3-27-2015

The Evaluation and Study of Modern Radiation Dosimetry Methods as Applied to Advanced Radiation Therapy Treatments Using Intensity Modulated Megavoltage Photon Beams

Cassandra Stambaugh
University of South Florida, ckirklan@gmail.com

Follow this and additional works at: https://scholarcommons.usf.edu/etd

Part of the Physics Commons

Scholar Commons Citation

This Dissertation is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.
The Evaluation and Study of Modern Radiation Dosimetry Methods as Applied to
Advanced Radiation Therapy Treatments Using Intensity Modulated
Megavoltage Photon Beams

by

Cassandra K. K. Stambaugh

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Applied Physics
with a concentration in Medical Physics
Department of Physics
College of Arts and Sciences
University of South Florida

Co-Major Professor: Vladimir Feygelman, Ph.D.
Co-Major Professor: Garrett Matthews, Ph.D.
Eduardo Moros, Ph.D.
Geoffrey Zhang, Ph.D.
Benjamin Nelms, Ph.D.

Date of Approval:
March 27, 2015

Keywords: Quality Assurance, VMAT, IMRT, ArcCheck, Delta^4

Copyright © 2015, Cassandra K. K. Stambaugh
DEDICATION

I would like to dedicate this work to my husband Nate and my son Will. Thank you for the encouragement, love and support you have provided me over the years to help me achieve my goals and dreams.
ACKNOWLEDGMENTS

I would like to first thank my advisor, Dr. Vladimir Feygelman, for the support and encouragement he has provided over the last few years as well instilling in me the need and desire to always do my very best. I would also like to thank my program director and committee member, Dr. Eduardo Moros, for always taking the time to help and listen. A special thanks is also due to him for achieving the tremendous task of getting our program CAMPEP accredited before I graduated.

I would also like to thank the rest of my committee, Dr. Geoffrey Zhang, Dr. Benjamin Nelms and Dr. Garrett Matthews for providing advice, teaching me a great deal and for serving on my dissertation committee. I would like to extend a special thanks to the physics group at Moffitt Cancer center and especially to Dr. Kujtim Latifi, Dr. Daniel Opp and Dr. Dylan Hunt for tremendous help and education in and out of the clinic.

Most importantly I would like to thank my family for the unwavering and continued support. Without their willingness to listen, advise and support, this work would not have been possible.

Finally, I would like to thank the Physics department at USF for their help in navigating the PhD process as well as their support for our Medical Physics program.
# TABLE OF CONTENTS

LIST OF TABLES .................................................................................................................. iv

LIST OF FIGURES .............................................................................................................. v

ABSTRACT ............................................................................................................................ vii

CHAPTER 1: INTRODUCTION ............................................................................................... 1

1.1 Background ....................................................................................................................... 1
1.2 IMRT & VMAT .................................................................................................................. 2
1.3 Patient Specific Quality Assurance (PSQA) ................................................................. 3
1.4 Dose Measurement Techniques ...................................................................................... 4
1.5 Dose Analysis Techniques .............................................................................................. 5
1.6 References ....................................................................................................................... 10

CHAPTER 2: MATERIALS & METHODS ............................................................................. 13

2.1 Delta^4 ............................................................................................................................. 13
   2.1.1 Hardware .................................................................................................................... 13
   2.1.2 Software .................................................................................................................... 15
2.2 ArcCheck ........................................................................................................................ 16
   2.2.1 Hardware .................................................................................................................... 16
   2.2.2 Software .................................................................................................................... 18
2.3 3DVH ............................................................................................................................. 18
   2.3.1 Measurement guided dose estimates to static patient voxels ............................... 19
   2.3.2 Measurement guided dose estimates to moving patient voxels ............................. 20
2.4 References ....................................................................................................................... 22

CHAPTER 3: SEMI-EMPIRICAL VMAT DOSE RECONSTRUCTION ON A
PATIENT DATASET USING DELTA^4 MEASUREMENTS .................................................. 26

3.1 Introduction ....................................................................................................................... 26
3.2 Treatment Planning and Delivery ................................................................................... 26
3.3 Volumetric Dose Reconstruction Methods .................................................................... 27
3.4 Specific Tests ................................................................................................................... 28
   3.4.1 Simple static fields ...................................................................................................... 28
   3.4.2 VMAT dose reconstruction on homogeneous phantoms ....................................... 29
      3.4.2.1 PMMA Delta^4 cylindrical phantom ................................................................. 29
      3.4.2.2 Water-equivalent Delta^4 cylindrical virtual phantom ................................... 30
   3.4.3 VMAT dose reconstruction on patient CT datasets ............................................... 31
3.5 Results ............................................................................................................................. 31
CHAPTER 4: DOSE INTERPLAY EFFECTS IN VMAT SBRT LUNG TREATMENT

4.1 Introduction.................................................................44
4.2 Materials and Methods.................................................46
  4.2.1 Treatment planning and delivery ................................46
  4.2.2 Motion kernels .....................................................47
    4.2.2.1 Original empirical motion kernels .......................47
    4.2.2.2 Simulated extended motion ..............................49
  4.2.3 Data collection and analysis ..................................50
    4.2.3.1 Dose to the moving target ...............................50
    4.2.3.2 Metrics .....................................................50
    4.2.3.3 Single and multiple fraction simulations ..........51
4.3 Results...........................................................................52
  4.3.1 Single fraction .....................................................52
  4.3.2 Multiple-fraction courses ....................................58
4.4 Discussion .....................................................................60
  4.4.1 Dose to the moving target ....................................60
  4.4.2 Assumptions .........................................................63
4.5 Conclusions...................................................................64
4.6 References....................................................................65

CHAPTER 5: 3DVH FOR HELICAL TOMOTHERAPY ......................70

5.1 Introduction..................................................................70
5.2 Methods.......................................................................70
  5.2.1 Measurement-guided dose reconstruction ................70
    5.2.1.1 The challenges of TomoTherapy .......................70
      5.2.1.1.1 Delivery dynamics (gantry) .......................71
      5.2.1.1.2 Delivery dynamics (MLC leaves) ...............71
      5.2.1.1.3 Sub-beam size and complexity ..................72
      5.2.1.1.4 Beam parameters in DICOM RT Plan ..........73
      5.2.1.1.5 Synchronization of time-resolved measurements to beam parameters ..................73
      5.2.1.1.6 Threading effect and measurement locations ....74
    5.2.1.2 The algorithm ...................................................75
      5.2.1.2.1 Synchronization of RT plan to time-resolved measurement points ........................76
      5.2.1.2.2 Organization into 3D dose "sectors" ............77
LIST OF TABLES

Table 3.1. Gamma analysis passing rates (%) comparing Delta4 directly measured (Detectors Only) or D4 Interpolated (Volumetric) dose distributions on the PMMA cylindrical phantom with ACPDP and TPS...............................34

Table 3.2. Volumetric Gamma analysis passing rates (%) comparing D4 PB with ACPDP and TPS on a water-equivalent cylindrical phantom and actual patient CT datasets........................................34

Table 5.1. Percent differences (local) between calculated and ion chamber (IC) measured Exit to Entrance and Exit to Center dose ratios for the various size static beams on the ArcCHECK Phantom.................................89

Table 5.2. Descriptive statistics of the percent ion chamber dose difference from the TPS and TPDP (without chamber correction). ........................................90

Table 5.3. Descriptive statistics of the γ-analysis percent passing rates, TPS vs. ArcCHECK ........................................................................................................91

Table 5.4. Descriptive statistics of the γ-analysis passing rates with different criteria combination, comparing TPDP with and without IC correction to the Delta4 measurements for the 1 cm jaw plans (N=18). Median Dose Difference: No IC: -1.3±1.0%; IC: -0.8± 1.0 % ......92

Table 5.5. Same as Table 5.4 but for 2.5 cm jaws (N=19). ........................................93

Table 5.6. Descriptive statistics of the γ-analysis passing rates with different criteria combination, comparing TPS vs. TPDP and TPS vs. Delta4 for the 2.5 cm jaw plans (N=19). .........................................................94
LIST OF FIGURES

Figure 1.1 A sample MLC configuration ................................................. 3
Figure 1.2 Sample dose profile ................................................................. 7
Figure 1.3 A sample Dose Volume Histogram .......................................... 8
Figure 1.4 A sample Cumulative Dose Volume Histogram ......................... 9
Figure 2.1 Physical Delta$^d$ with its planar boards and “X” configuration ........ 14
Figure 2.2 Physical ArcCheck with schematic of the helical diode configuration .... 17
Figure 3.1 Dose comparison arrangements .................................................. 30
Figure 3.2 Relative lateral dose profiles in X direction (cross-plane) for a 2x2 cm$^2$ MLC-defined field .......................................................... 32
Figure 3.3 Relative lateral dose profiles in direction (in-plane) for a bar pattern MLC-defined field .......................................................... 33
Figure 3.4 Absolute dose profiles for a Pancreas SBRT case ....................... 35
Figure 3.5 Absolute dose profiles for a Head and Neck case ....................... 36
Figure 3.6 Absolute dose profiles for a Lung SBRT case ............................ 37
Figure 4.1 The moving target is the GTV drawn on Phase 0% and rigidly translated according to the centroid position on each phase (two positions with maximum separation shown) ........................................ 48
Figure 4.2 Quantification of interplay ......................................................... 53
Figure 4.3 Low interplay scenario ............................................................... 55
Figure 4.4 Hypothetical substantial interplay scenario ................................ 56
Figure 4.5 Areas of differences above 5% for the dose distributions with 60 s motion period and 3 cm range, corresponding to the starting phases of 0 and 50% ..................................................... 57
Figure 4.6. Reduction in interplay effect in $D_{99\%}$ with the number of fractions presented per patient (X-axis groups). ................................................................. 59

Figure 5.1. The workflow diagram of the Tomo-PDP process. .................................................. 72

Figure 5.2. In this figure, the top and bottom row show two different 0.25 second intervals of MLC leaf arrangements vs. the entry diode doses (as a function of time) measured by the 3D dosimeter ................. 75

Figure 5.3. Visual representation of 3DVH dose reconstruction............................................. 80

Figure 5.4 Median percent dose-difference between the TPS and ArcCHECK for cylindrical targets of varying diameter, with 1 and 2.5 cm jaws............ 98

Figure 5.5. A) Central axial plane for the TG-119 prostate, 2.5 cm jaw plan, pitch 0.286. ................................................................. 103
ABSTRACT

The purpose of this work is to evaluate quasi-3D arrays for use with intensity modulated radiation therapy (IMRT) and volumetric modulated arc therapy (VMAT) and to determine their clinical relevance. This is achieved using a Delta4 from Scandidos and ArcCheck from Sun Nuclear and the associated software. While certain aspects of these devices and software have been previously evaluated, the main goal of this work is to evaluate the new aspects, such as reconstructing dose on a patient CT set, and extending the capabilities. This includes the capability to reconstruct the dose based on a helical delivery as well as studying the dose to a moving target using measurement-guided motion simulations.

It was found that Sun Nuclear’s ArcCheck/3DVH system exhibited excellent agreement for dose reconstruction for IMRT/VMAT using a traditional C-arm linear accelerator and stringent 2%/2mm comparison constraints. It also is a powerful tool for measurement-guided dose estimates for moving targets, allowing for many simulations to be performed based on one measurement and the target motion data. For dose reconstruction for a helical delivery, the agreement was not as good for the stringent comparison but was reasonable for the clinically acceptable 3%/3mm comparison. Scandidos’ Delta4 shows good agreement with stringent 2%/2mm constraints for its dose reconstruction on the phantom. However, the dose reconstruction on the patient CT set was poor and needs more work.
Overall, it was found that quasi-3D arrays are powerful tools for dose reconstruction and treatment plan comparisons. The ability to reconstruct the dose allows for a dose resolution comparable to the treatment plan, which negates the previous issues with inadequate sampling and resolution issues found when just comparing the diodes. The ability to quickly and accurately compare many plans and target motions with minimum setup makes the quasi-3D array an attractive tool for both commissioning and patient specific quality assurance.
CHAPTER 1:  
INTRODUCTION

1.1 Background

In radiation therapy, quality assurance (QA) is an essential aspect in assuring that the most accurate treatments are delivered to a patient. When a new machine is commissioned at a hospital, the parameters of the system, for example radiation beam symmetry and flatness, are tested to see if they comply within the manufacturer’s specifications and published guidelines. The machine is then periodically tested on a daily, monthly and yearly basis to make sure that it remains within the specifications. This ensures that the machine continues to deliver what is indicated by a plan.\(^{(1)}\)

For simple, traditional treatment methods (forward planned 2D/3D plans), the quality assurance on the machine level is largely sufficient. Individual treatments are delivered without further empirical testing because possible errors are considered acceptable\(^{(1)}\). However, for intensity modulated inversely planned treatments, the increased complexity along with high dose gradients in the vicinity of critical structures, result in dosimetric errors, which are less predictable and forgiving and as such, require empirical testing.

The main goal of radiation therapy is to deliver a prescribed dose to a target while minimizing the dose to the surrounding normal tissue\(^{(2)}\). As new techniques are developed to achieve this goal, the treatments become more complex and the
importance of having accurate dosimetry methods for both initial system commissioning and ongoing QA increases. Currently, it is essentially mandated in the United States that a patient specific quality assurance (PSQA) test be performed prior to each new treatment course. In this work, the focus will be on the 3D and 4D (time-resolved) dosimetry for inversely planned treatment techniques such as IMRT (intensity modulated radiation therapy) and VMAT (volumetric modulated arc therapy).

1.2 IMRT & VMAT

While in conventional radiation therapy the radiation portal conforming to the target is irradiated in a uniform fashion, in IMRT/VMAT the radiation beam intensity is modulated across the aperture. The exact profile shape is determined by a computer algorithm through a process known as inverse planning \(^3,4\). IMRT has the ability to spare the normal structures immediately surrounding the target, particularly for concave targets \(^5\). Beam intensity modulation can be achieved in two principal ways: by physical attenuators (compensators \(^6\)), or by dynamically changing the beam shape with the help of a multi-leaf collimator (MLC). The latter can be accomplished on a slice-by-slice basis (serial) \(^7\), and helical tomotherapy \(^8\) or volumetrically on a conventional C-arm linear accelerator equipped with an MLC \(^9\). The last method can be further broken down into step and shoot IMRT and dynamic (“sliding window”) IMRT. In step and shoot, the leaves remain static during irradiation (“shoot”) and the beam is turned off when the MLC leaves are moving (“step”). In dynamic IMRT, the beam is left on while the MLC leaves continuously move through the field. This changes the fluence of the beam during treatment allowing the dose to be conformed to the target \(^9\). While related to IMRT, VMAT differs from it in that during delivery of the beam, the gantry rotates
around the patient. The conformity of the dose to the target is controlled by the MLC shape as well as the gantry rotation speed and the dose rate, which can be all simultaneously varied \(^{(10)}\). VMAT has decreased treatment time and increased monitor unit (MU) efficiency compared to IMRT but the optimization time of the plan can be longer \(^{(10,11)}\).

![A sample MLC configuration](image)

**Figure 1.1** A sample MLC configuration

### 1.3 Patient Specific Quality Assurance (PSQA)

For these techniques, PSQA is performed to assure that three things are true. First, it is performed to make sure that the treatment planning system (TPS) accurately calculated the dose for the planned treatment. Second, it verifies that pertinent data has been appropriately transferred to the treatment system. Finally, the PSQA tests to see if
the delivery system is capable of delivering the fields as planned \(^{(12)}\). Since it is, for the most part, impractical to measure the dose in a patient, the plan is recalculated on a suitable phantom. The phantom can then be treated and the measured dose can be compared to the calculated TPS dose. It is assumed that if the planning system can accurately predict the dose to the phantom, it will also be correct for the patient dose.

1.4 Dose Measurement Techniques

There are different methods for measuring the dose in the phantom. Historically, the gold standard was to use an ion chamber to collect, at most, a few absolute point doses and radiographic film to sample the dose distribution in one, or at most a few, planes. While indispensable in original development and clinical implementation of IMRT, this method is rather time consuming, provides limited spatial information, and lacks real-time readout. While the ion chamber measurements are still required as part of the system commissioning, new QA devices and methods have been developed that make it possible to routinely determine the distribution and absolute dose in fewer iterations and less time \(^{(13)}\).

A two-dimensional (2D) array QA device consists of a number of individual radiation detectors (diodes or ion chambers) arranged on a plane. They are typically spaced 5-10 mm apart. Similar to film, the array can be placed in a plane of the phantom when the beam is delivered and it only gives dosimetric information in that one plane. However, instead of having just one absolute dose measurement, each detector records dose delivered and, after applying a correction factor, can provide an absolute dose reading throughout the array. The fundamental limitation of a single planar array is its inability to record the two-dimensional modulation information for the full range of
gantry angles. When the beam central axis is parallel to the array plane, the 2D dimensional modulation map degenerates into 1D. To counteract this limitation, the concept of a quasi-three dimensional QA device was introduced, where instead of a single plane, the detectors (typically diodes) are arranged, for example, on two orthogonal planes or a cylindrical surface. In either case, the full 2D modulation information is preserved for any beam angle. As a result, for IMRT, the dose can be compared on a beam-by-beam basis or as a composite plan (when all the beams are delivered at their planned gantry angles). For VMAT, only composite measurements are possible by definition, due to the continuous rotation of the gantry \(^{(12)}\). The quasi-3D arrays combine the ease of a composite dose measurement with the details of a field-by-field measurement.

