Validation and uncertainty analysis of a pre-treatment 2D dose prediction model

To cite this article: Jose A Baeza et al 2018 Phys. Med. Biol. 63 035033

View the article online for updates and enhancements.
Validation and uncertainty analysis of a pre-treatment 2D dose prediction model

Jose A Baeza¹, Cecile J A Wolfs¹, Sebastiaan M J G Nijsten² and Frank Verhaegen²

Department of Radiation Oncology (MAASTRO), GROW—School for Oncology and Developmental Biology, Maastricht University Medical Centre, Maastricht, Netherlands

¹ These authors are equally contributing first authors to this study.
² These authors are equally contributing last authors to this study.

E-mail: frank.verhaegen@maastro.nl

Keywords: pre-treatment verification, 2D dose prediction, uncertainty analysis, sensitivity analysis

Abstract

Independent verification of complex treatment delivery with megavolt photon beam radiotherapy (RT) has been effectively used to detect and prevent errors. This work presents the validation and uncertainty analysis of a model that predicts 2D portal dose images (PDIs) without a patient or phantom in the beam.

The prediction model is based on an exponential point dose model with separable primary and secondary photon fluence components. The model includes a scatter kernel, off-axis ratio map, transmission values and penumbra kernels for beam-delimiting components. These parameters were derived through a model fitting procedure supplied with point dose and dose profile measurements of radiation fields. The model was validated against a treatment planning system (TPS; Eclipse) and radiochromic film measurements for complex clinical scenarios, including volumetric modulated arc therapy (VMAT). Confidence limits on fitted model parameters were calculated based on simulated measurements. A sensitivity analysis was performed to evaluate the effect of the parameter uncertainties on the model output. For the maximum uncertainty, the maximum deviating measurement sets were propagated through the fitting procedure and the model. The overall uncertainty was assessed using all simulated measurements.

The validation of the prediction model against the TPS and the film showed a good agreement, with on average 90.8% and 90.5% of pixels passing a (2%,2 mm) global gamma analysis respectively, with a low dose threshold of 10%. The maximum and overall uncertainty of the model is dependent on the type of clinical plan used as input. The results can be used to study the robustness of the model.

A model for predicting accurate 2D pre-treatment PDIs in complex RT scenarios can be used clinically and its uncertainties can be taken into account.

1. Introduction

The trend towards complex photon beam delivery techniques in radiotherapy, combined with hypofractionation, intensifies the need for quality assurance methods. Pre-treatment dose verification is well established in clinical practice and several methods exist to verify a treatment prior to irradiating the patient. A differentiation can be made between the detection of delivery errors with a mechanical or data transfer origin, and the detection of dosimetric deviations introduced by commercial dose engines. While the former can be detected with simplistic setups and efficient dose calculations, the latter entails the use of sophisticated dose engines, more accurate than the commercial methods themselves, often at the expense of longer calculation times, such as full Monte Carlo simulations. Independent fluence models and dose engines have been designed to efficiently predict the treatment delivery in pre-treatment conditions (Jiang et al 2001, Fippel et al 2003, Van Esch et al 2004, Baker et al 2006, Chytyk and McCurdy 2009). Generally, the predicted dose is intended to be compared with a measurement.
This work presents a model that provides a two-dimensional (2D) absolute dose prediction in pre-treatment conditions, i.e. without a patient or phantom in the beam. This model is integrated in our electronic portal imaging device (EPID) dosimetry routine, therefore the prediction is provided at the level of the imaging panel and it is referred to as portal dose image (PDI), although the prediction is not limited to a unique plane.

The most obvious use of the predicted PDIs is for comparison with PDIs derived from EPID measurements (van Elmpt et al 2005). However, the utility of the prediction model is not limited to pre-treatment verification, but it can also be used for the in-house monitor unit check and the in vivo dosimetry routines. It is worth noticing that predicted PDIs can be obtained from the RTPLAN, but also from the trajectory logs, allowing a more accurate description of the irradiation.

The inclusion of more, and more accurate, imaging solutions during the course of treatment together with the automation of some of the most time consuming aspects in radiotherapy (RT) planning, increase the feasibility of adaptive radiotherapy (ART) approaches. However, the effective implementation of ART necessitates quantification of the uncertainties involved in the ART process and their propagation into the final dose calculation. Quantiﬁying these uncertainties is expected to enable the generation of effective and customized adaptation strategies. Therefore, to evaluate the presented model’s applicability as a veriﬁcation method, the model uncertainty and sensitivity analysis are provided.

In summary, the aim of this work is to present a 2D dose prediction model as an independent method for pre-treatment veriﬁcation. The model description is accompanied by its model ﬁtting process and validation. Furthermore, uncertainty and sensitivity analyses are performed to evaluate the model’s robustness.

