Large-scale in vivo EPID dosimetry

Rozendaal, R.A.

Publication date
2017
Document Version
Final published version
License
Other

Large-scale *in vivo* EPID dosimetry

Roel Rozendaal

Roel Rozendaal

Large-scale *in vivo* EPID dosimetry

Roel Rozendaal

Cover by Renate Stevens, `renatestevens.com`

Large-scale *in vivo* EPID dosimetry

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor

aan de Universiteit van Amsterdam

op gezag van de Rector Magnificus

prof. dr. ir. K.I.J. Maex

ten overstaan van een door het College voor Promoties ingestelde commissie,

in het openbaar te verdedigen in de Agnietenkapel

op dinsdag 7 november 2017, te 10.00 uur

door Roel Arthur Rozendaal

geboren te Leeuwarden

Promotiecommissie:

Promotor(es):	Prof. dr. M.B. van Herk	AMC-UvA
Copromotor(es):	Dr. ir. A. Mans	NKI-AvL
Overige leden:	Prof. dr. C.R.N. Rasch	AMC-UvA
	Prof. dr. A.G.J.M. van Leeuwen	AMC-UvA
	Prof. dr. M.W.M. Jaspers	AMC-UvA
	Dr. M.S. Hoogeman	Erasmus Universiteit Rotterdam
	Prof. dr. B.W. Raaymakers	Universiteit Utrecht
	Prof. dr. ir. A.L.A.J. Dekker	Universiteit Maastricht

Faculteit der geneeskunde

Contents

Chapter 1

Introduction

Radiotherapy is one of the main methods of cancer treatment, besides surgery and chemotherapy. It is estimated that about 50% of cancer patients will receive radiotherapy during the course of their disease [1]. Over the past two decades, radiotherapy has seen a steep increase in technological complexity of treatment preparation and treatment execution, moving from 2D to 3D/4D planning techniques and from rectangular field treatments via 3D-conformal to intensity-modulated radiotherapy (IMRT) and volumetric arc therapy (VMAT). As complexity increased, a demand for more and better quality assurance (QA) has developed. Indeed, some very serious incidents in radiotherapy have been reported [2, 3, 4, 5, 6]. Often, plan QA in radiotherapy is performed pre-treatment: performing dose measurements based on the patient's treatment plan before the treatment starts. Drawbacks of such pre-treatment QA are: (1) the actual patient geometry is not included in these measurements, making it difficult to estimate the effect of observed dosimetric deviations on the actual patient; (2) deviations from intended treatment which only occur during treatment, such as described in Mans et al. [7] and Ford et al. [8], cannot be detected; (3) pre-treatment verification requires an additional measurement, taking up valuable linac[1] time and increasing workload. The first issue is currently being addressed by radiotherapy QA system manufacturers by reworking the observed deviations to deviations on the patient anatomy (using the planning CT-scan); the last two issues are fundamentally unsolvable by means of pre-treatment verification.

Another method for dosimetric QA is performing *in vivo* dosimetry. This entails

[1] **Lin**ear **ac**celerator, linac for short, is the radiation source for treatment.

measuring dose during treatment delivery, reconstructing patient dose if needed, and comparing it to the intended dose. Several forms of *in vivo* dosimetry exist; diode detectors and electronic portal imaging devices (EPIDs) are most commonly applied. An EPID is a 2D detector, originally developed as an imaging device [9, 10] which can measure high-energy radiation. EPIDs are mounted opposite to the radiation source and hence detect the radiation after it has left the patient. It was soon noted that EPIDs are also suitable for dosimetry [11, 12], as the detected signal is proportional to the dose incident on the EPID. The first dosimetric application was exit-dosimetry [12, 13, 14, 15], where the dose leaving the patient is verified; calculating the dose in the radiological mid-plane (i.e., inside the patient) soon followed [16]. Different methods are currently employed for *in vivo* dosimetry: point-based [17, 18], comparison of predicted and measured portal images [19, 20, 21], and 3D dosimetry [22, 23, 24, 25]. This thesis will deal exclusively with a variant of the last method, 3D EPID *in vivo* dosimetry, based on a backprojection method developed at the Netherlands Cancer Institute (NKI). This method takes as input an EPID image of the radiation beam after having traversed through the patient and the (planning) CT-scan of the patient in question. After calculating the transmission through the patient using the CT-scan, the dose information contained in the EPID image is backprojected to yield the delivered dose in the patient geometry.

The NKI backprojection method had improved greatly since its inception in 1995, but clinical use has remained very limited for a long time. At the introduction of a new treatment planning system (TPS) in 2005, it was decided to implement EPID dosimetry clinically; first as a pre-treatment tool and soon as an *in vivo* dosimetry method replacing pre-treatment verification [26]. With expanding clinical use, several issues emerged. First, with an increasing number of treatments being verified, the clinical workload increased unacceptably. Second, the results obtained from the applied method for dose verification have no clear relation to the clinical relevance of observed deviations, making the inspection of found deviations a complex task. This led to the aim of this thesis: improving patient safety in radiotherapy by enabling the EPID *in vivo* dosimetry system to be used efficiently on a large-scale, and ultimately in real-time.

Figure 1.1 shows the development of the number of treatment plans verified by means of *in vivo* EPID dosimetry in our institute over the years. As manual verification of a VMAT treatment could easily take up to one hour, shortly after the introduction of VMAT the workload of performing *in vivo* EPID dosimetry became unacceptable. Improving the efficiency of the clinical workflow by automation therefore forms the first

part of this thesis.

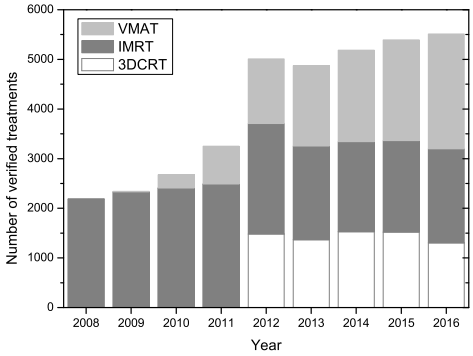


Figure 1.1: Number of treatment plans verified using *in vivo* EPID dosimetry procedure at the NKI.

For each verified treatment, the EPID *in vivo* dosimetry system will raise an alert if the verification result is out of tolerance. Figure 1.2 shows the development of alerted/unalerted treatments for IMRT and VMAT over the years; the number of treatment plans for which action was taken (e.g., informing the attending clinician, treatment replanning) is roughly constant at about 0.3% [7, 27]. It can be seen that the number of false positives is quite elevated, indicating a lack of specificity in the current clinical workflow. To explain this, a little background on the verification method is well in place.

Traditionally, radiotherapy treatment plan QA is evaluated using γ-analysis [28]; clinically, γ-analysis is employed for our EPID dosimetry method as well. In short, this method determines the shortest distance in space and dose between a measurement point and a reference dose distribution, as depicted in Figure 1.3. Both the distance in space and the distance in dose, more commonly referred to as a dose difference, need to be normalized; common normalization values are 3 mm (distance) and 3% (dose). The dose normalization can be chosen to be global, i.e., 3% of a fixed dose value such as the maximum dose or the prescribed dose, or local, i.e., 3% of the dose of the measurement point under consideration. Normalization values are usually chosen such that a γ-value of up to 1 is considered acceptable. Colloquially speaking, γ-analysis allows for larger dose differences when the local dose gradient is higher. A major drawback of γ-analysis is the lack of correspondence to clinically relevant information: it is a mathematical

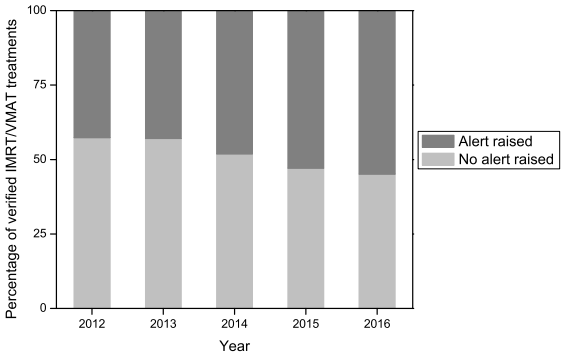


Figure 1.2: Amount of alerts raised over the years for IMRT and VMAT treatments.

technique to compare two dose distributions, but the relevance of observed deviations for the patient under consideration is unclear. Moreover, several studies [29, 30, 31, 32] have shown that γ-analysis can fail to detect clinically relevant dosimetric deviations in treatment plan delivery – depending on the tolerance levels employed, of course. Thus, in order to achieve a high sensitivity (i.e., avoiding false negatives), strict tolerance levels are used clinically, leading to a decreased specificity.

As our *in vivo* EPID dosimetry method is capable of reconstructing delivered dose in 3D, a solution for the drawbacks of γ-analysis emerges. The reconstructed 3D dose distribution can be evaluated by employing the common methodology for plan-evaluation used in the treatment planning stage. Specifically, this entails calculating dose-volume histogram (DVH) parameters for the reconstructed dose distribution and comparing these with the planned DVH parameters. Besides converting observed dosimetric variations into a deviation with a clear clinical interpretation, using DVH-based analysis for *in vivo* dosimetry verification might provide an additional benefit: reduction of false positives. This would reduce clinical workload dramatically, as any treatment which shows an unacceptable deviation at EPID dosimetry needs to be manually approved by a trained medical physicist.

Dose verification is done by computing several statistics (mean, near-maximum and percentage smaller than one), called indicators, on the histogram obtained from a γ-analysis. For each indicator, a tolerance level has been set for acceptability. The obvious solution for increasing specificity for γ-based dose verification would be to increase the

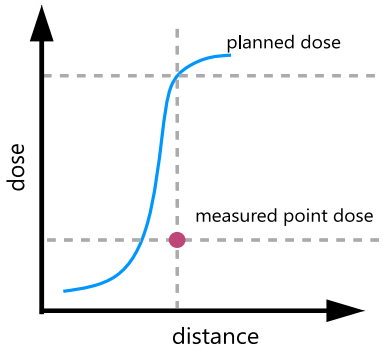


Figure 1.3: The concept of γ-analysis. The γ-value of the measured point dose equals the shortest distance between the point and the reference (planned) dose curve in dose and distance. This procedure is repeated for all measured dose points and can be done in 1D, 2D or 3D. Note that in this example the dose-difference is very large, but the shortest distance to the reference dose is small leading to a low γ-value.

tolerance levels for the indicators. But this is a delicate matter, precisely because the relation between γ-analysis results and clinical relevance is unclear! DVH-based dose verification provides a solution because the results obtained from DVH-analysis do provide clinically relevant information as these are actually used to evaluate the treatment plans.

Using DVH-analysis for *in vivo* dose verification is a relatively uncharted area; no institutes are yet known to employ this on a large scale in the clinic and only very few results have been published [33]. The second part of this thesis investigates the feasibility of applying dose-volume metrics for *in vivo* dose verification.

Finally, the current clinical EPID *in vivo* dosimetry method verifies the delivered dose after delivery of one or more fractions. This means that large, life-threatening errors that may occur during delivery can only be detected after the harm has been done. Such incidents are very rare, but are described in the literature, e.g. occurring due to failure to properly shape the treatment fields [2, 3, 4]. A natural extension of the *in vivo* dosimetry method would then be real-time verification of the delivery, i.e., dose verification during treatment. Key points of such a system are speed and the ability to verify the delivered dose in a time-resolved manner. Currently, the only clinically employed method for real-time dosimetry is based on the predicted/measured portal images method [34, 35]. A recently developed system for 3D EPID-based pre-treatment verification is also intended

to be developed into a real-time 3D *in vivo* dose verification system [36]. Extension of the existing backprojection method to a real-time dose reconstruction and DVH-based verification system is described in the final part of this thesis.

This thesis is organized as follows:

Chapter **2** shows how the EPID dosimetry method can be applied in an efficient and automated way such that all treatment plans of a large institute can be verified using *in vivo* EPID dosimetry.

Chapter **3** deals with the connection between γ-analysis and DVH-analysis for *in vivo* dosimetry of two treatment sites and techniques (head-and-neck (H&N) VMAT and lung IMRT).

Chapter **4** investigates the effect of daily anatomical variations as observed via pre-delivery cone-beam CT (CBCT) scans on planned and delivered dose to the target volume for H&N VMAT.

Chapter **5** presents a way of performing DVH-based dosimetry for additional verification measurements (i.e., without the patient being present).

Chapter **6** demonstrates a proof-of-concept for real-time 3D dose verification using an accelerated version of our back-projection method.

Chapter **7** analyses historical *in vivo* dosimetry DVH-analysis results for lung stereotactic body radiation therapy (SBRT) in a time-resolved manner.

Chapter **8** concludes with a summary and a general discussion of the presented work.

Enjoy!

Chapter 2

Automatic *in vivo* portal dosimetry of all treatments

I. Olaciregui-Ruiz
R.A. Rozendaal
B. Mijnheer
M. van Herk
A. Mans

Physics in Medicine and Biology 58 (22), 2013

Abstract

At our institution EPID (Electronic Portal Imaging Device) dosimetry is routinely applied to perform *in vivo* dose verification of all patient treatments with curative intent since January 2008. The major impediment of the method has been the amount of work required to produce and inspect the *in vivo* dosimetry reports (a time-consuming and labor-intensive process). In this paper we present an overview of the actions performed to implement an automated *in vivo* dosimetry solution clinically. We reimplemented the EPID dosimetry software and modified the acquisition software. Furthermore, we introduced new tools to periodically inspect the record-and-verify database and automatically run the EPID dosimetry software when needed. In 2012, 95% of our 3,839 treatments scheduled for *in vivo* dosimetry were analyzed automatically (27,633 portal images of IMRT fields, 5,551 portal image data of VMAT arcs, and 2,003 portal images of non-IMRT fields). The *in vivo* dosimetry verification results are available a few minutes after delivery and alerts are immediately raised when deviations outside tolerance levels are detected. After the clinical introduction of this automated solution, inspection of the detected deviations is the only remaining work. These newly developed tools are a major step forward towards full integration of *in vivo* EPID dosimetry in radiation oncology practice.

2.1 Introduction

Complicated treatment techniques such as intensity-modulated radiotherapy (IMRT) and volumetric-modulated arc therapy (VMAT) are now widely used to deliver a high dose to the tumor while minimizing the dose in the surrounding healthy tissue. In order to achieve accurate dose delivery with these techniques, a great deal of attention has to be paid to quality assurance (QA) of all steps in the planning and delivery process.

Not long after its clinical introduction for setup verification, it was realized that EPID images contain dose information. Particularly after the introduction of amorphous silicon-type of EPIDs, their use for dose verification was emphasized e.g. by Greer and Popescu [37]. A review of this topic was presented by van Elmpt et al. [38]. While patient setup verification using EPIDs has become a routine procedure in many hospitals, the dosimetric verification of treatments using EPIDs has not yet come to that stage. The main reason for this is that commercial tools have only recently become available. Several groups, however, developed their own software and currently perform EPID dosimetry clinically in a routine way with *in vivo* portal measurements (i.e. portal images acquired during treatment). For instance, Nijsten et al. [39] provide results obtained over a period of 24 months of more than 2,500 patient treatments. In Italy a national project recently started and will be applied in a large number of centers [40]. In our institution Mans et al. [7] performed an analysis of the clinical experience with EPID dosimetry on 4,700 patients. The results of these studies illustrate the importance of performing dosimetry with *in vivo* measurements in tracing errors in dose delivery that cannot be detected by means of pre-treatment QA measurements.

In The Netherlands Cancer Institute-Antoni van Leeuwenhoek Hospital an EPID-based dosimetry method has been developed for both pre-treatment and *in vivo* dose verification of IMRT and VMAT using a fast semi-empirical back-projection algorithm [41, 22] which allows for dose reconstruction within the patient (or phantom) in multiple planes parallel to the EPID. It is worth noting that we refer to our method as an *in vivo* dosimetry method because the dose is determined within the living from portal measurements that contain information about the radiation dose that has been received by the patient. *In vivo* dosimetry in radiotherapy means the measurement of the radiation dose received by the patient during treatment, opposed to ex vivo or in vitro dose measurements made either before or after the treatment using a phantom to represent the patient, as elucidated elsewhere [42].

Since January 2008 we have routinely applied this method to the verification of all

treatments with curative intent (99% IMRT). In 2009 the method was extended to VMAT treatments [23]. Pre-treatment verification is performed only for high dose, single-fraction treatments, when large field sizes are involved that will damage the EPID electronics or when the EPID would hit the couch during the delivery of a VMAT arc. Between January 2008 and August 2011, over 120,000 portal images of 6,800 patients that had an IMRT treatment, and portal image data of 2,000 VMAT arcs of 575 patients were acquired and analyzed. All these analyses were performed manually using our software. For that purpose, two part-time radiotherapy technicians (~1.5FTEs) were dedicated to the production of "dosimetry report" in which the results of the verification are presented for quick inspection. In this inspection process, the clinical relevance of the detected deviations (i.e. verification results outside clinical criteria) must be determined. The major impediment of the method has been the amount of manual work required to produce and inspect the *in vivo* dosimetry reports (a time-consuming and labor-intensive process) and to manage the clinical workflow.

An ideal *in vivo* EPID dosimetry solution should be able to verify all delivered fractions automatically (i.e. without human interaction). This involves the following challenges:

1. automatic tools that decide which fractions need to be analyzed for which treatments;

2. automatic acquisition of portal image data for these fractions;

3. automatic production of dosimetry reports;

4. automatically raising alerts and scheduling actions when deviations outside tolerance levels are detected.

In August 2011 we started a project to implement such an automated *in vivo* dosimetry solution in our institution. The purpose of this article is to present an overview of the steps that were taken in the process.

Together with the automation work, an ideal *in vivo* EPID dosimetry solution should raise alerts only when clinically relevant delivery errors occur. However, this is not yet the case in our method where most of our currently detected deviations have little (or no) clinical significance and end up being manually approved. This topic will be covered in the discussion section.

2.2 Materials and Methods

2.2.1 Equipment

Nine linear accelerators (SL20i, Elekta, Crawley, UK) are used at our institution: eight equipped with a standard MLC (1 cm leaf width) and one with a Beam Modulator MLC (4 mm leaf width). All are equipped with a PerkinElmer RID 1680 AL5/Elekta iViewGT amorphous silicon EPID. This EPID includes a touch guard and a 1 mm thick copper plate on top of the scintillation layer. On eight out of nine EPIDs an extra 2.5 mm thick copper plate was used as additional build-up material. This modification has little impact on image quality [43]. The sensitive area of the EPID is 41×41 cm^2, and the EPID is situated at 160 cm distance from the linac focus. Photon beams of 6, 10 and 18 MV were used.

2.2.2 EPID commissioning

Our back-projection model reconstructs a 2D dose distribution in an arbitrary plane parallel to the EPID within a patient or a phantom, i.e. it converts pixel values in the EPID to absolute dose values in the reconstruction plane. Our algorithm uses measurement-based input data with parameters that need to be commissioned upfront for each combination of linac, EPID, and beam energy. For more details see Wendling et al. [41].

2.2.3 IMRT and VMAT dose verification

Treatment plans are calculated using Pinnacle (V9.2, Philips Medical Systems, Eindhoven, The Netherlands). All dose distributions are calculated using Pinnacle's adaptive convolution-superposition algorithm. Each treatment is verified with *in vivo* EPID dosimetry and, as a result, there are as many dosimetry reports as delivered treatments. For IMRT verification, the comparison of the EPID-reconstructed and planned dose distributions is done by a 2D γ-evaluation method using a dose difference of 3% of the maximum planned dose and a distance-to-agreement of 3 mm (3%/3 mm). Pass-fail alert criteria are based on the mean γ-value, the near-maximum (1%) γ-value and the percentage of points with $\gamma \leq 1$ within the 20% isodose line of the planned dose distribution in the dose-reconstruction plane [7]. The height of the reconstruction plane is determined by the coordinates of the Dose Reconstruction Point (DRP). Although the coordinates of the DRP are treatment specific, in most cases they correspond to the

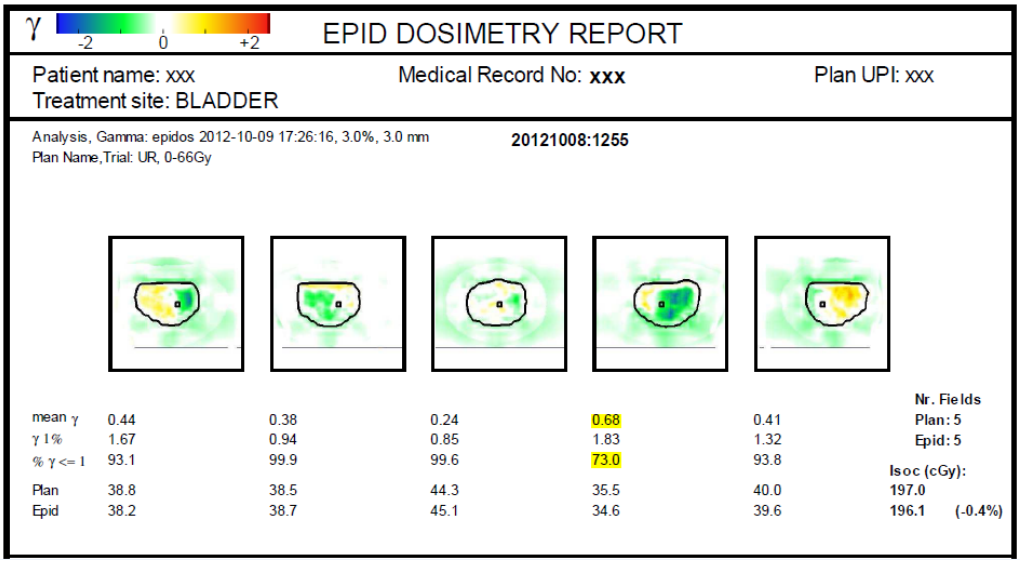


Figure 2.1: *in vivo* EPID dosimetry report showing the results of the verification of one fraction of a 5-field IMRT treatment of a bladder cancer patient. Warning level alerts were raised for the γ-values of one of the five beams (yellow highlighted values).

isocentre. For VMAT verification, a 3D dose reconstruction and a 3D γ-evaluation is performed. γ-statistics are reported within the 50% isodose surface of the planned dose distribution [23]. The difference between the measured and predicted dose at the DRP for the whole fraction (DRP) is also used as an alert criterion. For multi-fraction treatments, the composite weighted average of all fractions, the so-called low-γ average (in between median and minimum) [26] is calculated to separate random from systematic errors. For each beam, a 'signed' gamma image is shown, i.e., the gamma value is multiplied with the sign of the dose difference. The color scale is such that yellow-red indicates overdosage, while green-blue indicates underdosage. A treatment is automatically approved when the clinical criteria are met for the low-γ average, otherwise a medical physicist has to review and manually approve the report. Our method allows two levels of alert criteria: warning (strict) and error (less strict). Each treatment is classified as either 'approved', 'warning' or 'error', depending on the gamma and dose difference results. An example of an *in vivo* IMRT dosimetry report for a fraction of a bladder treatment is shown in Figure 2.1.

2.2.4 Automatic tools to decide which fractions need to be verified

Intelligence is required to decide which fractions need to be analyzed for which treatment. This automation was implemented in a dedicated set of tools; the so-called "*in vivo* EPID dosimetry workspace". The main task of this workspace is to periodically read the daily delivery schedule from the record-and-verify (R&V) system MOSAIQ (IMPAC Medical Systems Inc, Sunnyvale, CA, USA) and automatically determine if the next *in vivo* fraction should be verified. If so, they indicate to the radiotherapy technicians (RTTs) at the linac that portal images (arcs) must be acquired. These tools run in a dedicated server and verify the deliveries at all linacs. Figure 2.2 presents a schematic diagram showing the clinical workflow of our automated *in vivo* EPID dosimetry implemention.

In our department, the vast majority of treatments with curative intent are performed with IMRT or VMAT. The routine clinical procedure so far has been to perform *in vivo* dose verification during the first three fractions for normal fractionation schemes, and all fractions for hypo-fractionated schemes (fraction dose $\geq$ 5 Gy). Currently all treatments are analyzed automatically. For non-IMRT treatments only the first fraction is analyzed. It should be noted that in our institution all non-IMRT treatments are also simulated by means of a CT-scan, thus allowing the use of the same EPID verification software and procedures as developed for IMRT/VMAT treatments.

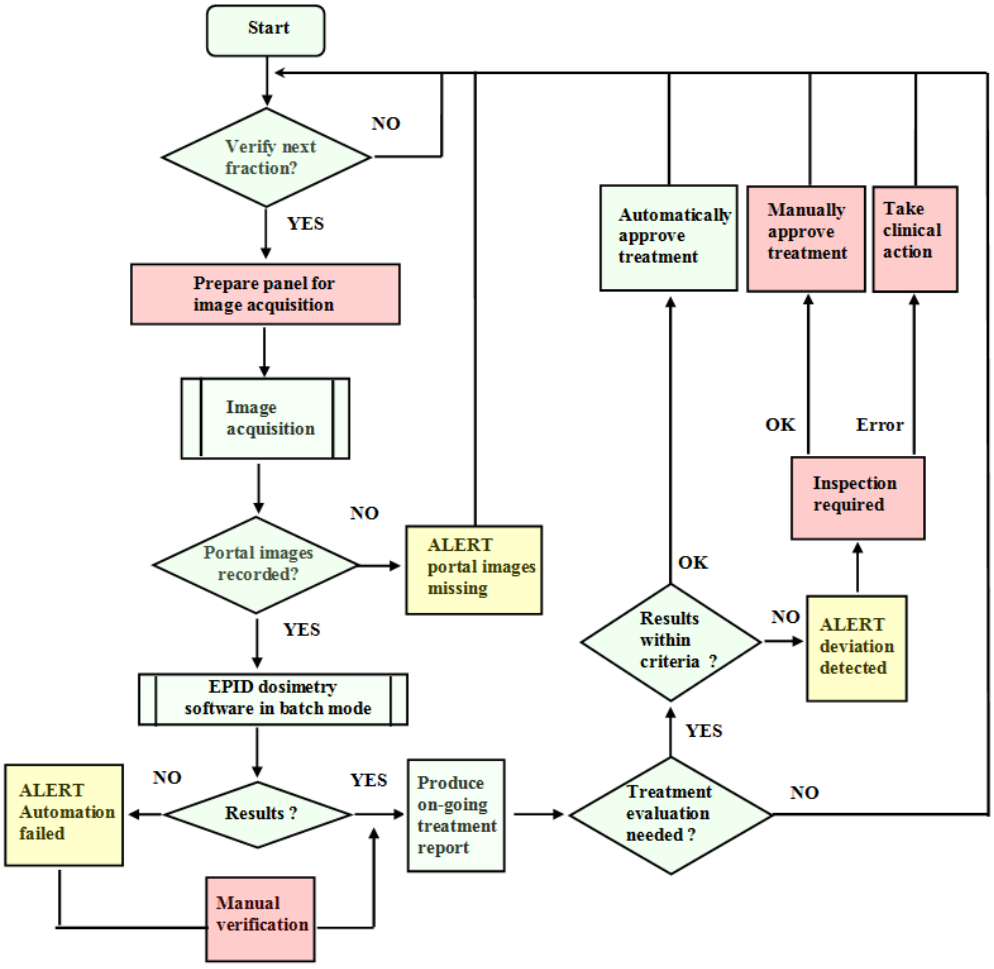


Figure 2.2: Schematic diagram showing the clinical workflow for our automated *in vivo* EPID dosimetry implementation. The EPID dosimetry workspace reads the daily delivery schedule from MOSAIQ and calls the EPID verification software in batch mode as soon as a treatment fraction has been delivered and portal images have been recorded. Results are available a few minutes after delivery of each fraction and alerts are immediately raised when deviations are detected. The light-green boxes in the workflow represent the natural flow of automated events, the light-yellow boxes represent alerts and the light-red boxes represent scheduled actions requiring human intervention. The only remaining manual action in this workflow is "prepare panel for image acquisition". The "image acquisition" sub-process is explained in Section 2.2.5. and the "EPID dosimetry software in batch mode" in Section 2.2.6.

2.2.5 Automatic acquisition of portal image data

EPID images are acquired with in-house developed software. Images are recorded at a resolution of 512×512 pixels and then re-sampled to 256×256 pixels yielding an effective pixel size of 1 mm at the isocentre plane. The detector is acquiring free-running at 2.5 fps without synchronization with the MV pulses. For IMRT, a frame-averaged image and the respective number of frames are stored. However, the acquisition of an accumulated image would not be appropriate for VMAT verification since gantry angle-resolved dosimetric information is essential to perform the back-projection. In this case, the acquisition software saves every detector frame separately (in a 'movie') and the gantry angle position is recorded via an ICOM connection and stored with every recorded frame [23]. The acquisition software uses the couch position, as reported through the ICOM connection, to set the image *type* as 'open' when the portal image was acquired in the absence of a couch.

When the treatment is exported from Pinnacle to DICOM as RTPLAN, extra information is stored for each of its fields (beams or arcs) using Pinnacle scripts. This information is extracted from the DICOM header and then stored in the image acquisition database to allow automatic coupling of each portal image with its corresponding treatment beam in the TPS. A treatment plan is associated with a Unique Plan Identifier (UPI) upon import into MOSAIQ.

The process of image acquisition is now also automated as indicated in Figure 2.3. Through an ICOM connection, the image acquisition software reads the selected field and UPI from the treatment machine. If this field is also found in the acquisition database and the intensity of the panel is higher than a given threshold, then the acquisition software proceeds to record and save the average portal image (or movie in the case of VMAT), using two frames prior to and after the time period that the threshold is exceeded. The following information is associated to each image: acquisition date and time, Patient ID, accelerator name, field ID, number of frames, image type, beam energy, wedge presence, treatment site, Pinnacle plan, trial and beam and several other treatment parameters.

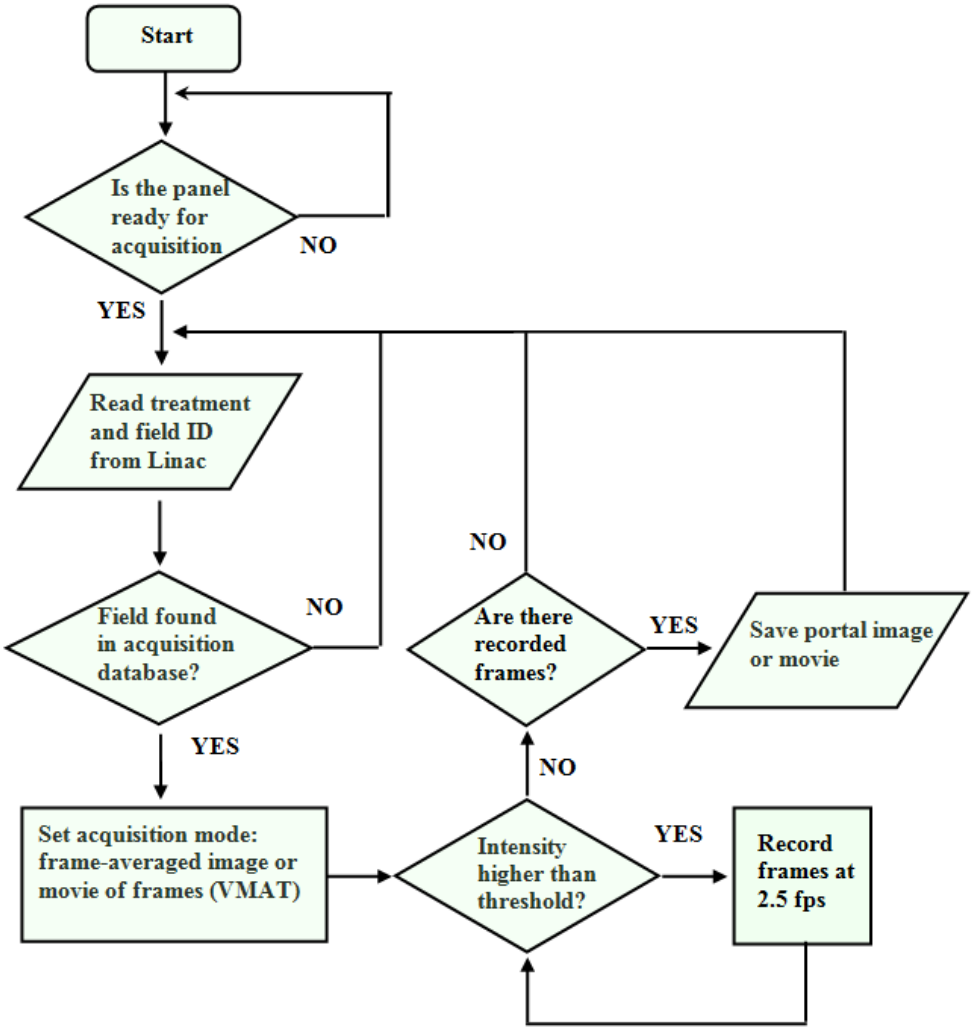


Figure 2.3: Schematic diagram of the automated image acquisition process.

