Thesis for the degree of Doctor of Philosophy

Radiochromic Film Dosimetry: Protocol and Model Selection

Ignasi Méndez Carot

Under the supervision of:
Dr. Domingo Granero Cabañero, Eresa-Hospital General de València
Asst. Prof. Primož Peterlin, Institute of Oncology Ljubljana

Doctorat en Física
Departament de Física Atòmica, Molecular i Nuclear
UNIVERSITAT DE VALÈNCIA
València, March 2018

A mis padres

Agraïments

Aitana començarà prompte l'escola. A Tilen encara li queden un parell d'anys sense llibres ni deures. D'alguna manera, la meua escola termina amb esta tesi. Al món actual, mai no acabes de formar-te. Però, dins del sistema educatiu, el doctorat es l'última fita del camí. En el meu cas, un camí de més de trenta anys: escola, institut, facultat, un primer tast del doctorat i, ara sí, açò que lliges.

No tinc cap dubte de que l'educació que tinc li la dec als meus pares, al seu sacrifici i a la seua espenta. El meu germà em renyirà, de broma, però em renyirà, si no li agraïsc el seu esforç. Per tant, gràcies també als meus germans, i a Petra, per aguantar-me. Així mateix, gràcies a aquells que ja no estan però se sentirien orgullosos de veure este treball.

Ara bé, la persona que més m'ha ajudat en esta última etapa, que em va estendre la mà per encetar-la i m'ha guiat fins a completar-la, és el meu tutor de tesi: el professor Facundo Ballester. Sense ell, tal volta, mai no hauria tornat a encarrilar esta via. Ell i els meus directors de tesi, els doctors Domingo Granero i Primož Peterlin, han dedicat molt de temps i paciència endreçant este manuscrit. En el cas de Primož, a més, el reconeixement és doble, com a coautor junt a Vida, Rihard, Andrej, Božo, Željko i Aljaša, d'alguns dels articles ací presents. Ells, i en general tots els meus companys a l'Onkološki, són coparticipants d'esta tesi. A tots els que m'han ajudat, a tots vosaltres,

gràcies!

Ignasi Méndez Carot
Trzin, Gener 2018

List of Publications

This thesis is based on the following appended papers:

Paper 1. I. Méndez, V. Hartman, R. Hudej, A. Strojnik and B. Casar. *Gafchromic EBT2 film dosimetry in reflection mode with a novel plan-based calibration method.* Medical physics 40.1 (2013), 011720. *2016 Impact Factor: 2.617*

Paper 2. I. Méndez, P. Peterlin, R. Hudej, A. Strojnik and B. Casar. *On multichannel film dosimetry with channel-independent perturbations.* Medical physics 41.10 (2014), 011705. *2016 Impact Factor: 2.617*

Paper 3. I. Méndez. *Model selection for radiochromic film dosimetry.* Physics in medicine and biology 60.10 (2015), 4089. *2016 Impact Factor: 2.742*

Paper 4. I. Méndez, Ž. Šljivić, R. Hudej, A. Jenko and B. Casar. *Grid patterns, spatial inter-scan variations and scanning reading repeatability in radiochromic film dosimetry.* Physica Medica 32.9 (2016), 1072-1081. *2016 Impact Factor: 1.990*

Conference Presentations

Portions of this thesis were presented at the following meetings and courses:

I. Méndez, V. Hartman, R. Hudej, P. Peterlin, A. Strojnik, A. Šarvari and B. Casar. *A novel method for EBT2 radiochromic film dosimetry.* ESTRO 31 Meeting (Barcelona 2012). *Poster presentation*

I. Méndez. *Un nuevo método para la dosimetría con películas radiocrómicas.* III Congreso Conjunto SEFM-SEPR (Caceres 2013). *Oral presentation*

I. Méndez. *New developments in radiochromic film dosimetry.* 6th Alpe-Adria Medical Physics Meeting (Budapest 2014). *Oral presentation*

I. Méndez. *Dosimetría con películas radiocrómicas. Uso de la aplicación web Radiochromic.com.* Curso SEFM de Imagen en Radioterapia (Pamplona 2014). *Invited talk*

I. Méndez. *Dosimetría con película radiocrómica: densidad óptica vs. densidad óptica neta.* IV Congreso Conjunto SEFM-SEPR (Valencia 2015). *Poster presentation*

I. Méndez, Ž. Šljivić, R. Hudej, A. Jenko and B. Casar. *The repeatability of the scanner in radiochromic film dosimetry.* 1st European Congress of Medical Physics (Athens 2016). *Oral presentation*

I. Méndez. *Curso actualización: Dosimetría con película radiocrómica.* V Congreso Conjunto SEFM-SEPR (Girona 2017). *Invited talk*