2. Materials and methods

2.1. Theoretical model

The presented prediction model is an exponential point dose model with separable primary and secondary (scatter) components, initially inspired by the work published by Baker et al (2006) and Hounsell and Wilkinson (1997). The model includes an input ﬂuence matrix that accounts for the dosimetric impact of the beam limiting components, extending the point dose method to a 2D dose prediction. This model has been developed in full scatter conditions, and provides the dose prediction in a plane within a homogeneous virtual phantom (ﬁgure 1).

The model deﬁnes the dose at the beam central axis (CAX) of a circular ﬁeld with radius $r$, normalized to the dose of the largest ﬁeld, in this work 40 cm, as

$$
\frac{D_{CAX}(r)}{D_{CAX}(r_{\text{max}})} = \frac{D_{CAX}(r)}{D_{CAX}(r_{\text{max}})} = 1 - a e^{-b r} \tag{1}
$$

where $a$ and $b$ are empirical parameters that establish an initial steep dose variation near the CAX, which can be attributed to the inﬂuence of the ﬂattening ﬁlter (FF), and an extended tail with radius $r$ that could be related to other scattering effects, as suggested by Hounsell and Wilkinson (1997) in a similar model. Theoretically, the model allows for the separation of primary dose $D_{p,CAX}$ and secondary dose $D_{s,CAX}$.

$$
D_{p,CAX} = 1 - a \tag{2}
$$

$$
D_{s,CAX}(r) = a(1 - e^{-b r}) \tag{3}
$$

$D_{p,CAX}$ is obtained by setting $r$ to 0 in (1), and $D_{s,CAX}$ is obtained by subtracting $D_{p,CAX}$ from $D_{CAX}$. For an arbitrary irregular treatment ﬁeld, the secondary scatter dose at the CAX can be obtained by integrating the individual differential contributions of $N$ circular segments of radius $r$ and area $\Delta A$.

$$
D_{s,CAX} = \frac{a b \Delta A}{2\pi} \sum_{i=1}^{N} \frac{e^{-b r_i}}{r_i} \tag{4}
$$

To expand this model by Baker et al (2006) from CAX to a planar prediction, a scatter kernel accounting for all scatter sources was derived from (4) and a 2D ﬂuence map incorporating the dosimetric characteristics of the beam delimiting components was included. The total dose can be calculated in a Cartesian coordinate system by

$$
D = D_p + D_s = [(1 - a) \cdot \Phi(x, y) + K_s \otimes \Phi(x, y)] \cdot C \tag{5}
$$

where the photon ﬂuence $\Phi$ is scaled by the primary component in the ﬁrst term, and convolved with a scatter kernel $K_s$ in the second term. $C$ is a normalization constant that converts monitor units to absolute dose.

The ﬂuence is deﬁned as

$$
\Phi(x, y) = OAR(x, y) \cdot MU \cdot \Phi(\theta) \cdot \Phi_{c,jaw/s,MLC} \cdot T(x, y) \tag{6}
$$

where the off-axis ratio map OAR is a radially symmetric primary dose distribution that includes the effects of the FF and the off-axis beam softening, MU is the number of monitor units prescribed by the treatment planning system (TPS) and the other terms are correction factors.
\( \mathcal{R} \) represents a rotation by collimator angle \( \theta \). The \( S_{\text{c,jaws}} / S_{\text{c,MLC}} \) ratio corrects the fluence for the relative difference between jaws and the MLC apertures, as the jaw shielding effect before the MLC has been demonstrated to have a non-negligible effect in dose delivery, especially for highly modulated treatments (Swinnen et al 2017). Finally, the transmission map \( T \) is a projection of the linac head geometry at the prediction plane. \( T \) is calculated using different transmission factors and penumbra kernels per beam delimiting component (X-jaws, Y-jaws, and MLC):

\[
T(x, y) = \left\{ \begin{array}{l}
\left\{ \frac{1}{T_{\text{JAW}, x}} |_{y(x, y) \in \text{JAW}_x} \right\} \otimes K_{P,\text{JAW}, x}(x) \cdot \left\{ \frac{1}{T_{\text{JAW}, y}} |_{y(x, y) \in \text{JAW}_y} \right\} \otimes K_{P,\text{JAW}, y}(y)
\end{array} \right.
\]

where \( T_{\text{JAW}, x} \) and \( T_{\text{JAW}, y} \) are transmission values for the cross-line and in-line jaw pairs respectively, and \( K_{P,\text{JAW}, x} \) and \( K_{P,\text{JAW}, y} \) are penumbra kernels for these jaws. \( T_{\text{MLC}} \) and \( T_{\text{TG}} \) are transmission values for the MLC, and \( K_{P,\text{MLC}, x} \) and \( K_{P,\text{MLC}, y} \) their penumbra kernels in the cross-line and in-line direction respectively. For every transmission factor a dedicated transmission map based on its projected geometry is created. The maps are convolved with their corresponding kernels and superimposed, resulting in an initial normalized fluence estimation.