1.5 Dose Analysis Techniques

Once the dose at multiple points is measured, a metric should be introduced to compare it to the planned dose. For simple point dosimetry, a percentage difference comparison is adequate. While this can be extrapolated to comparing multiple point doses throughout a phantom, this method tends to be unreliable in high dose gradient areas. In these areas, it is more favorable to use a distance to agreement (DTA) metric. This is the distance from the point of interest to the closest point of matching dose. With both of these methods, the comparison is reported in the number of points passing a certain criterion (ex. 3% percent difference and 3 mm DTA). Gamma analysis is a method proposed by Low et al\(^{(14)}\) to appropriately combine these analyses into a comprehensive comparison. Gamma is given at every point by the equation

\[
\gamma(r_m) = \min[\Gamma(r_m, r_c)] \forall r_c
\]
Where

\[
\Gamma(r_m, r_c) = \sqrt{\frac{r^2(r_m, r_c)}{\Delta d_M^2} + \frac{\delta^2(r_m, r_c)}{\Delta D_M^2}}
\]

The terms \(r_m\) and \(r_c\) are the reference and evaluated points (measured vs calculated) and \(\Delta d_M\) and \(\Delta D_M\) are the DTA and dose-difference criteria respectively. The dose difference and distance are given by

\[
\delta(r_m, r_c) = D_c(r_c) - D_m(r_m) \quad r(r_m, r_c) = |r_m - r_c|
\]

where \(D_c(r_c)\) and \(D_m(r_m)\) are the doses at the measured and calculated points respectively. The dose difference can compare the evaluated point to the equivalent point in the reference plan (local dose-difference) or it can be compared to a maximum reference dose (global dose-difference). In low dose regions, the local dose-difference can produce inflated values since the local dose difference criteria (ex. 3% of the local dose) produces a smaller denominator. The global dose difference has the benefit of having the same dose difference criteria for all points (ex. 3% of the max dose) but when comparing only parts of the plan instead of the volumetric dose, it is not always clear what value should be used for the max dose.

The value of \(\Gamma\) is found for a region around each point and the smallest value becomes \(\gamma\) for the measure point (see Figure 1.2). A threshold can be set which is a dose value below which the percentage difference and DTA statistics would be ignored. This aids in the speed of calculations but can have an affect on the overall passing rate. Gamma analysis is repeated for each point in the reference set. A point is said to pass if the gamma value at that point is less than one for a given set of criteria (e.g.
3%/3mm)\(^{(14)}\). The TG-119 report used the statistical process analysis method to provide expected ranges for the gamma analysis passing rates with 3%/3mm criteria, which have become *de-facto* industry.

![Sample dose profile](image)

**Figure 1.2** Sample dose profile.

Given the reference curve (red) and the evaluated curve (blue), the minimized gamma would be dominated by the dose difference at point A. At point B the gamma function minimization would be dominated by the DTA. In the transition from the two extremes of low dose to high dose gradient, gamma analysis has the benefit of analyzing both dose difference and DTA to achieve a true minimum.

While gamma-analysis is a reasonable tool for system commissioning, where the level of agreement between the calculated and measured dose across the entire phantom volume is of primary interest, its ability to provide clinically relevant comparisons for PSQA’s is unclear. In addition to failure to detect systematic errors when the 3%/3mm metric is used\(^{(15,16,17,18)}\) spatially limited gamma-analysis shows counter-intuitive, if any, correlation with the clinically relevant dose-volume metrics. To be more intuitive and clinically relevant, volumetric (3D) comparison of measured versus planned dose is desirable. To reduce the vast amount of spatial dose information to a
comprehensible set of numbers, the volumetric dose to a region of interest (ROI) is presented in a form of histogram of dose vs. volume receiving this dose (a dose-volume histogram, or DVH).\(^{(19,20,21)}\) The process starts with a differential histogram where the dose values from each voxel in the volume are binned and the frequency data is reported with the dose ranges on the horizontal axis and the number of voxels on the vertical.

![A sample Dose Volume Histogram](image)

Figure 1.3  A sample Dose Volume Histogram

The shape of the histogram is highly dependent on the dose bin size chosen. The more clinically utilized DVH is the cumulative one. It is calculated by summing the differential dose starting at the dose of interest, D, up to the max dose, Dmax.
The cumulative DVH displays the number of voxels in a volume, which receive at least a dose $D$. Therefore, ROI in a plan that are targets will ideally have a rectangular shape where 100% of the target will receive the desired dose and then 0% will receive a higher dose. Alternatively, critical structures ROI’s would ideally have a cumulative DVH that drops immediately to zero.

Figure 1.4 A sample Cumulative Dose Volume Histogram

In order to obtain reconstructed DVH’s, 3D dose must be reconstructed which can be done using the quasi-3D arrays described in the following chapter.
1.6 References


17. Zhen et al. Moving from gamma passing rates to patient DVH-based QA metrics


CHAPTER 2:
MATERIALS & METHODS

2.1 Delta

2.1.1 Hardware

Delta (ScaniDos AB, Uppsala, Sweden) was the first commercially available quasi-3D dosimeter array. It is a 22 cm diameter Poly(methyl methacrylate) (PMMA) cylindrical phantom broken into four pie shaped sections with three detector boards, creating an X shape in the axial direction. Each board is 10 mm thick and consists of 9.5 mm PMMA and 0.5 mm fiberglass printed circuit board with diodes soldered to copper conductors. There are a total of 1069 p-type cylindrical silicone diodes, each with a 1 mm diameter and a 0.05 mm thickness. The diode nominal sensitivity is approximately 5 nC/Gy but it varies somewhat from diode to diode. The diodes are arranged in a rectangular pattern on a 20 x 20 cm² plane (the main board) and two 20 x 10 cm² planes (the wings). In the central 6 x 6 cm² region, the diodes are spaced 0.5 cm apart, while everywhere else they are 1 cm apart.

The boards are removable from the cylinder and the PMMA is milled out near the diodes to accommodate them, which creates small air cavities around the diodes. The boards are angled at +50° from the vertical for the main board and -40° for the wings. This allows for added flexibility to make sure the beam is not parallel to the boards. The orientation of the Delta needs to be indicated in the software. The electrometers are
connected through Ethernet cables to the portable network switch and the switch is connected to the dedicated network card in the control PC.

The device receives a trigger signal from the accelerator right before the dose pulse. The electrometers then integrate the charge just before, during and after the dose pulse, which improves signal to noise ratio and adds a temporal component. Several pulses are recorded, packaged and sent to the PC. For rotational treatments, an independent inclinometer can be mounted on the gantry to provide gantry angle information. This allows dose per control point to be examined.

Figure 2.1 Physical Delta$^4$ with its planar boards and “X” configuration
2.1.2 Software

Taking raw readings and applying correction factors produces the measured dose. During calibration, the dose at the reference detector location is calculated by the TPS in the measurement geometry. This value is scaled by the accelerator output and supplied to the software as the measured dose at the reference detector. The specific diode sensitivity across the boards is also collected and thus a correction factor can be applied to account for the varying sensitivity. There is also a directional correction factor that corrects for the beam angle and the diodes' position along the longitudinal axis of the phantom. Rotational, depth and field size correction are also applied on a segment-by-segment basis to each diode based on the conditions. A temperature correction can be applied to account for any temperature fluctuations between calibration and measurement. In its basic implementation, the measured dose at the detector positions is compared to the planned dose (on the Delta$^4$ cylindrical phantom) extracted from the DICOM RT DOSE object transmitted from the TPS.

Along with a gamma analysis comparison, the Delta$^4$ software has an option to compare DVH. This is done by using a volumetric interpolation inside the Delta$^4$ phantom. This 3D volumetric dose reconstruction on the phantom is fairly straightforward and was previously validated.$^{(1,2,3)}$ The Delta$^4$ calculates the 3D dose depending on what planned data are available. It is either calculated using TPS data or using PDD (percentage depth dose) data. For the TPS method, the dose is calculated by taking the known planned dose along each beam ray and renormalizing it using the ratio between the planned dose and the measured dose at the point of intersection between the ray and the detector board. If control point data is missing for an arc plan,
then PDD data is used to calculate the 3D dose. This method requires depth dose
distributions to be entered into the Delta\textsuperscript{4} software. Here, the PDD is normalized to the
measured dose for every control point and detector position. The renormalized depth
doses for all detector positions compose the 3D dose for that control point. This is
repeated for each control point and the summation is compared to the planned dose
through gamma analysis. The criteria for passing can be controlled. By transferring the
defined regions of interest (ROI) from the patient to the phantom, the DVHs between the
planned and reconstructed doses can be compared. The gamma analysis passing rates
can be spatially filtered by the ROI. This optional software package called Delta\textsuperscript{4DVH}
gives the user more control when trying to determine if dose errors are clinically
relevant.

Finally, a new comparison option, Delta\textsuperscript{4DVH} Anatomy, has been recently
released. This option calculates the delivered dose to the real patient based on the
measurement data and patient’s CT images and compares this dose to the TPS
planned dose. The process of this calculation method will be examined in greater detail
later in this work.

2.2 ArcCheck

2.2.1 Hardware

Similar to the Delta\textsuperscript{4}, the ArcCheck phantom (Sun Nuclear Corp, Melbourne, Fl)
is a quasi-3D dosimetry system. However, instead of the X configuration of the Delta\textsuperscript{4},
the ArcCheck is a doughnut shaped phantom with an outer diameter of 26.6 cm and an
inner diameter of 15.1 cm. A total of 1386-point diodes form a 2D array on a curved
plane 10.4 cm from the center of the phantom. These diodes form a helical pattern and
are spaced 1 cm apart along both the cylindrical length and the circumference. The detector array is nearly cylindrically symmetrical which reduces the rotational response dependence. Also, the design of the diode arrangement increases the apparent detector density in the beams eye view (BEV) by shifting the exit diodes compared to the entrance ones. The detector array length is 21 cm and the PMMA phantom surrounding the array is 32.4 cm long. An additional 11.9 cm is composed of electronics giving the device an overall length of 44.3 cm. There is an option to use a PMMA plug in the cavity of the device to eliminate the inhomogeneity of the device. Instead of having an external inclinometer like the Delta\(^4\), ArcCheck has a virtual inclinometer that has been shown to improve the accuracy of the absolute dose measurement\(^{(4)}\).

Figure 2.2  Physical ArcCheck with schematic of the helical diode configuration
2.2.2 Software

Once the raw data is gathered by the phantom, correction factors are applied to obtain the measured dose.

\[ D_l = R_i N_{ref}^T C_{RR}^T C_{rel}^T C_{FS}^T C_{pos}^T \]

\( R_i \) is the raw reading. The absolute dose calibration factor for a normalization (reference) diode in the array, at a known accelerator dose rate, \( N_{ref}^T \) is applied, which converts the raw reading to dose based on a previous calibration. This is followed by a temperature correction, array correction for the accelerator pulse repetition rates, sensitivity correction per diode, field size correction per diode and finally position correction for each diode with respect to the radiation beam.

Once the measured data is available, the ArcCheck software can compare the measured and planned dose and report the gamma analysis. This can be done in a relative or absolute manner and the gamma analysis can be reported locally or globally.

2.3 3DVH

An extension of the software for ArcCheck beyond gamma analysis is 3DVH. 3DVH is a dose delivery QA software application that is capable of comparing 3D dose and DVH’s in the patient (as opposed to a phantom). It can also be used to compare a measured dose from ArcCheck to the calculated dose-to-phantom. 3DVH can display metrics such as DVH, ROI statistics and volumetric dose review as well as the conventional gamma analysis passing rates. This allows for a more in depth comparison. The software also allows the high-density 3D calculated dose-on phantom to be viewed in its entirety, instead of just the 2D plane.
2.3.1 Measurement guided dose estimates to static patient voxels

The 3DVH system was designed to use “planned dose perturbation” (PDP) for estimating the dose delivered to the patient. This is done by incorporating the beam-by-beam phantom doses back into the patient’s images, structures, and TPS dose by using the dose differences found between the device measurement and the TPS dose calculation. The differences are then projected back into the TPS 3D dose calculation to obtain an estimate of the actual delivered 3D dose distribution\(^\text{(5)}\). For VMAT, ArcCheck records the dose in 50 ms intervals and the virtual inclinometer is used to synchronize the dosimeter with the gantry position and planned control points. The beam is discretized into sub-beams with fixed gantry angle but variable fluence due to the MLC motion during the time allotted for the sub-beam. The relative volumetric dose is calculated for each sub-beam on a virtual phantom by a convolution engine. The dose grids are then morphed based on the measured data and summed to produce a high-density 3D absolute dose matrix on the cylindrical phantom. A ratio is taken between this reconstructed dose and the planned one and then the planned dose on the patient is perturbed by the found ratio. This yields the measurement-guided estimate of the 3D dose on a patient. This can be summarized by the following:

\[
D_{\text{Pat}}^{\text{ACPDP}}(r_i) = D_{\text{Pat}}^{\text{ACPDP}}(r_i) \frac{D_{\text{Phant}}^{\text{ACPDP}}(r_i)}{D_{\text{Phant}}^{\text{TPS}}(r_i)}
\]

where \(D_{\text{Pat}}^{\text{ACPDP}}\) is the patient dose estimated by ACPDP, \(D_{\text{Pat}}^{\text{TPS}}\) is the patient dose calculated by the TPS, and \(D_{\text{Phant}}^{\text{ACPDP}}\) and \(D_{\text{Phant}}^{\text{TPS}}\) are the doses in phantom reconstructed by PDP on the ArcCHECK (ACPDP) and calculated by the TPS, respectively and has been thoroughly validated on both homogenous and heterogeneous phantoms\(^\text{(6,7,8)}\).
2.3.2 Measurement guided dose estimates to moving patient voxels

While software such as Delta\textsuperscript{4} and 3DVH have evolved to better analyze patient specific QA’s, this may not be enough as the interplay between the dynamic delivery and intrafraction organ motion could potentially alter doses accumulated by different structures\textsuperscript{(9,10,11)}. Therefore, this motion may need to be incorporated into patient dose reconstruction.

There have been motion studies performed previously, either with motion tables\textsuperscript{(12)} or virtual TPS-based simulations of the interplay effects\textsuperscript{(13,14)}. These methods are time consuming and tend to look at the dose to the phantom. There have also been methods presented for 4D dose reconstruction. However, these methods are either very complex\textsuperscript{(15)} or are impractical for clinical use\textsuperscript{(6,5)}. Therefore, it is desirable to achieve measurement-guided VMAT dose reconstruction to moving patient voxels from a known motion kernel and the static phantom measurement data. This is achieved by treating the motion as another perturbation to the ACPDP model.

As described above, the VMAT 3D dose to a static patient can be estimated by applying a phantom measurement-guided perturbation to the TPS-calculated dose grid. The fraction dose to any voxel in the presence of motion, assuming the motion kernel is known, can be derived in a similar fashion by applying a measurement-guided motion perturbation.

\[
D_{Mov}^{Pat}(r_i) = D_{St}^{Pat}(r_i) \frac{D_{Mov}^{Phant}(r_i)}{D_{St}^{Phant}(r_i)}
\]
where $D_{St}^{Pat}$ is the static ACPDP patient dose, $D_{St}^{Phant}$ is the static ACPDP phantom dose, and $D_{Mov}^{Phant}$ is obtained from

$$D_{Mov,i}^{Phant} = \sum_{m=1}^{M} \sum_{k=1}^{K} \frac{d_m(r_i)}{\Delta T_m} \Delta t_k$$

The term $d_m/\Delta T_m$ is the average dose rate in the phantom at the voxel position $r_i$, during the delivery of the $m^{th}$ dose grid in the 4D array, and $\Delta t_k$ is the smallest time interval considered in the summation.

The dose to the diodes in a helical phantom is recorded at 50 ms intervals and is transformed into a series of time-resolved high-density volumetric dose grids. A moving voxel is propagated through this 4D dose space and the fraction dose to that voxel in the phantom is accumulated. The ratio of this motion-perturbed, reconstructed dose to the TPS dose in the phantom serves as a perturbation factor, applied to the TPS fraction dose to the similarly situated voxel in the patient.

This approach was validated by the ion chamber and film measurements on four phantoms of different shape and structure (8): homogeneous and inhomogeneous cylinders, a homogeneous cube, and an anthropomorphic thoracic phantom. A 2D motion stage was used to simulate the motion. The stage position was synchronized with the beam start time with the respiratory gating simulator.