Contents

List of Abbreviations

3D-CRT	Three-Dimensional Conformal RadioTherapy
AAA	Anisotropic Analytical Algorithm
AIC	Akaike Information Criterion
ANOVA	ANalysis Of VAriance
CCD	Charge-Coupled Device
CHIP	Channel-Independent Perturbations
EDW	Enhanced Dynamic Wedge
IMRT	Intensity Modulated RadioTherapy
MAD	Mean Absolute Deviation
MM	Micke-Mayer
MOSFET	Metal-Oxide-Semiconductor Field-Effect Transistor
MU	Monitor Units
NOD	Net Optical Density
OD	Optical Density
OSL	Optically Stimulated Luminescence
PDF	Probability Density Function
PV	Pixel Value
RMSE	Root Mean Square Error
ROI	Region Of Interest
TLD	Thermoluminescent Dosimeter
TPS	Treatment Planning System
SAD	Source-Axis Distance
SRS	Stereotactic RadioSurgery
SSD	Source-to-Surface Distance
TIFF	Tagged Image File Format
TN	Truncated Normal
UD	Uniform Distribution
VMAT	Volumetric Modulated Arc Therapy
WM	Weighted Mean

Part I

Synopsis

Chapter 1

Synopsis

1.1 Introduction

Ionizing radiation is capable of removing electrons from matter it passes through. It takes two forms: directly and indirectly ionizing radiation. Directly ionizing radiation consists of charged particles (*e.g.*, electrons, protons, ions, etc), which interact with the medium through the Coulomb potential. Indirectly ionizing radiation involves neutral particles (*e.g.*, photons and neutrons), which release charged particles into the medium. These particles, in turn, ionize atoms through Coulomb interactions. Ionizing radiation has medical applications: diagnostic radiology and nuclear medicine use ionizing radiation in the diagnosis of disease, and radiation therapy in its treatment.

Absorbed dose is defined as the mean energy imparted by ionizing radiation to matter per unit of mass. An instrument that yields a reading in response to the absorbed dose is referred to as a radiation dosimeter. A dosimeter together with its reader is called a dosimetry system. Common types of dosimeters include water calorimeters, ionization chambers, thermoluminescent dosimeters (TLD), optically stimulated luminescence systems (OSL), silicon diodes, metal-oxide-semiconductor field-effect transistors (MOSFETs), alanine, diamond dosimeters, gel dosimeters, radiochromic films, etc.

Radiochromic films are based on an active layer composed of diacetylene monomers dispersed in a binder. In particular, all generations of Gafchromic films (Ashland Inc., Wayne, NJ) contain lithium salt of pentacosa-10,12-diynoic acid (LiPCDA) as the active monomer. Upon irradiation, the LiPCDA polymerizes, becoming increasingly dark with the absorbed dose [1–9]. Changes in the visible absorption spectrum result in different responses when films are digitized with flatbed scanners. As a consequence, the system comprising radiochromic films and a flatbed scanner can be used to measure two-dimensional dose distributions [10].

Each dosimetry system has advantages and disadvantages, making it suitable for different measuring conditions. High spatial resolution, near water-equivalence [1, 11] and weak energy dependence [12–17] mean that radiochromic film dosimetry is convenient for many applications in radiation therapy [18]. It is capable of delivering accurate dose measurements [19–22] despite being subject to several sources of uncertainty, such as the evolution of film darkening with post-irradiation time [23, 24], variations in the active layer thickness [4], the influence of humidity and temperature [2], noncatalytic and ultraviolet-catalyzed polymerization [25], the lateral artifact [26–28], dependency on the orientation of the film on the scanner bed [29], the cross talk effect [27], dependency on film-to-light source distance [30, 31], Newton rings [32], warming-up of the lamp [33, 34], inter-scan variations [30, 35], noise [36, 37], dust, scratches, etc.

The purpose of this work was to improve the accuracy of radiochromic film dosimetry. In order to do so, some of the perturbations that affect film dosimetry were analyzed, different protocols were compared, corrections were proposed and dosimetry models were selected.

1.2 Paper 1. Gafchromic EBT2 film dosimetry in reflection mode with a novel plan-based calibration method

1.2.1 Effect of surrounding film

Paper 1 addressed several open questions concerning a dosimetry system formed by Gafchromic EBT2 films and an Epson Expression 10000XL flatbed scanner (Seiko Epson Corporation, Nagano, Japan) operated in reflection mode. Richley *et al.* [38] had previously reported that the pixel value (PV) of a film fragment scanned in reflection mode depends on the transparency of the media surrounding it. However, they did not find any dependency in transmission mode. Paper 1 investigated the effect of surrounding film by measuring the variations in the PV of a film fragment when other fragments irradiated with different doses were positioned next to it. Contrary to Richley *et al.* [38], no effect of surrounding film was noticed in this work. In light of recent publications, the results of both Paper 1 and Richley *et al.* [38] regarding the surrounding film effect should be called into question. Their conclusions could be impacted by the repeatability of the scanner [30, 35], which requires an unexposed film fragment to correct it. In transmission mode, van Battum *et al.* [27] found that the surrounding area effect, which they termed *the cross talk effect*, is only relevant for small regions (*i.e.*, radius < 5 mm) in combination with high dose gradients.