To model the MLC leaves two regions are differentiated: (1) an inner part, at the center of the leaf in the in-line direction, and (2) a tongue-and-groove (TG) region at the leaf edges. The size of these theoretical regions is defined by the TG width. To account for the rounded leaf tips a dosimetric leaf gap (DLG) width is introduced in the projected geometry of the leaf pairs that are closed.

The transmission equation (7) is intended for linac models with one MLC and two pairs of jaws above the MLC, such as the TrueBeam (Varian Medical Systems, Palo Alto, CA). The model can be adapted to predict dose images for other types of linacs.

### 2.2. Model fitting

The model fitting process comprises several independent optimization procedures in cascade, fed with point dose and dose profile measurements, as shown in figure 2. Ideally, all parameters should be derived independently from a customized experiment. However, most parameters cannot be isolated. As detailed below, to address the interdependency among parameters, some of them need to be initially approximated and others require iterative optimization.

The presented model was implemented for both TrueBeam Millennium™ 120 Leaf MLC, and High Definition 120 Leaf MLC (Varian Medical Systems). The TrueBeam Millennium MLC, or non-STx, contains 40 inner leaf pairs of 5 mm width and 20 peripheral leaf pairs of 10 mm width allowing a maximum static field of 40 \( \times \) 40 cm\(^2\) at 100 cm from the source. The TrueBeam High Definition, or STx, MLC contains 32 inner leaf pairs of 2.5 mm width and 28 peripheral leaf pairs of 5 mm width allowing a maximum static field of 40 \( \times \) 22 cm\(^2\).

The fitting process needs to be executed for every working mode, i.e. for every beam energy, fluence mode (e.g. FF or flattening filter free (FFF)), and for different desired source-to-detector distances (SDD). For pretreatment purposes, an SDD of 100 cm is ideal but for transit dosimetry, a distance of e.g. 150 cm can be used.
2.2.1. Input measurements
To obtain point dose and dose profile measurements the prediction model’s reference conditions were reproduced in the Blue Phantom 2001 water tank (Scanditronix Wellhofer, Schwarzenbruck, Germany). Measurements were obtained with a CC13 ionization chamber (Scanditronix Wellhofer) with an active volume of 0.13 cm$^3$. No geometrical or spectral corrections were applied to the detector readout. Ideally, all model fitting inputs should have been obtained experimentally, but in this study, some of them were calculated with the TPS (Eclipse; Varian Medical Systems). The measurement set originally used for fitting is further referred to as $D_0$, and its details are outlined in table 1.

2.2.2. Scatter kernel derivation
The first part of the model to be derived was the scatter kernel $K_s$. To calculate $K_s$, the parameters $a$ and $b$ were fitted to output factor measurements at the CAX as in equation (1). To measure the output factors, point dose measurements of square fields of different sizes were obtained (see table 1). These fields were delimited by the jaws, and normalized to the largest field size.

2.2.3. Off-axis ratio derivation
To derive the OAR map, a rotationally symmetric map resulting from a diagonal profile of a 40 × 40 cm field measured with a CC13 was used. The OAR corrects for the primary dose as in (6) and it is independent of the scatter component as in (5), therefore an iterative process was used to adjust the OAR in every iteration using $K_s$ to recalculate the dose map and to compare it with the diagonal profile, until convergence. Transmission factors and penumbra kernels were set to zero.

2.2.4. Jaw modelling
The dosimetric impact of every jaw pair was modelled individually to fit dose profile measurements of fields delimited with the corresponding jaw pair. To optimize the transmission factor and the penumbra kernel together, the simplex search method described by Lagarias (1998) and built-in MATLAB 2010 (MathWorks Inc., Natick, MA, USA) was used. The optimization process aimed to minimize the number of pixels failing a 2% dose.
difference, 2 mm distance to agreement global gamma evaluation (Low et al 1998) between the predicted and measured dose profiles.

2.2.5. MLC modelling
As presented in equation (7), two sets of parameters were independently used to model the MLC in the in-line and cross-line directions. The in-line profile part of the model needed the determination of the transmission factors of the inner and TG region of the leaves ($T_{MLC}$ and $T_{TG}$, respectively), the TG region width, and the in-line penumbra kernel for the leaf $K_{P,MLC,y}$. To determine all parameters simultaneously the complexity of input measurements needed to be increased as shown in figure 3. The cross-line MLC leaf penumbra kernel $K_{P,MLC,x}$ was determined using simple cross-line profiles delimited with the MLC. The optimization method used to determine the MLC-related parameters consisted again of minimizing the pixels failing the gamma evaluation using MATLAB’s nonlinear programming solver described in 2.2.4.