The ion chamber errors predominantly under 2% put the accuracy of the method on par or better than the published values acceptable for empirical motion studies (2-5%) (11,15). The film vs. motion-perturbed 3DVH $\gamma$-analysis results (3%/3mm with global dose-error normalization) are within the range currently considered acceptable for IMRT/VMAT plans delivered to the static phantoms ($\geq$ 90% passing rate) (16,17,18). This
should be considered a very good agreement for a moving phantom, considering the additional uncertainties introduced by the motion.

Based on this study (8), the dose to a moving voxel in a patient can be estimated for a VMAT delivery by performing a single QA measurement with a cylindrical phantom and applying two consecutive perturbations, based on that measurement, to the TPS-calculated dose to the patient. The first one accounts for the differences between the planned and delivered static doses (6), while the second quantifies the effects of motion and interplay between a moving target and a dynamic dose delivery.

To be useful in practice, two modules now need to be added to the prototype proof of principle software described in this chapter. One is the interface to streamline the input of the realistic motion kernels. Another is the analysis tool to generate and statistically analyze accumulated dose differences when multiple fractions with random starting motion phases are considered. This would allow an estimate for any given patient on how interplay would affect clinical tumor volume (CTV) dose coverage, compared to the treatment plan based on the ITV (internal target volume) to PTV (planning target volume) expansion.

2.4 References


CHAPTER 3:

SEMI-EMPIRICAL VMAT DOSE RECONSTRUCTION ON A PATIENT DATASET USING DELTA⁴ MEASUREMENTS¹

3.1 Introduction

Direct comparison of the planned and deliverable dose volume histograms (DVHs) exhibits higher sensitivity and specificity than gamma analysis. It is also expected to be more clinically meaningful for both physicians and physicists. Therefore, Scandicos released a software module called Delta⁴ Anatomy, which allows semi-empirical dose reconstruction on the patient CT based on the phantom measurements from their biplanar diode array dosimeter. This chapter presents the results of the initial tests of the system, focusing on its performance in VMAT patient dose reconstruction.

3.2 Treatment Planning and Delivery

Dose calculations were performed with Pinnacle treatment planning system (TPS) v 9.2 (Philips Radiation Oncology Systems, Fitchburg, WI) using Collapsed Cone Convolution algorithm. The test plans were arranged in an order of increasing complexity, starting with profile comparisons for simple static fields and progressing to VMAT dose comparisons first on a homogeneous cylindrical phantom, and then on a patient CT dataset. For the VMAT tests, five plans previously treated at our institution

¹ This chapter is partially reproduced from previously published work. The author of this dissertation is the first author of the previously publish work. All figures are used with copyright permission. See Appendix A and B.
were selected: three Head and Necks of varying complexity with conventional fractionation (2 Gy/fraction), one Pancreas stereotactic body radiation therapy (SBRT) plan (6 Gy/fraction) and one Lung SBRT (10 Gy/fraction). All VMAT plans were calculated on a 2.5 mm dose grid with 4° control point (CP) angular increment. All plans employed a 6MV beam from a TrueBeam linear accelerator equipped with a 120-leaf Millennium MLC (Varian Medical Systems, Palo Alto, CA).

3.3 Volumetric Dose Reconstruction Methods

As discussed previously, Delta⁴ can compare the planned to measured dose through direct measurement at the diodes’ locations, volumetric interpolation inside the Delta⁴ phantom (further referred to as D4 Interpolation), and dose reconstruction on an external patient CT dataset (further referred to as D4 PB). While the initial two have been previously validated (14-16,20,21) the accuracy of the third method is the main subject of this investigation and therefore the details of the dose reconstruction are further discussed.

For each CP, the most likely photon fluence is deduced from the phantom measurement. This fluence is then used to recalculate the dose on the patient CT dataset with a pencil beam (PB) algorithm. To perform this calculation, the beam must be first characterized in the Delta⁴ software. Specifically, percent depth doses on a water phantom and in-air relative output factors (S₂) are required for a set of field sizes, typically from 2x2 to 20x20 cm². The patient CT, dose, and structures are imported from the TPS as DICOM RT objects. A generic CT to density conversion table is used for PB calculations, with no option to override CT densities. Once the calculation is complete, the patient reconstructed 3D dose grid can be compared to the planned one.
with a set of standard analysis tools, such as dose profiles, 3D γ-analysis, and DVH comparisons for selected structures. We will call this dose reconstruction method “D4 PB”.

The primary comparison method we used to volumetrically evaluate accuracy of the D4 PB reconstruction algorithm is measurement-guided dose reconstruction by 3DVH software based on the ArcCHECK diode array measurements (ACPDP), which was described previously.

3.4 Specific Tests

While the primary goal was to compare D4 PB to ACPDP, the D4 Interpolation and TPS data were also collected and analyzed when appropriate.

3.4.1 Simple static fields

Two simple MLC-defined field arrangements were used: a 2x2 cm$^2$ square and a bar pattern (a set of 2 cm wide openings separated by 2 cm areas of closed leaves)$^{(22)}$. In each case a scan at 100 cm source to surface distance and 10 cm depth was obtained in a water tank with a Model PFD-3G diode (IBA Dosimetry GmbH, Schwarzenbruck, Germany). The diode’s sensitive volume diameter is 2 mm. The bar pattern was scanned in the Y direction (in-plane), while the square field was scanned in the X direction (cross-plane), corresponding to the MLC leaf movement direction. For the latter, the scan position was offset from the central axis by 2.5 mm to scan in the middle of the leaf. The same plans were delivered in a standard fashion$^{(12,18)}$ to the calibrated Delta$^4$ and ArcCHECK dosimeters. D4 PB and ACPDP reconstructions were performed on a CT scan of the rectangular Plastic Water (CIRS Inc., Norfolk, VA)
phantom ("patient"). The diode scans were compared with the relative dose profiles extracted from the D4 PB and ACPDP reconstructions, and TPS calculations.

In addition, the D4 Interpolation and ACPDP reconstructions were performed on a PMMA cylinder. The profiles from those data and the TPS-calculated profiles were compared at the 100 cm source to axis distance. All calculations and reconstructions were done on a 2 mm grid.

3.4.2 VMAT dose reconstruction on homogeneous phantoms

3.4.2.1 PMMA Delta⁴ cylindrical phantom.

The phantom density value relative to water was set to 1.147 in the TPS and the dose was calculated. Delta⁴ and ArcCHECK VMAT measurements were performed. For each plan, ACPDP reconstruction was performed on a cylindrical PMMA phantom ("patient"). First, the samples of the ACPDP and TPS 3D dose grids were compared to the directly measured Delta⁴ dose at the diode's locations. Then the ACPDP and TPS volumetric doses were compared to the D4 Interpolation 3D dose grid (Figure 3.1A). All comparisons here and elsewhere in the chapter used γ-analysis with local (at the evaluated point) dose-error normalization. Passing rates with both 3%/3mm and 2%/2mm threshold combinations are reported. Dose points receiving less than 10% of the maximum dose were excluded from evaluation, as customary.(3) All γ-analyses were performed in 3D, as implemented in the Delta⁴ software (February 2013 release). This was the only readily available option since there is no DICOM RT DOSE object export capability with Delta⁴. Representative dose profiles were exported to illustrate crucial areas of disagreement.
Figure 3.1. Dose comparison arrangements. A: Delta\textsuperscript{4} dose is either directly measured (at the detectors) or interpolated (D4 Interpolation) inside the native PMMA phantom. The ACPDP (reconstructed) and TPS (calculated) dose distributions compared to the Delta\textsuperscript{4} are also on the same PMMA cylinder; B: Delta\textsuperscript{4} Pencil Beam reconstruction (D4 PB), ACPDP reconstruction, and TPS calculation are compared on the water-equivalent cylindrical phantom; C: Same as B but comparisons are done on the corresponding patients' CT datasets.

### 3.4.2.2 Water-equivalent Delta\textsuperscript{4} cylindrical virtual phantom.

Since the D4 PB algorithm uses a generic CT to density conversion table, we decided to do the comparisons of D4 PB with ACPDP on a unit density cylindrical phantom, to eliminate the additional uncertainty associated with the CT number to density assignment. The virtual Delta\textsuperscript{4} phantom supplied by the manufacturer was
modified programatically so that the Hounsfield units were uniformly set to zero. This phantom was used as “patient” for D4 PB and ACPDP reconstructions, and also for the TPS calculation (Figure 3.1B). The resulting volumetric dose grids were compared three ways as described above.

3.4.3 VMAT dose reconstruction on patient CT datasets

The same dose reconstruction and comparison procedure were used, except an appropriate patient CT dataset was substituted for the Delta⁴ phantom for each case (Figure 3.1C).

3.5 Results

3.5.1 Simple static fields

A set of cross-plane (X) beam profiles in water or water-equivalent material is presented in Figure 3.2. As expected, a water phantom diode scan essentially overlays the TPS calculation (Figure 3.1A), since the MLC model optimization was based on a series of such scans. While ACPDP- reconstructed profile exhibits reasonable agreement with the water scan, the D4 PB penumbra shape is substantially different. Figure 3.2B shows similar profiles reconstructed or calculated on a homogeneous PMMA phantom. While the ground truth diode scan is not available for such configuration, Figure 3.1A demonstrates that the TPS profile can serve as a good approximation. A proven method of Delta⁴ volumetric dose reconstruction, D4 Interpolation, produces much better agreement in the penumbra region with the TPS and ACPDP profiles. Figure 3.3 demonstrates a similar trend for the profiles taken in the in-plane (Y) direction for a series of rectangular MLC openings (a bar pattern). D4
Interpolation again reproduces the true profile shape better than D4 PB. By looking at Figure 3.2A and Figure 3.3A, one would *a priori* expect substantial errors in D4 PB composite dose reconstructed for modulated beams comprised of small segments.

Figure 3.2. Relative lateral dose profiles in X direction (cross-plane) for a 2x2 cm$^2$ MLC-defined field. **A**: SSD 100 cm, depth 10 cm. The diode scan was taken in water, ACPDP and D4 PB reconstruction and TPS calculation are all on a Plastic Water phantom; **B**: SAD 100 cm, depth 11 cm in PMMA (12.6 cm water equivalent). ACPDP and D4 Interpolated reconstructions, and TPS calculation are all on a cylindrical homogeneous PMMA phantom.
Figure 3.3. Relative lateral dose profiles in direction (in-plane) for a bar pattern MLC-defined field. **A**: SSD 100 cm, depth 10 cm. The diode scan was taken in water, ACPDP and D4 PB reconstruction and TPS calculation are all on a Plastic Water phantom; **B**: SAD 100 cm, depth 11 cm in PMMA. ACPDP and D4 Interpolated reconstructions, and TPS calculation are all on a cylindrical homogeneous PMMA phantom.

### 3.5.2 VMAT dose reconstruction

#### 3.5.2.1 PMMA Delta⁴ cylindrical phantom.

The γ-analysis passing rates are presented in Table 3.1. One can see good agreement between ACPDP and both directly measured and reconstructed (interpolated) Delta⁴ doses. The average γ passing rate with a rather stringent 2% (local normalization)/2mm criteria combination exceeds 90%.
Table 3.1. Gamma analysis passing rates (%) comparing Delta4 directly measured (Detectors Only) or D4 Interpolated (Volumetric) dose distributions on the PMMA cylindrical phantom with ACPDP and TPS. (See Figure 3.1A) The mean values for five VMAT cases are presented with standard deviations and ranges.

<table>
<thead>
<tr>
<th>Comparison region</th>
<th>Delta4 vs. ACPDP</th>
<th>Delta4 vs. TPS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\gamma(3/3) \leq 1$</td>
<td>$\gamma(2/2) \leq 1$</td>
</tr>
<tr>
<td>Detectors Only</td>
<td>99.1±0.6 (98.5-100)</td>
<td>94.1±4.7 (87.0-100)</td>
</tr>
<tr>
<td>Volumetric</td>
<td>98.2±1.3 (97.0-100)</td>
<td>92.8±3.9 (89.5-99.2)</td>
</tr>
</tbody>
</table>

3.5.2.2 Water-equivalent Delta4 cylindrical virtual phantom.

The results for D4 PB reconstruction on a homogeneous cylindrical phantom are quite different from the D4 Interpolation method results on a similar dataset (first data line in Table 3.2). Comparison with both ACPDP and TPS show the average agreement rate dropping by more than 10 percentage points for each comparison and threshold combination.

Table 3.2. Volumetric Gamma analysis passing rates (%) comparing D4 PB with ACPDP and TPS on a water-equivalent cylindrical phantom and actual patient CT datasets (See Figure 3.1B and C). The mean values for five VMAT cases are presented with standard deviations and ranges.

<table>
<thead>
<tr>
<th>Comparison Dataset</th>
<th>Delta4 vs. ACPDP</th>
<th>Delta4 vs. TPS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\gamma(3/3) \leq 1$</td>
<td>$\gamma(2/2) \leq 1$</td>
</tr>
<tr>
<td>Water Cylinder</td>
<td>88.6±6.8 (81.2-96.1)</td>
<td>72.4±8.4 (62.1-81.1)</td>
</tr>
<tr>
<td>Patient CT</td>
<td>81.2±8.6 (70.4-90.4)</td>
<td>64.6±8.4 (56.5-74.7)</td>
</tr>
</tbody>
</table>

3.5.2.3 Patient CT datasets.

Comparisons between D4 PB and ACPDP/TPS on the patient datasets show further deterioration of agreement compared to the homogeneous water-equivalent phantom (Table 3.2). The mean $\gamma$ passing rates are about 81 and 64% for the 3%/3mm and 2%/2mm criteria combinations, respectively, indicating substantial disagreement.
Absolute dose profiles presented for three cases in Figure 3.4, Figure 3.5 and Figure 3.6 further illustrate this progressive deterioration of dosimetric agreement.

Figure 3.4. Absolute dose profiles for a Pancreas SBRT case. A: Dose distribution and a Left-Right (X) profile location relative to it. B: Reconstruction/calculation on the Delta^4 cylindrical PMMA phantom – D4 Interpolation vs. ACPDP vs. TPS; C: Reconstruction/calculation on the water-equivalent Delta^4 phantom - D4 PB vs. ACPDP vs. TPS; D: same as C but on the patient dataset. Gamma passing rates are presented comparing volumetric dose distributions from ACPDP and an appropriate D4 reconstruction method, on the corresponding phantoms.

While D4 PB differs from D4 Interpolation primarily in the high gradient regions (penumbra) on the homogeneous phantom (compare Figure 3.4, Figure 3.5 and Figure 3.6 B to C), substantial disagreement with both ACPDP and TPS is observed in the relatively flat, high dose areas on the patients’ CT datasets (Figure 3.4, Figure 3.5, and...
Figure 3.6 D). This disagreement is reflected in the low gamma analysis passing rates (D4 PB vs. ACPDP) noted in the figures for individual cases.

Figure 3.5. Absolute dose profiles for a Head and Neck case. A: Dose distribution and an Antero-Posterior (Z) profile location relative to it. B: Reconstruction/calculation on the Delta4 cylindrical PMMA phantom – D4 Interpolation vs. ACPDP vs. TPS; C: Reconstruction/calculation on the water-equivalent Delta4 phantom - D4 PB vs. ACPDP vs. TPS; D: same as C but on the patient dataset. Gamma passing rates are presented comparing volumetric dose distributions from ACPDP and an appropriate D4 reconstruction method, on the corresponding phantoms.
3.6 Discussion

The D4 PB semi-empirical dose reconstruction method was evaluated with a variety of dosimeters and methods. First, the dose profiles for small MLC-defined fields were examined against water phantom scans with a small dosimeter. Volumetric comparisons were based on the measurement-guided dose reconstruction with an independent dosimeter (ACPDP). This 3D dose reconstruction method is by now thoroughly validated.\textsuperscript{18,19,21} Furthermore, in this work it was additionally shown to
provide a satisfactory level of agreement on a homogeneous geometric phantom at the 2%/2mm level, with both direct Delta\textsuperscript{4} measurements at the diodes’ locations and previously validated volumetric D4 Interpolation (Table 3.1). Finally, while it is inappropriate to use TPS calculations as a sole benchmark for evaluating accuracy of a dosimetry system, a previously validated TPS can provide additional confirmation.

All tests point towards the fact that the D4 PB algorithm produces substantial systematic errors. It cannot accurately predict the penumbra shape for small MLC-defined fields (Figure 3.2A and Figure 3.3A), which is crucial for correct reconstruction of the modulated beams. D4 Interpolation, using exactly the same measurement data, produces a much better agreement. Therefore, the differences in Figure 3.2A and Figure 3.3A are not due to the resolution limitations of the 5 mm diode spacing, but rather to the method of fluence extraction and/or dose recalculation in the D4 PB algorithm. There is nothing in the standard PB dose calculation algorithm that should inherently prevent it from modeling the penumbra in homogeneous water-equivalent media reasonably well. The observed discrepancies are due to the algorithm implementation. No user-adjustable parameters are available to tune the shape of the penumbra.