1.2.2 EBT2 film response homogeneity

Initial EBT2 lots presented excessive response heterogeneities [4, 39], which raised doubts about the suitability of EBT2 films for accurate absolute dosimetry. This work examined film response homogeneity, concluding that different film areas presented significantly different responses both prior to and following irradiation. Film heterogeneity was still present even when net optical density (NOD) was employed as response, which means that, in order to correct film heterogeneity, more advanced methods than the use of NOD are necessary.

1.2.3 Sensitometric curves and lateral corrections

Different models for sensitometric curves and lateral corrections were compared in Paper 1. The best models according to the maximum likelihood estimation were selected. A polynomial of order four was chosen for the sensitometric curve. With respect to the lateral correction, two different functions obtained the lowest root mean square errors (RMSEs):

$$v = a_1(x - x_c) + a_2(x - x_c)^2 + \hat{v} \tag{1.1}$$

$$v = a_1(x - x_c) + a_2(x - x_c)^2 + \hat{v}(1 + a_3(x - x_c) + a_4(x - x_c)^2) \tag{1.2}$$

Here $\hat{v}$ is the PV before correction, x is the coordinate of the pixel in the axis parallel to the lamp, x_c is the x coordinate of the center of the scanner, v denotes the corrected PV, and a_i are fitting parameters.

Eq.(1.1) was selected according to the Akaike information criterion (AIC) [40]. Interestingly, the lateral correction of eq.(1.2) was later recommended by Lewis *et al.* [41]. For the range of doses being studied in Paper 1 (lower than 4 Gy) no improvement was found using eq.(1.2) with respect to the absolute correction of eq.(1.1). Furthermore, Saur *et al.* [42] also proposed absolute corrections for the lateral artifact. However, the lateral artifact is more pronounced for higher doses [26], which may require the model examined by Lewis *et al.* [41].

1.2.4 Plane-based calibration

To calibrate a film lot, reference doses should be associated with the responses of the dosimetry system. The most frequent calibration method in the literature [10, 36, 43, 44] employs fragments irradiated with known doses. The response measured in a region of interest (ROI) of a fragment is associated with the dose of the fragment. Paper 1 introduced a novel plan-based calibration method. In the plan-based method, a film is irradiated with a known 2D dose distribution. The response of a pixel is associated with the dose on this pixel. Since the reference 2D dose distribution can be either planned with a treatment planning system

(TPS) or measured with another dosimeter, the plan-based method was renamed as *plane-based method* in Paper 3.

The plane-based method is an alternative to the calibration with fragments. A disadvantage of the plane-based method is that the uncertainty of the reference doses is higher than with fragments. However, the calibration sample is more representative, the time required for calibration is shorter and it allows the simultaneous fitting of sensitometric curves and lateral corrections.

In this work, the reference dose distribution was a 60° Enhanced Dynamic Wedge (EDW) field of dimensions 20×20 cm^2 covering the range of doses of interest (from 75 cGy to 400 cGy). It was calculated with a TPS Eclipse v10.0 (Varian, Palo Alto, CA, USA) using the anisotropic analytical algorithm (AAA), and exported as a DICOM RT dose array with a resolution of 0.59 mm/px, which was bilinearly interpolated to register to the film. In this way, a set of 1100×1300 data points was obtained. Each point consisted of the coordinates of the pixel, the reference dose calculated with the TPS, and the PVs for all three color channels prior to and following irradiation. Sensitometric curves and lateral corrections for the channels were calculated, minimizing the RMSEs of the differences between the reference doses and the film doses. Minimization was conducted by a genetic algorithm. Film doses were calculated as the weighted mean dose of all three color channels. The mean square errors of each channel calibration were used as weights.

Film doses were compared with TPS doses employing the global gamma analysis with tolerances of 4% 3 mm and excluding points with less than 20% of the maximum dose. A set of seven different cases based on the IAEA TECDOC-1583 [45] tests for TPS commissioning was used to select between plane-based calibration and calibration with fragments, and between three-channel dosimetry and red-channel dosimetry. We concluded that three-channel dosimetry was superior to red-channel dosimetry, and that plane-based calibration was a feasible alternative to calibration with fragments.

1.3 Paper 2. On multichannel film dosimetry with channel-independent perturbations

1.3.1 Channel-independent perturbations

The multichannel model proposed in Paper 1 was the weighted mean of the color channels. This model assumes that there is no correlation between the channel doses. Micke *et al.* [46] considered that variations in the thickness of the active layer, artifacts, heterogeneities in the scanner response and other disturbances produced correlations between the channels, which could be modeled by channel-independent perturbations (CHIP). Mayer *et al.* [47] derived a closed-form solution to obtain the dose assuming channel-independent perturbations.

It was implicit in both cases that the probability density function (PDF) of the perturbation was a uniform distribution, since all possible perturbations were equiprobable. Paper 2 formulated a framework to develop and examine CHIP models with different PDFs of the perturbation.