2.3. Model validation
The prediction model was validated against the Eclipse TPS and radiochromic EBT3 film (Ashland Inc., Lexington, USA) for treatment fields in FF mode used to treat three patients. It should be noticed that the validation with the TPS is limited as some of the data used to fit the model was calculated with the TPS. The verified fields were (1) two static fields of a hybrid lung plan, (2) two arcs of a prostate plan, and (3) two arcs of a head-and-neck (H&N) plan. Details of these treatment fields are provided in table 2. To validate the model, verification plans were created setting the gantry angle to 0 degrees for all control points.

2.3.1. Validation against dosimetric film
Absolute dosimetry was performed with radiochromic EBT3 films. A customized protocol adapted from Lewis et al (2012) and Mathot et al (2014), was followed. Two film pieces, one non-irradiated and another irradiated in absolute dosimetry reference conditions, were used to scale the calibration curve for every measurement. Films were placed at 100 cm from the source in a phantom consisting of $40 \times 40$ cm$^2$ slabs of solid water RW3, with a nominal density of 1.045 g cm$^{-3}$ (PTW, Freiburg, Germany). The SSD was set to 95 cm and 15 cm of solid water was added underneath the film.

The films were scanned using an Epson 11000XL document scanner (Shinjuku, Tokyo, Japan) in transparency mode at a spatial resolution of 50 dpi. A 4 mm glass plate was placed between the film and the lid of the scanner. The software FilmQA Pro 5.0 (Ashland Inc.) was used to optimize the film positioning and to convert grayscale images into absolute dose values using its triple-channel correction to diminish scanner and film artefacts (Micke et al 2011). An in-house method to correct for the non-uniform lateral response of the scanner based on the work by Lewis and Chan (2015) was applied.

2.3.2. Validation against the TPS
To validate the prediction model against the TPS, verification plans were created in Eclipse for the selected fields. The plans were calculated with Acuros 11.0.31 as dose engine (Varian Medical Systems), setting the grid to 1 mm,

<table>
<thead>
<tr>
<th>Fitted parameters</th>
<th>Detector / method</th>
<th>Measurements</th>
<th>Field size (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a and b</td>
<td>CC13</td>
<td>Point dose</td>
<td>3, 6, 10, 15, 20, 25, 30, 40</td>
</tr>
<tr>
<td>OAR map</td>
<td>CC13</td>
<td>Point dose</td>
<td>10, 40</td>
</tr>
<tr>
<td>$K_{P,JAW,y}$</td>
<td>Eclipse (dose calculations)</td>
<td>Point dose</td>
<td>10, 20, 40</td>
</tr>
<tr>
<td>$T_{JAW}$</td>
<td></td>
<td>In-line profiles (jaw-defined fields)</td>
<td>10, 20</td>
</tr>
<tr>
<td>$K_{P,JAW,x}$</td>
<td>Eclipse (dose calculations)</td>
<td>Point dose</td>
<td>10, 20, 40</td>
</tr>
<tr>
<td>$T_{JAW}$</td>
<td></td>
<td>Cross-line profiles (jaw-defined fields)</td>
<td>10, 20</td>
</tr>
<tr>
<td>$K_{P,MLC,y}$</td>
<td>Eclipse (dose calculations)</td>
<td>Point dose</td>
<td>10, 40, penumbra patterns, MLC transmission</td>
</tr>
<tr>
<td>$T_{MLC}$</td>
<td>Eclipse (dose calculations)</td>
<td>Point dose</td>
<td>10, 20, 40</td>
</tr>
<tr>
<td>$T_{TG}$</td>
<td></td>
<td>In-line profiles (MLC-defined fields)</td>
<td>Penumbra patterns, MLC transmission</td>
</tr>
<tr>
<td>$K_{P,MLC,x}$</td>
<td>Eclipse (dose calculations)</td>
<td>Point dose</td>
<td>10, 20, 40, closed leaves (DLG width)</td>
</tr>
<tr>
<td>DLG width</td>
<td></td>
<td>Cross-line profiles (MLC-defined fields)</td>
<td>10, 20, closed leaves (DLG width)</td>
</tr>
</tbody>
</table>
on a virtual $40 \times 40 \times 40$ cm$^3$ homogeneous phantom with Hounsfield units equal to zero (mass density 1.01 g·cm$^{-3}$ and relative electron density 1.012). The plane at 5 cm depth was exported and compared with the model prediction.

2.3.3. Dose comparison
The 2D predicted dose was compared with both TPS and film measurements for the six fields, using an in-house developed gamma evaluation method (Persoon et al 2011, Podesta et al 2014). Different global dose difference and distance to agreement criteria were used, i.e. (1%, 1 mm), (2%, 2 mm) and (3%, 3 mm). Film and TPS dose matrices were considered as reference to evaluate the gamma function. The results of the gamma analysis were expressed as gamma pass rates (i.e. the fraction of pixels with $|\gamma| > 1$), which were calculated for pixels with doses above 10% of the maximum dose. Dose profiles were also obtained.