D4 PB also shows fairly poor agreement with ACPDP and TPS when volumetric VMAT dose distributions are compared on a homogeneous cylindrical phantom (Table 3.2, Figure 3.4, Figure 3.5, Figure 3.6 C). At the same time, D4 Interpolation using the same exact measurement data produces much better agreement with both the ACPDP reconstruction and TPS calculation (Table 3.1, Figure 3.4, Figure 3.5, Figure 3.6 B). Once again, this points towards the D4 PB dose reconstruction algorithm
implementation, as opposed to the hardware design, being the cause of disagreement. We can speculate that if the modeled incident fluence was interpolated on a finer grid, using the methods already available in the Delta4 reconstruction software, better agreement on a homogeneous phantom should be readily achievable.

Finally, the agreement between D4 PB and ACPDP/TPS is even worse when volumetric comparisons are performed on the patients’ CT datasets (Table 3.2, Figure 3.4, Figure 3.5, Figure 3.6 D). The exact cause cannot be ascertained without the intimate knowledge of the algorithm implementation, which is not available to us. We can only note that since the dose in the high dose area can be either high (Figure 3.4 D) or low (Figure 3.5, Figure 3.6 D), it is likely that more than one process is responsible for the disagreement, such as inaccurate penumbra modeling (Figure 3.2 A and Figure 3.3 A), and perhaps dose scaling with radiological depth. The use of a generic CT to density conversion table could be a contributing factor for the latter. Unlike in the homogeneous phantom, substantial inaccuracies are inherently expected in the lung, particularly at the tumor/lung interface due to the well-known limitations of the Pencil Beam dose calculation algorithm.\(^{(23)}\)

### 3.7 Conclusions

In summary, ACPDP, TPS and D4 Interpolation agree reasonably well at the 2%/2mm level. On the other hand, D4 PB, based on the exact same Delta\(^4\) measurements as D4 Interpolation, shows poor agreement with ACPDP, TPS, and water scans. Heterogeneous CT datasets present the biggest challenge. Modern electronic dosimetry arrays are sophisticated systems comprised of hardware, firmware, and software. An additional level of complexity is added when the measured dose on
relatively sparse detectors is used to reconstruct a high-resolution volumetric dose grid throughout the phantom. Even more complex can be the next step of dose reconstruction on the patient CT dataset. While this approach has the potential to provide for more intuitive and clinically useful evaluation of the patient-specific end-to-end tests,\(^{(7,10)}\) every system must be thoroughly tested before clinical use.

### 3.8 References


CHAPTER 4:

DOSE INTERPLAY EFFECTS IN VMAT SBRT LUNG TREATMENT²

4.1 Introduction

The challenge of tumors and other organs moving during radiation therapy is well documented and the history and current state of motion management techniques, as pertains to photon therapy, were recently reviewed by Korreman.(¹)

The effects of motion on the tumor dose can be broadly divided into the gradient (motion blur) and interplay effects.(²) In conventional 3D conformal radiation therapy delivered in many fractions of 1.8 – 2.0 Gy each, the goal is to deliver a homogeneous tumorcidal dose to the target. The gradient effect, historically first noted for this type of treatment, occurs with the occasional tumor excursion outside of the volume irradiated to a high, homogeneous dose. If the proper beam apertures are selected to encompass the entire volume potentially occupied by the target, the tumor will always be irradiated to the prescribed, relatively homogeneous dose, regardless of its motion path within the irradiated volume. However, in ablative type treatments such as stereotactic body radiation therapy (SBRT),(³) the target dose is deliberately highly inhomogeneous, with the cooperative lung treatment protocols specifying the maximum dose in the target volume between 15 and 40% higher than the prescription, RTOG 0813(⁴) being but one

² This chapter is partially reproduced from previously published work. The author of this dissertation is the first author of the previously publish work. All figures are used with copyright permission. See Appendix A and B.
example. Under these circumstances, the moving tumor dose will depend on the exact nature of the motion, even when the tumor does not venture outside the properly defined target volume. This is still a purely spatial effect and it would take place regardless of whether the dose distribution is produced by the static beam apertures or by a dynamic delivery technique such as intensity modulated radiation therapy (IMRT) or volumetric modulated arc therapy (VMAT). However, the second effect, the interplay, has a temporal component and is associated only with dynamic deliveries. Respiratory motion can interfere with the dynamically changing beam parameters, most notably the MLC segments’ shapes, to alter the target dose. The analogy between this process and the classic wave behavior, including the relative frequency dependence and the possibility for either constructive or destructive interference was aptly noted.\(^{(1,5)}\) For the homogeneously irradiated target volume, the interplay effect is relatively easy to define as any deviation of the tumor dose from the planned value in the presence of the dynamic delivery and motion. However, with the inhomogeneous target dose distribution, distinguishing between the gradient and interplay effects in an experiment for dynamic deliveries is more challenging. Since the gradient effect is purely spatial, while the interplay has a temporal component, the logical way to isolate the interplay effect is to analyze the dosimetric effect of the varying starting phase of the motion relative to the beam start time.\(^{(6)}\)

The motion effects in the presence of the IMRT delivery have been studied by many authors.\(^{(2,6-14)}\) Bortfeld et al\(^{(7)}\) predicted that the interplay effect would, for the most part, average out with the large number of fractions. However, it is not automatically the case with hypofractionated radiosurgery-type course of treatment (1-5
fractions). While target motion effects in IMRT have been studied extensively with both experimental and theoretical methods, VMAT is a newer modality, and so there are substantially fewer studies of interplay in VMAT,\(^{6,9,12,13}\) particularly for hypofractionated dose regimens. The theoretical studies, while elucidating, do not necessarily rely on the actual delivery sequences, as the linear accelerator controller may change the VMAT delivery timing compared to the treatment planning system (TPS) estimate.\(^{15}\)

Previous experimental studies have typically used labor-intensive physical motion stages, while reporting motion-induced dose changes to a relatively small number of points, or a single plane. In the absence of full volumetric analysis, it is not possible to evaluate the dose-volume parameters most important for SBRT, such as the near-minimum target dose.

In this chapter, it is shown how the methodology developed in section 2.3.2 can be applied to isolate and quantify the possible interplay influence on the target DVH between the respiratory motion and dynamic VMAT delivery for lung SBRT treatments.

### 4.2 Materials and Methods

#### 4.2.1 Treatment planning and delivery

Ten consecutive patients previously treated with VMAT lung SBRT were selected, starting with the most recent one at the time of writing and working backwards. Each plan dealt with a single lesion. Half of the patients had their tumors in the upper lung lobes and the other half in the lower. All patients were treated with 6MV photons on Varian Trilogy or TrueBeam linear accelerators (Varian Medical Systems, Palo Alto, CA) equipped with 120-leaf Millennium MLCs (5 mm leaf width in the central region). The beam characteristics of the different machines were closely matched (<1% differences
in dose distributions). Four-dimensional helical CT (4DCT) datasets obtained on a Brilliance Big Bore scanner (Philips Medical Systems, Cleveland, OH) and reconstructed with a 3 mm slice width were used for treatment planning. The untagged low-pitch scan was used for dose calculations. The institutional practice for VMAT lung SBRT is to use 2-3 approximately half-arcs, avoiding beam entrances from the contralateral side. All patients were treated in 5 fractions to 50 or 60 Gy prescribed to 95% of the planning target volume (PTV). The internal target volume (ITV) was defined as the union of the visible gross tumor volumes (GTVs) with no added margins for microscopic extension, from all the different phases of 4DCT. The PTV is derived from the ITV by a uniform 5 mm expansion. Treatment planning was done on Pinnacle treatment planning system (TPS) v. 9.2 (Philips Radiation Oncology Systems, Fitchburg, WI). Dose calculations were performed on a 3 mm grid and 4° VMAT control points (CP) spacing\(^\text{(15)}\) with the Collapsed Cone Convolution algorithm. For the target motion effect analysis, first the original treatment plans were used. Additionally, those plans were modified to accommodate the expanded simulated motion, as described below.

4.2.2 Motion kernels

4.2.2.1 Original empirical motion kernels

The current institutional practice is to allow VMAT treatments for lung SBRT tumors that move less than 1 cm. The average tumor excursion was 4±2 mm (1SD), with the range of 1 to 8 mm. For the purposes of this study, the tumor motion is assumed to be a 3D rigid body translation\(^\text{(17)}\). Therefore the tumor motion pattern (motion kernel) can be defined as a plot of the tumor centroid position against time,
discretized in practice according to the ten 4DCT phases. The tumor was defined as a physician-drawn GTV on a single 4DCT phase (GTV0 by convention, see Figure 4.1).

Figure 4.1. The moving target is the GTV drawn on Phase 0% and rigidly translated according to the centroid position on each phase (two positions with maximum separation shown). Static ITV is the union of all target volumes, and the PTV is the uniformly expanded ITV. The minimum and maximum ITV dose voxels are shown to illustrate the gradient motion effect.

The GTV0 volumes ranged from 0.9 to 17.2 cm$^3$ (average 6.8±5.1(1SD) cm$^3$).

Using deformable registration in the Mirada RTx software package v.1.2 (Mirada Medical Ltd, Oxford, UK), this contour was propagated to the remaining 9 phases. The centroids of the resulting volumes were taken as the tumor centroid locations on every phase, providing the spatial component of the original empirical motion kernel. The temporal component was based on the periodic tumor motion assumption, with the average period of the breathing trace from the 4DCT (average among the patients 3.5±0.8 (1SD) sec, range 2.9–4.6 sec).
A union of the original tumor volumes situated on all phases following the centroid excursion represented the ITV. Since the ITVs on which the original PTVs and plans were based were obtained by differing procedures, the tumor volume was slightly adjusted so that the boundaries of the new ITV were within 1 mm of the original plan ITV.

4.2.2.2 Simulated extended motion

Additional motion kernels simulated larger motion (2 or 3 cm total target excursion). For each patient, the original motion trajectory was proportionately expanded to result in the total tumor displacement of 2, and then 3 cm. When necessary, the direction of the motion was adjusted to avoid unrealistic patterns such as tumor incursion into the chest wall. The predominantly Superior-Inferior (SI) motion direction was maintained. A new, expanded, ITV and PTV were constructed as described above, and the patients were re-planned to ensure the new PTV prescription dose coverage at the 95% volume level. Since the 2 or 3 cm tumor motion is not reflected in the planning CT densities, the ITV density was manually overridden with the value of 0.8 relative to water. The necessity and validity of this approximation are elaborated in the Discussion section.

The modified treatment plans used the same objectives as the original ones, but additionally, the RTOG intermediate dose compactness and symmetry criteria (50% dose volume to PTV volume ratio, and maximum dose 2 cm from the PTV in any direction)\(^{(4)}\) were enforced. This resulted in higher modulation and, subsequently, dose inhomogeneity in the expanded ITVs compared to the original ones, which, if anything, should accentuate the findings. The maximum PTV dose was 15 to 40% higher than the
prescription, conforming to the RTOG protocols specifications.\(^4\) These plans were evaluated for motion effects first with the original 4DCT motion periods, and then with the 60 s period. While not realistic, making the motion period comparable to the delivery time helps to evaluate the causes of the interplay effect.

### 4.2.3 Data collection and analysis

#### 4.2.3.1 Dose to the moving target

Measurement-guided dose estimate to a moving target in a patient is based on a research version of the commercial 3DVH software package. As described in the previous chapter, the dynamic (4D) volumetric dose on the patient can be estimated by perturbing the patient TPS dose distribution based on the phantom measurements.\(^{18,20,21}\) In this study, the QA method described above was used to confirm the agreement between planned and delivered 3D dose distributions on the patients. Gamma analysis with rather stringent 2%/2mm (local dose-error normalization) criteria was used.

#### 4.2.3.2 Metrics

The main metric of significance for lung SBRT is the near minimum dose \(D_{\text{min}}\) to the target.\(^{17}\) In the prior relevant publications, the minimum dose encompassing 99% of the ROI \(D_{99\%}\) was used to represent \(D_{\text{min}}.\)\(^{13,17}\) On the other hand, ICRU report 83\(^{23}\) recommends \(D_{98\%}\) as the near minimum dose surrogate. Therefore we investigated both and evaluated correlation between them. Similarly, \(D_{1\%}\) and \(D_{2\%}\) were used as the near maximum dose \(D_{\text{max}}\) metrics. Finally, the mean doses \(D_{\text{mean}}\) were also recorded. When necessary, the dosimetric indices for the moving target are reported as differences from the corresponding metrics for the static ITV, thus eliminating the influence of the fraction
prescription dose and the differences in the individual plans. Throughout the chapter, the term “target” will be used to describe the moving tumor with its shape represented by GTV0. The term “ITV” will always refer to a static contour that is the union of the target volumes on all 4DCT phases.

4.2.3.3 Single and multiple fraction simulations

Single fraction.

For a single fraction, the interplay effect can be defined deterministically as the maximum difference in the moving target dose metrics calculated for one fraction for every possible pair of the available starting phases. Our methodology allows us to define and measure interplay according to its fundamental characteristic— the difference in dose caused by the change in the motion starting phase in relation to the beam start time\(^{(2)}\) without resorting to surrogates\(^{(8,10,12)}\). For each motion kernel with a corresponding plan, a single fraction motion simulation was run 6 times, corresponding to every other starting phase (0, 20, … 80%) plus 50%, out of 10 available from 4DCT. The maximum spread in dosimetric indices is reported as an aggregate average between all patients, with corresponding ranges.

The subsequent analysis for the single fraction simulation is contingent upon the interplay magnitude. For those motion kernels where the interplay effect is small, the data for all starting phases can be averaged (i.e. interplay neglected) and, for a single fraction, the gradient effect can be quantified in isolation. The target dose indices are reported as differences from the corresponding static ITV values. The averages and ranges over the patient population are presented. When the interplay effect is appreciable, only a combined (gradient and interplay) motion effect can be easily
discerned experimentally. The ranges of the target dosimetry indices in comparison with the ITV are presented for individual patients.

**Multiple-fraction courses**

Of the two motion effects, only the interplay depends on the number of fractions. Therefore, only when significant interplay is found for a single fraction is an attempt to quantify the effect of fractionation warranted. Unlike for a single fraction, only a statistical estimate can be made. A simulation of 3- and 5-fraction treatment courses was run 5 times each, with random starting phases assigned to each fraction for each simulation. The range of dosimetry indices' values between the different runs is a measure of the interplay effect which is expected to diminish as the number of fractions increases. The statistical multiple-fraction ranges for $D_{99\%}$ were compared to the deterministic single-fraction interplay-induced range caused by the different starting phases.

### 4.3 Results

#### 4.3.1 Single fraction

As a basic initial quality control measure, it was demonstrated that the 3D comparison between the measurement-reconstructed and TPS-calculated dose distributions with 2%/2mm gamma analysis for original plans resulted in >90% passing rate for each patient, with an average of 94.4±2.7(1SD)%. It was also determined that the changes in $D_{99\%}$ and $D_{98\%}$, as well as $D_{1\%}$ and $D_{2\%}$ correlated very closely for all plans and motion kernels (correlation coefficient > 0.99 when compared across the patient sample). Therefore only $D_{99\%}$ and $D_{1\%}$ data are presented in detail, for clarity and brevity.
In Figure 4.2, the interplay effect for a single fraction is essentially quantified. The maximum percentage differences in dosimetric indices, between any two of the starting phases, are plotted as averages for the patient population with the corresponding ranges.

Figure 4.2. Quantification of interplay. Maximum percentage spread of the dosimetric indices for the moving target between any two possible starting phases. The mean values across 10 patients are presented with the corresponding ranges. The groups (X axis) refer, respectively, to the original motion kernel from 4DCT, motion range increased to 2 or 3 cm with original period, original range with 60 s period, and increased ranges with 60 s period. Within each group, the order is $D_{99\%}$, $D_{\text{mean}}$, $D_{1\%}$.

For the original motion kernels obtained from 4DCT ($\leq 8$ mm motion, 3-5 s period), the effect is negligible (<0.2%). With the motion range increased to 2 or 3 cm, the maximum difference between the phases increases slightly: averaged between the patients, the
D\(_{99}\) interplay-induced spread is <1\%, with the maximum range of 2.2\%. Therefore, for the studied plans, the interplay effect is minimal within the clinical range of motion parameters. Once the motion period is artificially increased to 60 s, the interplay effect begins to increase, reaching its maximum for the 3 cm motion. At that point, the maximum difference in D\(_{\text{min}}\) between two starting phases can be >20\%.

Since for the realistic motion periods the dosimetric differences between the starting phases are small (Figure 4.2), for any given patient, the moving target doses can be averaged between all the phases (i.e. interplay neglected) before comparing to the static ITV. The results are plotted in Figure 4.3, representing predominantly the gradient effect.