If the perturbations are small, we can apply a first-order Taylor expansion of the dose in terms of the perturbation. Therefore, a general channel-independent perturbation model can be described by:

$$D(r) = D_k(r) + \dot{D}_k(r)\Delta(r) + \epsilon_k(r) \tag{1.3}$$

where k represents the color channel (*i.e.*, red (R), green (G) or blue (B)), $\dot{D}_k$ is the first derivative of D_k with respect to the response, and $\epsilon_k(r)$ is an error term, which accounts for the difference between the true absorbed dose $D(r)$ and the measured dose after correction of the perturbation $\Delta(r)$ at point r.

In this study, three different PDFs of $\Delta(r)$ were examined: normally distributed perturbation, truncated normal distribution and uniform perturbation. Aside from describing the perturbation, it is also necessary to know the PDF of $\epsilon_k(r)$. We assumed normally distributed error terms with σ_k^2 variance.

The estimated absorbed dose d (*i.e.,* the most likely value of D) and its type B uncertainty σ_D were derived:

$$d = \frac{A\beta - \gamma\delta}{A\alpha - \gamma^2} \tag{1.4}$$

and

$$\sigma_D = \sqrt{\frac{A}{A\alpha - \gamma^2}} \tag{1.5}$$

where

$$A = \frac{1}{\sigma_\Delta^2} + \sum_k \left(\frac{\dot{D}_k}{\sigma_k}\right)^2 \tag{1.6}$$

$$\alpha = \sum_k \frac{1}{\sigma_k^2} \tag{1.7}$$

$$\beta = \sum_k \frac{D_k}{\sigma_k^2} \tag{1.8}$$

$$\gamma = \sum_k \frac{\dot{D}_k}{\sigma_k^2} \tag{1.9}$$

$$\delta = \sum_k \frac{D_k \dot{D}_k}{\sigma_k^2} \tag{1.10}$$

and σ_Δ^2 is the variance of the perturbation.

1.3.2 Selection of protocol and CHIP model

Four CHIP models were compared: the Weighted Mean model, the model proposed by Micke *et al.* [46] and Mayer *et al.* [47], another model with Uniform Distribution of perturbations $\Delta(r)$, and a model with Truncated Normal distribution of $\Delta(r)$. The Weighted Mean model assumes that $\Delta(r)$ is zero. The Micke-Mayer model implicitly assumes that the PDF of $\Delta(r)$ is a uniform distribution and the variances of the error terms (σ_k) are equal. In the Uniform Distribution and Truncated Normal models, σ_k were considered different.

We also compared between EBT2 and EBT3 films, between transmission and reflection scanning modes, between 75±5 min and 20±1 h post-irradiation waiting-time windows, and between planned and measured reference dose distribution for the calibration.

The plane-based calibration method was followed. In order to extend the dose range and to reduce intralot variations, three films irradiated with 60° Enhanced Dynamic Wedge fields were employed. A genetic algorithm minimized the RMSEs of the differences between film doses for each channel ($D_k(r)$) and reference doses ($D(r)$). This optimization fitted the sensitometric curves and lateral correction parameters. The Uniform Distribution and Truncated Normal models also require the optimization of the parameters describing the PDF of the perturbation, which were fitted with the RMSE optimization of the differences between multichannel film doses (d) and reference doses.

Protocols and CHIP models were selected comparing film dose distributions with planned and measured dose distributions by means of global gamma analysis. A sample of 14 test cases considered representative of clinical dose distributions was tested.

No significant differences were found between transmission and reflection mode scanning, or between using EBT2 or EBT3 films. Short or long post-irradiation waiting-time windows were found equally accurate, which agrees with the conclusions made by Lewis *et al.* [19]. The Truncated Normal distribution multichannel dosimetry model provided the best agreement between film and reference dose distributions, as may be expected, considering that the other three models are either particular or limit cases of the Truncated Normal model. This model can be regarded as a metamodel, which minimizes the inherent uncertainty in the dose of the CHIP models being studied. Neglecting the correlations between color channels or, equivalently, considering that $\Delta = 0$, as does the Weighted Mean model, yielded worse results. Also, considering all perturbations as equally probable can produce unacceptable values of film dose. This is because the film dose uncertainty depends on the shape of the sensitometric curves. For instance, the film dose uncertainty of the model proposed by Micke *et al.* [46] and Mayer *et al.* [47] can be calculated as

$$\sigma_D = \sigma_k \sqrt{\frac{\sum_{k=1}^{n} \dot{D}_k^2}{n \sum_{k=1}^{n} \dot{D}_k^2 - (\sum_{k=1}^{n} \dot{D}_k)^2}} \tag{1.11}$$

The relationship between the dose absorbed by the film and the response of the dosimetry system is modeled in the calibration. The model with the lowest RMSE of the differences between film and reference doses is considered the most accurate. Consequently, lower RMSE in the calibration are expected to be correlated with lower values of $\overline{\gamma}$ and higher $\gamma_{<1}$ in the gamma analysis of the test cases. These correlations were found to be significant.