2.4. Uncertainty analysis
2.4.1. Uncertainty of fitted parameters.
The measurement set $D_0$ that is used for fitting the model parameters is, due to measurement uncertainty, only one potential input into the fitting procedure. From $D_0$, the fitted parameter set $p_0$ is derived, which is in turn only one potential representation of the true (unknown) parameter set $p_{true}$. This means that, had the measurements differed slightly and resulted in measurement set $D_x$, a different parameter set $p_x$ would have been derived, which would also be only one potential representation of $p_{true}$. Thus, there exists a distribution of possible parameter sets, meaning that each fitted parameter has its own uncertainty, which can be determined by establishing their confidence limits (Press et al 2007).

Confidence limits for the fitted parameters were established by simulating other possible measurement sets $D_j$ ($j = 1, \ldots, n$), which were within the measurement uncertainty of $D_0$. The model was then fitted using $D_j$ as input, leading to the distribution of possible parameter sets $p_j$. The simulations of $D_j$ were based on $D_0$ and its uncertainties. An overview of all measurements used as input for the fitting procedure is provided in table 1. The uncertainty of each of the inputs was considered to be 1%, and this value was taken as one standard deviation in a normal distribution, instead of as the maximum uncertainty (IAEA 2016). The analysis can similarly be repeated for any level of uncertainty.

The simulated measurements $D_j$ were randomly sampled from a normal distribution, with the mean being the actual measurements $D_0$ and the standard deviation being the uncertainty of those measurements. Each point dose measurement was sampled independently, while the profile measurements were sampled as a whole, i.e. they were scaled by a random factor which was dependent on the measurement uncertainty. For each original measurement, 100 measurements were simulated ($n = 100$).

From the distribution ($p_j - p_0$), the confidence limits of the fitted parameters were established. The 68%, 90% and 95% lower and upper confidence limits were determined by calculating the 16th and 84th, 5th and 95th, and 2.5th and 97.5th percentiles, respectively. For the 2D OAR map, this was done for each pixel separately.
2.4.2. Sensitivity analysis
A sensitivity analysis was performed to quantify the effect of the uncertainties of the fitted parameters on the output of the prediction model. The model was used to predict PDIs for the same clinical plans used for validation, while one fitted parameter of the prediction model was changed at a time. Each parameter was changed by its 68%, 90%, and 95% lower and upper confidence limits, and a new PDI was predicted after each change. The original PDI0 was compared with the PDIp resulting from running the model and changing a parameter. This comparison was made by calculating \( \frac{PDI_p - PDI_0}{\max PDI_0} \cdot 100 \), resulting in a dose difference for each pixel as percentage of the maximum dose in PDI0. Additionally, all PDIp were compared with PDI0 by a (2%, 2 mm) gamma analysis.

To evaluate the effect of the parameter changes in different areas of the PDIs, the dose differences and gamma analysis results were analyzed in several regions of the PDI. For static fields, the in-field region, the penumbra region and the out-of-field region were used. These regions were defined in PDI0 as pixels with a value >80% of the maximum dose, pixels with a value between 20% and 80% of the maximum dose, and pixels with a value <20% of the maximum dose, respectively. For the dynamic fields only in-field and out-of-field regions were considered, since there is no clear penumbra region for integrated dynamic fields. In these cases, in-field pixels had a value >50% of the maximum dose and out-of-field pixels had a value <50% of the maximum dose.

2.4.3. Uncertainty propagation
After establishing the uncertainty and sensitivity for each fitted parameter, the maximum uncertainty of the model was evaluated for each of the clinical cases presented. To this end, the fitting procedure was performed, with as input the \( D_j \) that deviated the most from \( D_0 \) (\( D_{\text{dev}} \)). The resulting parameter set \( p_j \) was then used to predict a PDI that contained the maximum uncertainty with respect to \( D_{\text{dev}} \): PDI(\( D_{\text{dev}} \)). \( D_{\text{dev}} \) was determined by calculating the root mean squared error (RMSE) for each \( D_j \) relative to \( D_0 \), specifically for the measurements for fitting the scatter kernel (resulting in \( a \) and \( b \)) and the OAR. PDI(\( D_{\text{dev}} \)) was compared with PDI0 by calculating dose deviations as in the sensitivity analysis \( \left( \frac{PDI(D_{\text{dev}}) - PDI_0}{\max PDI_0} \cdot 100 \right) \), and by gamma analysis.

2.4.4. Overall uncertainty
Last, the overall uncertainty was calculated. This was done by using all 100 \( p_j \) to predict PDIs, resulting in 100 PDIs per treatment field. These were all compared with PDI0 by calculating dose deviations as outlined in section 2.4.2. For each of the 100 PDIs, the median value of the dose deviations in each PDI region was determined. The overall uncertainty was taken as the mean of these 100 median dose deviations.

3. Results

3.1. Model fitting
The values obtained from the model fitting process for the STx MLC for the two energies used to validate the model are presented in table 3.