2.2.6 EPID dosimetry in batch mode: automatic production of dosimetry reports

As soon as a treatment has been delivered and portal images recorded, the tools of the *in vivo* EPID workspace call the *in vivo* EPID verification software to generate a report without user interaction. In Figure 2.4, the necessary inputs to reconstruct the dose within the patient are summarized. This section describes the human decisions that were automated and the specific protocols that had to be endorsed in the clinic in order to provide the EPID verification software with automatic access to the above-mentioned inputs.

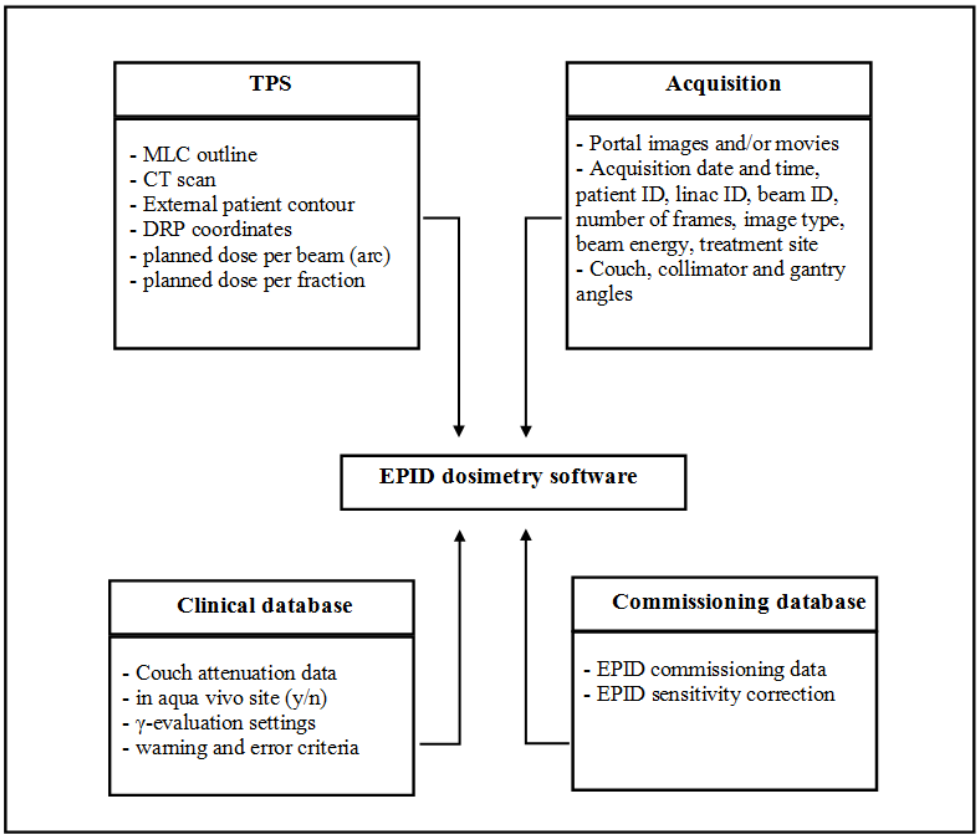


Figure 2.4: Inputs that must be available to the EPID dosimetry software to automatically reconstruct the dose within the patient in batch mode.

For every fraction the software inspects which portal images need to be analyzed.

Our IMRT model was originally designed to use measurements of images without a patient in the beam, 'open' images, in combination with the *in vivo* 'portal' images. In the past, the association of each portal image to its corresponding 'open' image was done manually by the technicians, a tedious task prone to errors. Now the software automates this process by inspecting the image *type* (open or portal), which is stored in the image header by our acquisition software. The problems with 'open' images are threefold: unavoidable extra-workload (~0.3 FTEs), treatment unit time (~1.5 hours per linac per week), and additional management. These difficulties can be overcome with a modification of our original model which removes the need for open images [44]. This modification is currently clinically in use for all treatments.

By using the acquisition information of the portal image (e.g., treatment unit, beam energy, date and time) the dose reconstruction parameters are automatically retrieved from the commissioning database. Within a fraction, images can have different sets of commissioning parameters, e.g. beams with different energies. Within a dosimetry report, diverse fractions can also have different sets of parameters, e.g. when delivered on different linacs. The EPID sensitivity is checked weekly and this information is also unequivocally accounted for in a similar way.

With the information in the acquisition database, the software automatically associates unequivocally a Pinnacle treatment, the accompanying CT-scan, the beam (or arc) and its dose distribution to each portal image or movie. The software uses the CT-scan data to calculate both the geometrical and radiological thickness of the patient and/or phantom. Since the thickness in beams-eye view depends on the gantry and couch angle, it is re-calculated for each reconstruction, i.e. for each IMRT portal image (assuming the gantry angles are different) and for each VMAT movie frame.

The EPID support arm allows for manual displacements in the detector plane to acquire images for off-axis beams. Besides, there is a small displacement due to gravity, the so-called 'flex', which is gantry angle dependent. These displacements determine the absolute EPID spatial location which is essential for an accurate dose reconstruction. In our current clinical practice, however, both shifts are unknown to the software. To overcome this problem, the EPID spatial location is estimated by matching the portal image contour to the beam outline of the field (read from the TPS). This applies to both IMRT portal images and VMAT movie frames.

Our back-projection method uses a water-based model, implying that the parameters are commissioned using homogeneous phantoms. For the *in vivo* verification of treatments involving highly inhomogeneous tissue volumes, the *in aqua vivo* method is

applied [45]. The software uses the acquisition information of the portal image and the clinical setting databases to know whether the *in aqua vivo* correction should be applied for the given treatment site. Currently this method is automatically used for the verification of esophagus, lung and some of the breast treatments. The extra work that this imposes on the planning department (i.e., the creation of a copy of the treatment plan without inhomogeneity corrections) has also been automated by Pinnacle scripts. The auto-association to this *in aqua vivo* plan was achieved by using naming conventions in the Pinnacle scripts.

Currently, mainly for the verification of IMRT breast treatments, other coordinates than those of the isocentre are used to determine more clinically relevant DRP coordinates (the isocentre is often located on a dose gradient or even outside the field). This process has also been automated by following strict naming conventions.

To calculate the thickness map, the patient contour must be obtained from the CT-scan. In the past, a rectangular clip box had to be adjusted manually to include the patient volume and exclude the treatment couch and immobilization devices. Auto-contouring algorithms were applied within this cropped CT volume to estimate the patients contour. However, reliable default values for this clip box are not possible since the height of the couch varies per treatment. Now the software automates this procedure by reading from the TPS the height of the couch and uses this parameter as the base of the clip box to make sure that the couch is excluded. For some treatment sites (head-and-neck, brain and breast) the auto-contouring software can occasionally be misled by the presence of immobilization devices since these can be still recognized as patient contour. For these sites, the external patient contour is delineated in the TPS and our software automatically uses this structure to mask the CT-scan data. Only the data that lies within this structure will then be used to calculate the required thicknesses.

Our back-projection model is commissioned in the absence of the treatment couch. Therefore, the portal dose has to be corrected for the possible extra attenuation when (part of) the beam traverses the couch before reaching the EPID. The software automatically checks in a database the couch type used for each given combination of linac, date, and treatment site and then retrieves the attenuation factor for the photon energy used. A single factor is applied to an image (or movie frame in the case of VMAT).

The gamma parameters and alert criteria for the given treatment site are also automatically retrieved from the clinical settings database. As a result of this automation work, the *in vivo* EPID verification software has all necessary inputs at hand for a complete dose reconstruction and report generation.

2.2.7 Automatically raising alerts and scheduling actions when deviations are detected

When all *in vivo* dosimetry data have been analyzed and the site-specific criteria have been successfully passed, the tools automatically approve the treatment, and the dosimetry report is stored. Otherwise an alert is raised and an action must be scheduled to review the detected deviation. Alerts are also raised if, for some reason, the portal images were not recorded or could not be automatically analyzed, see Figure 2.2.

2.3 Results

2.3.1 Automation success rate

Table 2.1 presents the success rate of the automation work during 2012. A total number of 27,633 portal images of IMRT fields (1873 treatments), 5,551 portal image data of VMAT arcs (1039 treatments), and 2,003 portal images of non-IMRT fields (927 treatments) were automatically analyzed. The *in vivo* EPID dosimetry workspace automatically produced dosimetry reports directly after delivery in 95% of the 6,416 IMRT and/or VMAT fractions that were scheduled for *in vivo* EPID dosimetry. Human intervention was needed for 5% of the fractions due to various types of errors in the automation process. The most common explanation for failures in our automation process arises from incorrect auto-associations with TPS data and/or image acquisition problems. Strict protocols must be followed in all steps of the clinical chain where the automation of the EPID dose verification software can be affected. For instance, there was a drop in success rate for VMAT in the second quarter of 2012 following a period when the *in aqua vivo* plan for lung treatments was not generated in a timely fashion. If we only look at the results of the last three months of 2012 the automation success rates were 98% and 97% for IMRT and VMAT respectively.

2.3.2 Automatic treatment approval rates

Treatments are "automatically" approved by our method when the verification results are within the clinical criteria regardless of whether the dosimetry reports were produced automatically or manually, see Figure 2.2. In 2012, the automatic *in vivo* treatment approval rates were 82% and 63% with the error and warning alert criteria respectively. Consequently, the total manual work after the introduction of our automated solution in

	# treatments	# fractions	# images or # movies	automation rate (%)	within criteria (%) error / warning
IMRT	1,873	3,309	27,633	96	80 / 50
VMAT	1,039	3,107	5,551	95	82 / 62
Non-IMRT	927	927	2,003	N.A.	87 / 85
Total	3,839	7,343	35,187	95	82 / 63

Table 2.1: Automation success rate and in vivo treatment approval rates during 2012. The automation success rate for IMRT prior to June 2012 is unknown because before that date some of the fraction reports were still produced manually. The automation success rate for non-IMRT treatments during 2012 is also unknown. The automatic treatment approval rate is presented as the percentage of treatments with verification results within the clinical tolerance levels.

2012 amounted to the production of reports for 5% of the treatments and the inspection of deviations for 18% and 37% of the treatments respectively.

2.4 Discussion

2.4.1 Automatic analysis as an essential tool for large-scale implementation of *in vivo* EPID dosimetry

Widespread implementation of *in vivo* EPID dosimetry has been so far hampered by the workload involved. To help minimize the related workload, the routine use of a large-scale clinical implementation of 2D or 3D *in vivo* dose verification should include fully automated data acquisition and data analysis. Reduction of workload is not the only advantage of an automated environment. A system that requires a large amount of manual input for the preparation of the dosimetry reports is error prone. Erroneous results in dosimetry reports are highly undesirable because they may lead to wrong conclusions. A solution that automatically provides the correct input data to the dose verification software greatly improves the reliability of the results. In our current *in vivo* EPID dosimetry workspace, the clinical goal to reach automatic success rates close to 100% is within reach. In our institution, about 3,000 IMRT/VMAT and about 2,500 non-IMRT treatments are expected to be delivered per year. With our existing protocols

(three *in vivo* measurements for IMRT/VMAT treatments and one *in vivo* measurement for non-IMRT treatments), and with an assumed automation success rate of 97%, about 350 fractions per year (one per day) still would require manual intervention to generate the dosimetry report. The estimated time needed to solve the problem that caused the failure and generate the dosimetry report is 5-10 minutes. The workload for the dosimetry report production work has become thus effectively 0 FTEs. Prior to the introduction of this automated solution, this workload amounted to 1.5 FTEs.

Automatic analysis may also allow a more frequent use of *in vivo* dose verification, for instance for fractions at which patient set-up imaging is performed or even for all fractions. The additionally obtained information can be used to examine inter-fractional trends and identify the causes of dose delivery differences, e.g. see Persoon et al. [46].

2.4.2 Reduce the number of clinically non-relevant deviations

The ideal automatic *in vivo* dose verification solution should raise alerts only when clinically relevant delivery errors occur. This is not yet the case in our method where deviations can be classified into three categories during the inspection process:

1. Limitations of our calculation model: The reconstructed EPID-based dose distribution was not accurate enough due to limitations in our current calculation model (such as tissue inhomogeneities, presence of cavities, insufficient accuracy in accounting for the extra attenuation caused by the couch and other immobilization devices, patient setup deviations, or dose-rate effects) or due to other sources of errors (such as bad EPID pixel quality, or wrongly acquired images).

2. Patient anatomy changes: These alerts cannot be explained by any of the reasons before-mentioned and the detected deviations are real. The presence of gas pockets, weight gain or loss and (recovery from) atelectasis are examples of anatomical changes leading to such alerts. The clinical relevance depends on the location of the changes relative to the target.

3. Other alerts: Delivery or planning errors; for instance errors due to output deviations of the linac, errors in the data transfer, or errors due to limitations of the dose calculation of the TPS.

The tasks of the inspection work are to filter out alerts due to model limitations from all other alerts, and to evaluate the clinical relevance of the detected real deviations. This

inspection work is far from straightforward and demands well-trained staff with thorough knowledge of both EPID dosimetry and clinical practice. This inspection work currently amounts to 0.25 FTEs. The reviewer decides for each alert if (and what) further action is required. For instance, when the observed deviation from an *in vivo* verification cannot be explained, the reviewer may request an extra phantom measurement to verify the treatment plan. The number of extra phantom measurements in 2012 was 106 cases (2.5% of cases). Clinical actions as a consequence of this *in vivo* verification work took place in 14 cases (0.4% of cases) which agrees well with previous clinical studies, see Mans et al. [7].

With our estimations of about 3,000 IMRT/VMAT and about 2,500 non-IMRT treatments to be delivered per year, our current alert criteria and the current approval rates (see section 2.3.2), around 900/1500 treatments are expected to show error/warning level deviations demanding expert inspection (~4/6 treatments per working day). Because most of these are found to be clinically irrelevant after the inspection work, efforts must still be spent to reduce the number of these falsely detected deviations. Model enhancement will be an important part of this effort, whereby the number of observed warnings per-site allows us to prioritize model development work.

The number of automatically approved treatments depends on the chosen set of clinical alert criteria. The stricter the criteria the larger the amount of deviations detected that demand inspection. For instance, if we were to have used error/warning criteria of 10% in %ΔDRP, as opposed to the current 5%/3%, then 98% of the total number of delivered fractions in 2012 would have been automatically approved. Deciding the per-site alert criteria is far from easy for *in vivo* dosimetry where, as opposed to pre-treatment verification, the aforementioned method limitations and anatomical changes play a significant role. A full discussion about alert criteria levels and their correlation with clinically relevant errors is outside the scope of this article.

2.5 Conclusions

In vivo EPID dosimetry has been fully automated and integrated in our clinical workflow where currently about 3,000 IMRT/VMAT and about 2,500 non-IMRT treatments are verified each year. This automation work removes the workload to generate *in vivo* dosimetry reports after delivery facilitating a large-scale clinical implementation of *in vivo* EPID dosimetry. Furthermore, an automated solution guarantees the timely production of dosimetry reports and greatly improves the reliability of the results.

The majority of the treatments are approved automatically directly after delivery, i.e. without any human intervention. Alerts are automatically raised and actions scheduled otherwise. The inspection of these deviations is currently the only remaining workload.

Because most of the detected deviations are found to be clinically irrelevant after the inspection work, efforts must still be spent to automatically reduce the number of these "falsely" detected deviations.

These newly developed tools are a major step forward towards full integration of *in vivo* EPID dosimetry in radiation oncology practice and open up the road to automatic *in vivo* verification of all fractions of IMRT and VMAT treatments.

Chapter 3

In vivo portal dosimetry for head-and-neck VMAT and lung IMRT: linking γ-analysis with differences in dose-volume histograms

R.A. Rozendaal
B. Mijnheer
M. van Herk
A. Mans

Radiotherapy and Oncology 112 (396), 2014

Abstract

Purpose

To relate the results of γ-analysis and dose-volume histogram (DVH) analysis for detecting dose deviations with *in vivo* dosimetry for two treatment sites.

Methods and materials

In vivo 3D dose distributions were reconstructed for 722 fractions of 200 head-and-neck (H&N) VMAT treatments and 183 fractions of 61 lung IMRT plans. The reconstructed and planned dose distributions in the PTV were compared using (a) the γ-distribution and (b) the differences in D_2, D_{50} and D_{98} between the two dose distributions. Using pre-defined tolerance levels, all fractions were classified as deviating or not deviating by both methods. The mutual agreement, the sensitivity and the specificity of the two methods were compared.

Results

For lung IMRT, the classification of the fractions was nearly identical for γ- and DVH-analysis (94% agreement) and the sensitivity and specificity were comparable for both methods. Less agreement (80%) was found for H&N VMAT, while γ-analysis was both less sensitive and less specific.

Conclusions

DVH- and γ-analysis perform nearly equal in finding dose deviations for lung IMRT treatments; for H&N VMAT treatments, DVH-analysis is preferable. As a result of this study, a smooth transition to using DVH-analysis clinically for detecting *in vivo* dose deviations is within reach.

3.1 Introduction

In vivo dosimetry can play an important role within the complex workflow of modern external photon beam radiotherapy by providing an efficient patient-specific end-to-end check [42]. Our clinic has introduced *in vivo* EPID dosimetry in clinical routine in 2008 for IMRT treatments [7], and in 2009 for VMAT treatments [23]. Using in-house developed software, *in vivo* EPID dosimetry is performed automatically for almost all treatments. Currently, about 90 new treatments are verified using *in vivo* EPID dosimetry each week in our department.

With *in vivo* EPID dosimetry, a dose distribution is reconstructed in the planning CT-scan using dose information (i.e., EPID images) obtained during treatment, which is subsequently compared to the planned dose distribution; usually with 2D or 3D γ-analysis [26, 47, 24]. The result of each comparison then needs to be interpreted, i.e. one needs criteria to determine whether a deviation was significant, and if so, whether this deviation is clinically relevant. In other words, would the treatment outcome, either in terms of local control or in terms of toxicity, be affected by the underlying cause of the observed difference? This issue is at the heart of *in vivo* dosimetry, as we care mostly about deviations which are clinically relevant. However, within our current verification system it is not easy to determine the relevance of detected deviations. The main reason for this is that γ-analysis provides a metric on the similarity of two dose distributions, which is a fundamentally different concept than determining whether a difference in dose distribution is clinically acceptable. To determine the relevance of an observed deviation, we revert to the experience of medical physicists to interpret the dose differences. Additionally, several studies have shown that γ-analysis using the γ-passrate has insufficient predictive power for pre-treatment IMRT and VMAT dose verification [48, 29, 31, 32, 30, 49]; this might also hold for *in vivo* dose verification. For these reasons we want to change our current dose evaluation protocol to a system where deviations are readily understandable by both clinicians and physicists, and have a clear relation with expected outcome data.

Because reconstructed 3D dose distributions are available, an obvious choice would be inspecting the corresponding DVHs. DVH-parameters derived from these DVHs would give insight in the relevance of a deviation for a specific patient and allow for an easier interpretation of an observed deviation. It is the purpose of this work to evaluate the differences in detected deviations by these two methods. In order to perform this evaluation, tolerance levels must be set for both methods of analysis. The relation be-

tween γ-analysis and DVH-analysis is assumed to be treatment-site specific; in this study, head-and-neck and lung cancer treatments are investigated.

3.2 Materials and methods

3.2.1 *In vivo* EPID dosimetry

For this study, 200 head-and-neck VMAT plans and 61 lung IMRT treatment plans of 202 patients treated in 2011 and 2012 were randomly selected. Treatment plans with field sizes too large to be imaged by the EPID or with couch positions prohibiting the EPID from being deployed were excluded, as no *in vivo* data was available.

All treatment plans were created using the Pinnacle treatment planning system (TPS) v9.0 or v9.2 (Philips Medical Systems, Eindhoven, The Netherlands) and consisted of a dual-arc setup (H&N VMAT) or 5–10 step-and-shoot beams (lung IMRT). Of the head-and-neck treatment plans, 186 were composed of two schemes: a regular scheme, with a total dose ranging from 24–46 Gy, and a boost scheme, with a total dose ranging from 6–24 Gy. Fraction doses ranged from 2–6 Gy. The remaining 14 head-and-neck treatment plans were simultaneous integrated boost (SIB) plans with a total maximum dose of 70 Gy given in fractions of 2 Gy. The lung treatment plans had a total dose ranging from 8-66 Gy and fraction doses ranging from 1.5–4 Gy.

All fractions were verified using 3D *in vivo* EPID dosimetry; the 3D delivered dose distribution is reconstructed using the recorded EPID images (continuous-mode for VMAT, integrated images per beam for IMRT) and the planning CT-scan, as described in [23, 22]. Dose reconstruction was done for each treatment arc or IMRT beam separately. After the reconstruction of the 3D *in vivo* dose distribution per arc (beam), the separate dose distributions are summed to yield the total fraction dose. Note that errors present in only one of the arcs (beams) of a fraction are not separately considered, which is motivated by our wish to verify the total dose received by the patient. The recorded EPID images from 722 H&N VMAT and 183 lung IMRT fractions belonging to these treatments were used.

The accuracy of the dose reconstruction method has been verified extensively by film and ionization chamber (IC) measurements [26, 41]. Comparison to IC measurements is also employed regularly by comparing both planned (TPS) and reconstructed (EPID) dose distributions to dose measurements performed using the PTW Octavius phantom and 729-IC-array (PTW, Freiburg, Germany).

The original back-projection algorithm as presented in Wendling et al. [22] is water-based, i.e., inhomogeneities are not taken into account. Therefore, the lung IMRT dose reconstructions were performed by using the *in aqua vivo* extension of the algorithm [45].

3.2.2 Reconstructed dose evaluation

Evaluation of the reconstructed dose distribution was done by comparing the reconstructed dose in the PTV with the planned dose in the PTV by both 3D γ-analysis and DVH-analysis. For head-and-neck SIB plans, evaluation was done, once for the boost-region (PTV_{boost}) and once for the non-boost-region of the PTV (PTV - PTV_{boost}). The DVH-analysis consists of calculating the difference between the reconstructed and planned value of D_2, D_{50} and D_{98}. These differences are denoted as ΔD_2, ΔD_{50} and ΔD_{98}; they are referred to as the DVH-analysis indicators. A reconstructed dose distribution is classified as having a deviation when at least one of these indicators is outside of its tolerance level (each indicator has its own tolerance level).

The γ-analysis method consists of calculating the 3D γ distribution followed by quantifiying this γ distribution by two indicators, the mean value ("γ-mean") and the 99th percentile ("γ-1%"). The calculation of the γ distribution is done with 3%/3 mm settings; the 3% value is relative to the maximum planned dose in the volume over which the γ-analysis is carried out (i.e., PTV, PTV_{boost} or PTV - PTV_{boost}). Similar to DVH-analysis, a deviation in the reconstructed dose distribution is defined to occur whenever when at least one indicator is outside tolerance.

As additional check the DVHs of the CTV or GTV were also compared in terms of ΔD_2, ΔD_{50} and ΔD_{98}. Not all treatment plans had the CTV or GTV structure separately available for automated analysis, the $\Delta D_2^{CTV/GTV}$, $\Delta D_{50}^{CTV/GTV}$ and$\Delta D_{98}^{CTV/GTV}$ could be determined for 140 H&N VMAT plans and for 60 lung IMRT plans.

3.2.3 Tolerance levels

Based on recommendations from ICRU Report 83 [50], tolerance levels for ΔD_{50} are set to $\pm 3.5\%$ with respect to the planned dose. As the DVH will be degraded by geometrical uncertainties and methodological inaccuracies, a larger tolerance level is expected to be needed for ΔD_2 and ΔD_{98}. As a starting point, the tolerance levels for ΔD_2 and ΔD_{98} will be set to twice those for ΔD_{50}, i.e., $\pm$ 7%, as an upper limit for clinically acceptable deviations. Finally, the tolerance levels will be adjusted by any systematic deviation present in ΔD_2, ΔD_{50} and ΔD_{98}: systematic deviations represent fundamental

and/or methodological limitations, which need to be addressed separately; the systematic deviation should not lead to an over- or underestimation of "true" deviations.

For γ-analysis, a first estimate of the tolerance level for γ-mean would be 3.5% / 3% $\approx$ 1.2 – the first-order estimate of the effect of an average 3.5% dose deviation (cf. ΔD_{50}) on γ-analysis. This value is slightly too relaxed, however, as the γ-analysis is carried out with a global dose criterion – although the dose distribution in the PTV is very uniform. Also, the 3 mm distance-to-agreement criterion will lower the γ-values. Therefore, a value of 1.0 is taken as the tolerance level for γ-mean. An estimate of the tolerance level for γ-1% is not so easily deduced from ΔD_2 and ΔD_{98} tolerance levels; based on our clinical experience [7, 26], a value of 3 is taken. This implies that no more than 1% of the voxels can have deviations larger than 9%/9 mm.

3.2.4 Agreement, sensitivity and specificity

The sensitivity and specificity of both dose evaluation methods can be determined by using a reference "true" classification of the analyzed fractions. As only the tolerance level for ΔD_{50} is related to an international recommendation, the classification of fractions by ΔD_{50} alone is taken as reference. This implies that (a) no false negatives will occur for DVH-analysis and (b) all deviations found by ΔD_2 and/or by ΔD_{98}, but not by ΔD_{50}, will be treated as false positives. Note that, with no false negatives present, the sensitivity of DVH-analysis is by definition equal to one. Finally, the agreement between the two methods is expressed as the percentage of fractions which has the same classification by both methods.

3.3 Results

3.3.1 Analysis indicator distributions

The results of the dose evaluation of the 905 analysed fractions are summarized in Figures 3.1 and 3.2. Figure 3.1 shows the distributions of γ-1% and γ-mean for head-and-neck and lung treatments. Clearly, better agreement is obtained between planned and reconstructed dose values for lung IMRT treatments than for H&N VMAT treatments. This is to be expected, due to several factors: (a) for VMAT, recorded EPID-frames are grouped ("binned") each 4-degrees, which will smear-out the reconstructed dose, (b) ghosting effects of the EPID are neglected which are more pronounced for VMAT than for IMRT, (c) the dose distribution is more complex for H&N VMAT treatments, (d)

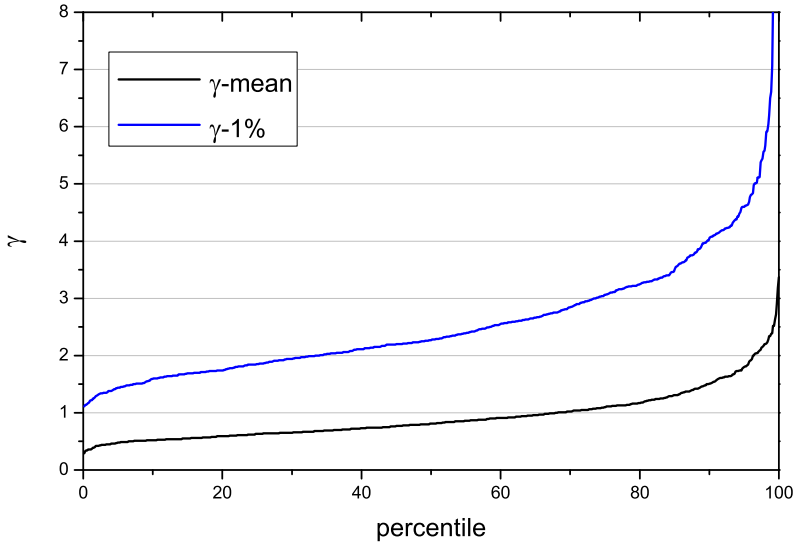
Dataset	Indicator	Median	Mean	Standard deviation	Tolerance level	Adjusted tolerance level
H&N VMAT	γ-mean	0.8	0.9	0.4	1.0	1.0
	γ-1%	2.3	2.7	1.7	3.0	3.0
	ΔD_2	2.7%	2.9%	4%	±7.0%	-4% 10%
	ΔD_{50}	-1.0%	-0.9%	3%	±3.5%	-4% 3%
	ΔD_{98}	-4.7%	-5.0%	6%	±7.0%	-12% 2%
Lung IMRT	γ-mean	0.5	0.7	0.8	1.0	1.0
	γ-1%	1.1	1.5	1.2	3.0	3.0
	ΔD_2	-0.7%	-0.9%	5%	±7.0%	-8% 6%
	ΔD_{50}	-0.7%	-0.9%	5%	±3.5%	-4% 3%
	ΔD_{98}	-1.5%	-1.8%	5%	±7.0%	-8% 6%

Table 3.1: Statistics (calculated for the PTV) and tolerance levels for the DVH-indicators ΔD_2, ΔD_{50}, ΔD_{98} and the γ-indicators γ-mean, γ-1%. $N = 722$ for H&N VMAT and $N = 183$ for lung IMRT. The adjusted tolerance levels for the DVH-indicators are equal to the specified tolerance level shifted by the systematic deviation for each indicator. The median value is taken as the systematic deviation. γ-indicator tolerance levels are not adjusted.

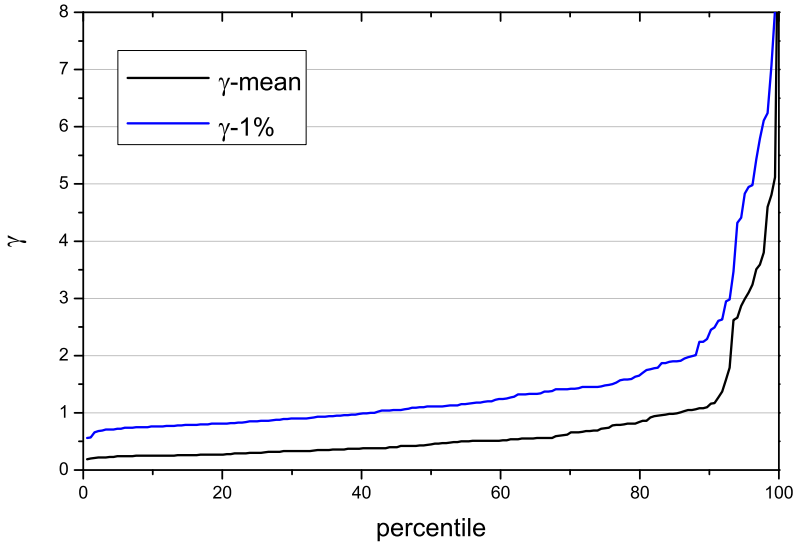
inhomogeneities are not accounted for in our model, but corrected for by using the *in aqua vivo* method for lung IMRT, but not for H&N VMAT. Figure 3.2 shows the distributions of ΔD_2, ΔD_{50} and ΔD_{98}. Again, results are better for lung IMRT treatments. Especially the distributions of ΔD_2 and ΔD_{98} for H&N VMAT treatments show large systematic deviations; the other distributions (ΔD_{50} for VMAT and ΔD_2, ΔD_{50} and ΔD_{98} for IMRT) show smaller systematic deviations. The median, average and standard deviation values of the DVH indicators are shown in Table 3.1. Table 3.2 shows these values for $\Delta D_2^{\mathrm{CTV/GTV}}$, $\Delta D_{50}^{\mathrm{CTV/GTV}}$ and $\Delta D_{98}^{\mathrm{CTV/GTV}}$.

3.3.2 Fraction classification

The fractions were classified using the tolerance levels specified. For DVH-analysis, the tolerance levels are adjusted for the observed systematic deviations as follows: the range is kept, but the center is shifted with the systematic deviation. Table 3.1 shows these tolerance levels adjusted for systematic deviations; the percentage of fractions outside tolerance levels (positives) is shown in Table 3.3. Figure 3.3 shows the classification

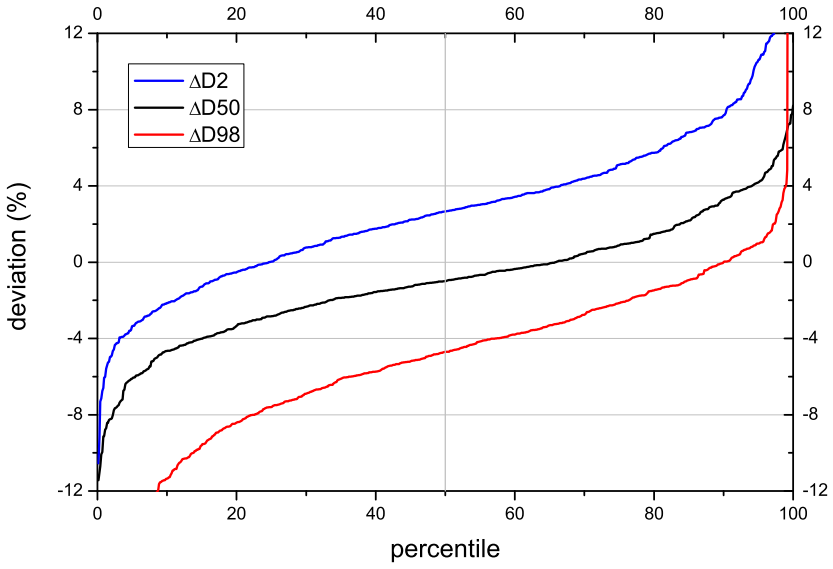


(a)

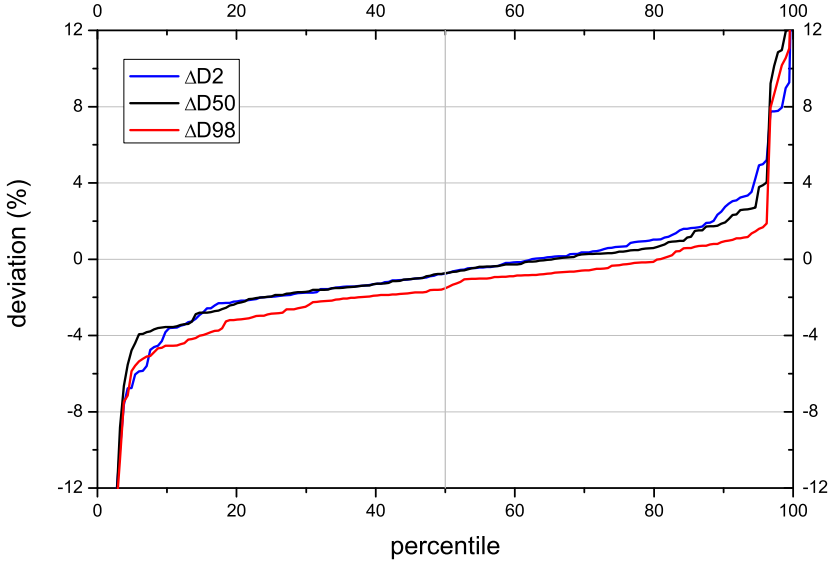


(b)

Figure 3.1: Distribution of γ-mean and γ-1% values by percentile, as obtained from 3D in vivo dose verification analysis of (a) 722 fractions of 200 clinical H&N VMAT treatment plans and (b) 183 fractions of 61 clinical lung IMRT treatment plans. The tolerance levels are set at γ-mean = 1 and γ-1% = 3.