1.4 Paper 3. Model selection for radiochromic film dosimetry

1.4.1 A General Perturbation Model

Channel Independent Perturbation Models are particular cases of the General Perturbation Model presented in Paper 3. According to this model, the dose absorbed by the film at point r, $D(r)$, can be expressed as:

$$D(r) = D_k(z_k(r) + \Psi_k(r, z_k)) + \Sigma_k(r) \tag{1.12}$$

where k represents the color channel, D_k is the dose directly obtained from the sensitometric curve (for channel k), $z_k(r)$ is the film-scanner response, $\Psi_k(r, z_k)$ corrects systematic local perturbations and $\Sigma_k(r)$ represents noise disturbances.

Noise disturbances consist of random perturbations and film-scanner damage (*e.g.*, dust, scratches). Random perturbations change or vanish between scan repetitions or between non-irradiated and irradiated film scans. Film-scanner damage causes large disturbances of the dosimetry response. Systematic local perturbations persist between scan repetitions (*e.g.*, active layer variations, the lateral artifact).

If the systematic local variations are small compared to the response, we can apply a first-order Taylor expansion of $D(r)$ in terms of $\Psi_k(r, z_k)$ in eq.(1.12):

$$D(r) = D_k(z_k(r)) + \dot{D}_k(z_k(r))\Psi_k(r, z_k) + \epsilon_k(r) \tag{1.13}$$

where $\dot{D}_k$ is the first derivative of D_k with respect to z_k, and $\epsilon_k(r)$ is an error term accounting for the difference between the true absorbed dose, $D(r)$, and the measured dose after correction of the perturbation.

CHIP models are particular cases of this General Perturbation Model, which assume that the perturbation $\Psi_k(r, z_k)$ is equal in all three color channels.

1.4.2 Model selection

Paper 3 expounds the mathematical background of the method followed in the course of this thesis in order to optimize the protocol and model for radiochromic film dosimetry. The model selection was intended to maximize the probability of film dose and true dose being equal.

According to Bayesian probability theory, the probability of an event is proportional to the evidence given the event times the prior probability of the event: $P(M|D) \propto P(D|M)\,P(M)$. In our case, the event was the dosimetry model and the evidence consisted of reference doses and measured film doses used for lot calibrations. The Pareto optimal choice for the prior is the universal weight based on the Kolmogorov complexity of the model [48, 49]. Nevertheless, since the Kolmogorov complexity is not finitely computable, the AIC [40] was chosen as a practical complexity-based model selection approach. According to the AIC, the most probable model is the one with the lowest AIC value: $AIC = 2\,c_M - 2\ln(P(D|M))$, where c_M is the number of parameters of the model. In order to calculate $P(D|M)$, we assumed that that the differences (ϵ_d) between the reference doses ($D(r)$) and the film doses ($d(r)$) were normally distributed. The number of parameters of the models was negligible compared to the size of the evidence data sample. Thus, the most likely dosimetry model was the one with the lowest RMSE between reference and film doses.

The data sample in Paper 3 included different lots of films with different dose ranges and amount of data. To give different lots the same weight, relative rather than absolute dose differences were employed. It was assumed that the relative differences between reference doses and film doses ($d(r)$) were also normally distributed:

$$\frac{\epsilon_d(r)}{d(r)} \sim \mathcal{N}(0, \sigma^2) \tag{1.14}$$

where σ was termed as the relative *uncertainty of the calibration*.

Information entropy is defined as the expected negative log-likelihood of a random variable. The (differential) information entropy of a model following Eq.(1.14) can be expressed as:

$$h(M) = \frac{1}{2}\ln(2\pi e\sigma^2) \tag{1.15}$$

The entropy of a model in a set of evenly weighted lots is the arithmetic mean entropy of the model. Maximizing the likelihood of a model is equivalent to minimizing the information entropy. Therefore, the dosimetry model with the lowest geometric mean of the calibration uncertainty for all lots being studied was selected.

1.4.3 Validation

Dosimetry models are approximations of the relationship between the dose absorbed by the film and the response of the dosimetry system. Besides, the calibration sample could be not sufficiently representative. For an additional validation of the model selection, film dose distributions were compared with planned dose distributions by means of global gamma-index analyses [50]. The tolerances were 3%, 3 mm and 2%, 2 mm, points with less than 30% of the maximum planned dose were excluded. Film doses were selected as the reference distributions and planned doses were the evaluation distributions in order to avoid noise artifacts [51] in the gamma results.

Lower calibration uncertainties were found to be significantly correlated with higher mean $\gamma_{<1}$ values in the gamma analysis. Furthermore, both approaches to the selection of the protocol and dosimetry model yielded the same conclusions.

For an additional validation, the statistical significance of the gamma analysis results was also studied.