The differences are small for the original motion kernels. Once the motion range is increased to 2-3 cm, while maintaining the original period, the mean near-minimum target to ITV dose ratios increase, while near-maximum ones decrease, characteristic of the “motion blur”. The balance of this dose change in the current study is such that the target D\(_{\text{mean}}\) is almost always higher than for the static ITV. There are only 3 instances out of possible 20 when the target D\(_{\text{mean}}\) for either 2 or 3 cm motion was lower than the static ITV one. This negative difference never exceeded 1\%.
Figure 4.3. Low interplay scenario. Percentage difference between the moving target dosimetric indices, averaged for one fraction between all possible motion starting phases, and the corresponding static ITV metrics. The data are reported as an aggregate average over all patients, with the ranges. The groups (X axis) refer, respectively, to the original motion kernel from 4DCT, motion range increased to 2 or 3 cm with original period, and original range with 60 sec period. Within each group, the data order is $D_{99\%}$, $D_{\text{mean}}$, $D_{1\%}$.

For the original motion amplitude with an artificial 60 s period, the motion-induced dose changes are still small. However, in comparing in Figure 4.3 the ranges for the original motion (<1 cm) to the original with increased period, the target near-minimum doses can now be slightly lower than for the static ITV, indicating some presence of the interplay. The maximum such difference observed for a single starting phase was -3.4% for $D_{99\%}$. 

55
Hypothetical substantial interplay scenario. The motion period is 60 s and the range is 2 cm (A) or 3 cm (B). Percentage differences are shown between the moving target and static ITV dosimetric indices, presented for individual patients as a range corresponding to all possible starting phases. The groups (X axis) refer to the individual patients. Within each group, the data order is $D_{99\%}$, $D_{\text{mean}}$, $D_{1\%}$.
With the 60 s period and 2-3 cm motion range, the interplay effect results in a substantial dependence of the fractional dose on the starting phase. Therefore, those data groups are omitted in Figure 4.3. Instead, the phase-dependent spreads of the single fraction target to ITV dosimetric indices differences are presented for individual patients in Figure 4.4. Those results vary greatly. For example, \( D_{99\%} \) for the moving target can be anywhere from 16\% low to 17\% high compared to the static ITV.

Figure 4.5. Areas of differences above 5\% for the dose distributions with 60 s motion period and 3 cm range, corresponding to the starting phases of 0 and 50\%. A sagittal cross-section is presented, with the contours corresponding to the maximum separation GTV positions, the ITV, and the PTV. The grey scale represents the ArcCHECK planned dose perturbation motion-blurred dose distribution obtained with 0\% starting phase.
As an example, locations of dose-differences ≥ 5% (a threshold used previously by Court et al (9)) between starting phases of 0 and 50%, with 60 s motion period and 3 cm range, are presented for a single patient in Figure 4.5. No differences above 2% are seen for this patient with the same motion range but with the normal period of 4.2 s.

4.3.2 Multiple-fraction courses

The reduction in the interplay effect with increased number of fractions is illustrated in Figure 4.6, with D$_{99\%}$ chosen as an example. The deterministic ranges of D$_{99\%}$ differences between the moving target and static ITV, due to the starting phase variation in a single fraction, are compared to a statistical spread due to five simulations, with random starting phases, of the 3- and 5-fraction courses. For the 2 cm, 60 sec simulated motion, the patient-population averages of ranges in Figure 4.6 were reduced from 7.1±4.2(1SD)% for a single fraction, to 3.7±1.7% and 2.2±1.0%, for the 3- and 5-fraction courses, respectively. For the 3 cm, 60 s motion, the corresponding numbers were 8.4±6.0%, 3.7±1.9%, and 2.8±1.6%. Repeat Measures Analysis of Variance (ANOVA) indicated that overall these differences in mean ranges were statistically significant between the three fractionation schemes (p<0.02). The post-test for individual pairs results in statistically significant differences between 1 and 3 or 5 fractions at the 95% confidence level, but no statistical difference between 3 and 5 fractions.
Figure 4.6. Reduction in interplay effect in $D_{99\%}$ with the number of fractions presented per patient (X-axis groups). The first data point in each group is the full range (deterministic) between all possible starting phases of the ratio of moving GTV0 $D_{99\%}$ to static ITV. The second and thirds ones are statistical means and ranges of the same ratios between five simulations with random starting phases of 3- and 5-fraction treatments, respectively. The last bar is the mean value for the large number of fractions (30) simulation. The motion period is 60 s and the range is 2 cm (A) or 3 cm (B).

The mean value for a large number of simulations (30) can also be compared to the 3- and 5-fraction averages. The results for 2 cm and 3 cm motion with 60 sec
periods are similar. The average D$_{99\%}$ target vs. ITV differences across ten patients deviated from the mean of 30 fractions by no more than 0.8±1.1% and 0.2±1.1% for the 3 and 5-fraction courses, respectively.

4.4 Discussion

4.4.1 Dose to the moving target

It is clear from Figure 4.2 that for the original 4DCT motion kernels (tumor excursion ≤ 8 mm, period 3-5 s), the interplay defined as the maximum spread of volumetric dosimetric indices between all possible starting phases is negligible. The studied interplay increases slightly with the motion range increase to 2-3 cm, reaching for D$_{99\%}$ an average value of 0.9% and the maximum of 2.2% among all the patients. A further slight increase is observed when the motion period is artificially set to 60 s. However, even in this hypothetical scenario, the D$_{99\%}$ is mostly higher for the target than the ITV, and is never lower by more than 1%. With the ITV dose being already higher than the prescription, the clinical effect would be negligible. This coincides with the conclusions from the computational study by Rao et al.$^{(13)}$ which used largely the same definition of interplay based on the near-minimum dose to the target. On the surface, an experimental RapidArc motion study by Court et al.$^{(8)}$ concludes that interplay can lead to appreciable changes in target dose. However, there are substantial methodology differences between the current study and that of Rao et al.$^{(13)}$ on the one hand, and Court et al.$^{(8)}$ on the other hand. In this work and Rao et al.$^{(13)}$, the primary metric of interest is the near-minimum dose to the target in the patient. In Court et al.$^{(8)}$, the dose differences are recorded in a single plane of a dosimetry array, making a DVH evaluation impossible. The metric was the percentage of the pixels in the target, which
never had dose-errors larger than a certain criterion (3-5%), regardless of the starting phase. It can be argued that for the lung SBRT target dosimetry the near-minimum dose is more clinically important than the exact dose distribution in the target.\(^{(17)}\)

With the differences in beam modulation caused by the planning objectives and ITV/PTV size changes when switching from the original to 2-3 cm motion, it is clear from Figure 4.2, that in order to observe a substantial interplay effect in the lung VMAT SBRT, two conditions must be met: the tumor excursion must exceed 1 cm, and the motion period must be comparable to the gantry rotational period. With the inverse of that period being one of the fundamental frequencies of the delivery system, this follows the observation by Kissick et al\(^{(5)}\) that the maximum interference (interplay effect) would occur when the machine fundamental motion frequency is comparable to the tumor fundamental motion frequency. The simulations by Li et al\(^{(11)}\) further confirm that the dose-error is a somewhat periodic function of both the delivery time and motion frequency.

Approximately two-thirds of the motion kernels had predominant tumor displacement in the SI direction, with the rest in the AP or LR directions. At the same time, 9 out of 10 patients had collimator angles between 8 and 15°, and one 30°. Thus a range of relative orientations of the leaf movement direction and tumor displacement\(^{(14)}\) was explored, with essentially the same results. Still, it must be emphasized that our findings of low interplay effect should not be automatically extended to sites and techniques other than lung SBRT. Different optimization schemes, planning constraints, modulation levels, target sizes and motion characteristics could lead to different results.
In Figure 4.4, the near-minimum target dose tends to be higher, and the near-maximum lower than the ITV. This is due to the bias induced by the gradient effect, which follows from simple geometrical considerations (Figure 4.1) and represents the motion-induced averaging, or “blur”. The interplay effect is superimposed on this gradient effect, resulting in the substantial range of the target dose values, depending on the starting phase. Interestingly, the mean target doses in Figure 4.4 tend to be higher than the mean ITV dose. This manifestation of the gradient effect is not general but is rather a function of the intersection of the patient-specific dose distribution with the tumor position probability density function.

The quantitative comparison between the single- and multiple-fraction interplay effects is not as rigorous because we are trying to compare a deterministic quantity with statistical samples. Only a small number of statistical simulations were performed because of the limited practical significance of simulating unrealistically high motion periods. However, qualitatively, the interplay effect decreased as the number of fractions is increased to 3 or 5, in agreement with theoretical simulations. For the data in Figure 4.6, the maximum standard deviation of $D_{99\%}$ among the patients for 5 runs was reduced from 2.9% for 3 fractions to 1.9% for five fractions. Averaged between all runs for all patients, the 3- and 5-fraction $D_{99\%}$ did not deviate from the 30 fraction value by more than 0.8±1.1% (1SD) and 0.2±1.2%, respectively. Similarly averaged $D_{mean}$ did not deviate from the 30-fraction value by more than 0.2±0.8%, even with the unrealistically high interplay for a single fraction.
4.4.2 Assumptions

A number of assumptions were required to make this study feasible. First, the tumor motion is approximated by a rigid body translation, as necessitated by the current 3DVH software design. This subject was studied previously by Admiraal et al.\textsuperscript{(17)} who concluded that the approximation was reasonable for a small tumor in the lung. Second, the tumor motion is assumed to be periodic. If anything, the periodic nature of motion is expected to exaggerate the interplay effect, and more random temporal motion patterns would not alter the findings for the realistic motion kernels (3-5 sec per cycle). The 2 and 3 cm simulated motion ranges can be characterized as “top of the range” and “exceedingly large”, respectively, for a lung tumor.\textsuperscript{(24-26)}

Finally, we had to assign a uniform density to the ITV for extended motion (2-3 cm) simulations. Even when 4DCT information is used for target definition, dose calculations are routinely performed on an average 3D CT dataset,\textsuperscript{(17, 27)} which can be either a mathematical average of the reconstructed phases or a physical slow (low-pitch), untagged scan. While it results in acceptable, if not negligible, dosimetric errors for real patient CT datasets, except perhaps at the PTV margins,\textsuperscript{(28)} in our case of simulated extended motion, the large portions of the ITV not visited by a real tumor, exhibit unrealistically low (lung) CT density. This interferes with the perturbation algorithm’s ability to accurately reconstruct dose to the moving target. Since in this study we are interested only in relative dose changes due to the target motion within the ITV, the exact average ITV density is not important, as long as it is reasonably high. We verified that the presented numerical results are negligibly (< 0.2%) affected by the ITV density.
density variations relative to water from 0.5 to 1.0, which should encompass the range of average ITV CT densities encountered in real life.

4.5 Conclusions

We applied a motion perturbation method to reconstruct the volumetric dose to a moving tumor during lung SBRT VMAT deliveries. The possible effect of interplay between the tumor motion and dynamic dose delivery was evaluated and separated from the gradient effect that depends only on the tumor position probability density function and 3D dose distribution, and is present regardless of the delivery dynamics. With the studied planning and treatment technique, for realistic motion periods (3-5 s), and regardless of the motion amplitude, the interplay has nearly no impact on the most important target dose metric in lung SBRT – the near-minimum dose. Thus, while motion management techniques may serve useful clinical purpose, such as reducing the amount of irradiated healthy lung, the target dosimetry alone does not warrant attempts to reduce the motion amplitude. The conclusions about the importance of the interplay can differ for different treatment sites, techniques and optimization algorithms, while the manifestation of the effect may depend on the data acquisition methods, analysis metrics, and the specific definition of the interplay.

The interplay effect was observed, for study purposes only, with the substantial tumor motion and unrealistically long period (60 s), comparable to the VMAT delivery time. The effect of motion on the target DVH for an individual patient can be further determined based on the experimental VMAT delivery pattern recorded during a routine QA session and 4DCT motion kernel. Future application of this method to non-rigid
motion is a matter of programming an interface that would track individual voxels through the 4D dose space in the presence of organ deformation.

4.6 References


CHAPTER 5:
3DVH FOR HELICAL TOMOTHERAPY

5.1 Introduction

The development and implementation of volumetric dose reconstruction to achieve more accurate and clinically relevant plan comparisons for IMRT and VMAT is essential. However, these are not the only areas that could benefit. Tomotherapy patient-specific QA could equally benefit from the volumetric reconstruction of the deliverable patient dose. While the general approach remains similar to VMAT, the unique characteristics of the TomoTherapy Hi Art treatment system (Accuray, Inc., Madison, WI) dose delivery required substantial, conceptual modifications of the algorithm. In this chapter, we describe this novel tomotherapy dose reconstruction method and validate it by sampling the produced 3D dose and comparing to independent measurements by an ion chamber and a biplanar array dosimeter. (12-14)

5.2 Methods

5.2.1 Measurement-guided dose reconstruction

5.2.1.1 The challenges of TomoTherapy

Helical tomotherapy presents many unique challenges that require a new variation of the ACPDP algorithm, which we will call Tomo-PDP (TPDP). Figure 5.1

---

3 This chapter is partially reproduced from previously published work. The author of this dissertation is the first author of the previously publish work. All figures are used with copyright permission. See Appendix A and B.
summarizes the workflow of the TPDP model to compare and contrast with a similar workflow diagram already presented for ACPDP.\(^8\) Some of the major measurement-guided dose reconstruction (MGDR) challenges unique to helical tomotherapy are listed below.

5.2.1.1.1 Delivery dynamics (gantry)

ACPDP reconstructs the dose for time-discretized sub-beams. A gantry rotation period for VMAT beams on a conventional linac is no less than 60 seconds. ACPDP divides a VMAT beam into sub-beams based on a gantry change threshold of ~2°, and as a result the typical ACPDP sub-beam spans on the order of 0.3 s. The AC measurement time resolution is 0.05 s, so a sub-beam is certain to include multiple measurement updates. On the other hand, TomoTherapy gantry rotation speed is much faster – up to five rotations/min, or a period of 12 s (30°/s). A sub-beam covering 2° thus spans less than 0.07 s, which is on the order of a single AC measurement update interval. Therefore, the gantry speed presents challenges for dose reconstruction and causes sensitivities to even the smallest errors in derived or simulated gantry angle compared to actual.

5.2.1.1.2 Delivery dynamics (MLC leaves)

TomoTherapy delivery is based on 51 intervals per rotation, spanning about 7.06°. During rotation through each interval, the leaves snap open and then closed symmetrically about the middle of the interval. The time they stay open depends on their “open fraction”, producing a “strobe” effect in dose delivery. The 7° interval span can happen as fast as 0.235 s, and for small open fractions the leaf open time can approach as low as 0.03 s. That is just a fraction of the measurement update period, thus
presenting another challenge unique to the delivery dynamics in relation to the measurement time resolution.

Figure 5.1. The workflow diagram of the Tomo-PDP process. The novel steps specific to TPDP are in bold. The RT Plan is first synchronized to the measurement (SYNC\(\text{T}\)) and the calculated time-resolved sub beams are grouped into sectors (PER SECTOR INTEGRATION). The sector doses are then morphed based on the measured entry and exit doses (MORPH-NORM [PER SECTOR]). These morphed dose grids are summed to give the cumulative 3D phantom dose (\(\Sigma\) DOSE-AC-HD-ABS [SECTOR]). Ion chamber measurements can be included to further correct the reconstructed dose. The resulting high-density dose grid can be used for dose comparisons on the phantom (VIRTUAL GEL) or further processed to produce measurement-guided dose reconstruction on the patient dataset. See Methods section II.A.2 for detailed explanations of each step of the process.

5.2.1.1.3 Sub-beam size and complexity

TomoTherapy MLC leaves move in the IEC Y direction and are considered binary, though naturally there is some latency as the leaves open and close. The leaf widths are 6.25 mm at the isocenter, and the allowed field lengths (Y jaw settings) are
approximately 1, 2.5, and 5 cm. Given the leaf dynamics and the strobe effect of all leaves being forced closed 51 times per rotation, at any given time the irradiated field can be very complex, consisting of small beamlets which are often discontiguous. The detector separation in the AC phantom is 10 mm along and across the diode helix. As a result, some TomoTherapy sub-beams intersect few diodes at entry and exit, thus limiting the capability of measurement-guidance if sub-beams are reconstructed as in conventional ACPDP, i.e. as a function of very small times corresponding to threshold geometry changes. This effect becomes more pronounced as the jaw setting decreases.

5.2.1.1.4 Beam parameters in DICOM RT Plan

The TomoTherapy MLC dynamics are not specified in the RT Plan as leaf positions over a number of control points as with IMRT or VMAT. Instead, 51 control points per rotation are provided but they do not explicitly define any MLC leaf positions; rather, a "sinogram" of the per-leaf open fractions is defined per control point, as a private DICOM tag. From the open fractions, the opening and closing times of each leaf are derived, for each of the 51 intervals per rotation. If there is non-negligible leaf latency, or device-specific MLC calibration/air pressure variation, it is not reflected in the RT Plan object.

