing and Computer-Assisted Interventions 2010

[49] Klinder T., Lorenz C., Ostermann, J. Free-breathing intra- and intersubject respiratory motion capturing, modeling, and prediction Proceedings SPIE Medical Imaging 2009

[50] Ionascu D., Jiang S.B., Nishioka S., et al. Internal-external correlation investigations of respiratory induced motion of lung tumors. Medical Physics 2007;34:3893-903.

[51] Ernst F., Bruder R., Schlaefer A., et al. Correlation between external and internal respiratory motion: a validation study. International Journal of Computer Assisted Radiology and Surgery 2011;7:483:92.

[52] Ernst F., Martens V., Schlichting S., et al. Correlating Chest Surface Motion to Motion of the Liver Using ϵ-SVR - A Porcine Study. Medical Image Computing and Computer-Assisted Intervention MICCAI 2009;5762:356-64.

[53] Wilson P.L. and Meyer J. A spring-dashpot system for modelling lung tumour motion in radiotherapy. Computational and Mathematical Methods in Medicine 2010;11:13-26.

[54] Ackerley E.J., Cavan A.E., Wilson P.L., et al. Application of a spring-dashpot system to clinical lung tumor motion data. Medical Physics 2013;40:021713-1:10.

[55] Poels K., Depuydt T., Verellen D., et al. A complementary dual-modality verification for tumor tracking on a gimbaled linac system. Radiotherapy and Oncology 2013;109:469-74.

[56] Kamino Y., Takayama K., Kokubo M., et al. Development of a four-dimensional image-guided radiotherapy system with a gimbaled X-ray head. International Journal of Radiation Oncology*Biology*Physics 2006;66:271-78.

[57] Takayama K., Mizowaki T., Kokubo M., et al. Initial validations for pursuing irradiation using a gimbals tracking system. Radiotherapy and Oncology 2009;93:45-49.

[58] Depuydt T., Poels K., Verellen D., et al. Initial assessment of tumor tracking with a gimbaled linac system in clinical circumstances: A patient simulation study. Radiotherapy and Oncology 2013;106:236-40.

[59] Kilby W., et al. The CyberKnife Robotic Radiosurgery System in 2010. Technology in Cancer Research and Treatment, 2010;9:433-52

[60] Ozhasoglu C., Saw C.B., Chen H., et al. Synchrony - Cyberknife Respiratory Compensation Technology. Medical Physics 2008;33:117-23.

[61] Sheng Y., Li S., Sayeh S., et al. Fuzzy and hybrid prediction of position signal in Synchrony Respiratory Tracking System. Proceedings of the Ninth IASTED International Conference 2007 ISBN Hardcopy 978-0-88986-675-1.

[62] Urschel H.C., Kresl J.J., Luketich J.D., et al. Treating tumors that move with respiration. by Springer 2007 ISBN 978-3-540-69885-2.

[63] Ernst F. Compensating for Quasi-periodic Motion in Robotic Radiosurgery. by Springer 2012 ISBN 978-1-4614-1911-2.

[64] Sayeh S., Wang J., Main W.T., et al. Respiratory motion tracking for robotic radiosurgery. Treating Tumors that Move with Respiration 2007:15-29.

[65] Hoisak J.D.P., Sixel K.E., Tirona R., et al. Correlation of lung tumour motion with external surrogate indicators of respiration. International Journal of Radiation Oncology*Biology*Physics 2004;60:1298-306.

[66] Lu W., Low D.A., Parikh P.J., et al. Comparison of spirometry and abdominal height as four-dimensional computed tomography metrics in lung. Medical Physics 2005;32:2351-57.

[67] Imura M., Yamazaki K., Shirato H., et al. Insertion and fixation of fiducial markers for setup and tracking of lung tumors in radiotherapy. International Journal of Radiation Oncology*Biology*Physics 2005;63:144247.

[68] Brown W.T., Wu X., Fowler J.F., et al. Lung metastases treated by CyberKnife image-guided robotic stereotactic radiosurgery at 41 months. South Medical JoSayehurnal 2008;101:376-82.

[69] Schweikard A., Shiomi H., Adler J. Respiration tracking in radiosurgery without fiducials. The International Journal of Medical Robotics and Computer Assisted Surgery 2005;1:1927.

[70] Schweikard A., Shiomi H., Fisseler J., et al. Fiducial-Less Respiration Tracking in Radiosurgery. Medical Image Computing and Computer-Assisted Intervention 2004;3217:992-99.

[71] Poels K., Depuydt T., Verellen D., et al. Improving the intra-fraction update efficiency of a correlation model used for internal motion estimation during real-time tumor tracking for SBRT patients: fast update or no update? Radiotherapy and Oncology 2014;under review.

[72] Depuydt T., Poels K., Verellen D., et al. Treating patients with real-time tumor tracking using the Vero gimbaled linac system: implementation and first review. Radiotherapy and Oncology 2014;under review.