1.4.4 Selection of protocol and dosimetry model

Paper 3 continued the optimization of the dosimetry protocol. In this work, the lateral correction model selected in Paper 1 was compared with the one proposed by Poppinga *et al* [52], both models were found to be equally accurate. After exploring whether lateral corrections are necessary when using multichannel models, it was found that lateral corrections improved the accuracy of the results, even for multichannel models. Models using the Truncated Normal perturbation were contrasted with models using a uniform distribution of perturbations with equal variance in the error terms (*i.e.*, the Micke-Mayer perturbation), which supported the recommendation of using Truncated Normal perturbations. It was also found that using NOD as the response of the dosimetry system gave better results than using PVs. However, this improvement did not produce significantly higher $\gamma_{<1}$(3% 3mm) values with triple-channel models. Finally, it was found that increasing the number of color channels did not necessarily yield more accurate film doses for the multichannel models being studied.

In total, 42 test cases were converted to doses by applying 44 different dosimetry models to each of them. The most accurate doses were found with the model applying lateral corrections, using NOD as the response, while combining all three color channels according to the CHIP with truncated normal perturbations.

It should be noted that the applicability of the conclusions is constrained by the limits of the test cases (*e.g.*, the dose range of 20-600 cGy) and the specific design of the models (*e.g.*, polynomial sensitometric curves).

1.5 Paper 4. Grid patterns, spatial inter-scan variations and scanning reading repeatability in radiochromic film dosimetry

1.5.1 Scanning protocol

Several elements of the dosimetry protocol had already been studied and selected in Papers 1-3. However, the scanning protocol remained practically the same from Paper 1 onwards. That is, films were scanned with an Epson Expression 10000XL flatbed scanner using the associated Epson Scan v.3.x software. They were scanned in portrait orientation (*i.e.*, the short side of the film was parallel to the scanner lamp). Before acquisitions, the scanner was allowed to warm up for at least 30 min. After that, and after long pauses, five empty scans were taken to stabilize the temperature of the scanner lamp. Five consecutive scans were made for each film, with the first scan discarded, while the resulting image was the average of the other four. The films were centered on the scanner with an opaque frame. Image processing tools were turned off. Images were acquired in 48-bit RGB mode (16 bit per channel) and saved as TIFF files. The only deviations in Paper 4 from the scanning protocol of Paper 1, apart from using an updated version of the Epson Scan software, were that in Paper 4 a 3 mm thick glass sheet was placed on top of the films to keep constant the distance between film and lamp, and transmission mode was used instead of reflection mode, bearing in mind that in Paper 2 no significant differences in accuracy were found between both modes.

The objective of Paper 4 was to examine the uncertainties related to the repeatability of the scanner in order to improve the scanning protocol.

1.5.2 Grid patterns

Calculating the average of several scans reduces the image noise and gives more accurate film dose distributions [53]. However, inter-scan variations give rise to different disturbances. One of them are grid patterns. Grid patterns are misleading grid artifacts which are rarely detected, but can occasionally appear in film dose distributions or gamma index analyses. In this work, grid patterns were found using 50, 72 and 96 dpi resolutions, but not with 150 dpi. They even emerge in the absence of transmitted light. The origin of grid patterns is the periodical variation in the scanner noise along both axes. Different resolutions present different periodical variations. The mean dose uncertainty of the scanner noise was found to be at least two times greater than the difference between the maximum and the minimum dose uncertainty that produces grid patterns. That is the reason why this artifact is usually not perceived. Yet, it is easily noticeable

when scanning with a resolution of 72 dpi. With this resolution, the grid pattern has a sinusoidal shape with a period of 8.5 mm.

1.5.3 Spatial inter-scan variability

Lewis *et al* [30] identified that the scanner response varies between scans, generating changes in the relationship between the dose absorbed by the film and the response of the dosimetry system. Unless corrected, these changes can produce relevant film dosimetry errors. Lewis *et al* [30] proposed correcting the response by means of an unexposed fragment scanned together with the film under analysis. They suggested calculating, for each scan, the mean PV in a reference region of interest of the unexposed fragment. This mean PV in the reference ROI (Ref ROI) was compared to a reference, namely the calibration, mean Ref ROI PV. The correction consisted of scaling the scanner response in PVs by the factor necessary to obtain in the Ref ROI the calibration PV. In Paper 4, this correction was referred as the *mean correction*, and expressed as:

$$M(i,j) = v(i,j)\left\langle \frac{M(i_{Ref},j_{Ref})}{v(i_{Ref},j_{Ref})} \right\rangle \tag{1.16}$$

where (i,j) represents the position of a pixel in the image (i is the row and j the column), M symbolizes the reference value of the pixel, v is the measured PV of the pixel, and (i_{Ref},j_{Ref}) is a pixel in the Ref ROI.

The mean correction is identical for every pixel of the scan, it is spatially invariant. We compared it with a spatially variant correction: the *column correction*. The column correction calculates a different correction for each position in the axis parallel to the scanner lamp. It considers that inter-scan variations depend on the individual charge-coupled device (CCD) detectors, which is why linear patterns perpendicular to the scanner lamp can be observed in film dose distributions. Column corrections compare the mean PV and the reference mean PV for each column in the Ref ROI:

$$M(i,j) = v(i,j)\left\langle \frac{M(i_{Ref},j)}{v(i_{Ref},j)} \right\rangle \tag{1.17}$$

Applying either the mean or the column response correction reduced the dose differences between repeated scans, to a larger extent with the column correction. If no response correction of the inter-scan variation was applied, systematic dose deviations larger than 1%, with respect to the reference dose distribution, ocurred in many scans. However, among the corrected images, no systematic deviation larger than 1% was found. Spatial inter-scan variations in the axis perpendicular to the scanner lamp were found to be negligible, except for the initial warm up scans.