5.2.1.1.5 Synchronization of time-resolved measurements to beam parameters

ACPDP relies on an accurate virtual inclinometer and robust synchronization logic, called “Sync(T)” in Figure 5.1, to assign absolute (accelerator) times to each control point in the RT Plan, which are required when processing each sub-beam. However, for TomoTherapy, given the dearth of diodes irradiated at any given short
time period and the high gantry speed, the AC virtual inclinometer cannot reliably derive actual gantry angle as a function of time from the AC signals. Instead, gantry angle vs. time must be taken directly from the RT Plan object, which is possible because TomoTherapy’s cumulative metersets are based on irradiation time (minutes) rather than monitor units. However, this causes a reliance on the coincidence of the RT Plan’s gantry angle vs. time relationship with the actual delivery, which turns out to be not sufficiently accurate at a time scale on the order of 0.05 s. Specifically, projected sub-beams (derived from the RT Plan as a function of time) do not show perfect temporal alignment with observed irradiated diode doses vs. time. This presents a challenge for measurement guidance. See Figure 5.2 for examples of synchronization issues.

5.2.1.1.6 Threading effect and measurement locations

Dose in regions radiating outward from the TomoTherapy isocenter is progressively harder to model due to the “threading effect” of divergent fan beams rotating in a helix (compare to Kissick et al\(^{(16)}\)). There is high sensitivity to Y jaw accuracy and penumbra modeling, for instance. The AC detector elements are located in the periphery, 104 mm from the isocenter, and dose distributions at that location are hard to model accurately from first principles. (This also holds true for the TomoTherapy primary treatment planning system (TPS), as will be shown later in this paper).
Figure 5.2. In this figure, the top and bottom row show two different 0.25 second intervals of MLC leaf arrangements vs. the entry diode doses (as a function of time) measured by the 3D dosimeter. Both 0.25 second time intervals are from the same irradiation of the TG119 Head and Neck plan. In each row, the entry diode doses accumulated in each 0.05 sec sub-interval are shown projected behind the nearest MLC open/closed positions derived directly from the sinogram data in the DICOM RT Plan. The sinogram and 4D delivery have been optimally synchronized in time. For row (A), all 0.05 second sub-intervals seem well-matched (MLC openings vs. diode dose exposure) except for the Ti + 0.15 sec sub-interval where according the sinogram all MLCs should be shut, while clearly the diodes are receiving dose during this time. The seemingly out-of-sync sub-interval is evidence of latent dose (from shutting leaves from prior control point) or leading dose (from opening leaves from the next control point) that results when absolute synchronization is not possible due to deviations in actual MLC positions from the expected positions defined by the sinogram data and control point time stamps. Row (B) illustrates a different kind of temporal challenge from the same irradiation. The pattern of nominal MLC arrangements vs. measured entry dose values suggests a 0.15 second time misalignment, which could be caused by actual delivery dynamics that do not match nominal as exported by the TPS in the RT Plan.

5.2.1.2 The algorithm

The TPDP method developed in this work attempts to address the challenges enumerated above with a number of tomotherapy-specific modifications to the conventional ACPDP method. Because this method has already been described, we will
describe only the steps that are new or modified to support tomotherapy. These steps are highlighted in Figure 5.1.

5.2.1.2.1 Synchronization of RT plan to time-resolved measurement points

The RT Plan’s control points have cumulative meterset values in units of time rather than monitor units, so synchronization of the control points to the time-resolved measurements is, at least at the surface, conceptually simple. Both are on an absolute time scale, and those two time scales simply need to be registered to each other. However, using the first “beam on” update from the measurement and mapping this to the start of beam delivery (i.e. first control point before cumulative meterset changes to the next) is not sufficient for either of two reasons: 1) for the first projection, MLC leaves start in the closed position, exposing no diodes or 2) the first MLC segments may be too small to directly expose any diodes. Either reason could cause a failure to meet threshold diode current that signals the “beam on” condition and introduce time inaccuracy; however, this method is used for the first pass to get within 0.5 second accuracy. To refine the synchronization, dose as a function of time is calculated for key diodes (moderate cumulative of at least ~50% of the maximum) using the sinogram data and the 3DVH convolution algorithm, for various time shifts in increments of 0.01 s. The eventual time sync derivation is based on the time registration giving the best dose agreement between the observed and the calculated dose at the critical diodes. Some plans can exceed the length of the AC sensitive volume. Such plans must be shifted longitudinally on the phantom (a “green laser shift”) so that the first or last beamlet projects onto the active volume. This allows the AC to register the beam on time, at the expense of only the ROIs within the AC sensitive volume being fully evaluated. While in
theory either shift direction would work, in reality the last beamlet has to end up in the active volume as an opposite shift direction would expose electronics to the direct beam. This is one of the system limitations.

5.2.1.2.2 Organization into 3D dose “sectors”

Time-resolved (four-dimensional, or 4D) sub-beams can be calculated in 3DVH for any integral number of projections per control point interval. One projection/control point proved far too coarsely discretized (i.e. 51 projections per rotation is inadequate for accuracy in the periphery\textsuperscript{(17)}). Other integer multiples (2, 3, etc.) of sub-beams per control point interval were tested, and two sub-beams/control point interval gave the best balance of accuracy and performance. As a result, there are 102 sub-beams per rotation and multiple thousands of sub-beams over the whole treatment, depending on the number of rotations. However, given the high speeds of the gantry and the motion of the table, even the best time synchronization (as described above and illustrated in Figure 5.2) results in a diode irradiation pattern that at times is out-of-sync (+/- 0.02 s) with some of the sub-beams’ MLC openings derived from the nominal sinogram patterns. These errors do not manifest in predictable patterns and are thus cannot be accounted for mechanistically. If per sub-beam measurement-based correction (“morphing”) was performed as in VMAT AC-PDP\textsuperscript{(8)} these minor time mismatches would cause unwarranted morphing of calculated sub-beam doses to fit 4D measurement patterns that do not match for any particular small time interval.

Therefore, before measurement guided dose morphing is performed, 4D sub-beam dose grids are grouped (i.e. summed) in two ways to lessen the effect of per-sub-beam synchronization vacillation: 1) over fixed spans in gantry angle per rotation, and 2) over
identical gantry angle spans over all rotations as the table is in motion. The resulting 4D dose summations are called “sectors” (see Figure 5.3). The advantage is that sector doses map nicely to entry and exit diode columns along the long axis of the AC phantom. To ensure a sufficient number of exposed diodes, the algorithm bins sub-beams according to gantry angle span equivalent to three control points (~21 degrees). It is important to remember that within a sector dose, all sub-beams’ dose grids were computed at their respective projection angles; the sectors serve only to group many thousands of sub-beam’s dose grids into a smaller set of dose grids for dose morphing to measurements, grids that are far less sensitive to small time sync imperfections.

5.2.1.2.3 Per sector dose morphing

For an error-free delivery and perfect input data (i.e. sinograms, gantry vs. time, etc.), there would be no need for morphing. There would only be a need for a simple scalar adjustment to fit the high-resolution, volumetric sector dose (in relative units at this point) to the absolute measured dose (Gy) accumulated over each sector’s associated time intervals. However, measurement-guidance implies that the high-resolution calculations are fit to the diode doses, thus requiring dose morphing. Each sector dose is morphed in a way similar to what conventional ACPDP uses for each sub-beam, the difference being that the morphing technique is applied not to a single sub-beam but to all that contributed to that sector. Each sector has natural entry and exit dose surface (see Figure 5.3) between which correction factors can be interpolated. In addition, a “loop morphing” is one along the IEC table Y-direction to ensure that if any errors are uniform over an AC axial loop, i.e. specific to a longitudinal position, that the
slab of dose straddling that loop is scaled accordingly. Normally this is a very minor correction.

5.2.1.2.4 Summation of 3D sector doses

After all three-dimensional (3D) sector doses are morphed and simultaneously calibrated as absolute, high-resolution dose grids, they are summed to give a cumulative absolute 3D phantom dose for the whole treatment.

5.2.1.2.5 Ion chamber correction of 3D phantom dose

The AC diodes are situated exclusively in the radial periphery of the phantom, with detectors displaced by ~10.4 cm from the isocenter. This is precisely where the threading effect of helical tomotherapy is most difficult to model and creates implicit challenges for measurement-guidance. To help allay such dependencies, the TPDP algorithm includes the ability to further correct (morph) a volumetric dose based on central ion chamber (IC) measurement. The AC can be fit with a “MultiPlug” insert that allows 25 fixed measurement options relative to the plug, and the plug can be rotated to produce many more options. Software helps guide the user to high dose, low gradient points which serve as the best measurement options, in that they are less susceptible to positioning or volume averaging errors. Upon finishing dose reconstruction using the diodes, the user can specify a measurement position on the central axial plane along with the absolute measured dose and IC active volume.
Figure 5.3. Visual representation of 3DVH dose reconstruction
A) Three high spatial resolution sub-beams (seen on the axial cross-section through the middle of the AC phantom) that are part of the group binned and added into a single “sector” which limits the impact of inevitable synchronization errors due to differences in actual vs. nominal Tomotherapy MLC and gantry dynamics. B) TPS cumulative dose for AC phantom. C) Tomo-PDP “virtual gel” cumulative dose after all sectors are morphed to match measurements and then summed and post-processed, resulting in a volumetric cumulative dose to compare to the TPS-calculated dose. (Note: In panels A and C, the AC exterior and detector surface cross-section are highlighted in white; in 3DVH, dose is reconstructed inside the detector surface and outside is left equal to the TPS.)

If the IC-measured dose is different from the TPDP estimate, then all dose voxels inside a longitudinal cylinder of radius \( r \) will be multiplied by the same IC/TPDP dose ratio. This radius \( r \) is defined as the radial distance from the phantom center to the measurement point. From the measurement point’s radius to the detector radius, correction values are tapered down linearly to be exactly 1.00 at the detector surface to ensure the dose exactly matches measurements at the detector cylindrical surface.
5.2.2 Validation of the algorithm

5.2.2.1 Static dose calculations and phantom density

Standard TomoTherapy treatment planning system (TPS) software does not allow dose calculations with the static table. While it is not necessary for routine use of the system, the ability to perform basic dose calculations with static-gantry and rotational beams without additional complications of helical delivery is helpful for dosimeter calibration and evaluation. To that end, we employed an Accuray standalone dose calculator similar to that described by Kissick et al.\textsuperscript{(16)} It allows dose calculation for any delivery plan, whether generated through the normal treatment planning process (helical) or manually. To verify that the beam models were identical in the clinical and standalone dose calculator software, three helical plans with different jaw settings (1, 2.5, and 5 cm) were developed in the clinical software, transferred to the dose calculator and recalculated. The corresponding dose distributions were indistinguishable at the 1%/1mm level.

Both diode arrays used in this work are embedded in Poly-methylmetacrilate (PMMA) phantoms. In theory, relative electron density should be used to scale the physical dimensions of the phantoms in dose calculations.\textsuperscript{(18)} However, the TomoTherapy TPS uses the CT number to mass density conversion tables for inhomogeneity corrections. Furthermore, those tables are not employed directly but rather used as a basis to derive tissue-specific mass energy absorption coefficients.\textsuperscript{(19)} Under these circumstances, it is not clear what value should be best used to represent the PMMA and it was prudent to derive it empirically. To that end, static beams were projected on a homogeneous AC phantom and the dose was calculated with different
assigned phantom densities ranging from 1.15 (relative electron density of PMMA) to 1.19 (mass density). Relative dose was compared to ionization ratios in a PMMA phantom mimicking the AC dimensions but accommodating a 0.06 cm$^3$ A1SL (Standard Imaging, Middleton, WI) ion chamber. Comparisons were made for static vertical beams at the locations of the entrance and exit diodes (2.9 cm from the surface) and at the center. The dose profiles were sampled along the table movement to find the maximum signal. The radiological distance between the exit and entrance measurement points is approximately 23.7 cm. The ratios were obtained for a number of field sizes ranging from 5×40 to 1×3.75 cm$^2$. For the 1 cm jaw setting, the entrance measurements with the A1SL chamber are unreliable due to the volume-averaging effect. However, center-to-exit comparisons were performed as the field size increased with beam divergence.

5.2.2.2 Helical array calibration

There are two valid ways to perform absolute calibration of the AC array. The manufacturer recommendation for TomoTherapy is to use an ion chamber with absolute calibration in dose to water in a flat Virtual Water™ phantom. The chamber is placed at the same distance from the source as the reference diode would be with the AC centered on the green lasers (74.6 cm) with the buildup equivalent to 2.9 cm of PMMA, which equals 3.3 cm of water-equivalent plastic. A vertical (gantry at 0°) static 5×40 cm$^2$ beam is used. The A1SL chamber correction for beam quality $k_q$ is at most 0.2% and is customarily ignored.

An alternative method, easily implemented for C-arm accelerators, is based on the premise that the TPS should be capable of an accurate dose calculation in a 10×10
cm² beam projected on a large, homogeneous cylindrical phantom. With the standalone dose calculator, the reference diode dose could also be determined through the static beam calculations. Both approaches were executed with the reference dose differing by no more than 1%. The calculational value was used.

The daily correction (cross-calibration) procedure was performed before every set of AC measurements. A rotational 5×40 cm² beam with a static table was calculated on the AC phantom. The irradiation time was 48 s and included four full gantry rotations. The standalone dose calculator can export DICOM RT Dose object but not the DICOM RT plan derived from the static treatment procedure defined at the operator console, outside of the clinical TPS. Angular corrections implicit in the AC software cannot be applied in absence of an RT Plan for TomoTherapy, so for these test data they were applied manually. If the beam aperture is constant and the number of gantry rotations is whole, the average dose-weighted angular correction factor can be applied equally to every diode without any approximation. After this factor (0.982) was applied, the measured dose in the central portion of the field was compared to the TPS, and a daily correction factor was introduced. This correction brought the median difference between the measured and calculated doses to zero. It must be noted that in essence this daily correction procedure renders the absolute calibration moot; it is an array equivalent of the daily ion chamber cross-calibration procedure, as it is irradiated in a simple phantom and reference geometry to obtain a correction factor that takes into account the daily accelerator output variations.

The important difference between the conventional and tomotherapy dose acquisition modes is the application of certain AC correction factors. The field size
correction is ignored because it is negligible for the small fields (<5cm) and has not yet been studied for the “very small” (<1.5 cm)\(^{(21)}\) ones. The angular correction for tomotherapy measurements was introduced in 2014 (SNC Patient v. 6.5). Since the virtual inclinometer\(^{(15)}\) does not perform well with tomotherapy, the gantry angle is derived from the gantry positions and the elapsed irradiation times known from the RT Plan’s control points.

### 5.2.2.3 Test plans

One of the main challenges to overcome with TPDP is the limited number of diodes irradiated by each beamlet. This issue is clearly exacerbated as the jaw width is decreased. In this study we use the two smaller jaw settings: 2.5 cm and 1cm. The former is by far the most frequently used in our practice, while the latter is seldom used but presents a good stress test of the system. In total, we designed 19 test plans, planned and delivered with both 2.5 and 1 cm jaw settings. All cases were planned with pitch values conforming to the original 0.860/N formula.\(^{(16)}\) One of the plans turned out to be undeliverable with 1 cm jaws, so only 18 test plans were analyzed for this jaw setting. The 19 plans are described below.

#### 5.2.2.3.1 Cylindrical targets

Four types of cylindrical plans were developed. First, cylindrical targets of 2.5, 5, and 7.5 cm diameter and 10 cm long were planned to receive a uniform 2.0 Gy dose per fraction. The fourth plan in the cylinder series delivered a uniform 2.0 Gy dose to a 1cm thick cylindrical shell encompassing the AC diodes (10.4 cm from the isocenter). This plan resulted in a slowly varying, relatively uniform dose inside the shell, suitable for ion
chamber measurements. We will call this case a 21 cm cylinder. Pitch values of 0.287
were used for these plans.

5.2.2.3.2 TG-119 plans

The second set of plans was comprised of four cases from TG-119(22) (C-shape, Head and Neck, Multi-Target, and Prostate), each planned according to the Report. Two sets of plans were made, one using pitch 0.286 and the other using loose pitch of 0.86, for a total of eight TG-119 test plans.

5.2.2.3.3 Clinical plans

Seven clinical plans were selected. Those were comprised of three head and neck sites (including two plans with two dose-levels and a re-irradiation plan), one gynecological pelvic plan (endometrial adenocarcinoma), one anal site with the primary target and lymph nodes irradiated to different doses, one abdominal (gallbladder), and one brain case. All clinical plans delivered 1.8 or 2.0 Gy per fraction to the primary target. Pitch values of 0.287 were used for these plans.

All plans were calculated on a “fine” Tomotherapy TPS dose grid, which translates into a voxel size between approximately 2.0 and 2.7 mm. Two dynamic quality assurance (DQA) procedures were calculated for each plan: one on the ArcCHECK and one on the Delta4 (ScandiDos AB, Uppsala, Sweden) phantom. For both, the RT Plan and RT Dose objects were exported for analysis as required.