(a)



(b)

Figure 3.2: DVH-indicator distributions by percentile for (a) H&N VMAT treatments and (b) lung IMRT treatments. The dashed, solid and dotted lines show the values of ΔD_2, ΔD_{50} and ΔD_{98}, respectively, as obtained for the 722 (H&N VMAT) and 183 (lung IMRT) analyzed fractions.

Dataset	Indicator	Median	Mean	Standard deviation
H&N VMAT	$\Delta D_2^{CTV/GTV}$	2.7%	2.6%	4%
	$\Delta D_{50}^{CTV/GTV}$	-0.3%	-0.6%	3%
	$\Delta D_{98}^{CTV/GTV}$	-1.8%	-2.5%	4%
Lung IMRT	$\Delta D_2^{CTV/GTV}$	-0.7%	-0.4%	5%
	$\Delta D_{50}^{CTV/GTV}$	-0.4%	-0.4%	5%
	$\Delta D_{98}^{CTV/GTV}$	-0.9%	-0.7%	6%

Table 3.2: Statistics for DVH-indicators calculated for the CTV ($N = 489$ for H&N VMAT, $N = 6$ for lung IMRT) or the GTV ($N = 0$ for H&N VMAT, $N = 92$ for lung IMRT). The CTV was used if present, otherwise the GTV if present. Results shown are for both the CTV and GTV results combined.

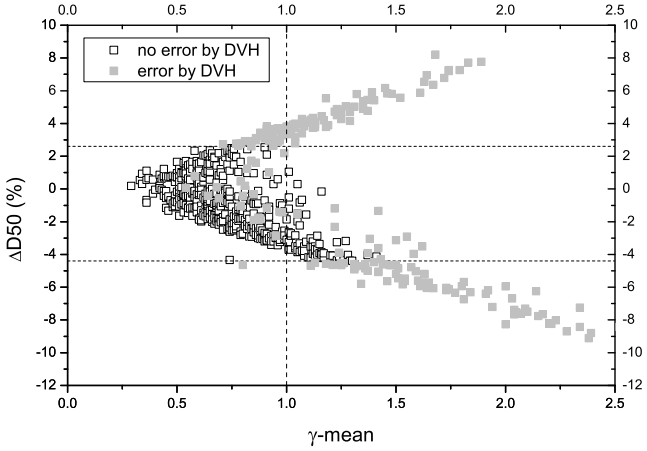
	Method	Analysis	# Negatives		# Positives			
Dataset	Agreement	Method	true	false	true	false	Sensitivity	Specificity
H&N VMAT	80%	γ	409 (57%)	36 (5%)	145 (20%)	132 (18%)	0.80	0.76
		DVH	506 (70%)	0	181 (25%)	35 (5%)	1	0.94
Lung IMRT	94%	γ	155 (85%)	2 (1%)	17 (9%)	9 (5%)	0.89	0.95
		DVH	164 (90%)	0	19 (10%)	0	1	1

Table 3.3: Agreement, classification rank (true/false, positive/negative), sensitivity and specificity for both γ- and DVH-analysis. The reference set for determining the classification was based on ΔD_{50} analysis: all true negatives are considered those fractions with ΔD_{50} inside tolerance levels, all true positives are considered those fractions with ΔD_{50} outside tolerance levels.

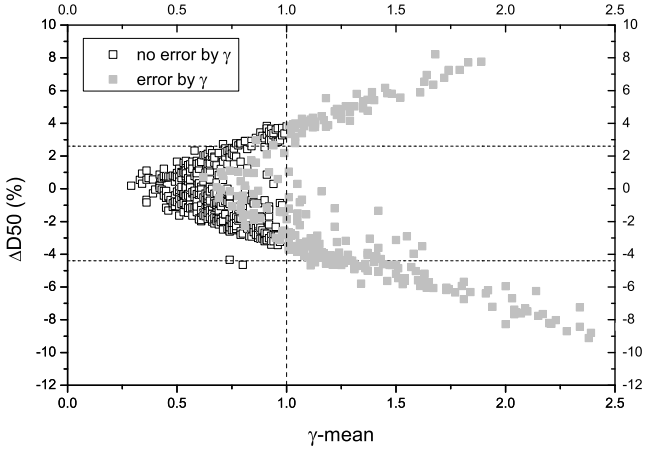
obtained by both methods and for both data sets as a function of ΔD_{50} and γ-mean. Indicated in this figure are the ΔD_{50} tolerance levels; points outside these tolerance levels are therefore those defining the reference set of all true positives.

3.3.3 Agreement, sensitivity and specificity

Table 3.3 shows the number of true negatives, false negatives, true positives and false positives for each analysis method. The sensitivity and specificity of each method is indicated as well. For H&N VMAT, analysis by DVH-indicators is clearly more specific than by γ-analysis. By contrast, DVH- and γ-indicator specificity is quite similar for lung IMRT. With the ΔD_{50} indicator as reference, the amount of false negatives is 8% for

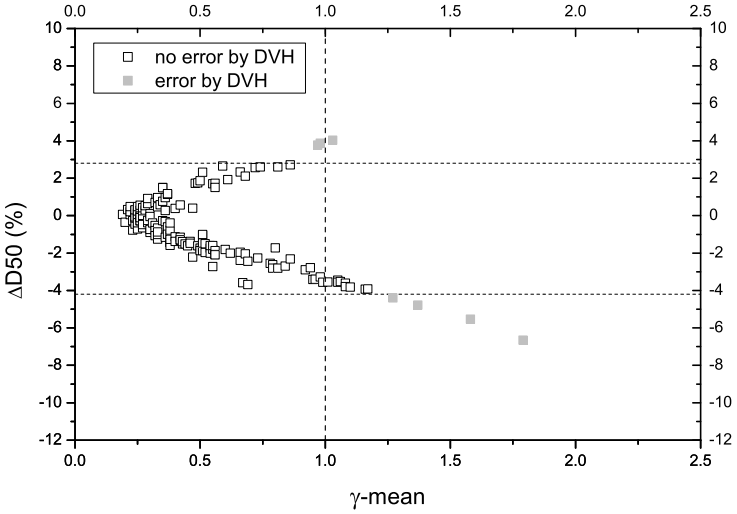


(a)

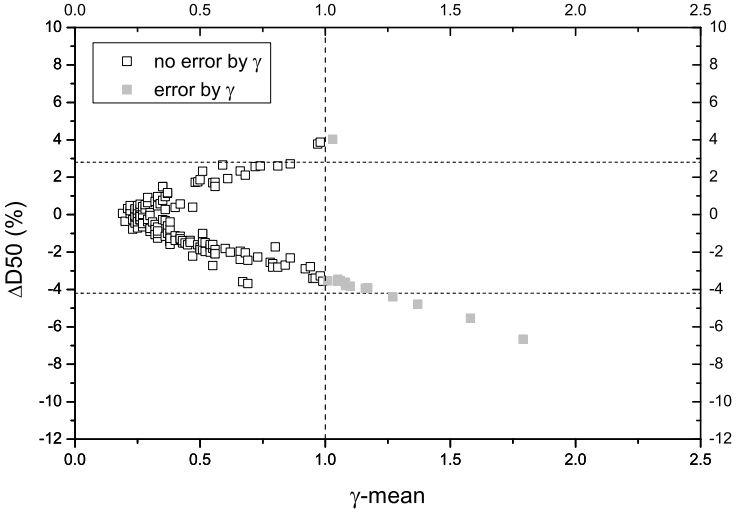


(b)

Figure 3.3: Relation between γ-mean and ΔD_{50} for analyzed H&N VMAT fractions ((a) and (b)) and lung IMRT fractions ((c) and (d)), with error/no error indication according to all DVH-indicators ((a) and (c)) or both γ-indicators ((b) and (d)). The horizontal dashed lines show the tolerance levels for the ΔD_{50} indicator; the vertical dashed line indicates the tolerance level for the γ-mean indicator. Filled symbols between the horizontal dashed lines in (a) and (c) show where the other DVH-indicators were outside tolerance levels; filled symbols left from the vertical dashed line in (b) and (d) show where γ-1% was outside tolerance level.



(c)



(d)

Figure 3.3: Relation between γ-mean and ΔD_{50} (cont.).

H&N VMAT by γ-analysis. For the lung IMRT data-set, the number of false negatives is less than 1%. The agreement between the two methods of analysis is 80% for H&N VMAT and 94% for lung IMRT.

3.4 Discussion

The observed DVH-indicator distributions show systematic deviations for the DVH-indicators ΔD_2, ΔD_{50} and ΔD_{98}, especially in case of H&N VMAT treatments. The largest contribution to these systematic deviations is intrinsic to the concept of using a PTV: as the dose falls of quickly outside the PTV, any random or systematic error in patient positioning, anatomy or dose reconstruction will lead to an underdosage and a decrease in PTV coverage. This effect is well known [51] and it is the reason why margins are used to create the PTV. The validity of this explanation can be seen by comparing DVHs of the planned and reconstructed dose in the CTV or GTV. As these are planned with a margin, they are embedded in a high-dose region which means that the effect of random and systematic errors should have very little influence on CTV/GTV coverage. Therefore, also $\Delta D_2^{\mathrm{CTV/GTV}}$, $\Delta D_{50}^{\mathrm{CTV/GTV}}$ and $\Delta D_{98}^{\mathrm{CTV/GTV}}$ were computed; the results are shown in Table 3.2. As expected, the systematic deviation in $\Delta D_{98}^{\mathrm{CTV/GTV}}$ is either no longer present (lung IMRT) or drastically reduced (H&N VMAT), indicating that the systematic deviation in ΔD_{98} is indeed caused by this intrinsic effect of reduced PTV coverage. Interestingly, for H&N VMAT treatments the effect of reduced PTV coverage is also seen in ΔD_{50}. ΔD_2 is not affected by this effect, as expected.

The remaining systematic deviations, the magnitude of which are well estimated by the systematic deviations in $\Delta D_2^{\mathrm{CTV/GTV}}$, $\Delta D_{50}^{\mathrm{CTV/GTV}}$ and $\Delta D_{98}^{\mathrm{CTV/\ GTV}}$, are explained by examining the effect of geometrical uncertainties (patient setup and anatomical changes) and methodological inaccuracies (due to ghosting, frame binning and/or inhomogeneities) on the reconstructed dose DVHs. Since the ΔD_{50} distribution is central and its average value is close to zero, we conclude that errors made in the dose reconstruction can be treated as random; random errors will cause a degradation of the planned DVH. As such, the systematic deviations seen in ΔD_{98} and ΔD_2 might partly be seen as a measure of how much uncertainties and inaccuracies are typically present for a particular treatment site.

In our experience, anatomical changes are the most common cause of both clinically relevant as well as irrelevant observed *in vivo* dose deviations [7]. It is therefore important to note that the planning CT-scan is used in the dose reconstruction algorithm,

which means that the larger the anatomical change, the less accurate the reconstructed dose distribution is – in general overestimating the deviation [44]. Although this lower accuracy is unrelated to the method of deviation detection, i.e., γ- or DVH-analysis, the implications for interpretation might be different as DVH-analysis states differences in dose, whereas γ-analysis does not. So, a review of observed deviations by a trained medical physicist will remain important for a correct interpretation of observed deviations.

This study is limited to verifying dose in the PTV, but a clinical implementation will need to include verification of dose in OARs and in the remaining volume – this is currently under investigation. An important remark in considering dose to OARs is that IGRT and *in vivo* dosimetry serve complementary goals: in our view, IGRT is used to verify the patient position and *in vivo* dosimetry is used to verify the dose delivery.

The reference used for determining what constitutes a "true" deviation – ΔD_{50} tolerance levels as specified in ICRU Report 83 – is a reasonable starting point for a reference, but will not detect all hot- or cold-spots in the PTV. The fractions ranked as "false positive" by DVH-analysis might thus be true positives, as apparently a hot- or cold-spot is present which does not affect the median dose to the PTV by more than $\pm 3.5\%$. In view of the scarce literature on 3D *in vivo* dosimetry, more work is urgently needed to define reference values for deviations observed during *in vivo* dosimetry. Key ingredients for such a recommendation would be tolerance levels for several indicators (such as ΔD_2, ΔD_{50} and ΔD_{98}) for the PTV, inclusion of the effects of reduced PTV coverage (the magnitude of which is related to the margins needed for obtaining the PTV), tolerance levels and indicators for dose distributions in OARs and for the remaining volume.

Table 3.3 shows that there is reasonable to good agreement between the two methods of analysis – indicating that a smooth transition in clinical practice from γ-analysis to DVH-analysis for 3D *in vivo* dosimetry is possible. It should be noted that the DVH-indicator tolerance levels are easily adapted to correct for the observed systematic deviations, but that this is not the case for γ-analysis – although the same deviations in dose are present. As such, it is not surprising that DVH-analysis outperforms γ-analysis especially for H&N VMAT, where the systematic deviations in ΔD_2, ΔD_{50} and ΔD_{98} are most pronounced. This illustrates the advantage of using DVH-analysis: the effect of missing PTV coverage is both easier seen and easier understood by using DVH-analysis than by using γ-analysis.

From a practical point of view, DVH-analysis is preferable because the systematic deviations (which are for the largest part intrinsic) are easily accounted for, in contrast to γ-analysis. Moreover, DVH-analysis finds less false positives, which has a clear clinical

advantage by reducing the number of treatments the medical physicist has to inspect. From Figure 3.3a it can be seen that the number of false negatives for H&N VMAT fractions found by γ-analysis can be reduced by lowering the tolerance level for γ-mean. However, it can also be seen from Figure 3.3a that this will lead to a steep increase in false positives. Finally, in contrast to γ-analysis, DVH-analysis of the PTV allows for asymmetrical tolerance levels. These might be of great interest when a slight underdosage of the PTV is considered more harmful than a slight overdosage. The use of asymmetrical tolerance levels will further increase the specificity of DVH-analysis.

3.5 Conclusions

In terms of detecting dose deviations DVH- and γ-analysis are comparable, but the results are treatment-site dependent. With the recommendations of ICRU Report 83 as a reference, γ-analysis is less sensitive (5% more false negatives) and less specific (15% more false positives) than DVH-analysis for head-and-neck VMAT treatments. For lung IMRT, the results are almost identical. DVH-analysis allows for easier interpretation of observed deviations than γ-analysis, and is easier to correct for systematic deviations. As a result of this reasonable (80%, for H&N VMAT) to large (94%, for lung IMRT) agreement, a smooth transition to using DVH-analysis clinically for detecting *in vivo* dose deviations is within reach.

Acknowledgements

The authors wish to thank Igor Olaciregui-Ruiz and Jan-Jakob Sonke for the many fruitful discussions.

Chapter 4

Impact of daily anatomical changes on EPID-based *in vivo* dosimetry of VMAT treatments of head-and-neck cancer

R.A. Rozendaal
B. Mijnheer
O. Hamming-Vrieze
A. Mans
M. van Herk

Radiotherapy and Oncology 116 (1), 2015

Abstract

Background and Purpose

Target dose verification for VMAT treatments of head-and-neck (H&N) cancer using 3D *in vivo* EPID dosimetry is expected to be affected by daily anatomical changes. By including these anatomical changes through cone-beam CT (CBCT) information, the magnitude of this effect is investigated.

Materials and Methods

For 20 VMAT-treated H&N cancer patients, all plan-CTs (pCTs), 633 CBCTs and 1266 EPID movies were used to compare four dose distributions per fraction: treatment planning system (TPS) calculated dose and EPID reconstructed *in vivo* dose, both determined using the pCT and using the CBCT. D_2, D_{50} and D_{98} of the planning target volume (PTV) were determined per dose distribution.

Results

When including daily anatomical information, D_2, D_{50} and D_{98} of the PTV change on average by $0.0 \pm 0.4\%$ according to TPS calculations; the standard deviation of the difference between EPID and TPS target dose changes from 2.5% (pCT) to 2.1% (CBCT). Small time trends are seen for both TPS and EPID dose distributions when using the pCT, which disappear when including CBCT information.

Conclusions

Daily anatomical changes hardly influence the target dose distribution for H&N VMAT treatments according to TPS recalculations. Including CBCT information in EPID dose reconstructions slightly improves the agreement with TPS calculations.

4.1 Introduction

Clinical experience has shown that in head-and-neck cancer patients treated with radiotherapy, anatomical changes often occur during the course of treatment [52, 53, 54, 55, 56, 57]. Both the tumor and organs at risk (OARs) may change in shape and size; this variability may be related to weight loss, fluid shift within the body, alteration in muscle mass and fat distribution, and the increase or decrease of the tumor volume. Several studies have described volumetric and positional changes of gross tumor volume (GTV) with repeat CT imaging at several time points during the course of radiotherapy [52, 53, 54, 55, 57]. A decrease of the GTV throughout the course of radiotherapy at a median rate of 1.8% per treatment day is described in Barker et al. [52]. Similarly, the volume of the parotid glands gradually decreased and their position changed during treatment.

These anatomical changes can result in a difference between the actual delivered dose compared to the planned dose. For instance, the study described in Wu et al. [54] showed that the mean dose of the parotid glands can increase up to 10% in a series of 11 patients evaluated with weekly repeat CT-scans. In another study, 15/23 (65%) patients benefited from adaptive planning, either due to inadequate dose to gross disease or to increased dose to OARs [55].

An important question is whether anatomical changes should be taken into account in our *in vivo* EPID dosimetry method. Using daily anatomical information, a more realistic reconstructed *in vivo* dose distribution would be obtained, with the potential clinical benefit of reducing false positive alerts. Our EPID dose reconstruction model overestimates the real difference between the actual delivered dose and the dose predicted by the TPS, when using the planning CTs, in case anatomical changes occur [44]. Also, the results of a previous study on target dose verification for head-and-neck VMAT treatments [58] indicated that the inclusion of daily anatomical information might be beneficial for our method.

The aim of this study is to quantify the effect of anatomical changes on our EPID-based *in vivo* dose verification results. As guidance for the expected change in dose due to these anatomical changes, TPS recalculations will be used. Also, the impact on the clinical practice of including CBCT information is investigated. This impact is characterized by the amount of alerts raised for a specific workflow, as in our clinical protocol all alerts need to be inspected by a medical physicist.

4.2 Materials and Methods

4.2.1 Patient and plan selection

For this study, 20 patients treated with VMAT for head-and-neck cancer were randomly selected. The patients were treated between September 2013 and May 2014; characteristics of the patients are shown in Table 4.1. The treatment plans were created using the Pinnacle treatment planning system (TPS) v9.6 (Philips Medical Systems, Eindhoven, The Netherlands) and consisted of a dual-arc technique without couch rotations. Of these patients, 10 were treated using a simultaneous-integrated-boost (SIB) technique and received 35×1.55 Gy for the elective region and 35×2.00 Gy for the boost region. The remaining 10 patients were treated using a sequential boost, with 9 patients having a 23×2.00 Gy scheme for the elective region and 12×2.00 Gy for the boost region; one patient, who was treated post-operatively, received 25×2.00 Gy (elective) combined with 8×2.00 Gy (boost). Two patients had been selected for adaptive radiotherapy (ART) and had their treatment plans adjusted during treatment course: one SIB-treated patient received a new treatment plan after 6 fractions, one non-SIB-treated patient received a new treatment plan after 30 fractions. For both patients, both the original and the adapted plans where included in this study. CBCTs and EPID movies of each arc were acquired for 633 out of the 700 total number of fractions. The 67 fractions not included had either no CBCT or no EPID movie available. At the start of each included fraction, the patient was positioned on-line using a multi-region-of-interest registration method [59] using the acquired CBCT.

4.2.2 Analysis of the dose distributions

As the CBCTs are not Hounsfield-unit-calibrated, they are not suited for dose calculations in the TPS directly. Instead, a modified-CT (mCT) was created from each CBCT and the corresponding pCT via deformable image registration. Using a b-spline method, a deformation vector field (DVF) was created to describe the anatomical changes between the pCT and the CBCT [60, 61]. The mCT was created by applying the DVF to the pCT; each mCT was then imported into the TPS and used to calculate the dose-of-the-day data set.

The EPID movies were used to reconstruct the *in vivo* dose distribution by using our 3D backprojection method [22, 23]. For each fraction, the *in vivo* dose distribution was reconstructed twice: once using the original pCT and once using the mCT data set.

Patient #	Tumor localization	Staging	GTV (cc)	ΔV_{max} (elective/boost)	Chemo
1	Oropharynx	T2N2a	41	11% / 5%	
2	Larynx	T3N0	7	7% / 8%	
3	Oropharynx	T3N2b	22	4% / 3%	Carboplatin weekly
4	Oropharynx	T2N2b	65	9% / 5%	Cisplatin 3-weekly
5	Parotid gland	T2N0	*	3% / 1%	
6	Oropharynx	T1N2b	26	5% / 3%	Cetuximab weekly
7	Larynx	T2N0	4	7% / 8%	
8	Larynx	T1N0	3	5% / 5%	
9	Oropharynx	T4N2b	45	10% / 9%	Cisplatin 3-weekly
10	Larynx	T2N0	1	5% / 4%	
11	Larynx	T2N0	21	7% / 7%	Cisplatin 3-weekly
12	Hypopharynx	T2N2b	30	7% / 7%	Cisplatin 3-weekly
13	Oropharynx	T2N2b	26	6% / 6%	Cisplatin 3-weekly
14	unknown	TxN2a	24	3% / 3%	
15	Hypopharynx	T2N1	6	8% / 8%	Cisplatin 3-weekly
16	Tongue	T2N0	13	4% / 4%	
17	Larynx	T2N0	2	4% / 4%	
18	Oropharynx	T3N1	13	7% / 7%	Cisplatin 3-weekly
19	Larynx	T1N0	0.3	1% / 1%	
20	Oropharynx	T1N2c	29	9% / 9%	Cisplatin daily

Table 4.1: Characteristics of included patients. The indicated GTV volume includes the primary tumor and affected lymph nodes. ΔV_{max} is the maximum observed change in relevant volume over the course of the treatment. Patient 5 was treated post-operatively; patients 10 and 20 were selected for adaptive radiotherapy and obtained a new treatment plan during treatment.

In total, four 3D dose distributions were evaluated per fraction: (1) TPS dose on pCT, referred to as TPS^{pCT}, (2) TPS dose on mCT (TPS^{mCT}), (3) EPID reconstructed dose on pCT ($EPID^{pCT}$) and (4) EPID reconstructed dose on mCT ($EPID^{mCT}$).

4.2.3 Target volume definition, dose evaluation, volume changes

Both the elective- and boost-PTV were used as volumes of interest for evaluating target dose. For the SIB treatment plans, however, the boost and elective region overlap in the same plan, leading to different DVH-characteristics compared to the sequential-boost plans. In order to facilitate comparisons between the SIB and non-SIB treated patients, the elective-PTV for the SIB plans is therefore defined as the total PTV minus the boost-PTV.

As the PTV is a region defined in room coordinates [62], it is by definition invariant to changes in patient anatomy. However, for the dose calculations using the mCTs, a problem arises when the irradiated patient volume shrinks in such a way that part of the PTV now lies outside the patient or in the build-up region. To accommodate for this anatomical change, a modified PTV (mPTV) was used for the evaluation. It was determined by taking the intersection of the original PTV and the patient volume minus 4 mm determined from the mCT. In the outer 4 mm, dose calculations by both the TPS and the EPID dosimetry method are inaccurate. Removing the outermost part of the patient volume for evaluation is in accordance with our clinical practice of generating treatment plans.

The dose distributions in the PTV and mPTV were characterized by D_2 (near-maximum dose), D_{50} (median dose) and D_{98} (near-minimum dose) as defined in ICRU Report 83 [50]. Differences in target dose between the four available dose distributions per fraction are then expressed as differences in D_2, D_{50} and D_{98}, denoted as ΔD_2, ΔD_{50} and ΔD_{98}, respectively. We will refer to ΔD_2, ΔD_{50} and ΔD_{98} as indicators; note that they can be calculated for any two dose distributions when the (m)PTV has been defined. The values of these indicators are calculated as relative differences for a combination of two dose distributions; e.g., ΔD_{50} of the $EPID^{pCT}$ compared to the TPS^{pCT} dose distribution would be calculated as (D_{50},$EPID^{pCT}$ - D_{50},TPS^{pCT})/ D_{50},TPS^{pCT}. This will then be denoted as ΔD_{50} of $EPID^{pCT}/TPS^{pCT}$.

Dose evaluation was done in several ways. First, the expected difference in delivered dose was evaluated by comparing TPS^{pCT} with TPS^{mCT}. With the assumption of having a properly functioning TPS, linac, CBCT-system and warping procedure, this will be a

measure of the real difference in dose delivered to the patient. Second, the change in deviations between TPS- and EPID-dose was determined by inspecting the differences between (a) EPID^{pCT} and TPS^{pCT} and (b) EPID^{mCT} and TPS^{mCT}. The effect on EPID dosimetry results alone is evaluated by comparing EPID^{pCT} to EPID^{mCT}. In order to explore time trends, all dose deviations will be presented as a function of time since the treatment start.

Only changes of a specific part of the patient volume are expected to have an effect on planned and reconstructed target dose. This volume will be referred to as relevant volume and is defined as the union of all axial CT-slices in which the PTV is present. The change of the relevant volume over time was determined. The effect of volume changes was also explicitly investigated for TPS calculations and EPID reconstructions by evaluating $\text{TPS}^{\text{pCT}}/\text{TPS}^{\text{mCT}}$ and $\text{EPID}^{\text{pCT}}/\text{EPID}^{\text{mCT}}$ as a function of change in relevant volume.

Finally, in a clinical implementation of an EPID-based dosimetry workflow, alerts are raised whenever an indicator is out of its tolerance level, which is $\pm 7\%$ for ΔD_2 and ΔD_{98}; $\pm 3.5\%$ for ΔD_{50}, as described in [58]. The number of alerts that would have been raised are determined for two workflows: pCT-based and mCT-based. The pCT-based workflow is equivalent to our current clinical practice, in which $\text{EPID}^{\text{pCT}}/\text{TPS}^{\text{pCT}}$ are compared. The mCT-based workflow is based on comparing EPID^{mCT} with TPS^{pCT} – not with TPS^{mCT}. This is because clinically, one wishes to compare the reconstructed dose with the originally intended dose distribution, TPS^{pCT}.

4.3 Results

For ΔD_2, ΔD_{50} and ΔD_{98} of $\text{TPS}^{\text{pCT}}/\text{TPS}^{\text{mCT}}$, all values are within [-1.7%; 1.9%], [-1.1%; 1.5%] and [-2.3%; 1.3%], respectively; the average values and standard deviations are shown in Table 4.2. Figure 4.1 shows the relative change in median dose, ΔD_{50}, of $\text{TPS}^{\text{mCT}}/\text{TPS}^{\text{pCT}}$ for both types of PTV. The data shown in Figure 4.1 indicate a very small increase in median PTV dose for both PTV types over time. Although small, the trend is statistically significant ($p \ll 0.01$) and positive. There is no clear distinction in results between boost- and elective-PTVs.

Relative changes in median PTV dose values of $\text{EPID}^{\text{pCT}}/\text{TPS}^{\text{pCT}}$ and $\text{EPID}^{\text{mCT}}/\text{TPS}^{\text{mCT}}$ dose distributions are shown in Figure 4.2a for PTV-boost and in Figure 4.2b for PTV-elective. The averages and standard deviations of the ΔD_2, ΔD_{50} and ΔD_{98} results for both comparisons are shown in Table 4.2. Indicator results for $\text{EPID}^{\text{pCT}}/\text{EPID}^{\text{mCT}}$

Data	Indicator	PTV elective	PTV boost	PTV boost + elective
$EPID^{pCT}$ / TPS^{pCT}	ΔD_2	1.4 (± 2.7)%	3.5 (± 2.9)%	2.3 (± 2.9)%
	ΔD_{50}	-1.1 (± 2.5)%	0.4 (± 2.4)%	-0.5 (± 2.6)%
	ΔD_{98}	-4.6 (± 3.7)%	-2.1 (± 2.6)%	-3.6 (± 3.5)%
$EPID^{mCT}$ / TPS^{mCT}	ΔD_2	0.8 (± 2.1)%	2.4 (± 2.2)%	1.5 (± 2.3)%
	ΔD_{50}	-1.1 (± 1.9)%	-0.4 (± 2.2)%	-0.8 (± 2.1)%
	ΔD_{98}	-4.2 (± 3.4)%	-2.5 (± 2.4)%	-3.5 (± 3.1)%
$EPID^{mCT}$ / TPS^{pCT}	ΔD_2	1.0 (± 2.2)%	2.6 (± 2.2)%	1.7 (± 2.3)%
	ΔD_{50}	-1.1 (± 2.1)%	-0.4 (± 2.2)%	-0.8 (± 2.2)%
	ΔD_{98}	-4.4 (± 3.4)%	-2.5 (± 2.5)%	-3.6 (± 3.2)%
TPS^{pCT} / TPS^{mCT}	ΔD_2	0.2 (± 0.4)%	0.2 (± 0.4)%	0.2 (± 0.4)%
	ΔD_{50}	0.0 (± 0.4)%	0.1 (± 0.4)%	0.0 (± 0.4)%
	ΔD_{98}	-0.2 (± 0.5)%	0.0 (± 0.5)%	0.0 (± 0.5)%

Table 4.2: Indicator results for EPID vs TPS, TPS vs TPS and EPID vs EPID results. Indicated values are means and standard deviations averaged over all 633 included fractions.