1.5.4 Scanning reading repeatability

The initial positioning of the lamp in the axis parallel to its movement, and the speed of the lamp vary between scans. The differences in the initial positioning were found to be lower than 0.1 mm. However, these differences grew with the distance from the initial position due to the variations in speed. In our experiments, we found differences of 0.7 mm at a distance of 20 cm from the initial position. Consequently, it can be stated that average scans are less accurate the further they are from the beginning of the lamp movement. The dosimetric impact of these variations is usually negligible. However, it can be relevant and actions should be taken in some studies, such as measurements of penumbras or plane-based calibrations with wedge fields.

1.6 Discussion, conclusions and future work

This thesis aimed to improve the accuracy of radiochromic film dosimetry. The main focus of research was the optimization of protocols and dosimetry models. The papers compiled here made several contributions to the body of knowledge in the field of radiochromic film dosimetry:

A General Perturbation Model for multichannel radiochromic film dosimetry with flatbed scanners was introduced. Channel Independent Perturbation Models, which are particular cases of the General Perturbation Model, were depeloped and analyzed. The estimated absorbed dose and its inherent uncertainty were obtained for CHIP models with normal, truncated normal and uniform probability density functions of the perturbation. The CHIP model proposed by Micke *et al.* [46] and Mayer *et al.* [47] falls within this scope. The implicit assumptions of the Micke-Mayer model were explained. It was noticed that all possible perturbations were equiprobable in this model, which makes it sensitive to the properties of the sensitometric curves and can lead to unacceptable uncertainties. The Micke-Mayer model was compared to the Truncated Normal perturbation model. The latter provided more accurate results. Better results were also found when applying lateral corrections, and with NOD instead of PV as the response of the dosimetry system. Lateral corrections yielded better accuracy even for multichannel models. However, using NOD was only found to be significantly better for single channel dosimetry. Using NOD was not enough to correct film heterogeneities. Reflection and transmission scanning modes did not yield significant differences. Employing EBT2 versus EBT3 film type or short versus long post-irradiation times did not yield any significant differences either. Commonly, dosimetry models and protocols in film dosimetry have been selected by means of gamma index comparisons. In this thesis, the selection based on the lowest RMSE of the calibration was proposed, mathematically justified, and validated with gamma comparisons. One should bear in mind that

the conclusions related to the selected models and protocols depend on the limits of the test cases (*e.g.*, the dose range) and the specific design of the models (*e.g.*, polynomial sensitometric curves in this work). For instance, the Absolute lateral correction model selected in this thesis was found to be as accurate as the model recommended by Lewis *et al.* [41]. Yet, this outcome was limited to doses lower than 4 Gy. This thesis also presented an alternative to calibration which fragments: plane-based calibration. This method is faster and simultaneously enables the computing of the lateral correction. Furthermore, the calibration sample is more representative, which makes it more robust against perturbations. However, the reference doses have higher uncertainty, it is computationally more intensive and can be affected by scanning reading repeatability. The scanning reading repeatability was introduced in this work. It was found that the initial positioning (in the axis parallel to its movement) and the speed of the scanner lamp vary between scans. Grid patterns, which were also discovered in this thesis, are artifacts caused by the periodical variation in the scanner noise along both axes. Another source of uncertainty produced by the repeatability of the scanner involves inter-scan variations. In this thesis, a novel column correction method was proposed to take into account the deviations of the individual CCD detectors.

Perturbation Models for radiochromic film dosimetry are based on the assumption that disturbances are small compared to the response. This assumption may be inappropriate (*e.g.*, the lateral artifact can considerably reduce the response away from the center of the scanner lamp) and cause important errors in the calculation of the dose. A greater knowledge of the sources of perturbation is key for the sake of improving film dosimetry. More accurate corrections for lateral artifacts, inter-scan variations, intra-lot variations, differences in post-irradiation time, etc. are directions for future work. Another source of uncertainty relates to film heterogeneities in the active layer. Employing NOD does not eliminate them. One of our current lines of research is focused on film heterogeneities. We have found correlations between irradiated and non-irradiated channels, which can be modeled with a Multivariate Gaussian probability density function of the pixel values, namely, the Multigaussian Model. The Multigaussian Model is being compared against other radiochromic film dosimetry models with encouraging results. Finally, another line of research for the author to pursue is the study of universal or generic sensitometric curves or corrections, which could enable easier, faster and more accurate dose measurements.