All plans were delivered on the TomoTherapy machine running Hi Art system software v. 4.2. Ours is a “hybrid” unit in that it has an older style Siemens linac coupled with an upgraded, fixed target.
5.2.2.4 Comparisons to the ion chamber

The ion chamber (IC) comparisons were performed on the PMMA AC phantom. The A1SL chamber was cross-calibrated against calculated dose in the center of phantom produced by a rotational, open 5×40 cm² beam with the static table. It was verified that the absolute IC dose agreed with the similar calculation near the center of a 30 cm diameter Virtual Water™ (“Cheese”) phantom to within 1%, essentially confirming the calibration comparison results. After cross-calibration, the test readings were taken with the MultiPlug inserted into the AC phantom, at the locations suggested by the 3DVH software as high-dose, low-gradient. Of 37 measurements, 10 required chamber movement to the off-center position. The ion chamber dose was compared to the TPS and 3DVH reconstruction without the IC normalization. The calculated/reconstructed values were reported as a mean dose to a 0.057cm³ sphere surrounding the point of interest, approximating the ion chamber volume. The same IC data were used as inputs to the optional 3DVH IC correction.

5.2.2.5 ArcCHECK dose analysis

The AC time-course measurement is used by TPDP, but the cumulative diode doses measured on the cylindrical surface can be also directly compared to the dose values extracted from TPS-calculated dose grid. Gamma analysis was employed to quantify the agreement between the dose distributions. Four sets of analysis criteria were used: 3%/3mm with global (G) and local (L) dose-error normalization, and the same for 2%/2mm. All dose analyses were in absolute dose. The lower dose cut-off threshold was set at 10% of the maximum dose. In addition, for each case the median
dose difference between the diodes and the TPS was correlated to the difference between the IC and the TPS.

5.2.2.6 Comparisons to a biplanar array dosimeter

5.2.2.6.1 Delta⁴ description, calibration and daily correction

After the point dose comparisons with an ion chamber, the next step in TPDP validation is comparing the reconstructed volumetric dose with the biplanar absolute dosimeter, the Delta⁴. As applied to TomoTherapy, the Delta⁴ system was validated and used for clinical patient QA in a large longitudinal series of 264 patient plans. It is important to note that in both papers, gamma analysis exclusively used 3% dose-error (global normalization)/3mm distance criteria. The TomoTherapy measurement mode has two important differences from the C-arm linac. First, no electrical trigger pulse is available from the accelerator and dose acquisition is triggered by sensing the radiation pulse itself. Second, the gantry angle as a function of time is unknown and therefore an average angular correction is applied to every diode, as opposed to the varying angle-specific series.

The reference dose for the Delta⁴ absolute calibration was obtained in a flat solid water phantom in the same fashion as for the AC. The only difference is that the calibration geometry requires the IC to be positioned at the isocenter (85 cm from the source) and the required buildup is 5.1 cm of solid water. The daily cross-calibration procedure is also essentially the same. The rotational 5×40 cm² beam with the static table was calculated on the Delta⁴ phantom and the differences between measurements and calculations in the central portion of the field were minimized by the application of
the multiplicative daily correction factor, which was then used for the subsequent measurements.

5.2.2.6.2 Specific tests

First, Delta\textsuperscript{4} diode measurements were compared to the TPS in a usual fashion, using the $\gamma$-analysis tools in Delta\textsuperscript{4} software (May 2014 version). Second, a TPDP DICOM RT Dose grid reconstructed on the Delta\textsuperscript{4} phantom was imported in the Delta\textsuperscript{4} software as a reference dose. Although the Delta\textsuperscript{4} and ArcCHECK phantoms are similar in cross-sectional shape, the TPDP dose had to be reconstructed using the patient dose pathway in 3DVH because the two phantoms differ in diameter. After spatial alignment, the differences between the measured and TPDP doses can be analyzed using the Delta\textsuperscript{4} software tools. For reference, TPDP was also compared to the TPS dose in 3DVH software, with the same four combinations of gamma analysis criteria.

5.2.2.7 Statistical analysis

Statistical analyses were performed with GraphPad Prism software (v. 6.0, GraphPad Software Inc., La Jolla, CA). Various tests were performed and the specifics of each one are provided along with the results, making the details easier to follow. If all the data series in a particular group passed the D’Augostino & Pearson normality test, the tests assuming Gaussian distribution of the data were selected. Otherwise, non-parametric statistical tests were employed. All $\pm$ errors quoted in the text correspond to one standard deviation.
5.3 Results

5.3.1 Phantom density

A relative density of 1.17 assigned to PMMA in the TPS produced the overall best agreement between the calculated and measured dose ratios. This value falls in between the PMMA relative electron and mass densities. When calculations were performed with this phantom density, excellent\(^{(24)}\) percent depth dose agreement between measured and calculated data was observed for a range of field sizes. The maximum difference was 1.2% of the local dose, at a substantial depth of 23.7 cm (Table 5.1). Based on these results, the uniform relative density of 1.17 was assigned in the TomoTherapy TPS to both the ArcCHECK and Delta\(^{4}\) phantoms.

Table 5.1. Percent differences (local) between calculated and ion chamber (IC) measured Exit to Entrance and Exit to Center dose ratios for the various size static beams on the ArcCHECK Phantom. Entrance and Exit points are 2.9 cm from the surface of the phantom.

<table>
<thead>
<tr>
<th>Field size Y×X (cm(^2))</th>
<th>Difference Calculated – IC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exit/Entrance ratio</td>
</tr>
<tr>
<td>5 × 40</td>
<td>0.0</td>
</tr>
<tr>
<td>5 × 3.75</td>
<td>1.0</td>
</tr>
<tr>
<td>2.5 × 6.25</td>
<td>1.2</td>
</tr>
<tr>
<td>2.5 × 3.75</td>
<td>1.2</td>
</tr>
<tr>
<td>1 × 6.25</td>
<td>--</td>
</tr>
<tr>
<td>1 × 3.75</td>
<td>--</td>
</tr>
</tbody>
</table>

5.3.2 Comparisons to the ion chamber

The descriptive statistics of the ion chamber comparisons to both the TPS and TPDP are presented in Table 5.2. The means of the columns corresponding to the two
different jaw settings were evaluated for statistically significant differences using the
paired t-test. While the TPS and IC agreed well for the 2.5 cm jaws, the 3.5% difference
in the high-dose low-gradient areas for the 1 cm jaws was a significant finding.\(^{(22)}\) Not
surprisingly, the t-test results were highly significant (P<0.0001). On the other hand,
TPDP (obviously without the IC correction in this case), shows approximately the same
average difference of about 1% for both 1 and 2.5 cm jaws. This TPDP difference
between the two jaw settings was not statistically significant, indicating that in both
cases the central point dose is on average reconstructed reasonably accurately from
the peripheral diodes readings. Individual TPDP vs. IC comparisons span the range
from -0.7 to +3.0%, compared to a range of -7.1 to 1.7% for TPS vs IC.

Table 5.2. Descriptive statistics of the percent ion chamber dose difference from the
TPS and TPDP (without chamber correction). N=19 for 2.5 cm jaw plans and N=18 for 1
cm plans. The means of columns with different jaws settings were compared with paired
t-test. Note: TPDP with chamber corrections are, by definition, 0%.

<table>
<thead>
<tr>
<th></th>
<th>Difference TPS – IC (%)</th>
<th>Difference TPDP – IC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Jaws 1 cm</td>
<td>Jaws 2.5 cm</td>
</tr>
<tr>
<td>Mean</td>
<td>-3.5</td>
<td>-0.3</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>1.5</td>
<td>1.4</td>
</tr>
<tr>
<td>Min</td>
<td>-7.1</td>
<td>-3.1</td>
</tr>
<tr>
<td>Max</td>
<td>-0.7</td>
<td>1.7</td>
</tr>
<tr>
<td>Lower 95% CI</td>
<td>-4.2</td>
<td>-1.0</td>
</tr>
<tr>
<td>Upper 95% CI</td>
<td>-2.7</td>
<td>0.4</td>
</tr>
<tr>
<td>Means different (P-value)</td>
<td>(Y(&lt;0.0001))</td>
<td>(N(0.7))</td>
</tr>
</tbody>
</table>

5.3.3 ArcCHECK dose analysis

The statistics of the γ-analysis of the AC vs. TPS dose are presented in Table
5.3. Due to the poor agreement at the 3%/3mm level, the data for 2%/2mm are omitted.
While the disagreement for the 1 cm jaws could be expected due to the previously described IC data bias, the lack of agreement for even the 3%G/3mm criteria combination was a significant finding for the 2.5 cm jaws setting.

The possible correlation between TPS-IC dose difference and median TPS-AC difference was studied with the non-parametric Spearman test. Spearman correlation coefficients were 0.61 for the 1 cm jaws (P=0.007) and 0.51 (P = 0.027) for the 2.5 cm jaws. The correlation between the peripheral and central doses is present but is not perfect, reflecting the fact that multiple factors contribute to the median TPS-AC difference.

Table 5.3. Descriptive statistics of the γ-analysis percent passing rates, TPS vs. ArcCHECK

<table>
<thead>
<tr>
<th>Median Dose Difference: 1 cm Jaws: -12.1±4.0 %; 2.5 cm Jaws: -7.8±2.2 % (1SD)</th>
<th>1 cm Jaws γ pass rate (%)</th>
<th>2.5 cm Jaws γ pass rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3%G/3mm</td>
<td>3%L/3mm</td>
</tr>
<tr>
<td>Mean</td>
<td>44.6</td>
<td>32.4</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>14.7</td>
<td>9.0</td>
</tr>
<tr>
<td>Min</td>
<td>22.1</td>
<td>19.6</td>
</tr>
<tr>
<td>Max</td>
<td>78.0</td>
<td>50.2</td>
</tr>
<tr>
<td>Lower 95% CI of mean</td>
<td>37.2</td>
<td>27.9</td>
</tr>
<tr>
<td>Upper 95% CI of mean</td>
<td>51.9</td>
<td>36.9</td>
</tr>
</tbody>
</table>

5.3.4 Comparisons between the TPDP, biplanar array, and TPS

In the next step, we compare TPDP dose reconstructed on the Delta4 “patient” (with and without the ion chamber correction) to the direct Delta4 biplanar measurements. The statistics for the γ-analysis comparisons are presented in Table 5.4 and Table 5.5 for the 1 and 2.5 cm jaw settings, respectively. The difference between
the results corrected and uncorrected with the IC point dose were analyzed with the non-parametric Wilcoxon matched-pairs signed rank test. For the cases planned with the 1 cm jaws, the IC correction leads to a small but statistically significant improvement in agreement between the TPDP and Delta⁴. However the overall agreement is not impressive for any criteria combinations more stringent than 3%G/3mm. Only 3%G/3mm in combination with the ion chamber correction produces the lower limit of 95% confidence interval (CI) above 90% of the Delta4 measurement points agreeing with TPDP. On the other hand, for the 2.5 cm jaws, the IC correction does not produce a statistically significant improvement in agreement between the TPDP and Delta⁴. However the overall agreement is much better, with the lower limit of the 95% CI reaching the 90% agreement rate for the 2%G/2mm criteria in combination with the IC correction.

Table 5.4. Descriptive statistics of the γ -analysis passing rates with different criteria combination, comparing TPDP with and without IC correction to the Delta⁴ measurements for the 1 cm jaw plans (N=18).

<table>
<thead>
<tr>
<th>Criteria Combination</th>
<th>NO IC</th>
<th>IC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median Dose Difference: No IC</td>
<td>-1.3±1.0%; IC: -0.8± 1.0 %</td>
<td></td>
</tr>
<tr>
<td>Std. Dev.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Min</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower 95% CI of mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper 95% CI of mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Means different (P-value)</td>
<td>Y(0.005)</td>
<td>Y(0.008)</td>
</tr>
</tbody>
</table>

\[\begin{array}{cccccc}
3\%G/3mm & 3\%L/3mm & 2\%G/2mm & 2\%L/2mm \\
\hline
\text{NO IC} & \text{IC} & \text{NO IC} & \text{IC} & \text{NO IC} & \text{IC} & \text{NO IC} & \text{IC} \\
\text{Mean} & 91.8  & 95.2  & 83.0  & 87.1  & 78.3  & 83.8  & 67.5  & 75.1  \\
\text{Std. Dev.} & 8.4   & 6.7   & 11.8  & 10.2  & 13.7  & 12.0  & 14.6  & 13.5  \\
\text{Min} & 74.3  & 76.7  & 64.8  & 66.9  & 55.1  & 59.8  & 47.4  & 51.9  \\
\text{Max} & 100.0 & 100.0 & 99.2  & 99.2  & 97.4  & 99.5  & 94.1  & 95.2  \\
\text{Lower 95\% CI of mean} & 87.7  & 91.9  & 77.2  & 82.0  & 71.5  & 77.8  & 60.3  & 68.4  \\
\text{Upper 95\% CI of mean} & 96.0  & 98.6  & 88.9  & 92.1  & 85.1  & 89.8  & 74.7  & 81.8  \\
\text{Means different (P-value)} & Y(0.005) & Y(0.008) & Y(0.002) & Y(<0.001) \\
\end{array}\]
Table 5.5. Same as Table 5.4 but for 2.5 cm jaws (N=19).

Median Dose Difference: No IC: -0.6±0.6%; IC: -0.3±0.8 %

<table>
<thead>
<tr>
<th>γ-analysis pass rate (%)</th>
<th>TPDP vs. Delta⁴</th>
<th>NO IC</th>
<th>IC</th>
<th>NO IC</th>
<th>C</th>
<th>NO IC</th>
<th>C</th>
<th>NO IC</th>
<th>IC</th>
</tr>
</thead>
<tbody>
<tr>
<td>3%G/3mm</td>
<td></td>
<td>98.8</td>
<td>99.1</td>
<td>95.0</td>
<td>95.7</td>
<td>91.6</td>
<td>93.4</td>
<td>84.1</td>
<td>86.0</td>
</tr>
<tr>
<td>3%L/3mm</td>
<td></td>
<td></td>
<td></td>
<td>1.5</td>
<td>1.8</td>
<td>3.9</td>
<td>4.8</td>
<td>4.2</td>
<td>7.1</td>
</tr>
<tr>
<td>2%G/2mm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>84.3</td>
<td>68.4</td>
<td>71.6</td>
<td>57.6</td>
</tr>
<tr>
<td>2%L/2mm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>99.9</td>
<td>99.9</td>
<td>96.8</td>
<td>98.9</td>
</tr>
</tbody>
</table>

Although not reaching significance, the agreement levels improve consistently with the application of the IC correction. Therefore in the last test, the IC-corrected results were used when comparing the TPDP and Delta⁴ dose grids to the TPS. The γ-analysis agreement rates are presented side by side for the TPDP vs. TPS and Delta⁴ vs. TPS comparisons.
Table 5.6. Descriptive statistics of the γ-analysis passing rates with different criteria combination, comparing TPS vs. TPDP and TPS vs. Delta⁴ for the 2.5 cm jaw plans (N=19).

Median Dose Difference: TPS vs TPDP: 4.4±1.7%; TPS vs Delta⁴: 0.6± 0.6 %

<table>
<thead>
<tr>
<th>γ-analysis pass rate (%)</th>
<th>TPS vs. TPDP</th>
<th>TPS vs. Delta⁴</th>
</tr>
</thead>
<tbody>
<tr>
<td>3%G/3mm</td>
<td>Mean</td>
<td>98.1</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>2.6</td>
<td>2.6</td>
</tr>
<tr>
<td>Min</td>
<td>91.1</td>
<td>88.7</td>
</tr>
<tr>
<td>Max</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Lower 95% CI of mean</td>
<td>96.9</td>
<td>97.8</td>
</tr>
<tr>
<td>Upper 95% CI of mean</td>
<td>99.4</td>
<td>100.0</td>
</tr>
<tr>
<td>Means different (P-value)</td>
<td>Y (0.03)</td>
<td>Y(0.0004)</td>
</tr>
</tbody>
</table>

The data for the 2.5 cm jaws setting are provided in Table 5.6. Because the 1cm jaws plans showed a large mean disagreement between the calculated and measured point doses in the high-dose low-gradient regions, the dose distribution agreement data cannot be meaningfully discussed. The results of the comparisons between TPDP and Delta⁴ vs. the TPS are mixed. Depending on the criteria, the differences range from non-significant to highly significant based on the Wilcoxon test (Table 5.6). The smallest difference in mean passing rates is observed for the 3%G/3mm criteria (0.9% in favor of the Delta⁴) while the largest one is 13% for the 2%L/2mm analysis. The larger differences in the passing rates are associated with the local dose-error normalization,
indicating that the differences between the TPDP and the Delta\textsuperscript{4} tend to occur more in the lower dose region.

5.4 Discussion

5.4.1 Comparisons to the ion chamber

While for the 2.5 cm jaws the average difference between the TPS and IC was minimal (-0.3±1.4%), the result for the 1 cm jaw (-3.5±1.5%) was clearly outside of the generally accepted ±1.5% range demonstrated in TG-119\textsuperscript{(22)}. In an attempt to find the immediate cause of this discrepancy we turned to the individual static beam measurements. Comparison of the measured (IC) and calculated dose at the center of a cylindrical phantom demonstrated that for the 1 cm jaw settings the measured dose was consistently 5.5 to 6% higher than the TPS prediction, depending on the number of the open MLC leaves. Since the IC dose is higher than the TPS, the discrepancy cannot be blamed on the chamber volume averaging effect in a narrow beam. ArcCHECK diode readings at the beam entrance agreed with the IC dose at the phantom center. Apparently the modulated helical delivery dynamics reduce the error to 3.5% on average, but the significant bias remains. As mentioned in the TG-148 Report\textsuperscript{(20)} in the course of routine QA, the user does not have access to the calculated static beam dose distribution to compare to the IC reading.