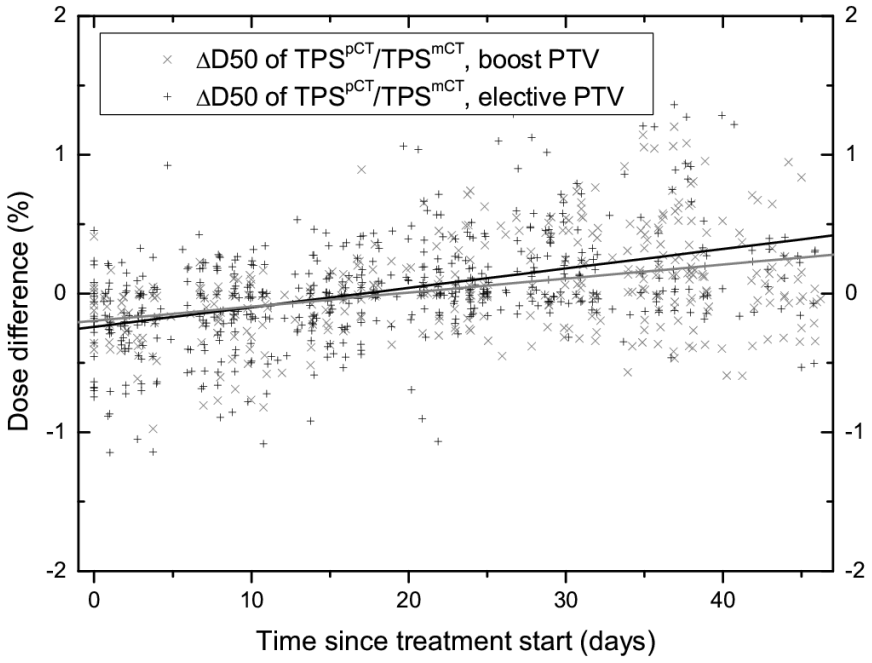
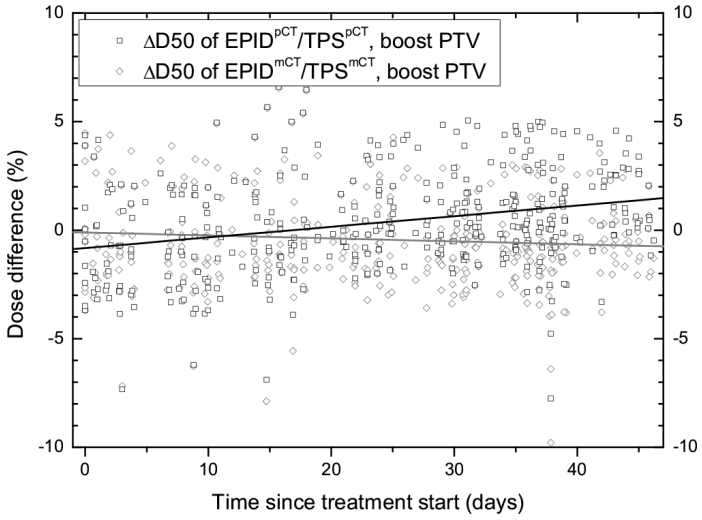


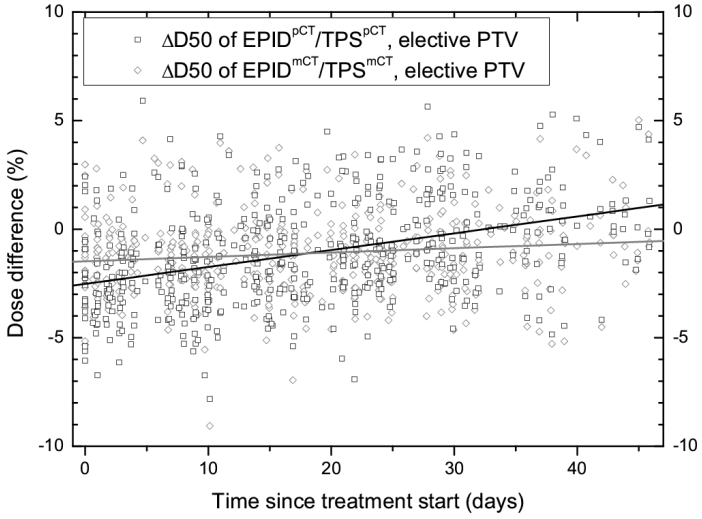
Figure 4.1: TPS^{pCT} / TPS^{mCT} dose differences as a function of time, expressed as ΔD_{50} for both the elective- and boost-PTV. Solid lines are linear fits to elective (black) and boost (gray) PTV datapoints; the slope of both fitted lines is positive with a p-value $\ll 0.01$.

are $0.6(\pm1.4)\%$, $0.3(\pm1.4)\%$ and $0.0(\pm1.7)\%$ for ΔD_2, ΔD_{50} and ΔD_{98}, respectively, for type of PTV combined.

The $EPID^{pCT}/TPS^{pCT}$ median PTV dose difference shows a slight yet statistically significant ($p \ll 0.01$) increase over time; after recalculation using the mCT for both the reconstructed and the TPS dose (i.e., $EPID^{mCT}/TPS^{mCT}$), the trend disappears. In accordance with these findings, the change in relevant volume ΔV decreases over time. At the end of the last week of treatment, the relevant volume has shrunk by $3.5(\pm0.2)\%$ on average, compared to its value at the start of treatment. Figure 4.3 shows the effect of change in relevant volume on ΔD_{50} of TPS^{pCT}/TPS^{mCT} and $EPID^{pCT}/EPID^{mCT}$,



(a)



(b)

Figure 4.2: Median (D_{50}) PTV dose differences EPID/TPS, using either pCT or mCT for both dose distributions, as a function of time for (a) boost PTV and (b) elective PTV. The linear fit to pCT-ΔD_{50} has a positive slope with $p \ll 0.01$; the linear fit to mCT-ΔD_{50} has a slope which is not significantly different from zero, i.e., the slope could be either negative or positive.

for both types of PTV combined. ΔD_{50} changes by 0.091(±0.004)% and 0.51(±0.01)% per percent volume change (%/%ΔV) for TPS^{pCT}/TPS^{mCT} and $EPID^{pCT}/EPID^{mCT}$, respectively.

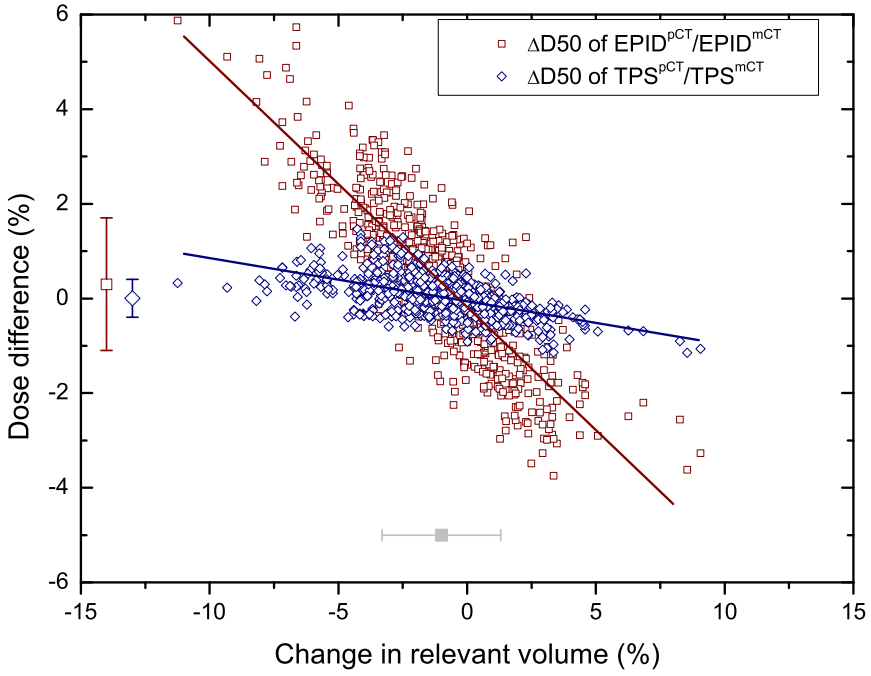


Figure 4.3: ΔD_{50} of $EPID^{pCT}/EPID^{mCT}$ and TPS^{pCT}/TPS^{mCT} as a function of change in relevant volume. Solid lines are fits to the data; indicated next to the axes are the average values of ΔD_{50} of $EPID^{pCT}/EPID^{mCT}$ (open red square), ΔD_{50} of TPS^{pCT}/TPS^{mCT} (open blue diamond), and ΔV (solid gray square). The error bars on the average values indicate one standard deviation.

After employing the tolerance levels for the indicators, the number of fractions for which an alert is raised was reduced from 156 (17%) for the pCT-based workflow to 101 (11%) for the mCT-based workflow.

4.4 Discussion

4.4.1 General

This study evaluates the impact of including daily anatomical changes for H&N VMAT treatments on a clinically applied EPID-dosimetry method. It also provides insight in changes in TPS-calculated target dose due to anatomical changes, which is in accordance with other work [63]. Though the EPID-dosimetry results are specific for our method, the TPS-results are expected to be generally valid. In contrary to the assumptions made when starting this study, the – statistically significant – effect on EPID-reconstructed dose distributions is of smaller magnitude than the spread in the data. A clear clinical benefit in terms of number of raised alerts has been found nonetheless.

The choice for reporting dose to the PTV was a priori based on simplicity: as the PTV is defined in room-coordinates, it is independent of anatomy variation. Hence, no residual uncertainties from patient positioning and deformable registration propagate into the PTV. As the PTV-coverage remains intact when including daily anatomical information, we can infer that the CTV-coverage is still achieved, giving an a posteriori justification for the use of the PTV. The dose to OARs is of paramount importance in radiotherapy; as a next step, we plan to investigate the effect of anatomical changes on EPID dosimetric verification of OAR doses. However, in contrast to the PTV dose, proper OAR dose estimation would require detailed knowledge of anatomical changes which complicates the evaluation significantly compared to the current solution.

4.4.2 Expected change in delivered dose

A striking finding of this study is that the difference between intended and expected delivered target dose for this group of head-and-neck patients treated with VMAT is very small indeed, even if the irradiated volume changes considerably. Though the criteria upon which it is decided whether a deviation in expected delivered dose is clinically relevant are up to debate, it is clear that the impact of daily anatomical changes on target dose, as judged by changes in PTV dose, is limited. Of course, the indicators used are position independent; i.e., the position of hot- and/or cold-spots inside the PTV might change without the indicators changing value. This is not considered to be an issue: the quality of PTV coverage is assumed to be well determined by these indicators alone, in accordance with ICRU [50] recommendations.

4.4.3 Effect of anatomical changes on EPID dosimetry results

The average values of the indicators do not change substantially when including daily anatomical information. Note that the systematic deviations in the indicator values have been observed and explained before [58]; these are due to residual setup errors and the use of the PTV for dose evaluation. The precision of the dose reconstructions does improve somewhat: the standard deviation in ΔD_{50} decreases from 2.4% to 2.2% (PTV boost) and 2.5% to 2.1% (PTV elective). Comparable changes are seen in the standard deviations of ΔD_2 and ΔD_{98}; the average of the standard deviations of ΔD_2, ΔD_{50} and ΔD_{98} changes from 3.0% to 2.5%.

The EPID-based dose reconstruction is ~5 times more sensitive to volume changes than TPS calculations, as can be seen from Figure 4.3. This enhanced sensitivity with respect to TPS-calculations is attributed to the use of the CT-scan in the dose reconstruction method. The transmission through the patient is an input to our dose reconstruction method, and it is calculated using a CT-scan [23]. If the patient anatomy during irradiation differs from the CT-scan used for dose reconstruction, a deviation will be introduced. The magnitude can be estimated by assuming exponential attenuation of the beam through the patient and considering a change in thickness. For symmetrical thicknes changes, the effect on the dose is then proportional to $\exp(\mu \cdot d/2)$, with μ the attenuation coefficient and d the change in thickness. By approximating the relevant volume as a cylinder and assuming a radial thickness change, we find an effect of 0.24%/%ΔV based on the data used for this study. This effect is additional to any real changes in delivered dose; combined with the expected real difference in delivered dose of 0.091%/%ΔV, the sensitivity of EPID dose calculations to volume changes is then estimated to be 0.33%/%ΔV. Clearly, the order of magnitude is in accordance with the observed dependence on volume changes (0.51%/%ΔV); yet it also indicates that the precise location of the volume changes and the distribution of monitor units over the control points need to be considered for a precise calculation.

Even so, the volumetric changes are not the largest contributor to the total uncertainty (spread) in reconstructed *in vivo* dose, σ^{ivd}. Assuming independent uncertainties, σ^{ivd} can be related to the volume-dependent uncertainty σ^{volume} and other (volume-independent) uncertainties σ^{other} via the equation $(\sigma^{\mathrm{ivd}})^2 = (\sigma^{\mathrm{volume}})^2 + (\sigma^{\mathrm{other}})^2$. Table 4.2 shows that $\sigma^{\mathrm{ivd}} = 2.6\%$ and $\sigma^{\mathrm{other}} = 2.1\%$ (ΔD_{50}, both types of PTV), leading to $\sigma^{\mathrm{volume}} = 1.5\%$. This corresponds well to the observed spread in ΔD_{50} of $\mathrm{EPID}^{\mathrm{pCT}}/\mathrm{EPID}^{\mathrm{mCT}}$ of 1.4%. The impact on EPID dosimetry workflow is character-

ized by the number of raised alerts, which shows a 6%-point drop when switching to an mCT-based workflow. Though this change is substantial, the tolerance levels used for ΔD_{50}, $\pm 3.5\%$, are quite close to the remaining uncertainty in reconstructed dose, 2.1%. In this view, one might argue that the tolerance levels used are too strict in view of the uncertainty of the method. Changing the ΔD_{50} tolerance levels to $\pm 5\%$ would decrease the number of raised alerts to 78 (8%) for the pCT-based workflow. Though this appears to be more in accordance with the uncertainties in the dose reconstruction method, this relaxation of the tolerance levels may lead to an unacceptable increase in false negative rate, i.e would mask true dose deviations of the same magnitude ($\pm 5\%$). A better approach could be to use the mCT-based workflow to take morphological changes into account when they are larger than for example 5% in relevant volume. This is currently being investigated.

For detecting gross errors in treatment delivery, relaxed tolerance levels should suffice. However, higher accuracy is needed to find smaller systematic errors in treatment planning or delivery to justify the use of *in vivo* EPID dosimetry as a replacement for pre-treatment dose verification. Despite a per-fraction agreement of 5-7% between EPID reconstructed and TPS predicted dose, this is still attainable by employing aggregate analyses over larger sets of delivered plans: as the deviations in reconstructed dose are random for different plans, the standard deviation in the average value will decrease by $1/\sqrt{N}$ (N the number of samples). As typical numbers for N would range from 10-20 per day, daily aggregate analyses would have an estimated accuracy of ~2%; weekly analyses would be capable of reaching accuracies of <1%.

As anatomical variations only account for a small part in the spread in *in vivo* dosimetry results, other factors influencing the uncertainty in the EPID dose reconstructions must be at play. Part of the spread may be explained by linac output variations, but as the linac output is checked daily, this should not account for more than ~0.5%. MLC leaf positioning errors might be another source of uncertainties. The tolerance of the MLC leaves in our institute is 0.5 mm, the effect of this uncertainty on the delivered dose is yet to be determined. Residual setup errors are 1–3 mm, depending on location [64]; the deformable image registration method applied to head-and-neck anatomies has been shown to be highly accurate (<1 mm) and to have a precision of 2–3 mm (one standard deviation) [61].

Phantom measurements showed that the reproducibility of the D_{50} of EPID-based verifications of a single H&N VMAT treatment is about 1.6% (data not shown). Remaining explanations are limitations of our dose verification model, especially neglecting

ghosting effects of the EPID [43], reduced accuracy for small fields and the amount of scatter from the patient reaching the EPID. From CBCT reconstruction it is known that the assumption in our model of homogeneous scatter from the patient is violated in the case of H&N treatments, as the shoulders are much closer to the EPID than the neck. Clearly, *in vivo* dosimetry of a complex technique such as H&N VMAT treatments is quite challenging for a rather simple back-projection method such as ours, resulting in an overall uncertainty of about 2% (one standard deviation). Work is in progress to reduce this uncertainty.

4.5 Conclusions

Useful insight in the effect of daily anatomical changes on target dose for VMAT treated head-and-neck cancer patients has been obtained. The effect on TPS-calculated dose correlates welll with changes in relevant volume, though the absolute magnitude of the effect on target dose is limited. 3D *in vivo* EPID dosimetry is circa 5 times more sensitive to these anatomical changes than TPS calculations; the anatomical changes account for 1.5% in the total uncertainty of reconstructed dose. Including daily anatomical information would substantially lower the number of raised alerts in our clinical dose verification workflow.

Chapter 5

Virtual patient 3D dose reconstruction using in air EPID measurements and a back-projection algorithm for IMRT and VMAT treatments

I. Olaciregui-Ruiz
R.A. Rozendaal
R.F.M. van Oers
B. Mijnheer
A. Mans

Physica Medica 37 (49), 2017

Abstract

Purpose

At our institute, a transit back-projection algorithm is used clinically to reconstruct *in vivo* patient and in phantom 3D dose distributions using EPID measurements behind a patient or a polystyrene slab phantom, respectively. In this study, an extension to this algorithm is presented whereby in air EPID measurements are used in combination with CT data to reconstruct "virtual" 3D dose distributions. By combining virtual and *in vivo* patient verification data for the same treatment, patient-related errors can be separated from machine, planning and model errors.

Methods and materials

The virtual back-projection algorithm is described and verified against the transit algorithm with measurements made behind a slab phantom, against dose measurements made with an ionization chamber and with the OCTAVIUS 4D system, as well as against TPS patient data. Virtual and *in vivo* patient dose verification results are also compared.

Results

Virtual dose reconstructions agree within 1% with ionization chamber measurements. The average γ-pass rate values (3% global dose/3mm) in the 3D dose comparison with the OCTAVIUS 4D system and the TPS patient data are 98.5±1.9%(1SD) and 97.1±2.9% (1SD), respectively. For virtual patient dose reconstructions, the differences with the TPS in median dose to the PTV remain within 4%.

Conclusions

Virtual patient dose reconstruction makes pre-treatment verification based on deviations of DVH parameters feasible and eliminates the need for phantom positioning and re-planning. Virtual patient dose reconstructions have additional value in the inspection of *in vivo* deviations, particularly in situations where CBCT data is not available (or not conclusive).

5.1 Introduction

The introduction of intensity modulated RT (IMRT) and volumetric modulated arc therapy (VMAT) was accompanied by additional verification of individual patient treatments, as recommended by several organizations [65]. These patient specific checks are generally performed pre-treatment by using phantom-detector array combinations, film or portal dosimetry. Although pre-treatment verification using in phantom dosimetry helps in determining whether the treatment can be delivered as expected to a stable anatomy, its usefulness is known to be limited in detecting clinically relevant errors in the delivery of the treatment to the patient [66, 67]. Furthermore, several studies have shown that γ-pass rate has insufficient predictive power for pre-treatment dose verification [29, 30, 32] and have indicated the necessity to assess the delivered dose to a patient using other metrics such as dose-volume histograms (DVHs) of the target volume and organs at risk. Alert criteria based on deviations in DVH parameters have in addition a clearer relation with treatment planning data, while being easier to interpret by most radiotherapy staff than γ-evaluation data.

At the Netherlands Cancer Institute (NKI), pre-treatment verification of clinical plans is performed using in phantom EPID dosimetry with a slab polystyrene phantom [27]. Pre-treatment verification based on deviations of DVH parameters related to the patient anatomy is, by definition, not feasible with in phantom dosimetry. A second disadvantage of our in phantom EPID-based dosimetry method, which also applies to other pre-treatment dose verification methods, is that it requires additional clinical time due to the need for phantom re-planning and positioning. The main purpose of this study is to use in air EPID measurements to reconstruct dose distributions within the patient anatomy. This will eliminate the need for phantom positioning and re-planning and will make pre-treatment verification based on deviations of DVH parameters feasible.

Non-transit EPID dosimetry has been in use for some time in IMRT pre-treatment verification [68, 69, 70, 71, 72, 73]. For patient dose reconstruction in 3D, in air EPID measurements are generally used to derive the actual fluence delivered by the accelerator, which may then be used as input to a dose engine that generates the patient 3D dose distribution. Van Zijtveld et al. [74] used their clinical treatment planning system (TPS) for the dose engine of their reconstruction method and the system was clinically evaluated for 17 IMRT treatments of different sites. The use of the same TPS for dose reconstruction and patient treatment planning may hide, however, potential inaccuracies in the TPS dose calculation algorithm. Aiming at a TPS-independent dose calculation,

Van Elmpt et al. [75] used an in-house developed Monte Carlo dose engine and their model was used in the pre-treatment verification of 9 lung cancer patients treated with a 3D conformal technique and 5 head-and-neck cancer patients treated with a step-and-shoot 7-field IMRT technique. A similar approach is followed by a commercially available system, Dosimetry Check (Math Resolutions, Columbia, MD, USA), where the fluence is entered into an independent pencil beam type dose calculation algorithm [76, 77]. Recently, a GPU-accelerated collapsed cone convolution technique was explored for dose reconstruction in phantom or in CT simulation data sets using in air EPID measurements by Zhu et al. [36]. Results for the verification of a head-and-neck, two lung and one prostate IMRT plan showed good agreement with TPS calculated dose distributions using 3%/3mm criteria.

In this study, a novel approach is taken. Rather than estimating the energy fluence delivered by the accelerator and then using it as input to a dose engine for a forward dose calculation, the use of an extension of our transit back-projection algorithm was investigated. This new extension to the algorithm allowed us to use in air EPID measurements in combination with planning CT data to calculate patient 3D dose distributions, i.e., virtual patient dose distributions. These virtual patient dose distributions can be used for both IMRT and VMAT pre-treatment verification.

A unique advantage of using the same algorithm for virtual and *in vivo* reconstruction would be that patient-related errors may be separated from machine, planning and model errors by combining virtual and *in vivo* patient verification data for the same treatment. This circumvents a common limitation of *in vivo* dose verification methods which is the inability to discriminate changes in the measured *in vivo* dose distribution due to the variation in the fluence incident on the patient from changes due to anatomical variations within the patient. In this study, we illustrate the usefulness of this approach with some clinical examples. It should be noted that, in principle, the methods mentioned earlier [75, 76, 36] are also capable of combining pre-treatment with *in vivo* 3D dose verification results for the same patient. However, to the best of our knowledge these groups have not yet published results of those types of studies. Finally, this is one of the first published studies presenting patient-specific pre-treatment VMAT verification results in 3D using in air EPID measurements. Another practical advantage of this approach would be that no extra commissioning work is required prior to the introduction of patient virtual dose reconstruction in our clinic. In this study, we present the modifications made to our algorithm to allow for virtual dose reconstruction and the results of the assessment of its accuracy.

5.2 Materials and methods

5.2.1 Virtual patient dose reconstruction algorithm

Our back-projection algorithm requires the primary portal dose distribution, i.e. the dose component at the EPID level which results from radiation coming directly from the radiation head of the accelerator.

For *in vivo* patient dose reconstruction, the transit algorithm uses *in vivo* EPID measurements to determine the primary portal dose distribution behind the patient. For in phantom dose reconstruction, the transit algorithm uses EPID images acquired behind the phantom to determine the primary portal dose distribution behind the phantom.

For virtual patient dose reconstruction in this study, the new virtual algorithm uses in air EPID measurements in combination with the CT data of the patient to predict the primary portal dose distribution behind the patient. In a similar way, the new virtual algorithm can use in air EPID measurements in combination with the CT data of the phantom to predict the primary portal dose distribution behind the phantom. Figure 5.1 shows schematically these four situations. The algorithm then uses this primary portal dose distribution to reconstruct the dose distribution in any plane parallel to the EPID [41, 22]. By iterating this 2D reconstruction in multiple planes, the dose distribution in 3D can be obtained.

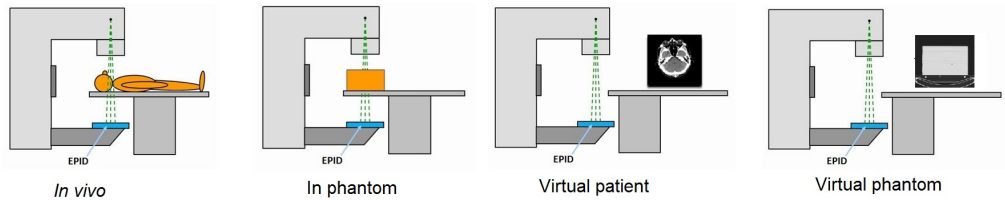


Figure 5.1: *In vivo*, in phantom, virtual patient and virtual phantom EPID measurement configurations to determine the primary portal dose distribution at the EPID level. *In vivo* and in phantom dose reconstructions use our current clinical transit algorithm. Virtual dose reconstructions use the new extension of the algorithm.

In the following, we use the notation X_{ij} for a quantity X at pixel ij of the EPID. The line from the target of the linac to this pixel ij constitutes the geometrical back-projection line ij. Note that the coordinate system for dose reconstruction is fixed to the gantry.

In our algorithm, the *in vivo* primary portal dose distribution is calculated from the pixel values $PV_{ij}^{EPID,patient}$ of EPID images measured behind the patient by correcting for the sensitivity matrix S_{ij}, applying the EPID dose response D_r, removing the scatter within the EPID with convolution kernel K_{ij}^1 and blurring the EPID signal with convolution kernel K_{ij}^2, correcting for the couch attenuation C_{ij} and removing the scatter from the patient to the EPID, SC_{ij}:

$$Pr_{ij}^{EPID,in\ vivo} = \left(\left(PV_{ij}^{EPID,patient} \cdot S_{ij} \cdot D_r\right) \otimes^{-1} K_{ij}^1 \otimes K_{ij}^2\right) \cdot C_{ij} \cdot SC_{ij}. \quad (5.1)$$

If the treatment is delivered without the patient in the beam, the in air primary portal dose distribution can be calculated in a similar way, but without the corrections for the couch attenuation and the patient scatter:

$$Pr_{ij}^{EPID,in\ air} = \left(\left(PV_{ij}^{EPID,in\ air} \cdot S_{ij} \cdot D_r\right) \otimes^{-1} K_{ij}^1 \otimes K_{ij}^2\right). \quad (5.2)$$

The primary transmission at the EPID level $T_{ij}^{primary,CT}$ can be calculated using the radiological thickness of the patient t_{ij}), the linear attenuation coefficient of water for a specific beam energy μ and the beam hardening coefficient σ [44]. The radiological thickness is determined by ray-tracing through the patients planning CT-scan, i.e., by calculating a Digitally Reconstructed Radiograph (DRR). The numerical values for μ and σ are determined during the commissioning process of the portal dosimetry system fitting the equation

$$T_{ij}^{primary,CT} = \exp\left(-\mu \cdot t_{ij} + \sigma \cdot (t_{ij})^2\right) \quad (5.3)$$

against the ratio of equations 5.1 and 5.2 for a set of measurements behind phantoms of different thicknesses.
In this study, the algorithm was adapted to use $Pr_{ij}^{EPID,in\ air}$ and $T_{ij}^{primary,CT}$ to estimate the *virtual* patient primary portal dose distribution by combining equations 5.2 and 5.3:

$$Pr_{ij}^{EPID,virtual} = Pr_{ij}^{EPID,in\ air} \cdot T_{ij}^{primary,CT}. \quad (5.4)$$

The rest of the algorithm remains the same regardless of whether equation 5.1 or 5.4 is used to determine the primary portal dose distribution. Note that in air EPID measurements contain information about the dose to be delivered to the patient, including

machine and/or planning errors. *In vivo* EPID measurements contain additional information about the dose that has been delivered to the patient, including patient-related errors such as anatomical changes and/or setup issues.

The parameters of our model are determined using water-based kernels and consequently, the model is expected to work most accurately in reconstructions within slab phantoms and within water-based patients. For dose verifications of sites involving (large) tissue heterogeneities, e.g. lung, esophagus and breast, the *in aqua vivo* approach [45] is always used.

5.2.2 Accelerator, EPID image acquisition and dose verification

Measurements were performed on various SL20i linear accelerators (Elekta, Crawley, UK) using 6 and 10 MV photon beams. The linacs are equipped with a PerkinElmer RID 1680 AL5 amorphous silicon EPID (Elekta iViewGT). Treatment plans were generated with the clinical version of our TPS (Pinnacle V9.10, Philips Medical Systems, Eindhoven, The Netherlands).

For IMRT verification, our in-house acquisition software averages the total signal of all EPID frames between beam-on and beam-off into one accumulated portal image, which is stored with the number of frames in the image header. For virtual dose reconstruction, the new extension of the algorithm used the accumulated portal image of in air measurements in combination with CT data to estimate the accumulated virtual primary portal dose distribution corresponding to each field. The reconstructed 3D dose distributions of all the fields were summed together to obtain the reconstructed 3D dose distribution of the delivered IMRT fraction.

For VMAT verification, cine-mode image acquisition is used and separate EPID frames are continuously being acquired during delivery. Each recorded frame is associated with a gantry angle. The EPID verification software groups these frames in bins, where each bin contains the sum of all frames acquired within a certain gantry-angle range. For virtual dose reconstruction, the algorithm used in air measurements in combination with CT data to estimate the virtual primary portal dose distribution corresponding to each bin. The reconstructed 3D dose distributions of all bins were then summed to obtain the 3D dose distribution of each VMAT arc. The reconstructed 3D dose distributions of the arcs were summed together to obtain the reconstructed 3D dose distribution of the delivered VMAT fraction. Details of our VMAT verification procedure can be found elsewhere [23].

In our clinical practice, the comparison of 3D dose distributions is performed by γ-evaluation [28, 78] using as γ-criterion a dose difference of 3% of the maximum planned dose, and a distance-to-agreement of 3 mm. For SBRT treatments, a dose difference of 3% of the prescribed dose was used because a maximum dose of 160% is allowed for these treatments. This is how 3D dose distributions have been compared throughout this study, i.e. virtual dose reconstruction vs. OCTAVIUS, virtual dose reconstruction vs. TPS and *in vivo* dose reconstruction vs. TPS. Results were presented using the mean γ-value (γ-mean), the percentage of points with γ-value less than one (γ-pass rate) and the near maximum γ-value (γ-1%), the 99^{th} percentile of the γ-distribution. A fast algorithm was used to speed up the computation of γ-distributions in 3D [79]. The results were calculated within the volume surrounded by the 50% isodose surface.

5.2.3 Verification of virtual dose reconstruction against transit dose reconstruction

The equivalence between the virtual back-projection algorithm and the transit back-projection algorithm was proven with measurements performed with a stable anatomy, i.e. by comparing in phantom and virtual phantom reconstructed dose distributions. This comparison serves also as an initial dosimetric verification of virtual dose reconstruction since the transit algorithm has been extensively verified for IMRT using film and point dose ionization chamber (IC) measurements [22], and for VMAT using point dose IC measurements and a 2D IC-array [23]. A 20 cm thick phantom consisting of 30x30 cm^2 polystyrene slabs (pslab phantom) was chosen for these measurements. In phantom and in air EPID measurements were obtained for 20 square fields and for 31 IMRT fields (7 clinical plans of 5 tumor sites). The sizes of the square fields were 3x3 cm^2, 5x5 cm^2, 10x10 cm^2, 15x15 cm^2 and 20x20 cm^2 and were created in Pinnacle using 6MV and 10MV beam models. The measurements were performed on the same linac and on the same date to reduce reproducibility uncertainties. For the square fields, cross-plane dose profiles and percentage dose-depth curves through the isocenter were obtained and compared. For the IMRT fields, reconstructed in phantom and virtual phantom 2D dose distributions at the isocenter were compared by γ-evaluation using as γ-criterion a local dose difference of 2%, and a distance-to-agreement of 2 mm. The γ-results were calculated within the area surrounded by the 20% isodose line.

5.2.4 Verification of virtual dose reconstruction against dose measurements

The virtual back-projection algorithm was then verified directly with dose measurements. First, the dose at 10 cm depth in a water phantom, positioned at an SSD of 90 cm, was measured with a calibrated Semiflex IC (PTW, Freiburg, Germany) on the central axis of 9 square fields with sides ranging from 3 cm to 23 cm. Additionally, cross-plane dose profile for the 10x10 cm^2 square fields were obtained using a PTW scanning water phantom at a measuring depth of 10cm. 6 MV and 10 MV photon beams were used in these measurements. For virtual reconstruction, in air EPID measurements were performed on the same linac on the same day. For 3D dose verification, absolute dose measurements for 5 IMRT and 5 VMAT treatments of different tumor sites were performed using the OCTAVIUS 4D system, i.e. a OCTAVIUS 4D phantom in combination with the OCTAVIUS 1500 2D ion chamber array (PTW, Freiburg, Germany). In air EPID measurements of these 10 clinical treatments were performed with the same linac on the same day. The virtual 3D dose distributions were reconstructed on a OCTAVIUS CT scan with a homogeneous insert having the same dimensions as the 2D array. Finally, the OCTAVIUS 3D dose volumes were imported via DICOM into our EPID dosimetry verification software for a direct dose comparison with the virtual 3D dose distributions.

5.2.5 Verification of virtual patient dose reconstruction against TPS

The virtual back-projection algorithm for patient 3D dose reconstruction was further investigated by comparison with the corresponding patient planned data in the TPS. To this purpose, in air EPID measurements of 25 IMRT treatments (5 tumor sites) and 50 VMAT treatments (9 tumor sites) were performed and used to reconstruct virtual patient 3D dose distributions within the planning CT data of the treatment. These treatments were randomly selected among the clinically available data. The obtained virtual patient dose distributions were verified against the planned dose distribution using γ-evaluation and the difference in the dose-volume histogram (DVH) parameter PTV-D_{50} , i.e., the median dose to the PTV. For 10 treatments (each corresponding to a different tumor site) the difference in the dose-volume histogram (DVH) parameter OAR-D_2 , i.e. the near-maximum dose, to serial organs at risk such as the spinal cord, heart, bladder, and rectum, was also calculated.

5.2.6 Comparison between virtual and *in vivo* patient dose verification results

Differences between virtual and *in vivo* dose verification results for 16 IMRT treatments (3 tumor sites) and 27 VMAT treatments (7 tumor sites) were obtained. The in air and *in vivo* EPID measurements were performed on the same linac and, whenever possible, on the same date to avoid reproducibility uncertainties such as variation in the panel dose response characteristics, and/or intra-linac variabilities. The obtained patient dose distributions were verified against the planned dose distribution using γ-evaluation and the difference in the dose-volume histogram (DVH) parameter PTV-D_{50}.