3.2. Model validation
Figure 4 shows in-line and cross-line dose profiles for the three dose modalities (prediction, TPS and film) and for one field from each clinical case. In all cases, the prediction model slightly overestimates the dose compared with the TPS and film, particularly in the high dose regions and for the H&N case.

Table 4 summarizes the results of the gamma analyses for the model validation against film and the TPS. Gamma pass rates for (2%, 2 mm) and (3%, 3 mm) criteria show good agreement for all fields. Passing rates for H&N fields are lower than for prostate and lung fields.

Figure 5 shows the (2%, 2 mm) gamma evaluation maps for a prostate field, which seems to indicate the overestimation of the high dose regions by the prediction model when compared with the TPS, which confirms the results seen in the dose profiles.

3.3. Uncertainty analysis
3.3.1. Uncertainty of fitted parameters
Figure 6 shows the results for the parameters \( a \) and \( b \) for the repeated random fitting process, whereas table 3 provides the confidence limits for the fitted parameters for one accelerator type and two beam energies. The latter shows that the uncertainty is quite large for most parameters. It should be noted that the value of some of these parameters is around zero. Hence, a large change results in a value that is still very small, and will likely have little influence on the model output. What is also evident from table 3, is that some parameters have a clear bias towards a positive or negative difference (e.g. \( K_{\text{PMLC}} \)), despite the symmetric simulated variation of \( D_j \) around \( D_0 \). Except for a few cases, the parameters show similar trends for different accelerator types (STx versus non-
Table 3. Model fitting results and 95% confidence intervals for 6 MV and 10 MV beam energies, STx MLC, FF fluence mode, and 100 cm SDD.

<table>
<thead>
<tr>
<th>Fitted parameters</th>
<th>6 MV</th>
<th>10 MV</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>0.2883 (0.2516–0.3788)</td>
<td>0.2466 (0.2066–0.3531)</td>
</tr>
<tr>
<td>$b$ (cm$^{-1}$)</td>
<td>0.1607 (0.1238–0.2402)</td>
<td>0.1827 (0.1347–0.3097)</td>
</tr>
<tr>
<td>$K_{P,JAW,y}$</td>
<td>0.3543 (0.1695–0.4653)</td>
<td>0.4307 (0.2853–0.5538)</td>
</tr>
<tr>
<td>$T_{JAW,y}$</td>
<td>$3.96 \cdot 10^{-8}$ ($2.13 \cdot 10^{-10}$–$4.67 \cdot 10^{-8}$)</td>
<td>$9.27 \cdot 10^{-9}$ ($2.85 \cdot 10^{-11}$–$4.64 \cdot 10^{-9}$)</td>
</tr>
<tr>
<td>$K_{P,MLC,y}$</td>
<td>0.2178 (0.1516–0.2220)</td>
<td>0.2268 (0.2096–0.3966)</td>
</tr>
<tr>
<td>$T_{MLC}$</td>
<td>$9.27 \cdot 10^{-9}$ ($2.85 \cdot 10^{-11}$–$4.64 \cdot 10^{-9}$)</td>
<td>$2.13 \cdot 10^{-7}$ ($1.43 \cdot 10^{-9}$–$9.86 \cdot 10^{-7}$)</td>
</tr>
<tr>
<td>$K_{P,MLC,x}$</td>
<td>0.3179 (0.1695–0.3788)</td>
<td>0.4307 (0.2853–0.5538)</td>
</tr>
<tr>
<td>$T_{JAW,x}$</td>
<td>$3.96 \cdot 10^{-8}$ ($2.13 \cdot 10^{-10}$–$4.67 \cdot 10^{-8}$)</td>
<td>$9.27 \cdot 10^{-9}$ ($2.85 \cdot 10^{-11}$–$4.64 \cdot 10^{-9}$)</td>
</tr>
<tr>
<td>$K_{P,MLC}$</td>
<td>0.2671 (0.1544–0.3914)</td>
<td>0.4988 (0.1945–0.5189)</td>
</tr>
<tr>
<td>$T_{TG}$</td>
<td>0.0098 ($6.22 \cdot 10^{-6}$–$0.0157$)</td>
<td>0.0124 ($4.29 \cdot 10^{-4}$–$0.0187$)</td>
</tr>
<tr>
<td>TG width (cm)</td>
<td>0.0689 (0.0656–0.1085)</td>
<td>0.0735 (0.0659–0.1019)</td>
</tr>
<tr>
<td>$T_{TG}$</td>
<td>0.0892 (0.0140–0.1603)</td>
<td>0.1499 (0.0225–0.2620)</td>
</tr>
<tr>
<td>DLG width (cm)</td>
<td>0.0689 (0.0656–0.1085)</td>
<td>0.0735 (0.0659–0.1019)</td>
</tr>
</tbody>
</table>

Figure 4. Cross-line and in-line profiles for one field of each of the clinical cases and for each method. Profiles were taken through the center of each field.