We hypothesized that with a loose pitch the results might be closer to the static beam. Comparison between the four TG-119 plans optimized with 0.287 and 0.860 pitch revealed a small difference in the average TPS-IC difference, -3.4±1.5% vs. -4.1±1.0%. While the direction of the change followed the hypothesis, it is not statistically significant and is not likely to be of practical importance. Unlike the TPS, the TPDP – IC
differences are essentially the same for both jaw settings. The mean error does not exceed 1.1±1.1%, which implies that on average the impact of the IC correction on the TPDP results should be moderate.

It must be noted, that the 1 cm jaw setting is very seldom used in our clinical practice. When it was used, the plans showed better than 90% γ(3%/3mm) passing rates in comparison with the Delta\(^4\). Moreover, in the current work 14 of 19 plans exhibited >95% agreement rate between the TPS and Delta\(^4\). This is yet another example of how the 3%/3mm criteria may hide a systematic commissioning error.\(^{(25)}\)

The possible remedy to eliminate the dose bias is under investigation. While this dose bias is of clinical concern, it is, ironically, beneficial for the current study, as it provides a built-in test of TPDP’s sensitivity to dose errors.

### 5.4.2 ArcCHECK dose analysis

The agreement between the TPS and AC diode readings is poor regardless of the jaw setting (Table 5.3). Since our results disagreed with a previous report by Bresciani et al.\(^{(26)}\), they were spot-checked with a second ArcCHECK unit. While there were expected minor differences in the measured datasets, the trend of the TPS substantially underestimating the peripheral dose remained intact. The combination of our results with those previously reported\(^{(26)}\) suggest that the peripheral dose is sensitive to the potential differences that are machine-specific, e.g. in the machine Y-jaw alignment. The median differences between the TPS and the diode readings for all detectors above the 10% threshold were 12.1± 4.0 and 7.8± 2.2% for the 1 cm and 2.5 cm jaw settings, respectively. The difference in the median values for the two jaws settings (4.3%) is in line with the ion chamber central dose differences (3.2%). This
suggests that the AC response is not noticeably affected by the field size, despite potential partial loss of the charged particles equilibrium in the 1 cm beam.

While the differences between our results and the previous report are noteworthy, they are not entirely unexpected. It was described previously how a small change in the arc aperture can lead to a substantial variation of the peripheral (ArcCHECK) dose for the conventional arcs.\(^{(27)}\) The situation is even more complicated in tomotherapy, as dosimetric effects of the small disagreements between the calculated and delivered fluence in the penumbra of the diverging helical beam on the periphery of the phantom have not been studied at all.

Just as with conventional arcs\(^{(27)}\), the difference in the peripheral dose between the TPS and AC is strongly dependent on the beam aperture width. Figure 5.4 shows how the median difference between the calculated and measured (AC) TomoTherapy dose-difference increases with the decrease in the cylindrical target diameter. For the 1 cm jaws, as the target diameter is reduced from 21 to 2.5 cm, the median dose-difference increases from -4.4 to -23.5%, or more than five-fold. For the 2.5 jaws the change is less dramatic but is still very clear (Figure 5.4). These consistent patterns of change further confirm that the observed differences are real and not the measurement artifact. As the target diameter decreases, the AC diodes location becomes “more peripheral” in that the accuracy of the dose calculation there becomes more and more dependent on the exact representation of the penumbra and out of field portions of the dose profile, which are typically the weakest points of any TPS beam model.
There are also subtle but important details that should be factored into comparing the current ArcCHECK results with Bresciani et al.\textsuperscript{(26)} First, the authors did not specify how the AC was calibrated and there are no ion chamber data to make an inference. In addition, it was assumed that an angular correction was applied to the diode readings. Such correction was not available for tomotherapy AC measurements in 2013, leading to the average measurement dose overestimation by about 1.8%. Unless the cross-calibration against the TPS was used, this could potentially substantially alter the passing rates with 1 or 2% dose-error criteria, and somewhat affect those using 3%. Finally, it is not specified if the “measurement uncertainty” feature in the AC software was turned on, as it was in TG-119\textsuperscript{(22)} for the planar array. If so, it effectively increases the dose-difference threshold by \(\sim 1\%\), leading to the artificially inflated agreement rates.
We also highlight an important caution: it might be tempting to use the plan-class specific methodology\(^{(28)}\) for cross-calibrating the AC against the TPS in a “simple” helical field. However, if a standard field described in TG-148\(^{(20)}\) delivering a uniform dose to a 8 cm diameter cylinder is used verbatim, the calibration of the AC can be altered greatly. To illustrate this point, we “cross-calibrated” the AC, introducing a single multiplicative correction factor that maximized agreement with the TPS for the 7.5 cm diameter cylinder. The resulting average TPS vs. AC agreement rate for the set of clinical and TG-119 2.5 cm jaw plans increased from 86.7.9±10.8 to 94.2±5.9%. This shows that the concept of the plan-class specific correction factor should be applied with caution in the peripheral areas, where the TPS may be less accurate. Developing correction factors based on helical delivery can lead to vastly different results from the static-table calibration/correction. While the passing rates with plan-class specific cross-calibration may be higher, they would hide the true difference between the TPS calculations and measurements on the periphery.

Regardless of the potential methodology differences, the disagreement with the previously published data served a useful purpose, as it led to the realization that the same relationship between the peripheral and central dose cannot be assumed across all TomoTherapy systems. As a result, the IC correction technique was developed to mitigate any potential differences given that TPDP uses a standard dose model for all TomoTherapy machines, though clearly for our test data, the required extra correction was modest.
5.4.3 Comparisons between the TPDP, biplanar array, and TPS

For these studies on our TomoTherapy unit, the IC correction leads to a modest improvement to the agreement between the TPDP and Delta\(^4\). The effect is larger for the 1 cm jaws, consistent with the notion that the narrow field exposes fewer diodes to the high dose at any given time, thus complicating the MGDR process. For the 2.5 cm jaws, the IC-corrected TPDP dose distributions compare favorably with the Delta\(^4\) at the 3\%G/3mm level (99.1±1.8\%). The average passing rates get progressively lower as the analysis criteria tighten up (Table 5.5). With the lower 95\% confidence interval (CI) of mean ≥90\% as a cutoff point, the agreement can be considered satisfactory at the 3\%L/3mm and 2\%G/2mm levels. With 2\%L/2mm criteria the agreement is suboptimal. This is in contrast to the VMAT data, which showed over 95\% agreement rates between 3DVH and the Delta4 at the 2\%L/2mm level for the three TG-119 datasets. We attribute this difference to the necessary approximations introduced in the tomotherapy version of MGDR compared to VMAT, as described in the Methods section. Dose reconstruction is even more challenging with the 1 cm jaw setting. As a result, the lower 95\% CI of the mean agreement rate between the TPDP and Delta\(^4\) is above 90\% only for γ(3\%G/3mm). Therefore for the jaw setting of 1 cm we cannot ascertain that the QA results using TPDP would be meaningful with the criteria tighter than 3\%G/3mm.

While this level of accuracy may be suboptimal relative to MGDR for IMRT and VMAT,\(^25\) the 3\%G/3mm evaluation criteria are still widely employed and with the exception of Ref.\(^26\) were used in every publication dealing with TomoTherapy dosimetric accuracy. Specifically, the known reports of the Delta\(^4\) validation and use with TomoTherapy were based on the γ(3\%G/3mm) criteria. This is perhaps a reflection
of the more general problem that it is hard to find a dosimeter to reliably sample
TomoTherapy dose distribution with, say, 2%L/2mm accuracy level. This affects the
precision of both new dosimeter validation and TPS dose verification. Great care must
be taken to achieve even 3% dose readout accuracy with radiochromic film\(^{(29)}\). Planar
electronic arrays suffer from uncorrected angular dependence reducing their
accuracy\(^{(30)}\). A chamber based array with a phantom cavity partially compensating for
the angular dependence could be successfully compared to the TomoTherapy TPS at
the 5%/3mm level in one publication\(^{(31)}\) and 3%/3mm in another.\(^{(32)}\) Another planar
chamber array required up to 3% plan-specific correction factor to achieve acceptable
agreement with the TPS, again at the 3%G/3mm level.\(^{(33)}\) This was attributed to the
“uncertainties of TomoTherapy delivery”. Chamber-based arrays have limited spatial
resolution, reducing their usefulness in the high gradient areas.\(^{(34)}\) Volumetric
radiochromic plastic or gel dosimeters which provide the most comprehensive way of
evaluating 3D dose reconstruction are typically used with global 3% dose-error
criteria.\(^{(9,35)}\) The AAPM Report on the QA for helical tomotherapy (TG-148)\(^{(20)}\)
recommendation for evaluating agreement between calculated and measured
TomoTherapy dose distribution is \(\geq 90\% \gamma(3\%G/3\text{mm})\) passing rate for all plans. Based
on the ion chamber and Delta\(^4\) results in the current work, years of patient-specific QA
experience, and the results of the RPC end-to-end test on the anthropomorphic head
phantom,\(^{(36)}\) we believe that the studied TomoTherapy unit is commissioned in
accordance with the current standard of practice, excluding the 1 cm jaw setting that
needs corrective action. Supporting this assertion is the fact that the lowest
\(\gamma(3\%G/3\text{mm})\) agreement rates between TPDP or Delta4 and the TPS was 91.1% and
88.7%, respectively, with the means of 19 cases at 98.1±2.6% and 99±2.6%, respectively (Table 5.6). This is in stark contrast to the low passing rates – for any reasonable criteria combination – if one compares TPS to ArcCHECK diodes alone (Table 5.3).

Finally, as discussed in the previous section, the TPS vs. ArcCHECK (diodes only) analysis showed TPS calculations consistently low compared to the peripheral dose measured by the diode surface (radius 10.4 cm) which in turn caused low pass rates as seen in Table 3. However, TPS vs. TPDP passing rates (Table 6) were much higher, as the TPS errors in low dose regions were balanced by the more central voxels. This is further evidence that the location of the detectors in a 3D phantom can have a drastic impact on passing rates, with all else equal, as has been pointed out previously.\(^{(27)}\) We illustrate this with Figure 5.5. A TPS vs. TPDP difference map shows the many points different by more than 2% (local) are confined to the lower dose periphery, exactly where the diodes are. At the same time high dose, central regions agree well within 1%. This figure also partially illustrates why accurately calculating the dose at the diodes’ locations is so hard, as the dose distribution constitutes a complex streak pattern.
Figure 5.5. A) Central axial plane for the TG-119 prostate, 2.5 cm jaw plan, pitch 0.286. B) TPDP – TPS dose differences (red denotes TPDP > TPS by more than 2%) showing that the TPS calculations low in the periphery but within 1% centrally. This was a typical error pattern and is consistent with the very low TPS vs. ArcCHECK diode passing rates (Panel C), but higher passing rates for the same TPS dose grids compared to TPDP or Delta4, both of which are not limited to sampling peripheral dose. C) Red dots correspond to the diodes that fail 2%G/2mm gamma comparison against the TPS (43.9% passing rate).

Prior studies on MCPD(6,7,37,38) and ACPDP(8-11,25) have shown acceptable accuracy as verified by comparison to independent dosimeters, as well as using both in silico simulations of known errors (with known results) and independent measurement of induced-errors vs. the PDP estimates. At those previously established levels of accuracy, a physicist can have reasonable confidence in both the sensitivity and specificity of dose and DVH analysis using those methods. However, TPDP had to solve many challenges unique to helical tomotherapy, the most important of which were: 1) very small and complex
exposed areas as a function of time, given the narrow binary leaves relative to the semi-sparse diodes at entry and exit surfaces of the measurement array; and 2) extreme dynamics of the MLC leaves and the gantry speed which cannot be effectively discretized even by 50 ms measurement updates. These challenges were handled reasonably well, but the imperfections in Tomo-PDP could, in theory, themselves lead to “false negatives” in QA (e.g. not reporting an important error) or “false positives” (e.g. predicting an error that is not real). This is especially a concern if the region(s) of error are overlapping with target volumes and/or organ-at-risk volumes. So, one should keep this in mind and perhaps adopt more lenient DVH-matching criteria due to the potential uncertainty of the dose reconstruction method itself. We do not quantify the specificity and sensitivity of TPDP in this paper. First, it would be a subject for a separate full-length report. Second, if it were a goal, TomoTherapy is a rather closed system and inducing known errors for empirical verification would be quite difficult. Although TPDP is imperfect, it certainly highlighted some limitations of the TomoTherapy system in the ability to accurately predict the delivered dose. Other recent publications\(^{(39,40)}\) suggest similar limitations in the TomoTherapy TPS dose calculation. These errors are worth investigating at the level of commissioning of the system, and in this regard TPDP could be quite valuable. On a per-patient basis, the value of TPDP is that it is measurement-guided and thus 100% independent of the TPS dose engine and generates the error map on the patient dataset; as such, any large differences are worth investigating.
5.5 Conclusions

The volumetric measurement-guided dose reconstruction algorithm required substantial modifications to accommodate the much faster delivery dynamics of the TomoTherapy machines compared to the C-arm accelerator VMAT. The necessary approximations and hardware limitations lead to the somewhat reduced reconstruction accuracy. However the volumetrically reconstructed dose agrees very well with an independent array dosimeter and TomoTherapy TPS when gamma analysis with 3%G/3mm is used, which is consistent with the current standard of practice. The optional ion-chamber based correction results in a modest improvement in dose agreement. When properly calibrated in a static-table beam, the ArcCHECK alone samples only the peripheral dose and at least for the studied TomoTherapy unit can report poor agreement with the TPS for the plans that are in reality clinically acceptable.

5.6 References


CHAPTER 6:
CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

As a proof-of-concept, this work has demonstrated that when evaluated and used properly, electronic arrays can be a powerful tool for quality assurance and clinical studies for IMRT and VMAT using both a conventional C-arm linear accelerator and helical delivery. While varying viewpoints have been expressed \(^{(1,2,3)}\), the argument against the use of the electronic arrays based on resolution does not consider the volumetric reconstruction capabilities of these devices. Using quasi-3D arrays allow for quicker and easier dose comparisons with reasonable resolutions in comparison to traditional methods. They also allow for DVH comparison for the ROI’s, which enables the user to achieve a more clinically relevant plan comparison. The addition of the motion perturbation capabilities of 3DVH enables the user to easily analyze the effect any target motion on the overall dose without the need of cumbersome setups or long usage time on the treatment machine. These advantages point toward a need for further discussion of the role of electronic arrays in the commissioning of IMRT and VMAT as well as discussion on the best practice for plan comparison for patient-specific quality assurance.
6.2 Future Work

As treatment plans become significantly patient focused the need for easy to use and clinically relevant quality assurance grows. Even if the technology is available to achieve this, the system needs to be streamlined to make it user friendly and easy so that clinicians are willing to use it regularly. Also, with the development of real time tumor tracking \(^{(4,5)}\), the ease of simulating the dose to a moving target using 3DVH makes ArcCheck an interesting and potentially promising post-treatment quality assurance device.

6.3 References


APPENDIX A:

PUBLICATION LIST


Editor’s pick & top 10 download Feb 2013

APPENDIX B:
COPYRIGHT PERMISSIONS


DATE OF REQUEST: 02/6/2015

FROM:
Cassandra Stambaugh, 1038 Biltmore Place Lakeland, FL 33801

EMAIL ADDRESS: cstambaugh@mail.usf.edu

1. Permission is granted to:
Cassandra Stambaugh

2. Permission is requested to use the following material:

3. For what purpose:
Use in Dissertation.

Authors seeking permission must also notify the first author of the article from which permission is being sought.
Evaluation of semi-empirical VMAT dose reconstruction on a patient dataset based
on biplanar diode array measurements. *Journal of Applied Clinical Medical Physics*,
15(2).

Stambaugh, C., Nelms, B., Wolf, T., Mueller, R., Geurts, M., Opp, D., Moros,
reconstruction for helical tomotherapy. *Journal Of Applied Clinical Medical
Physics*, 16(2).
ABOUT THE AUTHOR

Cassandra Stambaugh received a Bachelor’s of Science degree in Mathematics and Physics from Bates College in Lewiston, Maine in 2008. She went on to receive her Masters of Science in Physics at Northeastern University in Boston, Massachusetts in 2010. Cassandra was admitted into University of South Florida’s Applied Physics program in 2012 and immediately focused on Medical Physics.

During the Ph.D. program, Cassandra published in many peer reviewed journals and earned the honor of editor's pick and top download twice from Medical Physics. She was also awarded the Fred L. and Helen M. Tharp summer research scholarship.