5.3 Results

5.3.1 Verification of virtual dose reconstruction against transit dose reconstruction

In phantom and virtual phantom cross-plane dose profiles through the isocenter for the 6MV square fields are displayed in Figure 5.2a. Similar results were obtained for the 10MV square fields. The average dose differences between in phantom and virtual phantom reconstruction at the isocenter were -0.4 $\pm$ 0.2%(1SD) and 0.4 $\pm$ 0.4%(1SD) for 6MV and 10MV square fields, respectively. Percentage depth-dose curves for the 10x10 cm^2 square fields (step interval of 2 mm) are presented in Figure 5.2b. The largest local dose differences were 0.2% and 0.4% for 6MV and 10MV, respectively.

The comparison results between reconstructed virtual phantom and in phantom dose distributions for the 31 IMRT fields indicated that the local dose differences were on average well within 2%/2 mm. The obtained average γ-pass rate and γ-mean values were 96.0 $\pm$ 2.5%(1SD) and 0.38 $\pm$ 0.13%(1SD), respectively. As an example, comparison results of an IMRT plan delivered to inguinal lymph nodes are shown in Figure 5.3 indicating that deviations mainly occur in the low dose regions.

5.3.2 Verification of virtual dose reconstruction against dose measurements

For the for the 9 square fields, the difference between the dose values measured with the ionization chamber at the isocenter in the water phantom and the corresponding values determined using virtual reconstruction, amounted to -0.5%$\pm$0.3% (1SD) and -

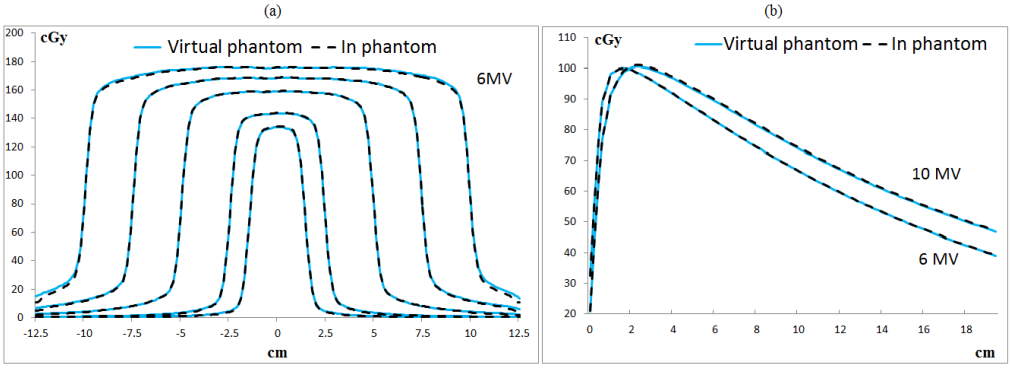


Figure 5.2: Figure (a) displays in phantom and virtual phantom cross-plane dose profiles through the isocenter for five square fields (200 MU) delivered to the 20 cm thick pslab phantom at SSD=90 cm. Figure (b) displays in phantom and virtual phantom percentage depth-dose curves for the 10x10 cm^2 field.

0.0%±0.4% (1SD) for the 6 MV and 10 MV beams, respectively. The maximum detected deviation was 1.1%. Figure 5.4a shows transverse LR dose profiles through the isocenter for the 6MV 10x10cm^2 field.

A summary of the comparison between the 3D dose distributions measured with the OCTAVIUS 4D system and the corresponding virtual dose distributions is presented in Table 5.1. The average γ-pass rate in the comparison of the 10 treatments was 98.5 ± 1.9 (1SD). The largest found dose difference at the isocenter was 2.6%. Transverse L-R dose profiles through the isocenter for the cases presenting the best and the worst agreement are presented in Figure 5.4b and Figure 5.4c, respectively.

#treatments	γ-mean	γ-1%	γ-pass rate	%ΔISOC
5 IMRT	0.37 ± 0.09	0.99 ± 0.16	98.7 ± 1.8	-0.3 ± 2.0
	(0.28, 0.51)	(0.85, 1.23)	(95.5, 99.8)	(-2.6, 1.8)
5 VMAT	0.36 ± 0.07	1.00 ± 0.21	98.3 ± 2.2	-0.1 ± 1.2
	(0.26, 0.45)	(0.79, 1.32)	(94.6, 99.9)	(-2.2, 1.1)

Table 5.1: Comparison between 3D dose distribution measured with the OCTAVIUS 4D system and the corresponding 3D dose distribution reconstructed with the virtual back-projection algorithm. Results are presented as AVG ± (1SD) with the range indicated between parentheses.

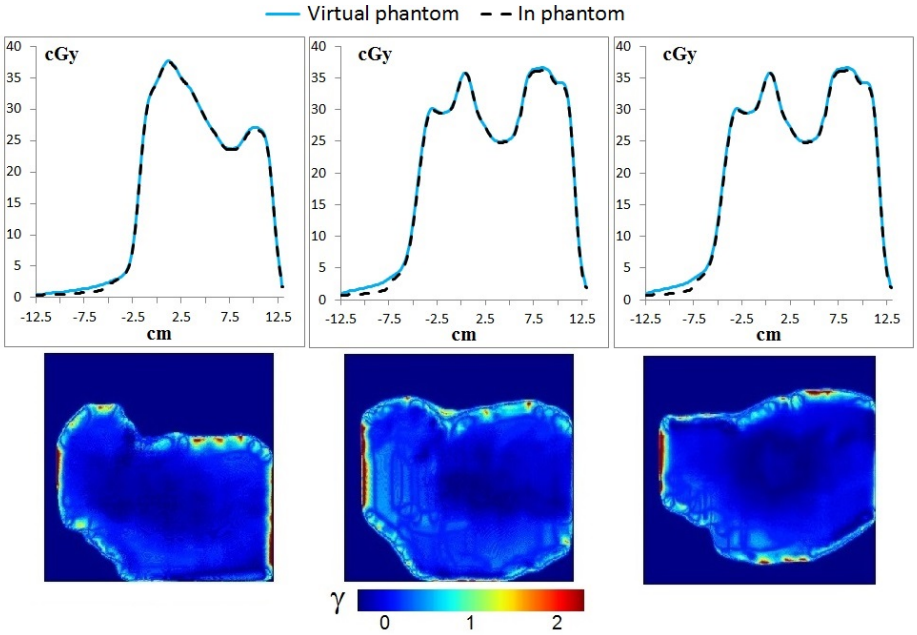


Figure 5.3: Comparison results between virtual phantom and in phantom dose reconstruction for three of the seven fields having the highest modulation of an inguinal lymph nodes IMRT plan delivered to the 20 cm thick pslab phantom. Cross-plane dose profiles and 2D γ-distributions (local 2%/2mm) through the isocenter are displayed for every field. The average dose difference at the isocenter and the average γ-pass rate values were -0.2 $\pm$ 0.3%(1SD) and 94.2 $\pm$ 1.6%(1SD) respectively.

5.3.3 Verification of virtual patient dose reconstruction against TPS

An overview of the results of the comparison between virtual and planned patient 3D dose distributions is shown in Table 5.2. In 95% of the treatments, the γ-mean values were lower than 0.53, the γ-1% values lower than 1.71 and the γ-pass rates higher than 92.5%. The difference between the virtual and the planned median dose to the PTV remained within 4% in all 75 treatments. Histograms with the γ-pass rates and the differences in the PTV-D_{50} values are displayed in Figure 5.4.

One head-and-neck VMAT case showed the lowest γ-pass rate (86.4%). This case presented also the highest γ-mean value (0.62) and the largest PTV-D_{50} difference (-3.9%). Another head-and-neck VMAT case presented the highest γ-1% value (2.64).

#treatments	γ-mean	γ-1%	γ-pass rate	%ΔPTV-D_{50}
25 IMRT	0.37 ± 0.08	1.04 ± 0.24	98.0 ± 2.6	-0.4 ± 1.1
	(0.24, 0.54)	(0.66, 1.51)	(92.3, 100)	(-2.6, 1.7)
50 VMAT	0.42 ± 0.08	1.27 ± 0.37	96.7 ± 2.9	-0.7 ± 1.3
	(0.29, 0.62)	(0.63, 2.64)	(86.4, 100)	(-3.9, 3.3)

Table 5.2: Comparison between virtual and planned patient 3D dose distributions. Results are presented as AVG ± (1SD) with the range indicated between parentheses.

Site	Technique	OAR1	%ΔOAR1-D_2	OAR2	%ΔOAR2-D_2
Lung	IMRT	Oesophagus	2.2	Spinal cord	2.8
Oesophagus	IMRT	Spinal cord	0.2	Heart	1.6
Rectum	VMAT	Bowel	0.7	Bladder	-4.6
Lung	VMAT	Oesophagus	-4.4	Spinal cord	-2.1
Prostate	VMAT	Rectal wall	1.8	Rectum	1.1
Bladder	VMAT	Bowel	0.9	Bowel ring	5.0
Cervix	VMAT	Bowel	4.6	Bladder	-0.2
Head & Neck	VMAT	Spinal cord	-0.9	Parotid gland	-1.3
Brain	VMAT	Brain stem	0.3	Eye	4.9
Stomach	VMAT	Liver	0.8	Spinal cord	-2.1

Table 5.3: Difference in the dose-volume histogram parameter OAR-D_2 obtained with virtual and planned patient 3D dose distributions for ten selected treatments.

These two cases will be discussed further in the discussion section.

An overview of the comparison between 20 OAR-D_2 values for 10 selected treatments is presented in Table 5.3. All results were within 5% with an average difference of 0.5 ± 2.7(1SD).

5.3.4 Comparison between virtual and *in vivo* patient dose verification results

It is apparent from the results presented in Figure 5.5 that *in vivo* patient reconstruction show worse agreement with the TPS than virtual patient reconstruction. The average increase in γ-mean value was 0.07±0.1(1SD) (-0.13, 0.37). The average absolute difference in %ΔPTV-D_{50} value was 1.1±1.0%(1SD) (0.1, 4.65). The two cases with the largest observed deviations will be discussed further in the Discussion section.

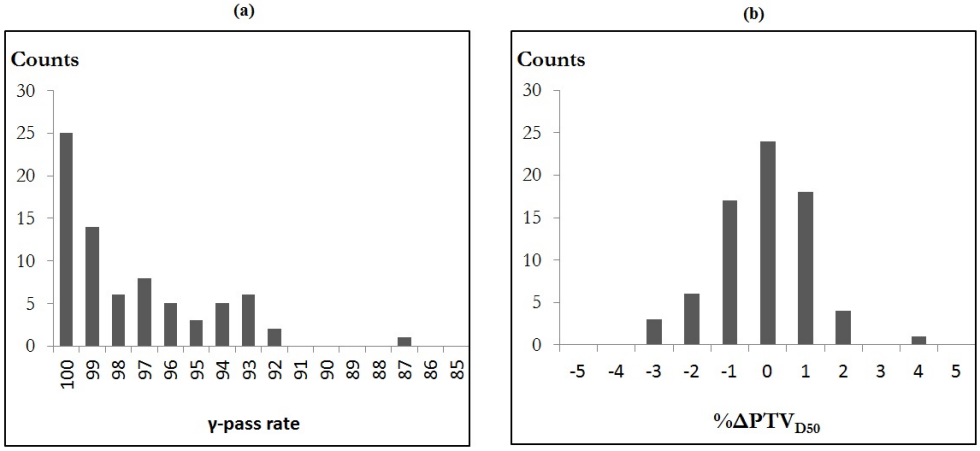


Figure 5.4: Figure (a) displays the distribution of γ-pass rate results in the verification of 75 virtual patient dose distributions corresponding to 25 IMRT and 50 VMAT treatments. Only one case showed a γ-pass rate lower than 90%. Differences in the dose-volume histogram parameter PTV-D_{50} between the virtual and the planned patient 3D dose distributions are shown in Figure (b).

5.4 Discussion

In this study, we have demonstrated how the virtual back-projection algorithm allows for patient dose reconstruction using in air EPID measurements of the treatment; hence making patient 3D dosimetry for IMRT and VMAT pre-treatment verification possible. The results presented in section 5.3.1 showed that the virtual and the transit back-projection algorithms yield dose distributions that are equivalent within 1%. Note that the reproducibility of the EPID signal readings has been measured to be $\pm 0.5\%$(1SD) [43].

The virtual dose reconstruction engine was further verified against direct dose measurements (calibrated Semiflex IC and OCTAVIUS 4D system) and against TPS planned patient data. In our opinion, the results presented in this study are considered to be sufficiently accurate to guarantee the correctness of virtual dose reconstructions. The presented γ-pass rate results are comparable to results published in recent studies for patient-specific pre-treatment IMRT and VMAT QA [80, 81, 82, 49]. This γ-pass rate metric has been commonly used by physicists when comparing measured and planned dose distributions as discussed in the AAPM TG-119 report [83, 84] which proposed

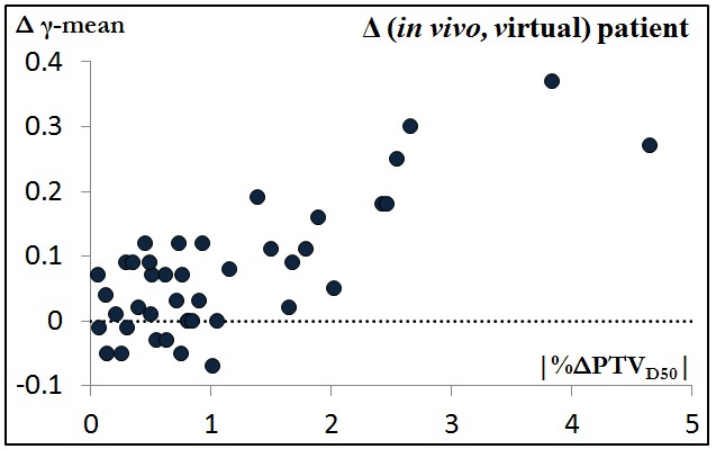
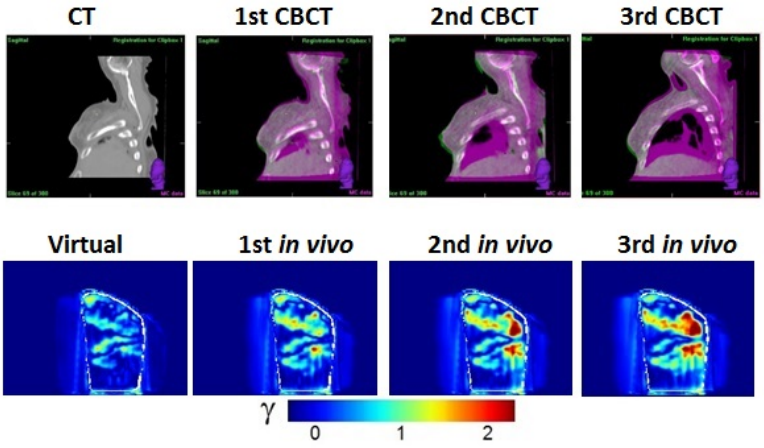


Figure 5.5: Scatter plot with the differences in γ-mean verification values and the absolute differences in %ΔPTV-D_{50} values obtained with virtual and *in vivo* dose reconstruction for 16 IMRT and 27 VMAT treatments.

action levels of 90%. To our knowledge, there are not yet standardized recommendations on acceptable tolerance levels based on deviations of DVH parameters. The observed deviations in this study, median PTV doses within 3-4% and near maximum OAR doses within 5%, compared well to results found in other studies [32, 38, 85]. These results give us confidence to implement virtual patient dose reconstruction in our clinic for patient specific pre-treatment verification of IMRT and VMAT treatments. Future work will include a (per-site) sensitivity and specificity study aiming at determining the most adequate alert criteria for virtual dose reconstruction in both gamma and DVH metrics.

Because the same algorithm is used, virtual and *in vivo* patient dose distributions are similarly affected by the machine, planning and model errors of the specific treatment. Since the effects of patient related errors such as anatomical changes and/or setup issues are specific to each fraction, the observed differences between virtual and *in vivo* dose distributions provide us with a first estimation of the magnitude of the patient-related errors for each delivered fraction. In this study, the largest difference in γ-mean values (0.37) corresponded to the third fraction of a 7-field IMRT lung treatment which presented a strong case of decrease in atelectasis. In Figure 5.6 it can be seen how the difference between *in vivo* and virtual patient results becomes larger as the changes in lung density increase. After the result of the 3rd *in vivo* verification the radiation oncologist was consulted.



γ-mean	0.46	0.58	0.72	0.83
γ- 1%	1.44	1.76	2.41	2.71
γ-pass rate	92.7%	84.6%	75.4%	68.7%
%ΔISOC	0.2%	0.8%	1.4%	2.8%
$\%\Delta PTV_{D50}$	-0.4%	1.0%	2.0%	3.4%

Figure 5.6: Top: left image shows the sagittal planning CT scan used for pre-treatment verification. The other images show the matching of the planning CT (purple) and sagittal CBCT (green) images made before the first three *in vivo* verifications of a 7-field IMRT lung treatment. Middle: TPS vs EPID γ-distributions for the pre-treatment (virtual) and the *in vivo* EPID measurements. Values are shown in the planning CT sagittal plane through the isocenter. The 50% isodose line is shown in white. Bottom: EPID verification results.

The case with the largest difference in PTV-D_{50} values (4.65%) corresponded to the first fraction of a cervix VMAT treatment. By visual inspection of the CBCT scan of that day, our IGRT therapists observed strong anatomical changes in bladder filling and gas pockets in this fraction, and the case was examined by a medical physicist before treatment was continued. The example presented in Figure 5.7 raised an *in vivo* alert in our clinic due to a high γ-1% value of 2.72. A comparable value of 2.64 in the virtual patient reconstruction results suggested that the effects of patient-related errors on the *in vivo* deviation were small in this case. The same in air measurements were used for virtual phantom reconstruction and a much lower value of γ-1% (1.13) was obtained, indicating that the machine delivered the plan correctly and that the TPS calculated the

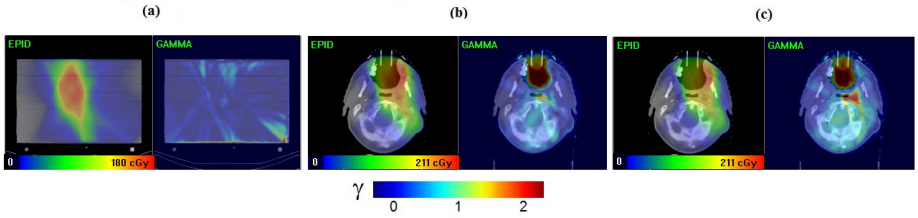


Figure 5.7: EPID dose and γ-evaluation distributions for (a) the virtual phantom reconstruction, (b) the virtual patient reconstruction and (c) the *in vivo* patient reconstruction of the VMAT head-and-neck treatment arc with the highest γ-1% value in Table 2. The specific geometry of the patient in this treatment is representative of the in-out-in limitation of our back-projection model and explains the high γ-1% values of 2.64 and 2.72 obtained in the virtual and in the *in vivo* patient reconstructions, respectively.

dose correctly within a phantom. The deviation was explained by a (rare) model error in the back-projection algorithm when reconstructing to this specific patient anatomy in which the beam goes in the patient, then out in an air cavity and then in tissue again. As a result, the EPID reconstructed dose was being wrongly deposited inside the mouth cavity. The case with the lowest γ-pass rate in Table 5.2 (86.4%) presented comparable virtual phantom and *in vivo* results of 87.1% and 86.7%, respectively. This suggested that the effects of patient-related errors and the effects of model (or TPS) errors when reconstructing to the patient anatomy were small. This H&N VMAT treatment delivered highly modulated fields at low dose rates. A possible explanation for this deviation might be a limitation of our model to accurately predict the dose under these conditions. However, in order to completely rule out possible machine or planning errors, measurements using another type of 3D verification, for instance using a phantom having an array of detectors, should in addition be performed.

In summary, besides using the virtual patient 3D dose reconstruction as an efficient pre-treatment verification method, virtual dose reconstructions prove to have additional value in the inspection process of *in vivo* deviations. The combination with *in vivo* dosimetry measurements offers a unique approach in separating patient-related errors from errors due to shortcomings in the back-projection algorithm or in the dose calculation of the TPS, errors in the transfer of the TPS data to the accelerator, or errors in the delivery of the plan itself. This is particularly true in situations where cone beam CT data is not available (or not conclusive).

5.5 Conclusions

A new extension of our back-projection algorithm allows us for the reconstruction of pre-treatment 3D dose distributions within the patient anatomy using EPID in air measurements of the treatment. Virtual patient dose reconstruction makes pre-treatment verification based on deviations of DVH parameters feasible and eliminates the need for phantom positioning and re-planning. In this study the accuracy of this new method was assessed with measurements of 25 IMRT and 50 VMAT treatments. The results presented in this study proved that virtual dose reconstruction is a valid solution for patient-specific pre-treatment verification of IMRT and VMAT treatments. A unique advantage of this approach is that patient-related errors can be separated from machine, planning and model errors by combining virtual and *in vivo* patient verification data for the same treatment.

Acknowledgements

The authors would like to thank Begoña Vivas Maiques for performing ionization chamber and OCTAVIUS dose measurements.

Chapter 6

Online 3D EPID-based dose verification: proof of concept

H. Spreeuw and R.A. Rozendaal[1]
I. Olaciregui-Ruiz
P. González
A. Mans
B. Mijnheer
M. van Herk

Medical Physics 43 (7), 2016

[1]H. Spreeuw and R.A. Rozendaal contributed equally.

Abstract

Purpose

Delivery errors during radiotherapy may lead to medical harm and reduced life expectancy for patients. Such serious incidents can be avoided by performing dose verification online, i.e., while the patient is being irradiated, creating the possibility of halting the linac in case of large over- or underdosages. The offline EPID-based 3D *in vivo* dosimetry system clinically employed at our institute is in principle suited for online treatment verification, provided the system is able to complete 3D dose reconstruction and verification within 420 ms, the present acquisition time of a single EPID frame. It is the aim of this study to show that our EPID-based dosimetry system can be made fast enough to achieve online 3D *in vivo* dose verification.

Methods

The current dose verification system was sped up in two ways. First, a new software package was developed to perform all computations that are not dependent on portal image acquisition separately, thus removing the need for doing these calculations in real time. Second, the 3D dose reconstruction algorithm was sped up via a new, multi-threaded implementation. Dose verification was implemented by comparing planned with reconstructed 3D dose distributions delivered to two regions in a patient: the target volume and the non-target volume receiving at least 10 cGy. In both volumes the mean dose is compared, while in the non-target volume the near-maximum dose (D_2) is compared as well. The real-time dosimetry system was tested by irradiating an anthropomorphic phantom with three VMAT plans: a 6 MV head-and-neck treatment plan, a 10 MV rectum treatment plan and a 10 MV prostate treatment plan. In all plans, two types of serious delivery errors were introduced. The functionality of automatically halting the linac was also implemented and tested.

Results

The pre-computation time per treatment was $\sim$180 s per treatment arc, depending on gantry angle resolution. The complete processing of a single portal frame, including dose verification, took 266$\pm$11 ms on a dual octocore Intel Xeon E5-2630 CPU running at 2.40 GHz. The introduced delivery errors were detected after 5 to 10 s irradiation time.

Conclusions

A prototype online 3D dose verification tool using portal imaging has been developed and successfully tested for two different kinds of gross delivery errors. Thus, online 3D dose verification has been technologically achieved.

6.1 Introduction

With the introduction of amorphous-silicon (aSi) EPIDs, the interest in EPID dosimetry has risen because of their favorable characteristics such as fast image acquisition, high resolution, digital format, and potential for *in vivo* measurements and 3D dose verification [38]. Several groups have implemented dosimetry using EPIDs for 3D pre-treatment verification [86] as well as for (3D) *in vivo* dose verification, e.g. [87, 24, 25, 88, 89]. Also, a real-time EPID-based verification system based on comparing predicted to measured portal images has been implemented clinically [35]. Furthermore, a quasi real-time system also based on comparing predicted to measured portal images, augmented with an isocenter dose comparison, has been used in clinical routine [90, 91]. In our study a new method is proposed, which combines all of these aspects: real-time EPID-based 3D *in vivo* dose verification. The aim of this feasibility study is to show that we are capable of reconstructing cumulative delivered 3D dose distributions in real time and generating triggers to stop the linac in case of a large deviation from the total planned dose distribution.

6.2 Methods and Materials

6.2.1 Clinical system overview

In the Netherlands Cancer Institute (NKI), aSi EPIDs (Perkin Elmer RID 1680 AL5/ Elekta iViewGT) mounted on SL20i and Agility linear accelerators (Elekta, Crawley, UK) are utilized for dose verification measurements. The software for the acquisition of portal frames was developed in-house and allows for the simultaneous extraction of gantry information, e.g., the gantry angle. A backprojection algorithm based on EPID transmission is applied for 3D *in vivo* dose verification of all treatments [87, 41, 22] which have treatment fields that fit on the EPID and would not cause collisions between the EPID and the couch. All dose reconstructions are compared with the 3D dose distributions calculated by the clinically used treatment planning system (Pinnacle 9.8, Philips Medical Systems, Eindhoven, The Netherlands).

In current clinical practice, the actual dose verification is done offline, thus after a fraction has been delivered. This implies that some very large delivery errors can not be corrected for in the remaining fractions by altering the treatment plan, which is especially relevant when organs at risk have been pushed close to their respective maximum tolerated doses. A complete overview of the portal dosimetry workflow at our

institution is given elsewhere [87, 27].

6.2.2 Online EPID-based 3D dose verification

Our prototype system for real-time dosimetry was designed not only with the aim to perform dose verification faster than the frame rate of our EPID (2.37 frames per second), but also to immediately halt the linac in case of serious delivery errors as depicted in Figure 6.1. During dose delivery, frames are acquired using the EPID acquisition software; acquired frames are saved to a network location. The real-time dose reconstruction software reads the frames, reconstructs the 3D dose distribution and performs dose verification. Whenever a dose deviation outside of tolerance levels is detected, a signal is sent to a linac monitoring software package. This monitoring software was also developed in-house and is capable of breaking an interlock chain, which halts the linac. The interlock chain is a physical chain of electrically-operated switches; the linac is physically incapable of generating radiation whenever one or more of these switches is not closed. On the linac used for testing, a custom-made switch was added to the vendor-supplied interlock-chain. It is this switch which was controlled by the monitoring software.

6.2.3 Accelerated dose reconstruction

Our 3D dose reconstruction method uses information derived from the planning CT-scan. As this data is independent of the acquired portal images, it is possible to compute these inputs to the algorithm before the dose delivery starts. A separate software package was written to generate these inputs. The dose reconstruction is dependent on the gantry angle, so inputs have to be generated for several gantry angles. It was decided to use a resolution of one degree, generating 360 sets of input data.

Additionally, the back-projection algorithm was an excellent candidate for multi-threading. This algortihm back-projects the dose at the EPID level to a stack of planes through the patient. As the scattered dose in the patient is modelled as lateral scatter in each plane, the total dose in each plane could be computed independent of the other planes, making a multi-threaded implementation straightforward. The implemented module for multi-threaded back-projection used as many threads as available logical CPU cores. After back-projection, the resulting dose distribution was in gantry-coordinates and needed to be rotated and resampled to machine coordinates. This transformation was re-implemented in a multi-threaded way as well. All multi-threaded code was implemented using OpenMP [92]; the Intel OpenMP runtime library 5.0 was used.

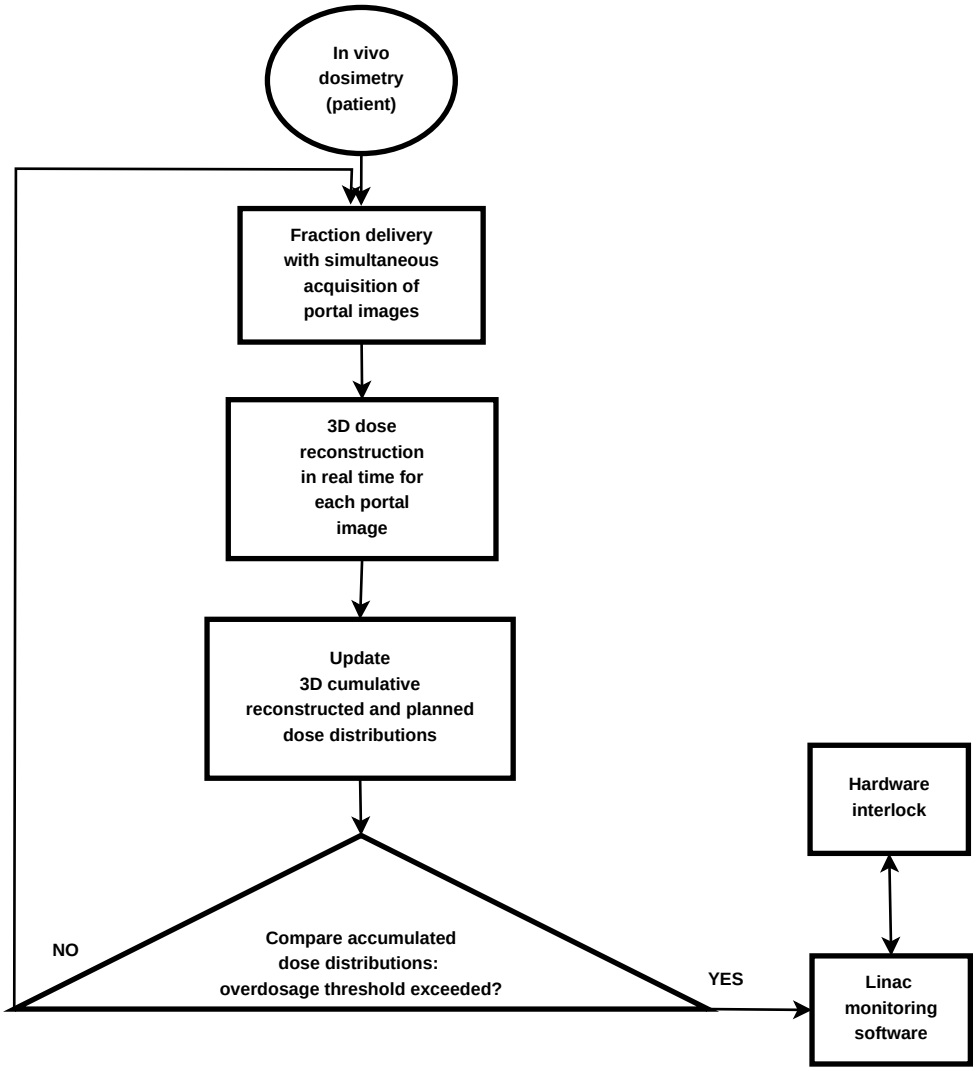


Figure 6.1: Real-time EPID-based *in vivo* dosimetry procedure, ideally performed for all fractions of a treatment and complementary to the more sensitive offline EPID-based *in vivo* dosimetry method used for the detection of smaller errors. On the right the communication flow for treatment discontinuation is depicted.

An extra constraint for the accelerated dose engine was backward compatibility with our existing dose reconstruction method. Given the same input data, both methods were required to reconstruct the same dose in every voxel to an accuracy of 1 ppm. A consequence of this requirement is that all previously published results [41, 26, 22, 23, 7, 58, 93, 27] are directly applicable to the accelerated dose reconstruction system. Specifically, it has been shown that the mean delivered dose in the PTV can be determined to an accuracy of about 3% for various treatment sites [58, 93].

6.2.4 Real-time dose comparison to detect gross errors

In the real-time dose verification system the linac will be halted when large over- or underdosages are detected. The method for quantifying a large dose deviation should preferably have a clear clinical interpretation, making the widely adopted γ-analysis [28] an unfavorable candidate. As the 3D *in vivo* dose distribution was calculated, an obvious choice would be to inspect dose parameters of specific subvolumes, similar to plan evaluation criteria. The presented real-time dosimetry method is capable of dose verification in any defined subvolume of the reconstructed dose distribution. In this proof of concept, the patient volume was divided into two subvolumes: target and non-target volume. The planning target volume (PTV) was used as the target volume, and the non-target volume consisted of the volume outside of the PTV receiving at least 10 cGy. The dose threshold of 10 cGy was added to prevent a large volume receiving a very low dose from masking overdosages. For the non-target volume, both local and global dose deviations were detected by inspecting the mean dose and the 98th percentile (D_2) in the dose-volume histograms (DVHs). It should be noted that the DVH-parameters selected for inspection do not affect the calculation speed: the complete DVH was calculated for each specified volume, making every DVH-parameter eligible for inspection. The dose distribution in the relatively small PTV was simply verified by inspecting only the mean dose. A priori, it is not obvious what would constitute a large enough deviation to halt the linac. As a starting point, tolerance levels for the mean dose in the PTV and D_2 of the non-target volume are set to 10% of the prescribed fraction dose. For the mean dose in the non-target volume, which is much lower than the previous two dose parameters, a tolerance level of 5% of the prescribed fraction dose was chosen.

The planned dose distribution, to which the reconstructed dose was compared, was specified per control-point - in contrast to the EPID frames which were read-out continuously during treatment. The reconstructed 3D dose distribution was compared to the

accumulated planned 3D dose distribution, which was updated every time the gantry reached the next control-point.