Table 4. Pass rates for the gamma analysis of the prediction versus the TPS and film measurement. (AP = anterior–posterior, PA = posterior–anterior, CW = clockwise, CCW = counter-clockwise).

<table>
<thead>
<tr>
<th>Case</th>
<th>Field</th>
<th>TPS (1%, 1 mm)</th>
<th>Film (1%)</th>
<th>TPS (2%, 2 mm)</th>
<th>Film (2%)</th>
<th>TPS (3%, 3 mm)</th>
<th>Film (3%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung</td>
<td>AP</td>
<td>53.3</td>
<td>45.7</td>
<td>90.9</td>
<td>81.1</td>
<td>98.7</td>
<td>92.2</td>
</tr>
<tr>
<td></td>
<td>PA</td>
<td>54.1</td>
<td>45.5</td>
<td>92.0</td>
<td>90.0</td>
<td>99.1</td>
<td>98.1</td>
</tr>
<tr>
<td>Prostate</td>
<td>CW</td>
<td>64.8</td>
<td>64.6</td>
<td>96.7</td>
<td>99.3</td>
<td>99.8</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>CCW</td>
<td>74.2</td>
<td>57.4</td>
<td>97.4</td>
<td>93.4</td>
<td>99.8</td>
<td>99.4</td>
</tr>
<tr>
<td>H&amp;N</td>
<td>CW</td>
<td>34.1</td>
<td>31.9</td>
<td>80.8</td>
<td>87.3</td>
<td>92.2</td>
<td>97.1</td>
</tr>
<tr>
<td></td>
<td>CCW</td>
<td>35.3</td>
<td>49.4</td>
<td>87.2</td>
<td>91.7</td>
<td>94.9</td>
<td>97.0</td>
</tr>
</tbody>
</table>
STx) and energies. The results for the other energies and accelerator types, as well as an overview of the relative differences for each parameter at all confidence limits are provided in the supplementary material (stacks.iop.org/PMB/63/035033/mmedia).

3.3.2. Sensitivity analysis
With the known uncertainties of each parameter for simulated $D_p$, the question is how these propagate through the model and which parameters are most sensitive to changes in the dataset used for fitting. Figure 7 shows the results of the sensitivity analysis for one field of each of the three clinical plans, for the 95% confidence limits of some of the fitted parameters. The results for all parameters and confidence limits are provided in the supplementary material (stacks.iop.org/PMB/63/035033/mmedia).

Each parameter was changed to its lower and upper confidence limits, leading to multiple PDI per changed parameter. These PDI were compared with PDI0 by calculating for each pixel the dose deviation as percentage of the maximum dose in PDI0. From these values, multiple boxplots per PDI were created: one for each region in the PDI, i.e. in-field, penumbra and out-of-field for static fields and in-field and out-of-field for dynamic fields. The boxplots thus represent the dose deviations for all pixels in a certain PDI region.

What can be derived from these results is that for the in-field regions, $a$ and $b$ are the most sensitive parameters, i.e. their uncertainties cause the largest dose deviations. The penumbra kernels are more sensitive in the static field than in the dynamic ones. In the static field their effect is, as can be expected, most pronounced in the penumbra region. The same holds for other beam-delimiting parameters such as the DLG width. In the out-of-field region, relative to the field’s maximum dose, none of the parameters is very sensitive to changes. Figure 7 also shows that the very large uncertainty in the jaw transmission parameters has very little influence on the resulting PDI. Together with the OAR map, whose uncertainty was also small, these seem to be the least sensitive parameters.

The results of the gamma analyses are very similar. As can be seen in figure 8, the gamma pass rates show the same patterns as the dose differences. Lower pass rates are seen for parameters $a$ and $b$ in the in-field region, and for the penumbra kernels in the penumbra region of the static field. For the transmission parameters and OAR, all pixels pass the gamma analysis for all cases.
3.3.3. Uncertainty propagation

To determine the maximum uncertainty for each of the clinical plans evaluated, PDIs were predicted using the parameter set fitted with the maximum deviating measurement set $D_{\text{dev}}$. Similar to figure 7, boxplots were created to visualize the dose differences (figure 9). Most notable in this figure are the large deviations in the in-field region of the H&N case, and the large spread of deviations in the penumbra region of the lung case. The prostate case shows the least variation. The dose deviations in the out-of-field region are small for all three cases.

For the prostate case, the comparison between PDI$_0$ and PDI$_{D_{\text{dev}}}$ by dose difference and by a (2%, 2 mm) gamma analysis are displayed in figure 10. As expected from figure 9, the in-field region shows the largest dose differences. When comparing with gamma analysis, many of the differences are not visible anymore. However, the area with the largest dose deviations does not pass the gamma analysis.

Figure 7. Boxplots for the dose deviations caused by changing each parameter separately and predicting a new PDI. The two colors represent the lower and upper 95% confidence limit to which each parameter was changed. For every color/confidence limit there are multiple boxplots, representing the different regions of the PDIs, i.e. in-field (left), penumbra (middle) and out-of-field (right) for static fields, and in-field (left) and out-of-field (right) for integrated dynamic fields.