6.2.5 Testing

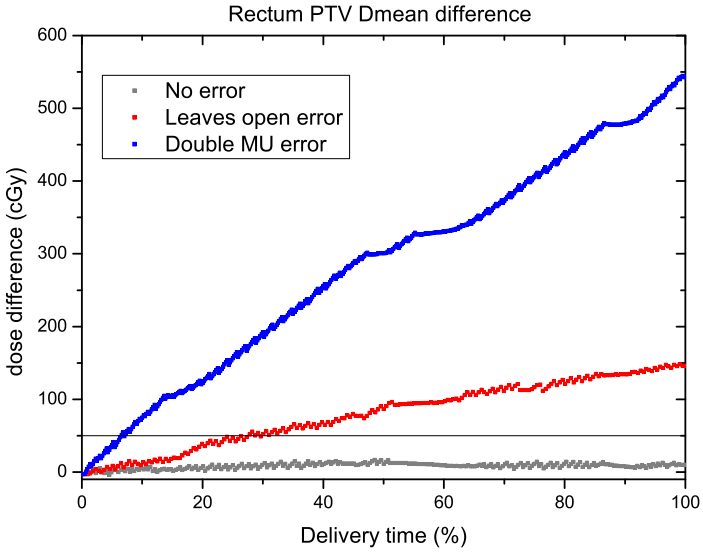
The new system was tested by irradiating three different VMAT plans on an anthropomorphic (Alderson) phantom: a dual-arc 6 MV head and neck (H&N) plan, a single-arc 10 MV rectum plan and a single-arc 10 MV prostate plan. All plans were specifically created for the Alderson phantom: a virtual tumor and organs at risk (OARs) were delineated for each case and treatment plans were created according to our clinical constraints by an experienced treatment planner. The H&N and prostate plans were to deliver 2 Gy per fraction; the rectum plan 5 Gy per fraction. Modified plans were created by introducing two types of serious delivery errors in these plans. First, the number of monitor units to be delivered per control point was doubled; second, all the moving MLC leaves and jaws were retracted 10 cm on both sides, creating a large open field.

All plans were irradiated and analysed in real time using the newly developed software. Before irradiation, the anthropomorphic phantom was positioned using the clinical, CBCT-based, protocol of our institute. The dose to the OARs at the detection threshold was determined after irradiation by using the same real-time dosimetry method, only now using the already recorded EPID movies as input. These recorded EPID movies consist of the exact same EPID frames recorded and analysed during plan delivery.

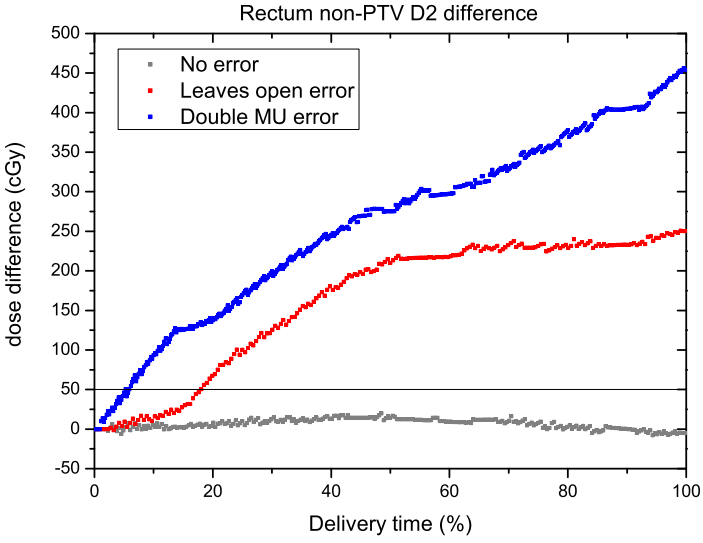
6.3 Results

6.3.1 Real-time dosimetry

Figures 6.2a, 6.2b and 6.2c show differences between planned and reconstructed dose distributions for the target and non-target volumes during delivery of the rectum VMAT plan. The introduced errors are clearly discernible. The other two VMAT plans showed similar results. Table 6.1 lists key features of the irradiated plans. For the introduced errors and using the proposed tolerance levels, the linac would have been halted after 5 to 10 s irradiation time; for the deliveries without errors, no deviations above the threshold were detected. The doses to OARs at the earliest detection times for each introduced error are listed in Table 6.2.

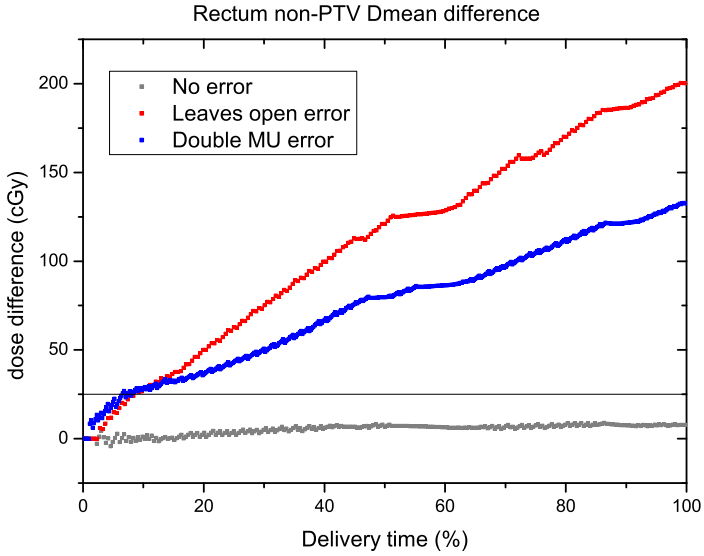


(a) PTV, D_{50}



(b) non-PTV, D_2

Figure 6.2: Difference between reconstructed and planned dose for the three scenarios (no error, leaves open error, double MU error) and three different ROI/indicator combinations, as a function of delivery time. The horizontal lines indicates the detection threshold.



(c) non-PTV, D_{mean}

Figure 6.2: Difference between reconstructed and planned dose for the three scenarios (cont.).

Site	Dose parameter	Tolerance (cGy)	Linac stopped after (s) leaves open	Linac stopped after (s) MU x2
H&N	PTV Dmean	20	16 (23%)	16 (22%)
	non-PTV D_2	20	22 (33%)	7 (9.2%)
	non-PTV Dmean	10	**5 (6.9%)**	**5 (6.5%)**
rectum	PTV Dmean	50	25 (28%)	12 (6.6%)
	non-PTV D_2	50	16 (18%)	**10 (5.2%)**
	non-PTV Dmean	25	**8 (8.8%)**	12 (6.6%)
prostate	PTV Dmean	20	8 (15%)	13 (13%)
	non-PTV D_2	20	6 (11%)	**9 (8.6%)**
	non-PTV Dmean	10	**5 (8.2%)**	**9 (8.6%)**

Table 6.1: Detection performance of the different dose parameters. Indicated is the irradiation time after which the tolerance level is reached, with the percentage of the full delivery time of the treatment arc in parentheses. When the linac would be halted for each introduced error is indicated in bold. In the no-error situation, the tolerance level is not reached.

6.3.2 Timing

An essential requirement for online dose verification is that it should be done in real time, i.e. an EPID-based dose measurement and verification should be completed before the next portal image is taken. The rate at which portal images are acquired (frame rate) depends on the acquisition mode, but is about 420 ms for the irradiations investigated in this research. Table 6.3 shows that the average per-frame processing time (266 ms) is well below the frame rate. Calculation and comparison of DVH-parameters for the PTV and all defined OARs did not significantly change the time needed for dose comparison. Generation of the pre-computed data is not time-critical: it only needs to be done once per treatment and can be done offline, well in advance of the treatment start. At one degree gantry angle resolution, generating the pre-computed data took about 3 minutes per treatment arc (0.5 s per gantry angle) on a standard workstation; this dataset was typically 1 to 2 gigabytes large.

The EPID-based 3D dose distribution was initially calculated on the conical backprojection grid, which rotates with the gantry since it is fixed to the beam. This grid was comprised of 256×256 pixel slabs parallel to the EPID. The number of slabs is dependent

Site	Organ	Parameter	Dose at detection threshold (cGy) leaves open	MU x2
H&N	base of tongue	D_{50}	19	6
	brainstem	D_0	22	16
	constrictor muscles	D_{50}	22	6
	larynx	D_{50}	21	6
	mandible	D_0	23	27
	oral cavity	D_{50}	16	7
	left parotid gland	D_{50}	16	16
	right parotid gland	D_{50}	8	1
	spinal cord	D_0	26	26
rectum	bladder	D_0	69	82
	bladder	D_{50}	56	65
	bowel area	D_0	75	87
prostate	anal sphincter	D_{50}	13	9
	bladder	D_0	13	39
	bladder	D_{50}	9	14
	bowel area	D_0	15	43
	rectal wall	D_0	15	28
	rectum	D_0	15	28

Table 6.2: Dose to organs at risk at the detection threshold for both introduced errors.

Pre-computed data read	3D dose calculation	Dose comparison	Total
38 ± 5	67 ± 3	161 ± 8	266 ± 11

Table 6.3: Processing times (ms) per portal image for online dosimetry, averaged over all 1960 frames of all nine performed irradiations (no error, leaves open error, double MU error for all three treatment plans). The PC used was equipped with a dual octocore (16 physical cores, 32 logical cores) Intel Xeon E5-2630 CPU, running at 2.40 GHz. Indicated uncertainty is one standard deviation.

on the gantry angle and varied between 51 and 83 for the arcs considered in this study, hence the number of nodes in this grid varied between 3.3×10^6 and 5.4×10^6 . The backprojection grid had a spacing of 1 mm $\times$ 1 mm in the midplane and 5 mm in the perpendicular direction. Each backprojected dose distribution was transformed using an affine transform which incorporated bilinear interpolation and added to the cumulative reconstructed dose distribution on the final computational grid in machine coordinates. In this study the final grids were equal to the TPS grids and had a voxel size of 2 mm in all dimensions.

6.3.3 An automated linac halt

The new software package for real-time 3D dose verification is able to send a signal to our in-house developed linac monitoring software. This monitoring software then breaks a hardware interlock to actually halt the linac. During testing of the newly developed system, the sending of this halting signal could be enabled or disabled at will. With the sending enabled, the linac was succesfully halted by the real-time 3D dose verification system for each of the treatment plans with large introduced errors included in this study. Note that in the results shown in Figures 6.2a-6.2c the linac obviously was not halted.

6.4 Discussion

It has been shown that real-time EPID-based 3D dose reconstruction and verification is possible using our back-projection system. The online dose verification system has been succesfully combined with a linac halting mechanism, enabling linac halts whenever deviations exceeding the detection threshold are found. The deviations introduced for testing purposes were intended to be large and easily discernible; the 5% and 10% tolerance levels used in this phantom study have proven to be able to detect the introduced deviations early during delivery. Table 6.2 shows that none of the OARs would have received high doses with the introduced delivery errors: the absolute overdoses are maximally in the order of 90 (rectum), 40 (prostate) and 30 (H&N) cGy. Assuming that such errors would have been detected at the first fraction, only minor – if any – modifications to the intended treatment plan would have been needed. A close inspection of the measured dose differences shows that the tolerance levels could have been even tighter, as figure 6.2c clearly shows. However, analysis of actual online *in vivo* data is needed to assess the variability in the no-error dose reconstruction when a patient is in the beam.

Dose verification in any subvolume of the patient, such as OARs, is supported by the method. However, finding appropriate tolerance levels for halting the linac is not trivial and will be subject of further investigations. This study has presented the detection of overdosages, but the system is equally capable of identifying underdosages.

A concern might be the impact of real-time dosimetry on the life expectancy of the EPID. Compared to offline dosimetry, the life expectancy of the EPID will be similar, as the EPID is used for both methods in the same way. In our clinic, where every treatment is verified using offline dosimetry, an EPID lasts on average for 32 months. A deterioration of the EPID image quality, not a dosimetric failure, is generally the reason for its replacement [27].

Compared to other EPID-based real-time treatment verification methods [35, 90, 91], the presented method is different in that the delivered 3D dose distribution is calculated and verified, instead of a comparison of portal images. Though both methods are suitable to detect the largest of deviations in dose delivery, verification of the delivered 3D dose distribution provides more insight in the relevance of detected deviations as they can be expressed in differences in DVH-parameters. Also, determination and interpretation of tolerance levels is more straightforward when discussing patient 3D dose distributions than, for example, pass rates at the EPID level, which is the method used by Woodruff [35]. For example, a constant dose deviation which is recorded at the EPID at a constant, off-axis position during delivery will appear as a deviation of constant magnitude when comparing predicted and measured portal images. In 3D, the observed dose deviation will not be constant but rather be distributed over the patient volume, which aids in determining the relevance of the delivered dose deviation. Even so, further investigations are necessary to gain insight in the sensitivity and specificity of both methods.

The presented method has been validated using VMAT treatments, but other treatment techniques (IMRT, 3D CRT) are equally well supported. In fact, the real-time dose verification software developed is unaware of the specific treatment technique used for dose delivery, it simply takes EPID frames and reconstructs dose.

6.5 Conclusions

It is shown that 3D planned dose distributions from VMAT treatments can be verified online, i.e., during treatment, by means of DVH-analysis. This was done using 3D EPID transit dosimetry in real time, i.e. performing a 3D dose verification faster than the portal frame acquisition rate. This enables linac halting without unnecessary delays, which was

demonstrated for two serious delivery errors and three VMAT treatment plans.

Acknowledgments

We thank René Tielenburg, Lennert Ploeger, Emmy Lamers, Martijn Barsingerhorn, Nelly Kager, Maarten Buiter and Peter Remeijer for extensive support during this study.

Chapter 7

Real-time 3D *in vivo* dosimetry for lung SBRT VMAT

R.A. Rozendaal
B. Mijnheer
P. González
I. Olaciregui-Ruiz
J.-J. Sonke
A. Mans
M. van Herk

Radiotherapy and Oncology *(under review)*

Abstract

Purpose

Currently, real-time 3D *in vivo* dosimetry is technologically feasible using EPIDs. Clinical introduction however, requires more than speed: it requires high specificity (robustness) and sensitivity (usefulness). It is the purpose of this study to investigate dose discrepancies observed during 3D *in vivo* dosimetry and establish tolerance levels for real-time 3D dosimetry for lung stereotactic body radiation therapy (SBRT) VMAT treatments and investigate the detectability of introduced errors.

Materials and Methods

EPID data recorded during delivery of 256 fractions of 105 lung cancer patients were included. All fractions were delivered acceptably according to our current clinical off-line QA protocols and delivered 15-18 Gy to the target volume. For each acquired EPID frame, a 3D *in vivo* dose distribution was reconstructed using the real-time method and compared to the cumulative planned dose distribution via dose-volume histogram (DVH) analysis. The maximum observed deviations in the plans without errors were used for determining tolerance levels; subsequently large errors, -20% – +100% in monitor unit (MU) and 6 – 18 mm in leaf-position, were introduced to demonstrate the sensitivity of the method.

Results

Mean dose to the target volume was accurate to at least 2.8% (one s.d.) during delivery. Outside the target volume systematic deviations up to 6% are present. Tolerance levels varied, depending on the region of interest, from 0 - 2% at the start of delivery to 9 - 26% at the end of treatment. Specificity of the method was 1 by definition; sensitivity varied from 0.96 to 1.

Conclusions

In this study we demonstrated the usefulness of a real-time *in vivo* dosimetry method for lung SBRT VMAT. The method combines high specificity with high sensitivity and shows that clinical introduction of real-time 3D DVH-based *in vivo* dosimetry for lung SBRT VMAT is feasible.

7.1 Introduction

Patient safety has a high priority in radiotherapy, a fact well illustrated by the wide range of quality assurance (QA) protocols and systems in use. Among these QA systems, real-time dosimetry offers the possibility of near-immediate detection of large delivery errors. Combined with an intervention system capable of interrupting the dose delivery of the linac, this allows for preventing harm to the patient in case of a serious error; several of which have been reported in the literature [94, 6, 5].

Currently, real-time dosimetry is still in an early stage of development. In the only clinical study published, Woodruff et al. [35] demonstrated real-time treatment verification based on comparing predicted portal images with portal images acquired during treatment delivery. Our goal is to take the real-time dose verification a step further by comparing - in real time - the delivered 3D dose distribution to the accumulating planned 3D dose distribution, as 3D dose verification provides more clinically insightfull information than dose deviations at the portal imager level. Technologically, real-time 3D dose reconstruction and verification has been shown to be feasible [95].

A clinical implementation requires more than just a fast system, however. Key requirements for practical real-time dosimetry are robustness and usefulness. The system needs to be robust in the sense of a high specificity, or, equivalently, a low false-positive rate. False positives in real-time dosimetry have serious consequences in the clinical workflow as it interrupts a treatment session leading to time loss and stress. Typically, high specificity can be accomplished by choosing tolerance levels which are not too strict, thus focusing on gross errors. At the same time though, the system needs to remain useful in the sense that it should have a high sensitivity, i.e., the ability to actually catch large errors on-line. Thus, the dominant issue for a clinical implementation of real-time dosimetry is the definition of tolerance levels. This is of course directly related to the accuracy of the method used and the motivation for this study.

7.2 Materials and Methods

7.2.1 Overview

To arrive at tolerance levels for real-time 3D *in vivo* dosimetry, we devised a simple method based on the use of historical data, i.e. EPID frames, acquired during routine treatment of patients. We applied the method for lung stereotactic body radiation therapy (SBRT) delivered using volumetric modulated arc therapy (VMAT); this treat-

ment site/technique combination was chosen based on clinical interest. Tolerance levels were determined based on the observed uncertainties. The sensitivity of the system with these tolerance levels was investigated for a series of errors introduced in the EPID frames before performing the dose reconstruction. Familiarity with our EPID-based back-projection dose reconstruction method is assumed in the remainder of this article. In short, the method takes an EPID frame, a set of model parameters and a CT-scan and calculates the reconstructed 3D dose distribution which is the *in vivo* dose distribution if the EPID frame was taken during treatment. The details of the method are extensively described elsewhere [95, 22, 23, 45]. Note that the *in aqua vivo* extension of the dose reconstruction model [45] is used, because the treatment site contains large inhomogeneities (i.e., lungs).

7.2.2 Method of tolerance level derivation

The basic principle of the method for tolerance level determination is to pick a set of EPID frames recorded during error-free dose deliveries and set the tolerance levels to the largest observed deviations between reconstructed and planned dose, optionally increased by a robustness-margin, i.e. an increase of the tolerance level of, say, 10 – 20%. This way, the specificity of the method is one by definition when applied to the data-set and the upper limit for the false positive rate (FPR) is on the order of $1/N$, with N the number of treatments included provided that the included treatments were chosen randomly. A complicating factor is that the dose deviations are not expected to be constant during dose delivery: random errors will average out over time and the dose deviations are thus expected to become smaller with delivered dose. Therefore, a single value as tolerance level would make the system quite insensitive and the tolerance levels were therefore made time-dependent instead. Dose deviations were inspected as a function of delivered dose, 0 – 100% in steps of 1%. Because dose deviations are only expected to become smaller with delivered dose and never larger, the tolerance levels were also required to be monotonic decreasing functions. Deriving tolerance levels in this manner creates a very robust system. The systems usefulness, however, can only be determined by analysing a different data-set, one which contains known errors. For 3D dose verification, dose-volume histogram (DVH)-based analysis; the sensitivity of the system for catching errors was investigated by retrospectively introducing both monitor unit (MU) and MLC leaf-position errors in the existing EPID frames as described in detail later in this section.

7.2.3 Plan selection and real-time dose reconstruction

The method described above was used to derive tolerance levels by analysing data from 105 lung SBRT VMAT plans delivered in 2014. Prescribed doses were 45 Gy (n = 4) for centrally located tumours and 54 Gy (n = 101) for peripherally located tumours, delivered in three fractions. Every treatment consisted of two 10 MV VMAT arcs, planned with the Pinnacle treatment planning system (TPS) (version 9.8, Philips Medical Systems, Eindhoven, The Netherlands). A non-homogeneous planning target volume (PTV) dose distribution was applied to increase the dose gradient outside the PTV. The maximum allowed planned dose in the PTV was 166% of the prescribed dose [96]. Cine-mode EPID images ("movies"), continuously acquired during treatment, were available for 256 fractions. EPID frames were acquired at a rate of 2.4 fps and the 3D dose distribution delivered to the patient was reconstructed and accumulated. The accumulation included both treatment arcs of every analysed fraction; i.e., the dose of the first treatment arc was summed in the analysis when verifying the dose of the second treatment arc. The planned dose distribution was specified per control point (CP) of the VMAT arc with a CP spacing of 4 degrees. This 4 degree spacing is quite wide for verification purposes, as the dose is delivered continuously. Therefore, the planned dose distribution was distributed evenly per degree for each CP. The accumulated planned dose distribution used for verification was updated as soon as the gantry had advanced by one degree. In contrast, the reconstructed dose distribution was updated with each EPID frame. On average, 817 3D dose distributions were reconstructed and verified per fraction. Dose reconstruction and verification was done on a 2x2x2 mm^3 dose-grid; a single 3D dose reconstruction and verification took ~200 ms; i.e., faster than the framerate.

7.2.4 3D dose verification

For 3D dose verification, we used DVH-based analysis instead of the widely adopted γ-analysis method [28]. This is because the correlation between γ-analysis results and clinical relevance is not obvious [29, 31, 30, 67]. With 3D *in vivo* dosimetry, delivered dose distributions can easily be analyzed in terms of DVH-parameters, providing a better indication to radiation therapy staff of the impact of a detected deviation. Three types of patient regions were analyzed: target, organs at risk (OARs) and the remaining patient volume (RPV). The target was defined by the planning target volume (PTV), OARs by their corresponding delineations, and the RPV by the entire volume receiving at least 5% of the prescribed fraction dose, as defined on the dose reconstruction, minus the PTV

and the OARs. The minimum dose requirement for the RPV prevents large, low-dose volumes from masking over-doses. Note that the RPV was updated with every EPID frame since the reconstructed dose distribution was also updated with every frame. For the PTV, differences were calculated in mean dose (D_{mean}) and near-minimum dose (D_{98}); for the RPV, both the near-maximum dose (D_2) and D_{mean} were used. The DVH-parameter used for dose verification of OARs depends on the type of organ. For organs generally considered to be serial, the maximum dose (D_{max}) was verified, for parallel organs the mean dose. The serial organs included in this study were brachial plexus, esophagus, great vessels, heart, main bronchus, rib, spinal cord, stomach and trachea; the only organ considered parallel was the lung. In accordance with our clinical plan evaluation, the normalized total dose (NTD) [97] was used instead of the physical dose for dose verification in the lungs, with $\alpha/\beta = 3$ Gy. Not every OAR was defined in each plan; OARs located far from the target volume are often not delineated. For tolerance levels, only OARs defined in at least 2/3 of the included treatments were considered. This was done to ensure enough data-points for tolerance level determination; note that not considering all OARs could make the system less sensitive for error detection, i.e., the performance is underestimated.

Deviations between reconstructed and planned dose parameters were assessed by determining $(D_x^{EPID}\text{-}D_x^{PLAN})/D_x^{TOTAL}$, where both D_x^{EPID} and D_x^{PLAN} are DVH-parameters calculated from the DVHs of their respective accumulated dose distributions, i.e., these are not constant during dose delivery: D_x^{PLAN} changed every gantry-angle degree; D_x^{EPID} changed with every acquired EPID frame. D_x^{TOTAL} is a DVH-parameter from the total planned dose distribution (for PTV and RPV) or planning constraints (for OARs). By calculating deviations relative to D_x^{TOTAL}, the relevance of an observed deviation is immediately known, which is not the case when calculating "plain" relative differences in DVH-parameters (i.e., $(D_x^{EPID}\text{-}D_x^{PLAN})/D_x^{PLAN}$) [1].

7.2.5 Introduction of errors

With the tolerance levels established, we wanted to test the sensitivity of the system by deliberately introducing errors in the EPID data of all included fractions and verifying whether these errors would have been detected for robustness-margins of 0%, 20% and

[1] Compare, for example, D_x^{EPID}=1 Gy, D_x^{PLAN} = 0.5 Gy and D_x^{TOTAL} = 18 Gy. $(D_x^{EPID}\text{-}D_x^{PLAN})/D_x^{PLAN}$ then equals 100%, which seems huge, whereas the relevance of this deviation is actually quite limited. $(D_x^{EPID}\text{-}D_x^{PLAN})/D_x^{TOTAL}$ = 2.2%, a much better indication of the relevance of this deviation.

50%. A complicating factor was that we wished to assess the sensitivity of the system for *in vivo* errors, which ruled out phantom measurements. The solution was to simulate the introduction of *in vivo* errors by changing the EPID frames before they entered the dose reconstruction algorithm. The first simulated error was a systematic over- or underdosage, called MU-error. As the dose on the EPID scales linearly with the linac output, this error was easily simulated by adjusting the recorded EPID frame intensity before starting the dose reconstruction. MU-errors of -20%, +20%, +40% and +100% were introduced. For a second type of error, we aimed at introducing an offset in the MLC leaf positions. To accomplish this, we used the fact that the position of the EPID during dose delivery is a separate input parameter of the dose reconstruction algorithm. Forcing this parameter to an erroneous value is equivalent to simulating a constant leaf-position offset in the dose delivery. Positional errors were introduced in the lateral position of the EPID and had a magnitude d of 6, 12, and 18 mm at the isocenter. As all included treatment plans used a collimator rotation of 20°, the leaf-offset in the direction of leaf-travel was $d \cdot \sin(20°)$ and the offset in the direction perpendicular to leaf-travel was $d \cdot \cos(20°)$. An introduced error was considered to be detected for a fraction when at least one of the EPID frames belonging to that fraction showed deviations exceeding the applied tolerance levels. By definition, the sensitivity only tells whether an introduced error is detected at all, it contains no information about the timeliness of the detection. Therefore, the timing of error detection was also investigated.

7.3 Results

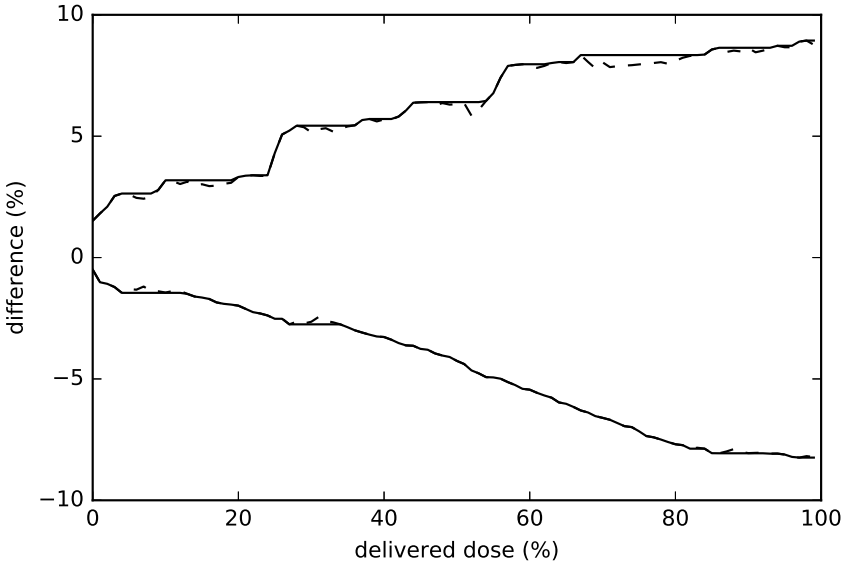
For comparison to regular off-line *in vivo* dosimetry, Table 7.1 shows results at complete delivery. It can be seen that good agreement between planned and delivered dose was found for the target volume, but systematic differences exist for dose to OARs and the RPV.

Figures 7.1a and 7.1b show the maximum observed differences in mean dose and near-minimum (D_{98}) dose to the PTV as a function of dose delivery. Apart from some minor fluctuations, the differences were indeed monotonically increasing (overdosages) and decreasing (underdosages). The 0% robustness-margin tolerance levels are indicated by the solid lines; they represent the tightest possible monotonically changing tolerance levels with zero false positives.

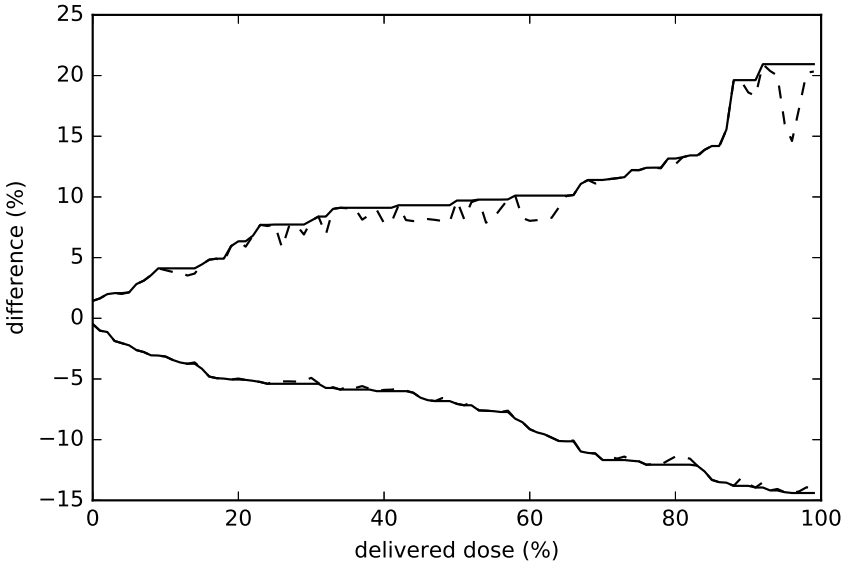
Similar results were obtained for the OARs, as illustrated in Figures 7.2a-7.2c, showing differences in dose to the lungs (NTD-D_{mean}), esophagus (D_{max}) and spinal cord (D_{max}).

Region	# of plans	Indicator	[Min, Max] (%)	[2nd, 98th] percentile (%)	Avg ± s.d. (%)
PTV	105	ΔD_{mean}	[-8.2, 8.9]	[-5.8, 7.8]	0.17 ± 3.1
PTV	105	ΔD_{98}	[-13, 20]	[-9.5, 10]	-0.50 ± 4.8
Lungs	105	NTD ΔD_{mean}	[-5.4, 10]	[-2.7, 6.9]	0.80 ± 2.1
Serial OARs					
brachial plexus	26	ΔD_{max}	[-8.5, 20]	[-4.0, 18]	5.1 ± 5.4
*esophagus**	105	ΔD_{max}	[-3.2, 18]	[-1.3, 9.4]	2.5 ± 2.8
*great vessels**	74	ΔD_{max}	[-8.9, 21]	[-5.7, 12]	2.1 ± 4.3
*heart**	99	ΔD_{max}	[-3.7, 17]	[-2.6, 7.9]	1.7 ± 2.6
main bronchus	62	ΔD_{max}	[-7.2, 21]	[-5.4, 14]	2.3 ± 4.6
ribs	13	ΔD_{max}	[-5.0, 6.9]	[-4.1, 6.1]	0.65 ± 2.8
*spinal cord**	105	ΔD_{max}	[-4.5, 20]	[-0.91, 17]	3.9 ± 3.9
stomach	4	ΔD_{max}	[-1.0, 5.2]	[-0.89, 5.1]	2.0 ± 2.2
trachea	4	ΔD_{max}	[1.1, 7.7]	[-1.2, 7.3]	3.6 ± 2.1
RPV	105	ΔD_2	[-6.5, 17]	[-5.5, 16]	2.5 ± 5.0
RPV	105	ΔD_{mean}	[-6.6, 21]	[-2.4, 20]	6.9 ± 6.4

Table 7.1: Dosimetric results after full treatment delivery. Serial OARs included in the tested tolerance levels are marked with an asterisk.

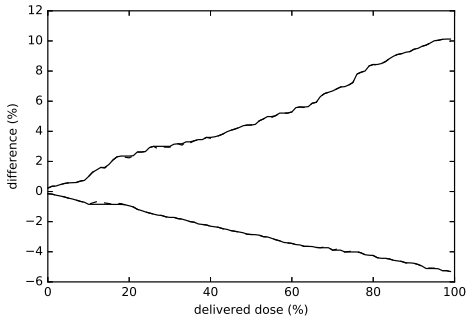


(a) PTV ΔD_{mean}

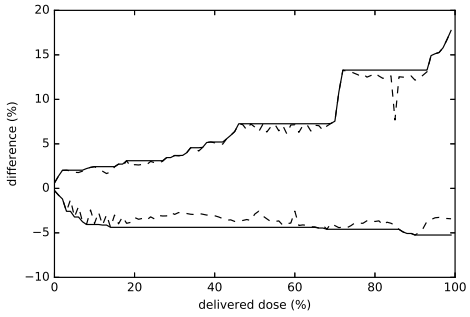


(b) PTV ΔD_{98}

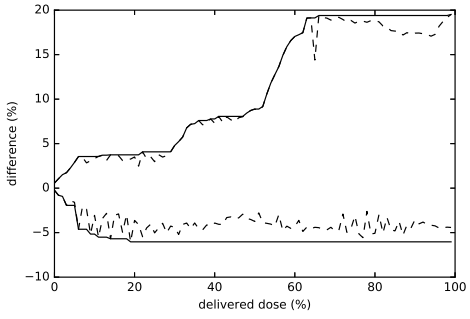
Figure 7.1: Observed maximal over- and underdosages in the mean and near-minimum dose to the PTV, relative to the total planned dose. The dashed lines indicate the actual maximum values, the solid lines the tightest possible monotonically changing bound, used to derive our tresholds.



(a) Lungs, NTD ΔD_{mean}



(b) Esophagus, ΔD_{max}



(c) Spinal cord, ΔD_{max}

Figure 7.2: Observed maximal over- and underdosages for three OARs, relative to the total planned dose for these OARs. The dashed lines indicate the actual maximum values, the solid lines the tightest possible monotonically changing bound.

Region	Indicator	Overdose tolerance [0%, 100%] (%)	Underdose tolerance [0%, 100%] (%)
PTV	ΔD_{mean}	[2, 9]	[0, -8]
PTV	ΔD_{98}	[1, 21]	[0, -14]
Lungs	NTD ΔD_{mean}	[0, 10]	N/A
Esophagus	ΔD_{max}	[1, 18]	N/A
Great vessels	ΔD_{max}	[1, 21]	N/A
Heart	ΔD_{max}	[2, 17]	N/A
Spinal cord	ΔD_{max}	[1, 20]	N/A
RPV	ΔD_{2}	[0, 19]	N/A
RPV	ΔD_{mean}	[0, 26]	N/A

Table 7.2: Initial (0%) and final (100%) tolerance levels for under- and overdosages for the included regions of interest, these thresholds are relative to the total planned dose per region. The tolerance levels between the initial and final values follow curves like the ones shown in Figures 7.2a-7.2c.

Note that the lower bounds are present only for completeness, underdosages to OARs were not considered a reason for halting the linac. Table 7.2 lists the initial and final tolerances for each ROI.

As the tolerance levels were set such that all dosimetric differences were encompassed, the number of false positives was equal to zero and the specificity was equal to one by definition. Table 7.3 shows the sensitivity of the method for all introduced errors for three different robustness-margin settings: 0, 20 and 50%. Also indicated is the average timepoint at out-of-tolerance detection, expressed as a percentage of the total treatment delivery showing longer delivery times when increasing the robustness margin. Sensitivity was quite high at 0.96 – 1, with introduced errors detected in almost all cases. Only at a robustness-margin of 50% the +20% MU error and the +6mm leaf errors were not detected in a few cases. Timeliness of the out-of-tolerance detection was also adequate: larger errors were detected at ~2%, as 1% of delivery typically takes ~3s, the largest introduced errors were detected within seconds.

	Robustness Margin					
	0%		20%		50%	
error	FN	Detected at (%delivery time)	FN	Detected at (%delivery time)	FN	Detected at (%delivery time)
-20%	0	6.0 ± 3.1	0	7.6 ± 3.3	0	9.9 ± 3.9
+20%	0	15 ± 10	0	24 ± 18	11 (0.96)	40 ± 27
+40%	0	5.7 ± 2.7	0	7.5 ± 3.7	0	10 ± 4.6
+100%	0	1.7 ± 0.6	0	2.1 ± 0.70	0	2.8 ± 1.1
+6 mm	0	5.6 ± 5.0	0	7.5 ± 6.3	2 (0.99)	12 ± 8.7
+12 mm	0	2.6 ± 2.0	0	3.1 ± 2.1	0	4.1 ± 2.8
+18 mm	0	1.8 ± 1.1	0	2.2 ± 1.2	0	2.7 ± 1.4

Table 7.3: Detectability of introduced errors: MU changes and leaf shifts. Each included fraction is regarded as an independent sample ($N = 256$). FN denotes the number of False Negatives; this number is followed by the sensitivity in parentheses if non-zero. When FN equals 0, the sensitivity equals 1. Due to the setup of the study, there are no false positives by definition. The average (± one s.d.) time of detection is indicated as well, as a function of total delivery time. Typically, 1% of delivery time takes about 3 seconds.

7.4 Discussion

In this study, the accuracy of a real-time 3D *in vivo* dosimetry method for a specific treatment site (lung) and technique (SBRT VMAT) was presented. To the best of our knowledge, our method is the only method capable of real-time 3D *in vivo* dosimetry. The results were used to derive robust tolerance levels; the usefulness of these tolerance levels was investigated by introducing two types of errors, each at different magnitude, and by examining the sensitivity of the method. Sensitivity and specificity analysis for *in vivo* EPID dosimetry has first been described by Bojechko and Ford [98]. Several important differences exist between the work of Bojechko and Ford and this study, however. First, this study starts from the premiss that there should be zero false positives. This excludes performing a full receiver-operating characteristic (ROC) analysis as the specificity will always be one by definition. This premiss was chosen for robustness, which is key to online dose verification, but much less so for offline dose verification. For the same reason, the introduced errors are of (much) larger magnitude than in the work of Bojechko and Ford. The obtained (dynamic) tolerance levels (Table 7.2) might seem quite relaxed when compared to "traditional", i.e. offline, tolerance levels for radiation dosimetry. It is therefore important to realize that the real-time tolerance levels are meant to detect the deviations in dose delivery which are so large that detecting them with off-line dosimetry, after completion of a single fraction, would imply unacceptable harm to the patient. In this light, a 10% overdose in the mean dose to the PTV can probably be regarded as acceptable for online dosimetry, whereas it remains unacceptable for offline dosimetry. In other words, online dosimetry should be regarded as an addition to offline dosimetry, not as a replacement. The thresholds are derived from the maximum observed deviations, which are on the order of 10%. Such deviations may seem rather large but one has to consider that many of these organs are partly in regions containing high dose gradients, such that small uncertainties in position can lead to large deviations, especially of maximum dose constraints. The presented results show some systematic overdosages with respect to the planned dose distributions outside of the PTV. These systematic deviations are either due to limitations of our dose reconstruction method or due to non-optimal modelling of the TPS in the penumbra region, or both; this is subject of further investigation. Still, the sensitivity to introduced errors showed that the current method is suitable to be used clinically for real-time DHV-based dose verification. Our method of DVH-based verification of 3D dose distributions has some limitations. The reconstructed dose distribution is calculated using the planning

CT-scan; i.e., any anatomical changes will lead to dose deviations of a potentially misleading magnitude, cf. [93]. Also, we are not considering displacements of the organs when evaluating the dose discrepancies. Real-time dosimetry only intends to verify the delivered dose and would not replace image guided and adapative radiotherapy. Another issue is the use of D_{max} for dose verification of serial OARs; this parameter is obviously quite sensitive to geometrical uncertainties in the dose reconstruction, especially in a region with a high dose gradient. The OARs receiving a dose close to their constraining dose will be in such a region, as they will be close to the target volume. One option could be using a near-maximum parameter (e.g., D_2) to mitigate this issue. However, we found that many deviations persist due to the small volume of several OARs (great vessels, brachial plexus). Another solution would be to use a margin around the OARs for determining the planned parameter value; this has as a drawback that it might mask true overdosages. All in all, we find that the results obtained for D_{max} without any OAR margin are acceptable.

7.5 Conclusions

In this study we demonstrated the robustness and usefulness of a real-time *in vivo* dosimetry method for lung SBRT VMAT. The method combines high specificity with high sensitivity and shows that clinical introduction of real-time 3D DVH-based *in vivo* dosimetry for lung SBRT VMAT is feasible.

Chapter 8

Discussion and conclusions

8.1 General discussion

In this thesis we aimed to improve patient safety in radiotherapy by employing *in vivo* EPID dosimetry efficiently on a large scale, and ultimately in real-time. This aim was met via the development of a fully automated system, allowing physicists to focus on dosimetric deviations instead of performing tedious manual labour. Also, feasibility and potential benefits of DVH-based *in vivo* dose verification were demonstrated. Based on these results, DVH-based dose verification is currently being implemented in the clinic. The final step in optimal patient safety through *in vivo* dosimetry, real-time dose verification, has been developed and is being implemented clinically now. Thus, the work presented in this thesis represents a major step forward in radiotherapy patient safety.

Interest for EPID based *in vivo* dosimetry has increased globally in the last years. Vendors of radiotherapy QA systems have released tools for *in vivo* dosimetry, such as DosiSoft [99], SunNuclear [100], IBA [101] and Elekta [102]. A driving force for this development has been national regulation; in France, for example, a form of *in vivo* dosimetry is mandatory for all treatments since 2015; this requirement followed a series of serious incidents [6]. Publications on the application of EPID *in vivo* dosimetry in a routine clinical setting are scarce. An Italian initiative for a large-scale employment of EPID *in vivo* dosimetry exists [91], but the applied method verifies dose at a single point, making the method difficult to compare to the method discussed in this thesis. In Maastro clinic (Maastricht, NL), 3D *in vivo* EPID dosimetry has been developed as well [25], but the clinical use seems limited – it does use DVH-based dose evaluation, though [33]. A second Maastro initiative is 2D EPID dosimetry by comparing predicted and measured portal images [103]. Although the portal images are acquired during treatment, the method does not strictly qualify as an *in vivo* dosimetry method since it does not perform dose determination inside the patient. An overview of the clinical experience with *in vivo* dosimetry at the NKI-AvL is given by Mijnheer et al. [27].

The basis for a large-scale clinical implementation of our EPID *in vivo* dosimetry system is a highly automated workflow; the process of automation requires a careful description of all parts of the clinical protocol associated with EPID *in vivo* dosimetry. Chapter 2 of this thesis demonstrated how this has been achieved. The proposed automated system has been used clinically since 2011 and has provided a strong reduction in clinical workload. Moreover, it has allowed for a considerable extension of the amount of treatments verified: initially, due to workload constraints, only IMRT and VMAT treatments were verified via *in vivo* EPID dosimetry. With the fully automated system in

place, all remaining treatments – mainly 3D-CRT – were included with a limited increase in workload. Hence, our clinic is now the fortunate position of having almost every treatment verified via *in vivo* dosimetry. The only exception is treatments where the fields exceed the active area of the EPID, to limit radiation damage of the electronics.

The second part of the thesis dealt with replacement of γ-analysis with DVH-based metrics. The usefulness of the widely adopted γ-analysis method has been shown to be questionable [67, 30, 31, 29]. It should be noted that these papers discuss the use of γ-analysis for pre-treatment measurements (i.e., using a phantom). As such, several vendors have developed methods for converting dose distributions measured in a phantom to a virtual dose distribution in the CT-scan of the patient [104, 105]. These methods are comparable to the method presented in chapter 5 as described below.

Separate quality control measurements are sometimes needed next to the *in vivo* measurements – for example, when a dosimetric deviation is observed *in vivo* and one wishes to examine the dose delivery under reference conditions. This is a challenge for DVH-based analysis: these measurements are usually performed using a phantom, which bears no resemblance to the patient geometry the plan was intended for. Using DVH-based analysis on such a phantom is pointless, as the relevant dose-information is intrinsically linked to the patient anatomy. Chapter 5 presents a novel method for reference condition measurements which uses the anatomy of the patient, making DVH-based analysis also available for this type of measurements. It represents a critical step in moving towards DVH-based dose evaluation, as reference condition measurements are an integral part of any clinical QA system.

A natural extension of the current *in vivo* dose verification methods is real-time dosimetry. After all, a serious delivery error would be caught only after the damage to the patient has been done using off-line dosimetry. A notable initiative in real-time dosimetry is the method of Woodruff et al. [35], which is already being tested clinically. However, it is not a proper *in vivo* dose verification method as it compares measured to predicted portal images. Such a method is well-suited to catch the largest of delivery errors, but as the errors become more subtle, it becomes harder and harder to estimate the clinical relevance of the observed deviations. This issue is resolved by moving towards 3D, DVH-based real-time dosimetry, for which the proof of concept has been demonstrated in chapter 6. It represents a major technical improvement: a full 3D dose reconstruction and evaluation is performed within 300 ms. But more importantly, it allows for a major step in patient safety.

Real-time 3D *in vivo* dosimetry does step into uncharted territory. First, DVH-based

dose evaluation for both the target volume and the OARs seems inevitable. How else to handle the situation of a linac halting with the patient on the treatment couch, when one can only guess at the clinical relevance of the observed deviation causing the halt? Second, real-time dosimetry implies time-resolved dosimetry, a new concept in radiotherapy QA. Chapter 7 evaluated time-resolved, DVH-based, dosimetric results of historic lung SBRT VMAT treatments. The presented work is a prerequisite for the clinical implementation of real-time *in vivo* dosimetry; it defines how tolerance levels can be set in a practical way.

The largest remaining clinical issue is the number of false alerts generated by the *in vivo* dosimetry system. For complex treatments (i.e., IMRT or VMAT), currently (2016) more than 50% of the verified treatments are being classified as having a minor or major deviation in delivered dose; the observed true positive rate is much lower, at about 0.3%. This thesis proposes to change the evaluation method for *in vivo* dosimetry from γ-analysis to DVH-analysis to resolve this issue. Chapter 3 has shown that tolerance levels for DVH-based evaluation of the target volume can be chosen such that there is reasonable correspondence between the results of both methods. At the same time, it provides an indication of the accuracy of our method for *in vivo* dose verification of the target dose distribution: about 3% (one s.d.). Hence, it has now become insightful to which deviations in target dose the γ-analysis results correspond. Relaxing the tolerance levels for γ-analysis is difficult, as γ-analysis lacks information about clinical relevance, though it could be considered using the results from chapter 3. A more promising approach is to use the DVH-analysis directly for dose evaluation. This will make the trade-off between sensitivity and specificity insightful: relevant deviations are defined to have a certain deviation in a DVH-parameter. By setting the tolerance levels accordingly, the sensitivity will equal one and the established accuracy of the method will determine the specificity. For example, setting tolerance levels to two standard deviations would lead to 5% positives[1], of which ~0.3 percentage point (p.p.) are expected to be true positives. The drop in false positive rate would be considerable (~45 p.p.). The downside is that the system would become insensitive to small deviations which are irrelevant for a specific patient but might indicate an underlying systematic problem. This might be mitigated by employing statistical analyses on the *in vivo* dosimetry results; this possibility will be discussed in the next section.

One of the treatment sites showing many dosimetric deviations is H&N. Anatomical changes during treatment are common for H&N cancer patients and are in clinical

[1]Strictly speaking, this is only true when a single indicator is being used.

practice often cited as a cause for an observed *in vivo* dose deviation. Chapter 4 of this thesis has put this to the test, using DVH-analysis to gain insight in the clinical relevance of observed deviations. Surprisingly, the target dose distribution as seen in room coordinates is hardly influenced by anatomical variations, even if these are considerable. This study has demonstrated that our current γ-based dose evaluation method mainly raises alerts for statistical variations. Another important result is the quantification of the over-estimation of dosimetric deviations when neglecting the daily patient anatomy and using the anatomy of the planning-CT instead. From calculations it was assumed that anatomical changes would lead to an over-estimation of dosimetric deviations, but this had not yet been experimentally verified. Chapter 4 has shown that the calculated over-estimation agrees well with a first-order approximation of the changing anatomy. Finally, the statistical power of large-scale dosimetry has been illustrated: even a very minor effect of the changing anatomy could be quantified significantly. This result implies that employing statistical analyses for large-scale *in vivo* dosimetry will be useful for detection of small systematic deviations while patient-specific tolerance levels are relaxed.

8.2 A broader view of radiotherapy QA

Developments in DVH-based analysis lead to the question whether all QA measurements in radiotherapy should be verified via DVHs. Although one might intuitively expect an unconditional positive answer to this question, there are several details to be noted.

Let us start by making a distinction between QA and QC[2]: quality assurance aims at making sure that quality goals will be met in general, quality control aims at verifying whether the quality goals are met for a specific case. As an example, commissioning of a linac is part of QA: the goal is that the linac will meet the quality goals (i.e., delivering the right dose at the right spot) for any dose delivery. Patient-specific measurements[3] are part of QC: the goal is to verify that the linac meets the quality goals for the specific treatment plan in question.

The EPID *in vivo* dosimetry method discussed in this thesis is thus part of the QC system. Obviously, QA is a broader and more difficult subject than QC. Performing

[2]This distinction has not been made yet in this thesis; it is customary in radiotherapy to refer to both QA and QC as QA. This custom has been used in this thesis as well; it will only be disregarded in the remainder of this chapter.

[3]Usually referred to as patient-specific QA in radiotherapy.

QA with DVH-based methods seems awkward: DVH-based results are suited for QC, as these describe the relevance of dosimetric deviations for a specific patient. Therefore, it makes sense in radiotherapy to hold on to specific, patient-independent QA measurements which should be verified using an evaluation metric such as γ-analysis (note that the distance metric is introduced for geometrical errors in measurement equipment). As such, within the distinction between QA and QC, DVH-based dose evaluations are not the only solution needed in radiotherapy QA. For QC, however, it seems evident that DVH-based analysis should be employed. An interesting feature of these QC measurements develops when the QC system is applied on a large scale, such as the EPID *in vivo* dosimetry system described in this work: with enough measurements available, it becomes possible to incorporate the QC measurement results in the QA system. This is because small systematic deviations are not relevant for a specific patient, and thus can safely be ignored on a per-measurement basis, but are relevant for the population of patients being treated. Thus, statistical analysis of a large number of QC results can – and probably should – be part of the QA system. Such a practice is already employed in image-guided radiotherapy (IGRT), where the population mean of patient-setup measurements are used to show equipment or protocol errors. This observation was hinted at in chapter 4 of this thesis and will be an interesting way of extending the usefulness of the large-scale EPID *in vivo* dosimetric QC system.

8.3 Future work

As automation requires in-depth knowledge of the clinical protocol and custom-made extensions to the TPS and the record & verify (R&V) system, it is not (yet) easy to port our automated solution to other clinics. The *in vivo* dosimetry system itself without automation has been developed into a commercial package in cooperation with Elekta (Elekta Ltd., Crawley, UK) and is available for purchase since early 2016.

DVH-based analysis is a very promising technique for dose evaluation; however, one needs to bear in mind that the dose reconstruction and evaluation is done on the planning-CT. This means that the ROIs used for dose verification might correspond to a different region than assumed. For the target volume this is not so much of a concern, as IGRT will ensure the correct positioning of the target. For the OARs larger deviations could be expected; a study on the effect of daily anatomical changes on the dose to OARs (similar to the study described in chapter 4, but then including motion of the organs) should thus be done. Also, the dose reconstruction method used has been developed with

a focus on the high-dose region. It can be seen from the results presented in chapter 7 that systematic deviations are present outside of the target volume. The cause of these deviations – limitations of the dose reconstruction method or limitations of the TPS – should be investigated.

8.4 Conclusions

Improving patient safety through large-scale *in vivo* EPID dosimetry has been demonstrated. The results presented in this thesis allow for real-time, 3D, DVH-based *in vivo* dose verification for all patients to be implemented clinically. This will lead a further increase in patient safety, reduced workload and better sensitivity and specificity for *in vivo* EPID dosimetry.

References

[1] http://www.iknl.nl/over-iknl/nieuws/nieuws-detail/2015/07/09/inzet-van-radiotherapie-op-vrijwel-optimaal-niveau-in-nederland.

[2] W. Bogdanich, "Radiation offers new cures, and ways to do harm," *New York Times*, Jan 2010.

[3] W. Bogdanich, "As technology surges, radiation safeguards lag," *New York Times*, Jan 2010.

[4] W. Bogdanich, "Case studies: when medical radiation goes awry," *New York Times*, Jan 2010.

[5] M. V. Williams, "Improving patient safety in radiotherapy by learning from near misses, incidents and errors," *British Journal of Radiology*, vol. 80, no. 953, pp. 297–301, 2007.

[6] S. Derreumaux, C. Etard, C. Huet, F. Trompier, I. Clairand, J. F. Bottollier-depois, B. Aubert, and P. Gourmelon, "Lessons from recent accidents in radiation therapy in France," in *Radiation Protection Dosimetry*, vol. 131, pp. 130–135, 2008.

[7] A. Mans, M. Wendling, L. N. McDermott, J.-J. Sonke, R. Tielenburg, R. Vijlbrief, B. Mijnheer, M. van Herk, and J. C. Stroom, "Catching errors with in vivo EPID dosimetry," *Medical Physics*, vol. 37, no. 6, pp. 2638–44, 2010.

[8] E. C. Ford, S. Terezakis, A. Souranis, K. Harris, H. Gay, and S. Mutic, "Quality control quantification (QCQ): A tool to measure the value of quality control checks in radiation oncology," *International Journal of Radiation Oncology*Biology*Physics*, vol. 84, no. 3, pp. e263–e269, 2012.

[9] M. van Herk and H. Meertens, "A matrix ionisation chamber imaging device for on-line patient setup verification during radiotherapy," *Radiotherapy and Oncology*, vol. 11, no. 4, pp. 369–378, 1988.

[10] H. Meertens, M. van Herk, J. Bijhold, and H. Bartelink, "First clinical experience with a newly developed electronic portal imaging device," *International Journal of Radiation Oncology*Biology*Physics*, vol. 18, no. 5, pp. 1173–1181, 1990.

[11] M. Essers, B. R. Hoogervorst, M. van Herk, H. Lanson, and B. J. Mijnheer, "Dosimetric characteristics of a liquid-filled electronic portal imaging device," *International Journal of Radiation Oncology*Biology*Physics*, vol. 33, no. 5, pp. 1265–1272, 1995.

[12] M. C. Kirby and P. C. Williams, "The use of an electronic portal imaging device for exit dosimetry and quality control measurements," *International Journal of Radiation Oncology*Biology*Physics*, vol. 31, no. 3, pp. 593–603, 1995.

[13] M. Essers, J. H. Lanson, G. Leunens, T. Schnabel, and B. J. Mijnheer, "The accuracy of CT-based inhomogeneity corrections and in vivo dosimetry for the treatment of lung cancer," *Radiotherapy and Oncology*, vol. 37, no. 3, pp. 199–208, 1995.

[14] R. Boellaard, M. Van Herk, H. Uiterwaal, and B. Mijnheer, "Two-dimensional exit dosimetry using a liquid-filled electronic portal imaging device and a convolution model," *Radiotherapy and Oncology*, vol. 44, no. 2, pp. 149–157, 1997.

[15] M. Kroonwijk, K. L. Pasma, S. Quint, P. C. M. Koper, A. G. Visser, and B. J. M. Heijmen, "In vivo dosimetry for prostate cancer patients using an electronic portal imaging device (EPID); demonstration of internal organ motion," *Radiotherapy and Oncology*, vol. 49, no. 2, pp. 125–132, 1998.

[16] R. Boellaard, M. Essers, M. Van Herk, and B. J. Mijnheer, "New method to obtain the midplane dose using portal in vivo dosimetry," *International Journal of Radiation Oncology*Biology*Physics*, vol. 41, no. 2, pp. 465–474, 1998.

[17] A. Piermattei, A. Fidanzio, G. Stimato, L. Azario, L. Grimaldi, G. D'Onofrio, S. Cilla, M. Balducci, M. A. Gambacorta, N. Di Napoli, and N. Cellini, "In vivo dosimetry by an aSi-based EPID," *Medical Physics*, vol. 33, no. 11, pp. 4414–4422, 2006.

[18] A. Piermattei, A. Fidanzio, L. Azario, F. Greco, A. Mameli, S. Cilla, L. Grimaldi, G. D'Onofrio, B. G. Augelli, G. Stimato, D. Gaudino, S. Ramella, R. D'Angelillo, F. Cellini, and L. Trodella, "In patient dose reconstruction using a cine acquisition for dynamic arc radiation therapy," *Medical and Biological Engineering and Computing*, vol. 47, no. 4, pp. 425–433, 2009.

[19] K. Chytyk and B. M. C. McCurdy, "Comprehensive fluence model for absolute portal dose image prediction," *Medical Physics*, vol. 36, no. 4, pp. 1389–1398, 2009.

[20] K. Chytyk-Praznik, E. VanUytven, T. a. VanBeek, P. B. Greer, and B. M. C. McCurdy, "Model-based prediction of portal dose images during patient treatment," *Medical Physics*, vol. 40, no. 3, p. 031713, 2013.

[21] J. L. Bedford, I. M. Hanson, and V. N. Hansen, "Portal dosimetry for VMAT using integrated images obtained during treatment," *Medical Physics*, vol. 41, no. 2, p. 021725, 2014.

[22] M. Wendling, L. N. McDermott, A. Mans, J.-J. Sonke, M. van Herk, and B. J. Mijnheer, "A simple backprojection algorithm for 3D in vivo EPID dosimetry of IMRT treatments," *Medical Physics*, vol. 36, no. 7, pp. 3310–21, 2009.

[23] A. Mans, P. Remeijer, I. Olaciregui-Ruiz, M. Wendling, J.-J. Sonke, B. Mijnheer, M. van Herk, and J. C. Stroom, "3D Dosimetric verification of volumetric-modulated arc therapy by portal dosimetry," *Radiotherapy and Oncology*, vol. 94, no. 2, pp. 181–7, 2010.

[24] W. van Elmpt, S. Nijsten, S. Petit, B. Mijnheer, P. Lambin, and A. Dekker, "3D in vivo dosimetry using megavoltage cone-beam CT and EPID dosimetry," *International Journal of Radiation Oncology*Biology*Physics*, vol. 73, no. 5, pp. 1580–7, 2009.

[25] W. van Elmpt, S. Petit, D. De Ruysscher, P. Lambin, and A. Dekker, "3D dose delivery verification using repeated cone-beam imaging and EPID dosimetry for stereotactic body radiotherapy of non-small cell lung cancer," *Radiotherapy and Oncology*, vol. 94, no. 2, pp. 188–194, 2010.

[26] L. N. McDermott, M. Wendling, J.-J. Sonke, M. van Herk, and B. J. Mijnheer, "Replacing pre-treatment verification with in vivo EPID dosimetry for prostate IMRT," *International Journal of Radiation Oncology*Biology*Physics*, vol. 67, no. 5, pp. 1568–77, 2007.

[27] B. J. Mijnheer, P. González, I. Olaciregui-Ruiz, R. A. Rozendaal, M. van Herk, and A. Mans, "Overview of 3-year experience with large-scale electronic portal imaging device-based 3-dimensional transit dosimetry," *Practical Radiation Oncology*, vol. 5, no. 6, pp. e679–e687, 2015.

[28] D. A. Low, W. B. Harms, S. Mutic, and J. A. Purdy, "A technique for the quantitative evaluation of dose distributions," *Medical Physics*, vol. 25, no. 5, pp. 656–661, 1998.

[29] B. E. Nelms, H. Zhen, and W. A. Tome, "Per-beam, planar IMRT QA passing rates do not predict clinically relevant patient dose errors," *Medical Physics*, vol. 38, no. 2, pp. 1037–44, 2011.

[30] H. Zhen, B. E. Nelms, and W. A. Tome, "Moving from gamma passing rates to patient DVH-based QA metrics in pretreatment dose QA," *Medical Physics*, vol. 38, no. 10, pp. 5477–89, 2011.

[31] P. Carrasco, N. Jornet, A. Latorre, T. Eudaldo, A. Ruiz, and M. Ribas, "3D DVH-based metric analysis versus per-beam planar analysis in IMRT pretreatment verification," *Medical Physics*, vol. 39, no. 8, pp. 5040–9, 2012.

[32] M. Stasi, S. Bresciani, A. Miranti, A. Maggio, V. Sapino, and P. Gabriele, "Pretreatment patient-specific IMRT quality assurance: a correlation study between gamma index and patient clinical dose volume histogram," *Medical Physics*, vol. 39, pp. 7626–34, 2012.

[33] L. C. G. G. Persoon, A. G. T. M. Egelmeer, M. C. Ollers, S. M. J. J. G. Nijsten, E. G. C. Troost, and F. Verhaegen, "First clinical results of adaptive radiotherapy based on 3D portal dosimetry for lung cancer patients with atelectasis treated with volumetric-modulated arc therapy (VMAT)," *Acta Oncologica*, vol. 52, no. 7, pp. 1484–1489, 2013.

[34] T. Fuangrod, H. C. Woodruff, E. van Uytven, B. M. C. McCurdy, Z. Kuncic, D. J. O'Connor, and P. B. Greer, "A system for EPID-based real-time treatment delivery verification during dynamic IMRT treatment," *Medical Physics*, vol. 40, no. 9, p. 091907, 2013.

[35] H. C. Woodruff, T. Fuangrod, E. Van Uytven, B. M. McCurdy, T. van Beek, S. Bhatia, and P. B. Greer, "First Experience With Real-Time EPID-Based Delivery Verification During IMRT and VMAT Sessions," *International Journal of Radiation Oncology*Biology*Physics*, vol. 93, no. 3, pp. 516–522, 2015.

[36] J. Zhu, L. Chen, A. Chen, G. Luo, X. Deng, and X. Liu, "Fast 3D dosimetric verifications based on an electronic portal imaging device using a GPU calculation engine," *Radiation Oncology*, vol. 10, no. 1, p. 85, 2015.

[37] P. B. Greer and C. C. Popescu, "Dosimetric properties of an amorphous silicon electronic portal imaging device for verification of dynamic intensity modulated radiation therapy," *Medical Physics*, vol. 30, no. 7, pp. 1618–1627, 2003.

[38] W. van Elmpt, L. McDermott, S. Nijsten, M. Wendling, P. Lambin, and B. Mijnheer, "A literature review of electronic portal imaging for radiotherapy dosimetry," *Radiotherapy and Oncology*, vol. 88, no. 3, pp. 289–309, 2008.

[39] S. M. J. J. G. Nijsten, B. J. Mijnheer, A. L. A. J. Dekker, P. Lambin, and A. W. H. Minken, "Routine individualised patient dosimetry using electronic portal imaging devices," *Radiotherapy and Oncology*, vol. 83, pp. 65–75, 2007.

[40] A. Piermattei, F. Greco, L. Azario, A. Porcelli, S. Cilla, S. Zucca, A. Russo, E. D. Castro, M. Russo, R. Caivano, V. Fusco, A. Morganti, and A. Fidanzio, "A National project for in vivo dosimetry procedures in radiotherapy: First results," *Nuclear Instruments and Methods in Physics Research, Section B*, vol. 274, pp. 42–50, 2012.

[41] M. Wendling, R. J. W. Louwe, L. N. McDermott, J.-J. Sonke, M. van Herk, and B. J. Mijnheer, "Accurate two-dimensional IMRT verification using a back-projection EPID dosimetry method," *Medical Physics*, vol. 33, no. 2, pp. 259–273, 2006.

[42] B. Mijnheer, S. Beddar, J. Izewska, and C. Reft, "In vivo dosimetry in external beam radiotherapy," *Medical Physics*, vol. 40, no. 7, p. 070903, 2013.

[43] L. N. McDermott, R. J. W. Louwe, J. J. Sonke, M. B. van Herk, and B. J. Mijnheer, "Dose-response and ghosting effects of an amorphous silicon electronic portal imaging device," *Medical Physics*, vol. 31, pp. 285–295, 2004.

[44] R. Pecharroman-Gallego, A. Mans, J.-J. Sonke, J. C. Stroom, I. Olaciregui-Ruiz, M. van Herk, and B. J. Mijnheer, "Simplifying EPID dosimetry for IMRT treatment verification," *Medical Physics*, vol. 38, no. 2, pp. 983–92, 2011.

[45] M. Wendling, L. N. McDermott, A. Mans, I. Olaciregui-ruiz, J.-J. Sonke, J. Stroom, and M. V. Herk, "In aqua vivo EPID dosimetry," *Medical Physics*, vol. 39, no. 1, pp. 367–377, 2012.

[46] L. C. G. G. Persoon, S. M. J. J. G. Nijsten, F. J. Wilbrink, M. Podesta, J. A. D. Snaith, T. Lustberg, W. J. C. van Elmpt, F. van Gils, and F. Verhaegen, "Interfractional trend analysis of dose differences based on 2D transit portal dosimetry," *Physics in Medicine and Biology*, vol. 57, no. 20, pp. 6445–6458, 2012.

[47] L. N. McDermott, M. Wendling, J. Nijkamp, A. Mans, J.-J. Sonke, B. J. Mijnheer, and M. van Herk, "3D in vivo dose verification of entire hypo-fractionated IMRT treatments using an EPID and cone-beam CT," *Radiotherapy and Oncology*, vol. 86, no. 1, pp. 35–42, 2008.

[48] J. J. Kruse, "On the insensitivity of single field planar dosimetry to IMRT inaccuracies," *Medical Physics*, vol. 37, no. 6, pp. 2516–24, 2010.

[49] C. Wu, K. E. Hosier, K. E. Beck, M. B. Radevic, J. Lehmann, H. H. Zhang, A. Kroner, S. C. Dutton, S. A. Rosenthal, J. K. Bareng, M. D. Logsdon, and D. R. Asche, "On using 3D γ-analysis for IMRT and VMAT pretreatment plan QA," *Medical Physics*, vol. 39, no. 6, pp. 3051–9, 2012.