Figure 8. Gamma pass rates obtained by comparing PDI$_p$ with PDI$_0$ by a (2%, 2 mm) gamma analysis. The multiple bars per color/confidence limit represent the different regions of the PDI (from left to right: in-field, penumbra and out-of-field for static fields, in-field and out-of-field for dynamic fields).
3.3.4. Overall uncertainty

The overall uncertainty, incorporating all 100 $p_j$, is shown in Table 5. Comparing these overall uncertainties with the maximum uncertainties, the same trends are observed. The uncertainties in the H&N case and penumbra region of the lung case are largest, whereas those in the out-of-field regions are close to zero.

4. Discussion

A pre-treatment dose prediction model has been presented. The physical description of the model was complemented by detailing its fitting process and its required input measurements. The model underwent an
exhaustive uncertainty analysis to understand its robustness with regard to reasonable deviations of fitting data. This type of sensitivity analysis is rarely performed in RT modelling.

The results of the validation showed good overall agreement between the prediction and both TPS and radiographic film measurements. The largest deviations were found for the H&N case, probably due to the highly modulated nature of this case. This could indicate that the prediction model is less accurate for small fields. The role of penumbra regions and inter-leaf transmission is also expected to be more noticeable in highly modulated fields. Including small fields in the model fitting process may mitigate the overestimation of the prediction model compared with the TPS and film measurements.

Furthermore, the dose profiles seem to indicate a general overestimation when comparing the prediction model to the TPS. However, to confirm a systematic effect, the model needs to be validated for more clinical cases.

The uncertainty analysis of the fitted parameters shows that changing the measurements used for fitting can have a large effect on the values of some of these parameters. It should be noted that the measurement uncertainty used in this work might be too large. As the standard deviation of the normal distribution was considered to be 1%, differences in measurement values up to 3% were possible.

However, the sensitivity analysis shows that some very large changes in parameter values have very little effect on the resulting PDI. The parameters whose effect is most notable in the output of the model are $a$ and $b$. As expected, penumbra kernels have an impact particularly in the penumbra regions. Their sensitivity is larger for static fields, which have sharper penumbras than integrated dynamic fields. Out-of-field, the effect of any parameter deviation is negligible.

As many parameters show little sensitivity, it might be possible to convert some of these into constants. This would reduce the time needed for fitting, and would make the model simpler. An example of this are the transmission parameters, especially those for the jaws. Even though they have a very large uncertainty, changing their values barely has an effect on the model output.

Comparing parameter uncertainties between accelerator types, the parameters that show the most prominent differences are $K_{P,MLC}$, $K_{P,MLC}$, $T_{MLC}$, and the TG and DLG parameters. Not surprisingly, all these parameters are related to the MLC geometry, which is the main difference between the two accelerator types.

Propagating the uncertainties from fitting the model with a measurement set that deviates maximally from $D_0$ gives very different results depending on the clinical case. Especially in the H&N case, the deviation from PDI is in the in-field region is much larger than in the other two cases. It needs to be investigated further if this is caused solely by the parameter uncertainty, or if the larger proportion of small fields in this plan also contributes to the larger dose deviations. The wide spread of dose differences in the penumbra region of the lung case is likely due to the fact that it concerns a static field with a sharp penumbra. This is more sensitive to changes than the more spread-out penumbra of the integrated VMAT fields.

The uncertainties also propagated through in a (2%, 2 mm) gamma analysis. This shows that it is necessary to take the model’s uncertainty into account when developing clinical decision models based on the outcome of a gamma comparison (e.g. between a pre-treatment prediction and an EPID measurement). Based on the results of the uncertainty analysis, it may be necessary to establish the uncertainty per clinical plan. To test this hypothesis, further research including more clinical cases is needed. It may be possible to establish trends per treatment site or type of plan, e.g. for plans with a larger proportion of small fields, the uncertainty of the model could be higher.

5. Conclusion

To conclude, this work presented a pre-treatment prediction model, including a validation and uncertainty analysis. The validation shows the model can be used clinically, but more testing is needed, especially for small fields. Furthermore, its uncertainties can be taken into account when developing clinical decision models.

Acknowledgments

The authors would like to thank Dr Karen Zegers and Dr Mark Podesta for their insight and fruitful discussion. This study was partially funded by Varian Medical Systems (project: TimePort-III).

References


International Atomic Energy Agency (IAEA) 2016 Accuracy Requirements and Uncertainties in Radiotherapy (IAEA Human Health Series vol 31) (Vienna: IAEA)
Van Esch A, Depuydt T and Huyskens D P 2004 The use of an aSi-based EPID for routine absolute dosimetric pre-treatment verification of dynamic IMRT fields Radiother. Oncol. 71 223–34