[50] The International Commission on Radiation Units and Measurements, "Prescribing, Recording, and Reporting Intensity-Modulated Photon-Beam Therapy (IMRT). ICRU Report 83," *Journal of the ICRU*, vol. 10, no. 1, 2010.

[51] M. van Herk, P. Remeijer, C. Rasch, and J. Lebesque, "The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy," *International Journal of Radiation Oncology*Biology*Physics*, vol. 47, no. 4, pp. 1121–1135, 2000.

[52] J. L. Barker, A. S. Garden, K. K. Ang, J. C. O'Daniel, H. Wang, L. E. Court, W. H. Morrison, D. I. Rosenthal, K. S. C. Chao, S. L. Tucker, R. Mohan, and L. Dong, "Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system," *International Journal of Radiation Oncology*Biology*Physics*, vol. 59, pp. 960–970, 2004.

[53] P. Castadot, X. Geets, J. A. Lee, N. Christian, and V. Grégoire, "Assessment by a deformable registration method of the volumetric and positional changes of target volumes and organs at risk in pharyngo-laryngeal tumors treated with concomitant chemo-radiation," *Radiotherapy and Oncology*, vol. 95, pp. 209–217, 2010.

[54] Q. Wu, Y. Chi, P. Y. Chen, D. J. Krauss, D. Yan, and A. Martinez, "Adaptive Replanning Strategies Accounting for Shrinkage in Head and Neck IMRT," *International Journal of Radiation Oncology*Biology*Physics*, vol. 75, no. 3, pp. 924–932, 2009.

[55] P. H. Ahn, C.-C. Chen, A. I. Ahn, L. Hong, P. G. Scripes, J. Shen, C.-C. Lee, E. Miller, S. Kalnicki, and M. K. Garg, "Adaptive planning in intensity-modulated radiation therapy for head and neck cancers: single-institution experience and clinical implications," *International Journal of Radiation Oncology*Biology*Physics*, vol. 80, no. 3, pp. 677–685, 2011.

[56] C. Lee, K. M. Langen, W. Lu, J. Haimerl, E. Schnarr, K. J. Ruchala, G. H. Olivera, S. L. Meeks, P. A. Kupelian, T. D. Shellenberger, and R. R. Mañon, "Evaluation of geometric changes of parotid

glands during head and neck cancer radiotherapy using daily MVCT and automatic deformable registration," *Radiotherapy and Oncology*, vol. 89, no. 1, pp. 81–88, 2008.

[57] S. A. Bhide, M. Davies, K. Burke, H. A. McNair, V. Hansen, Y. Barbachano, I. A. El-Hariry, K. Newbold, K. J. Harrington, and C. M. Nutting, "Weekly Volume and Dosimetric Changes During Chemoradiotherapy With Intensity-Modulated Radiation Therapy for Head and Neck Cancer: A Prospective Observational Study," *International Journal of Radiation Oncology*Biology*Physics*, vol. 76, no. 5, pp. 1360–1368, 2010.

[58] R. A. Rozendaal, B. J. Mijnheer, M. van Herk, and A. Mans, "In vivo portal dosimetry for head-and-neck VMAT and lung IMRT: Linking γ-analysis with differences in dosevolume histograms of the PTV," *Radiotherapy and Oncology*, vol. 112, p. 396, 2014.

[59] S. van Beek, S. van Kranen, A. Mencarelli, P. Remeijer, C. Rasch, M. van Herk, and J.-J. Sonke, "First clinical experience with a multiple region of interest registration and correction method in radiotherapy of head-and-neck cancer patients," *Radiotherapy and Oncology*, vol. 94, no. 2, pp. 213–7, 2010.

[60] S. van Kranen, A. Mencarelli, S. van Beek, C. Rasch, M. van Herk, and J. J. Sonke, "Adaptive radiotherapy with an average anatomy model: Evaluation and quantification of residual deformations in head and neck cancer patients," *Radiotherapy and Oncology*, vol. 109, pp. 463–468, 2013.

[61] A. Mencarelli, S. R. van Kranen, O. Hamming-Vrieze, S. van Beek, C. R. Nico Rasch, M. van Herk, and J.-J. Sonke, "Deformable Image Registration for Adaptive Radiation Therapy of Head and Neck Cancer: Accuracy and Precision in the Presence of Tumor Changes," *International Journal of Radiation Oncology*Biology*Physics*, vol. 90, no. 3, pp. 1–8, 2014.

[62] The International Commission on Radiation Units and Measurements, "Prescribing, recording, and reporting photon beam therapy. ICRU Report 50," *Journal of the ICRU*, vol. 50, no. 1, 1993.

[63] P. Castadot, X. Geets, J. A. Lee, and V. Grégoire, "Adaptive functional image-guided IMRT in pharyngo-laryngeal squamous cell carcinoma: Is the gain in dose distribution worth the effort?," *Radiotherapy and Oncology*, vol. 101, pp. 343–350, 2011.

[64] S. van Kranen, S. van Beek, A. Mencarelli, C. Rasch, M. van Herk, and J. J. Sonke, "Correction strategies to manage deformations in head-and-neck radiotherapy," *Radiotherapy and Oncology*, vol. 94, no. 2, pp. 199–205, 2010.

[65] A. C. Hartford, M. G. Palisca, T. J. Eichler, D. C. Beyer, V. R. Devineni, G. S. Ibbott, B. Kavanagh, J. S. Kent, S. A. Rosenthal, C. J. Schultz, P. Tripuraneni, and L. E. Gaspar, "American society for therapeutic radiology and oncology (astro) and american college of radiology (acr) practice guidelines for intensity-modulated radiation therapy (imrt)," *International Journal of Radiation Oncology*Biology*Physics*, vol. 73, no. 1, pp. 9 – 14, 2009.

[66] S. F. Kry, A. Molineu, J. R. Kerns, A. M. Faught, J. Y. Huang, K. B. Pulliam, J. Tonigan, P. Alvarez, F. Stingo, and D. S. Followill, "Institutional Patient-specific IMRT QA Does Not Predict Unacceptable Plan Delivery," *International Journal of Radiation Oncology*Biology*Physics*, vol. 90, no. 5, pp. 1195–1201, 2014.

[67] E. M. Mckenzie, P. A. Balter, F. C. Stingo, J. Jones, D. S. Followill, and F. Kry, "Toward optimizing patient-specific IMRT QA techniques in the accurate detection of dosimetrically acceptable and unacceptable patient plans," vol. 41, p. 121702, 2014.

[68] W. Ansbacher, "Three-dimensional portal image-based dose reconstruction in a virtual phantom for rapid evaluation of imrt plans," *Medical Physics*, vol. 33, no. 9, pp. 3369–3382, 2006.

[69] C. Talamonti, M. Casati, and M. Bucciolini, "Pretreatment verification of IMRT absolute dose distributions using a commercial a-Si EPID," *Medical Physics*, vol. 33, no. 11, pp. 4367–4378, 2006.

[70] M. van Zijtveld, M. L. P. Dirkx, H. C. J. de Boer, and B. J. M. Heijmen, "Dosimetric pretreatment verification of IMRT using an EPID; clinical experience," *Radiotherapy and Oncology*, vol. 81, no. 2, pp. 168–175, 2006.

[71] B. E. Nelms, K. H. Rasmussen, and W. A. Tome, "Evaluation of a fast method of EPID-based dosimetry for intensity-modulated radiation therapy," *Journal of Applied Clinical Medical Physics*, vol. 11, no. 2, p. 3185, 2010.

[72] D. S. Sharma, V. Mhatre, M. Heigrujam, K. Talapatra, and S. Mallik, "Portal dosimetry for pretreatment verification of IMRT plan: a comparison with 2D ion chamber array," *Journal of Applied Clinical Medical Physics*, vol. 11, no. 4, p. 3268, 2010.

[73] D. W. Bailey, L. Kumaraswamy, M. Bakhtiari, H. K. Malhotra, and M. B. Podgorsak, "EPID dosimetry for pretreatment quality assurance with two commercial systems," *Journal of Applied Clinical Medical Physics*, vol. 13, no. 4, p. 3736, 2012.

[74] M. van Zijtveld, M. L. P. Dirkx, H. C. J. de Boer, and B. J. M. Heijmen, "3d dose reconstruction for clinical evaluation of {IMRT} pretreatment verification with an {EPID}," *Radiotherapy and Oncology*, vol. 82, no. 2, pp. 201 – 207, 2007.

[75] W. van Elmpt, S. Nijsten, B. Mijnheer, A. Dekker, and P. Lambin, "The next step in patient-specific QA: 3D dose verification of conformal and intensity-modulated RT based on EPID dosimetry and Monte Carlo dose calculations," *Radiotherapy and Oncology*, vol. 86, no. 1, pp. 86–92, 2008.

[76] W. D. Renner, M. Sarfaraz, M. A. Earl, and C. X. Yu, "A dose delivery verification method for conventional and intensity-modulated radiation therapy using measured field fluence distributions," *Medical Physics*, vol. 30, no. 11, pp. 2996–3005, 2003.

[77] G. Narayanasamy, T. Zalman, C. S. Ha, N. Papanikolaou, and S. Stathakis, "Evaluation of Dosimetry Check software for IMRT patient-specific quality assurance," *Journal of Applied Clinical Medical Physics*, vol. 16, no. 3, pp. 329–338, 2015.

[78] D. A. Low and J. F. Dempsey, "Evaluation of the gamma dose distribution comparison method," *Medical Physics*, vol. 30, no. 9, p. 2455, 2003.

[79] M. Wendling, L. J. Zijp, L. N. McDermott, E. J. Smit, J.-J. Sonke, B. J. Mijnheer, and M. van Herk, "A fast algorithm for gamma evaluation in 3D," *Medical Physics*, vol. 34, no. 5, pp. 1647–1654, 2007.

[80] T. Sanghangthum, S. Suriyapee, S. Srisatit, and T. Pawlicki, "Statistical process control analysis for patient-specific imrt and vmat qa," *Journal of Radiation Research*, vol. 54, no. 3, pp. 546–552, 2013.

[81] V. Chaswal, M. Weldon, N. Gupta, A. Chakravarti, and Y. Rong, "Commissioning and comprehensive evaluation of the arccheck cylindrical diode array for vmat pretreatment delivery qa," *Journal of Applied Clinical Medical Physics*, vol. 15, no. 4, 2014.

[82] X. Jin, H. Yan, C. Han, Y. Zhou, J. Yi, and C. Xie, "Correlation between gamma index passing rate and clinical dosimetric difference for pre-treatment 2d and 3d volumetric modulated arc therapy dosimetric verification," *The British Journal of Radiology*, vol. 88, no. 1047, 2015.

[83] G. A. Ezzell, J. W. Burmeister, N. Dogan, T. J. LoSasso, J. G. Mechalakos, D. Mihailidis, A. Molineu, J. R. Palta, C. R. Ramsey, B. J. Salter, J. Shi, P. Xia, N. J. Yue, and Y. Xiao, "Imrt commissioning: Multiple institution planning and dosimetry comparisons, a report from aapm task group 119," *Medical Physics*, vol. 36, no. 11, pp. 5359–5373, 2009.

[84] D. K. Mynampati, R. Yaparpalvi, L. Hong, H.-C. Kuo, and D. Mah, "Application of aapm tg 119 to volumetric arc therapy (vmat)," *Journal of Applied Clinical Medical Physics*, vol. 13, no. 5, 2012.

[85] S. Thirumalai-Swamy, C. Anuradha, l. M. Kathirve, G. Arun, and S. Subramanian, "Pretreatment quality assurance of volumetric modulated arc therapy on patient ct scan using indirect 3d dosimetry system," *International Journal of Cancer Therapy and Oncology*, vol. 2, no. 4, p. 20416, 2014.

[86] P. B. Greer, "3D EPID based dosimetry for pre-treatment verification of VMAT methods and challenges," *Journal of Physics: Conference Series*, vol. 444, no. 1, p. 12010, 2013.

[87] I. Olaciregui-Ruiz, R. Rozendaal, B. Mijnheer, M. van Herk, and A. Mans, "Automatic in vivo portal dosimetry of all treatments," *Physics in Medicine and Biology*, vol. 58, no. 22, pp. 8253–64, 2013.

[88] P. M. McCowan, E. Van Uytven, T. Van Beek, G. Asuni, and B. M. C. McCurdy, "An in vivo dose verification method for sbrtvmat delivery using the epid," *Medical Physics*, vol. 42, no. 12, 2015.

[89] E. Van Uytven, T. Van Beek, P. M. McCowan, K. Chytyk-Praznik, P. B. Greer, and B. M. C. McCurdy, "Validation of a method for in vivo 3D dose reconstruction for IMRT and VMAT treatments using on-treatment EPID images and a model-based forward-calculation algorithm," *Medical Physics*, vol. 42, no. 12, pp. 6945–6954, 2015.

[90] A. Fidanzio, A. Porcelli, L. Azario, F. Greco, S. Cilla, M. Grusio, M. Balducci, V. Valentini, and A. Piermattei, "Quasi real time in vivo dosimetry for VMAT," *Medical Physics*, vol. 41, no. 6, p. 062103, 2014.

[91] S. Cilla, D. Meluccio, A. Fidanzio, L. Azario, A. Ianiro, G. Macchia, C. Digesù, F. Deodato, V. Valentini, A. G. Morganti, and A. Piermattei, "Initial clinical experience with Epid-based invivo dosimetry for VMAT treatments of head-and-neck tumors," *Physica Medica*, vol. 32, no. 1, pp. 52–58, 2016.

[92] `http://www.openmp.org`.

[93] R. A. Rozendaal, B. J. Mijnheer, O. Hamming-Vrieze, A. Mans, and M. van Herk, "Impact of daily anatomical changes on EPID-based in vivo dosimetry of VMAT treatments of head-and-neck cancer," *Radiotherapy and Oncology*, vol. 116, no. 1, pp. 70–74, 2015.

[94] J. Shafiq, M. Barton, D. Noble, C. Lemer, and L. J. Donaldson, "An international review of patient safety measures in radiotherapy practice," *Radiotherapy and Oncology*, vol. 92, no. 1, pp. 15–21, 2009.

[95] H. Spreeuw, R. Rozendaal, I. Olaciregui-ruiz, A. Mans, B. Mijnheer, and M. V. Herk, "Online 3D EPID-based dose verification : proof of concept," *Medical Physics*, vol. 3969, no. 43, pp. 1–13, 2016.

[96] A. Holt, C. Van Vliet-Vroegindeweij, A. Mans, J. S. Belderbos, and E. M. F. Damen, "Volumetric-modulated arc therapy for stereotactic body radiotherapy of lung tumors: A comparison with intensity-modulated radiotherapy techniques," *International Journal of Radiation Oncology*Biology*Physics*, vol. 81, no. 5, pp. 1560–1567, 2011.

[97] J. V. Lebesque and R. B. Keus, "The simultaneous boost technique: the concept of relative normalized total dose," *Radiotherapy and Oncology*, vol. 22, no. 1, pp. 45–55, 1991.

[98] C. Bojechko and E. C. Ford, "Quantifying the performance of in vivo portal dosimetry in detecting four types of treatment parameter variations," *Medical Physics*, vol. 42, no. 12, pp. 6912–6918, 2015.

[99] "Dosisoft epigray." `https://www.dosisoft.com/patient-qa/in-vivo-epigray.html`.

[100] "SunNuclear PerFraction." `http://www.sunnuclear.com/solutions/patientqa/perfraction`.

[101] "IBA OmniPro-InViDos." `http://www.iba-dosimetry.com/complete-solutions/radiotherapy/in-vivo-dosimetry/omnipro-invidos`.

[102] "Elekta iViewDose." `https://www.elekta.com/radiotherapy/treatment-solutions/quality-assurance/iviewdose`.

[103] M. Podesta, S. M. J. J. G. Nijsten, L. C. G. G. Persoon, S. G. Scheib, C. Baltes, and F. Verhaegen, "Time dependent pre-treatment EPID dosimetry for standard and FFF VMAT," *Physics in Medicine and Biology*, vol. 59, no. 16, pp. 4749–4768, 2014.

[104] "SunNuclear ArcCheck & 3DVH." `http://www.sunnuclear.com/solutions/patientqa/arccheck3dvh`.

[105] "PTW Octavius 4D." `http://www.ptw.de/2403.html`.

Summary

Radiotherapy is a powerful method for treating cancer. It is also a very complex method, requiring many different systems and advanced machines to cooperate seamlessly. As such, Quality Assurance (QA) plays a very important role within radiotherapy; different QA systems are needed for different parts of the radiotherapy treatment chain. *In vivo* transit dosimetry takes in a unique spot in the range of QA tools available for radiotherapy: as it measures during treatment, it takes into account both the delivered dose and the patient anatomy. Two-dimensional dose measurements at the side where the radiation exits the patient allows for 2D and 3D reconstruction of the dose delivered to the patient, providing a means of directly comparing the delivered dose to the intended dose. Currently, these 2D measurements are done by using Electronic Portal Imaging Devices, or EPIDs, and used in the day-to-day practice of our clinic. The aim of this thesis was to improve on the patient safety in radiotherapy by enabling this EPID *in vivo* dosimetry system to be used efficiently on a large-scale, and ultimately in real-time.

Several EPID-based *in vivo* dosimetry initiatives exist worldwide and most are used clinically, albeit on a restricted scale. The situation at the NKI is different and unique: *in vivo* dosimetry is being applied clinically on a large scale. Almost all external beam photon-therapy treatments are being verified via *in vivo* dosimetry. Making such a large scale deployment possible was the topic of **Chapter 2**. This was the starting point of this thesis.

With the large scale application of *in vivo* dosimetry in place, several issues unique to verifying delivered dose to the patient come up. These revolve mostly around obtaining dosimetric results which are clinically relevant and investigating - maybe even improving - the sensitivity and specificity of the system. As by far the most dosimetric QA world-wide is done pre-treatment on a rectangular or cylindrical piece of plastic, it is not surprising that dosimetric results are presented in a way unrelated to the patient anatomy; usually via a method called γ-analysis. This method provides a metric for

quantifying differences between two dose distributions in terms of local dose differences and distance-to-agreement. With the delivered dose distribution known, however, it is possible to move to Dose-Volume Histogram (DVH)-analysis, resembling more closely the way the intended (or planned) dose distribution is evaluated. The relationship between γ- and DVH-analysis for *in vivo* dosimetry of several treatment sites was investigated in **Chapter 3**. The agreement between the two methods was better than anticipated, but DVH-analysis was shown to have an advantage over γ-analysis in terms of sensitivity.

A potential issue for *in vivo* dosimetry arises when the planning CT-scan is used for dose reconstruction. A patient's anatomy can change considerably over time, and using the anatomy as recorded on the planning CT could lead to inaccuracies in the reconstructed dose. If the patient's anatomy has changed between recording the planning CT and dose delivery, there are three different dose distributions to deal with: the intended dose distribution, the delivered dose distribution and the reconstructed dose distribution. Differences between these three were investigated in **Chapter 4**, where the actual patient anatomy during treatment delivery was taken into account by using the cone-beam CTs obtained right before dose delivery[1]. Despite changes in relevant[2] volume of up to ~10%, it was demonstrated that changes in the delivered dose to the target volume were negligible (though demonstrable). Changes in reconstructed dose were somewhat larger – in accordance with theoretical expectations – and using the cone-beam CT information for dose reconstruction did lead to a 30% drop in false positives, thus improving on the method's sensitivity.

Even with a large-scale clinical implementation of *in vivo* dosimetry, there remains a need for separate, non-*in vivo* dose-verification measurements, akin to typical pre-treatment phantom measurements. This is the case for example when a *in vivo* deviation is observed and one wishes to verify the linac's dose delivery without the patient being present. For a DVH-based dose-verification system, this poses a challenge since the DVHs are not well-defined on any geometry which is not identical to the patient's anatomy. A method to resolve this issue was presented in **Chapter 5**, where EPID measurements without a patient or phantom in the beam were used to obtain a dose distribution which would have been found *in vivo* if the patient's anatomy would have been *exactly* as recorded on the planning-CT. Hence the name 'Virtual QA': the patient is *virtually* present. The validity of this extension of the *in vivo* dosimetry method was demonstrated, as was the improved accuracy due to not having any uncertainties about the patient's

[1] These cone-beam CTs are used in clinical routine for patient positioning.

[2] Relevant for dosimetric calculations, see Chapter 4.

anatomy.

In the last part of the thesis we change gears slightly, focussing on a different topic in *in vivo* dosimetry: real-time dose verification. A drawback of current *in vivo* dosimetry solutions is its post-factum nature: the dose is verified only after a fraction has been delivered. In case of very large errors in dose delivery, this could be too late to save the patient from serious harm. A solution is real-time dosimetry, where dose is reconstructed and verified in real-time during delivery. **Chapter 6** describes the development and performance of such a real-time DVH-based dosimetry system, capable of interrupting dose delivery whenever large deviations are observed.

The obvious next step for real-time dosimetry is then clinical implementation. This is a delicate subject, mainly because of the unknowns around setting approriate tolerance levels for error detection. One wishes to be as strict as possible, but false positives in real-time dosimetry are highly undesirable as they would have a large impact on clinical workflow and patient comfort. A systematic method for setting tolerance levels is presented in **Chapter 7**; it is also applied to a specific type of treatment: lung stereotactic body radiation therapy (SBRT). It was demonstrated that the real-time dosimetry system for lung SBRT can be made both very sensitive and very specific by using time-varying tolerance levels.

All in all, it has been demonstrated how a clinical EPID *in vivo* dosimetry system can be used on a large scale; improving patient safety by enabling valuable QA-checks for nearly all patients. Further improvements are found in moving to DVH-based analysis and including the anatomy of the patient during treatment. Finally, *in vivo* real-time dosimetry has been made possible and a method for establishing tolerance levels – necessary for clinical introduction – has been developed. The usefulness of the EPID dosimetry system is further acknowledged by the development of a commercial version, in collaboration with Elekta, one of the two large manufacturers of linacs.

Samenvatting

Radiotherapie is een krachtige methode om kanker te behandelen. Het is echter ook een zeer complexe methode, die vereist dat een tiental computersystemen en geavanceerde machines vlekkeloos samenwerkt. Daarom speelt continu bewaken van de behandelkwaliteit – ook wel Quality Assurance of QA genoemd – een onmisbare rol. Op elk bestralingstoestel zit een beeldcamera gemonteerd, achter de patiënt. Men noemt dit een Electronic Portal Imaging Device, afgekort EPID. Met de EPID wordt de straling gemeten die door de patiënt heen komt, en hiermee kan ook een uitspraak worden gedaan over de bestralingsdosis die wordt afgegeven in de patiënt. Deze procedure wordt transit *in vivo* dosimetrie genoemd. In het spectrum van beschikbare QA-systemen neemt de transit *in vivo* dosimetrie een unieke positie in: er wordt een dosismeting verricht tijdens de behandeling, en dus wordt zowel de afgeleverde bestralingsdosis als de ligging en anatomie van de patiënt gecontroleerd. We gebruiken een methode om de afgegeven bestralingsdosis te schatten in drie dimensies. Zodoende kan de afgegeven dosis direct worden vergeleken met de geplande bestralingsdosis in de hele patiënt.

Dit proefschrift beschrijft de resultaten van onderzoek om de patiëntveiligheid binnen de radiotherapie te verbeteren door een efficiënte, grootschalige en uiteindelijk instantane (real-time) implementatie van transit *in vivo* dosimetrie.

Wereldwijd worden enkele EPID-gebaseerde *in vivo* dosimetrie systemen klinisch gebruikt, zij het op beperkte schaal. Op het NKI verkeren we in een unieke positie: *in vivo* dosimetrie wordt op grote schaal gebruikt in de kliniek: alle radiotherapie behandelingen die gecontroleerd kunnen worden door EPID dosimetrie, worden dat ook daadwerkelijk. Het mogelijk maken van deze grootschalige implementatie is beschreven in Hoofdstuk 2 van dit proefschrift.

Na de grootschalige introductie van *in vivo* dosimetrie verschoof het perspectief van het onderzoek. Met het geautomatiseerde systeem hoeven de fysici eigenlijk alleen nog afwijkende dosimetrische resultaten te bekijken, en het viel daarbij al gauw op dat onze

methode veel incorrecte foutmeldingen produceerde. Ook kwam er een diepere onderzoeksvraag naar boven, namelijk het onderzoeken van de klinische relevantie van een geconstateerde dosimetrische afwijking. In een traditioneel dosis-QA systeem wordt de afgegeven dosis op een stuk plastic geïnspecteerd, terwijl de *in vivo* methode de dosis in de patiënt schat. Het ligt voor de hand om daarbij de resultaten op een andere manier te beoordelen, namelijk aan de hand van de dosis aan verschillende organen. Hoofdstuk 3 laat zien dat er inderdaad duidelijke voordelen zijn aan zon beoordelingsmethodiek; waaronder een verminderd aantal incorrect foutmeldingen en een duidelijkere klinische relevantie.

Een ander belangrijk punt in de *in vivo* dosimetrie is dat een patiënt continu verandert denk aan gewichtsverandering en toe- of afname van zwellingen. Uiteraard heeft dit invloed op de dosimetrische resultaten, maar hoe groot deze invloed exact is, was nog niet goed bekend. In Hoofdstuk 4 is dit uitgezocht voor patiënten met hoofd-hals kanker. Interessant genoeg bleek het effect van dergelijke veranderingen voor de patiënt kleiner dan verwacht, maar was er wel ook een forse afname van het incorrecte foutmeldingen te zien zodra er voor anatomische veranderingen gecorrigeerd werd.

Ook met grootschalige *in vivo* dosimetrie blijft er behoefte bestaan aan controle meting van een bestralingsplan zonder dat de patiënt aanwezig is, bijvoorbeeld als de bestraling in slechts een paar fracties wordt gegeven. Het is natuurlijk niet mogelijk om de dosis in de patiënt te meten als de patiënt niet aanwezig is. Hoofdstuk 5 beschrijft een methode die uitkomst biedt: er wordt een dosis-meting gedaan zonder een patiënt (en ook zonder een stuk plastic), waarna de dosis zoals die afgegeven zou zijn aan de patiënt uitgerekend wordt. Zo wordt de EPID dosimetrie nog beter in het vinden van relevante afwijkingen.

Het laatste gedeelte van dit proefschrift beschrijft een nieuwe ontwikkeling in de *in vivo* EPID dosimetrie: instantane of "real-time" dosis verificatie. Dit houdt in dat de afgegeven dosis al tijdens de bestraling uitgerekend en gecontroleerd wordt. Bij – gelukkig zeer zeldzaam voorkomende – forse afwijkingen in de afgegeven dosis kan dan de behandeling stopgezet worden voordat er onherstelbare schade aan de patiënt wordt toegebracht. Hoe dit technisch voor elkaar is gebokst staat beschreven in Hoofdstuk 6. Dan gaan we nog een stap verder: hoe zou je een dergelijk systeem in de kliniek kunnen introduceren? Hoofdzaak hierbij is het juist afstellen van de "alarmbellen" om (1) zo snel mogelijk te behandeling af te breken als er een fout optreedt, en (2) zo min mogelijk behandelingen onnodig te onderbreken. Deze vraag is bestudeerd voor een specifiek type behandeling waarbij in korte tijd een hoge dosis wordt gegeven (long stereotaxie). Een

systematische methode om de balans tussen deze twee eisen te vinden is beschreven en toegepast in Hoofdstuk 7.

Tot slot is het nog noemenswaardig dat één van de twee grote leveranciers van bestralingstoestellen, Elekta, de op het NKI ontwikkelde EPID *in vivo* dosimetrie methode in een commerci´eel produkt heeft verwerkt, iViewDose. De samenwerking met Elekta onderstreept het nut en belang van *in vivo* dosimetrie in de klinische praktijk.

List of publications

Included in this thesis

- I. Olaciregui-Ruiz, R.A. Rozendaal, B. Mijnheer, M. van Herk, A. Mans, *Automatic in vivo portal dosimetry of all treatments*, Physics in Medicine and Biology 58 (22), 2013

- R.A. Rozendaal, B. Mijnheer, M. van Herk, A. Mans, *In vivo portal dosimetry for head-and-neck VMAT and lung IMRT: linking γ-analysis with differences in dose-volume histograms*, Radiotherapy and Oncology 112 (396), 2014

- R.A. Rozendaal, B. Mijnheer, O. Hamming-Vrieze, A. Mans, M. van Herk, *Impact of daily anatomical changes on EPID-based in vivo dosimetry of VMAT treatments of head-and-neck cancer*, Radiotherapy and Oncology 116 (1), 2015

- H. Spreeuw and R.A. Rozendaal[1], I. Olaciregui-Ruiz, P. González, A. Mans, B. Mijnheer, M. van Herk, *Online 3D EPID-based dose verification: proof of concept*, Medical Physics 43 (7), 2016

- I. Olaciregui-Ruiz, R.A. Rozendaal, R.F.M. van Oers, B. Mijnheer, A. Mans, *Virtual patient 3D dose reconstruction using in air EPID measurements and a back-projection algorithm for IMRT and VMAT treatments*, Physica Medica 37 (49), 2017

- R.A. Rozendaal, B. Mijnheer, P. González, I. Olaciregui-Ruiz, J.-J. Sonke, A. Mans, M. van Herk, *Real-time 3D in vivo dosimetry for lung SBRT VMAT*, Radiotherapy and Oncology (under review)

[1] H. Spreeuw and R.A. Rozendaal contributed equally

Not included in this thesis

Within medical physics

- H. Spreeuw, R.A. Rozendaal, P. Camargo, A. Mans, M. Wendling, I. Olaciregui-Ruiz, J.-J. Sonke, M. van Herk, B. Mijnheer, *Portal dosimetry in wedged beams*, Journal of applied clinical medical physics 16 (3), 2015

Outside of medical physics

- R. Meppelink, R.A. Rozendaal, S.B. Koller, J.M. Vogels, P. van der Straten, *Thermodynamics of Bose-Einstein-condensed clouds using phase-contrast imaging*, Physical Review A 81, 2010
- R. van Rooij, J.S. Borbely, J. Simonet, M.D. Hoogerland, K.S.E. Eikema, R.A. Rozendaal, W. Vassen, *Frequency metrology in quantum degenerate Helium: direct measurement of the* $2\ ^3S_1 \rightarrow 2\ ^1S_0$ *transition*, Science 333, 2011

Curriculum Vitæ

My name is Roel Rozendaal, and I was born on November 3rd, 1979 in Leeuwarden, The Netherlands. Already as a teenager I spent lots of my spare time in combining physics and computer science – optimizing computer simulations compiled by a 16-bit C-compiler with 32-bit assembly code, to get the maximum performance from my trusty 386. This continued after graduating from the Prædinius Gymnasium (Groningen, NL) in 1997 with a mixture of CS and physics studies at Groningen and Utrecht University. A side-job as a programmer at a successful IT start-up in Utrecht led to a break from university, and allowed me to dive first into software development and later into software architecture and project management. Exciting as it was, I really wanted to finish my studies and so I returned to Utrecht University in 2005. After graduation in experimental atomic physics in 2007 and spending some more time in atomic physics, I joined the NKI in 2010. It was again a combination of physics and IT; first in a mostly clinical setting and then in a mostly research setting. The research part of my work in EPID dosimetry has culminated in this thesis.

Dankwoord

Bij deze wil ik graag iedereen die direct of indirect aan de totstandkoming van dit proefschrift heeft bijgedragen bedanken.

Ok, ok. Misschien nog een *paar* mensen in het bijzonder.

Igor, **Hanno**, **Patrick**: pure magie. **Marcel** (even briljant als koppig), **Anton** (beste baas ooit), **Ben** (gentleman pur sang), **Jan-Jakob** (speuluh): het wetenschappelijke fundament. **Diedie**: zou jij..? **Nienke** (Nienkie!), **Tessa** (dochterlief), **Yenny** (even kletsen?), **Catarina** (surf's up!): mijn lievelingen. **Zeno**, **Iban**, **Bruno**, **Tom**, **Sander**: let's book a trip to Turin via Renesse. **Jolien**, **Rosemarie**: lunch? **Natasja**, **Kleopatra** (how's the piano?): you've got free counselling for life. **Matthijs**, **Mustafa**: I've got an idea! **Vincent**, **Robbie**, **Arjen**, **Thijs**, **Carlo**: oude liefde roest niet. **Geert**, **Skad**i: Maastricht is echt te ver. **Yolanda**, **Renate** (GnD), **Marieke**, **Antonio**, **Stien**, **Robert**, **Angela**: lieve schatten, waar heb ik jullie toch aan te danken? **Donna**: coffee? **Papa**, **Mama**: waar ik uit besta. **Walter**, **Frieda**: thanks for keeping an eye on those two. **Maarten**, **Muriël**: het bloed kruipt waar het niet gaan kan.

En dan nog.. **Gina** en **Olivia**, jullie zijn mijn mooiste creaties ooit. Ik houd zoveel van jullie. Heb ik dat wel eens gezegd? En zullen jullie het nooit vergeten? Lieve **Jaimi**, hoe hebben we dat toch voor elkaar gekregen?! Ik denk dat onze meisjes het ultieme bewijs zijn dat wij samen horen te zijn. Ik hou van je.