Chapter 2

Sinopsi

2.1 Introducció

La radiació ionitzant és capaç d'eliminar electrons de la matèria que travessa. Pot prendre dues formes: directament o indirectament ionitzant. La radiació directament ionitzant consisteix en partícules carregades (*p.ex.*, electrons, protons, ions, etc.), que interactuen amb el medi a través del potencial de Coulomb. La radiació indirectament ionitzant està formada per partícules neutres (*p.ex.*, fotons i neutrons), que alliberen partícules carregades al medi. Aquestes partícules, al seu torn, ionitzen àtoms mitjançant interaccions de Coulomb. La radiació ionitzant té aplicacions mèdiques: la radiologia diagnòstica i la medicina nuclear utilitzen radiacions ionitzants en el diagnòstic de la malaltia, i la radioteràpia en el seu tractament.

La dosi absorbida es defineix com l'energia mitjana que imparteix la radiació ionitzant a la matèria per unitat de massa. L'instrument que produeix una lectura en resposta a la dosi absorbida s'anomena dosímetre. Un dosímetre juntament amb el seu lector formen un sistema de dosimetria. Tipus de dosímetres comuns són els calorímetres d'aigua, les cambres d'ionització, els dosímetres termoluminescents (TLD), els sistemes de luminescència estimulats òpticament (OSL), els díodes de silici, els transistors d'efecte de camp metall-òxid-semiconductor (MOSFET), l'alanina, els dosímetres de diamant, els dosímetres de gel, les pel·lícules radiocròmiques, etc.

Les pel·lícules radiocròmiques es basen en una capa activa composta de monòmers de diacetilè dispersos en un aglomerant. En particular, totes les generacions de pel·lícules Gafchromic (Ashland Inc., Wayne, NJ) contenen sal de liti d'àcid pentacosa-10,12-diinòic (LiPCDA) com a monòmer actiu. Després de la irradiació, la LiPCDA es polimeritza, enfosquint-se de manera creixent amb la dosi absorbida [1–9]. Els canvis en l'espectre d'absorció de llum visible resulten en diferents respostes quan les pel·lícules es digitalitzen amb escàners plans. Com

a conseqüència, el sistema que comprèn pel·lícules radiocròmiques i un escàner pla es pot utilitzar per mesurar distribucions bidimensionals de dosi [10].

Cada sistema de dosimetria té avantatges i inconvenients que el fan adequat per a diferents condicions de mesura. Les pel·lícules radiocròmiques presenten una alta resolució espacial, són aproximadament equivalents a aigua [1, 11] i tenen una dependència feble en l'energia [12–17], per tot això són escaients en moltes aplicacions de radioteràpia [18]. A més, poden proporcionar mesures precises de la dosi [19–22] tot i estar subjectes a diverses fonts d'incertesa, com ara l'evolució de l'enfosquiment de la pel·lícula amb el temps d'irradiació [23, 24], les variacions del gruix de la capa activa [4], la influència de la humitat i la temperatura [2], la polimerització no catalitzada i ultraviolada [25], l'artefacte lateral [26–28], la dependència amb l'orientació de la pel·lícula sobre l'escàner [29], l'efecte de diafonia (*cross talk effect*) [27], la dependència amb la distància a la font de llum [30, 31], els anells de Newton [32], l'escalfament de la làmpada [33, 34], les variacions entre escanejos [30, 35], el soroll [36, 37], la pols, les ratllades, etc.

El propòsit d'aquest treball era millorar l'exactitud de la dosimetria amb pel·lícula radiocròmica. Amb aquest objectiu, es van analitzar algunes de les pertorbacions que afecten la dosimetria amb pel·lícula, es van comparar diferents protocols, es van proposar correccions i es van seleccionar models de dosimetria.

2.2 Article 1. Gafchromic EBT2 film dosimetry in reflection mode with a novel plan-based calibration method

2.2.1 Efecte de la pel·lícula circumdant

L'Article 1 abordava diverses preguntes obertes sobre el sistema de dosimetria format per pel·lícules Gafchromic EBT2 i un escàner pla Epson Expression 10000XL (Seiko Epson Corporation, Nagano, Japó) operat en mode reflexió. En un estudi anterior, Richley *et al.* [38] van trobar que el valor de píxel (PV) en un fragment de pel·lícula escanejat en mode reflexió depèn de la transparència dels mitjans que l'envolten. No obstant això, no van trobar cap dependència en mode transmissió. L'Article 1 investigava l'efecte de la pel·lícula circumdant mitjançant la mesura de les variacions del PV en un fragment de pel·lícula quan es col·loquen al seu costat altres fragments irradiats amb diferents dosis. Contràriament a Richley *et al.* [38], en aquest treball no s'observà cap efecte de la pel·lícula circumdant. A la llum de publicacions recents, cal posar en dubte tant els resultats de l'Article 1 com de Richley *et al.* [38] respecte a l'efecte de la pel·lícula circumdant. En tots dos casos, les conclusions podrien veure's afectades per la repetibilitat de l'escàner [30, 35], que cal corregir amb un fragment de pel·lícula

no exposat. En mode transmissió, Van Battum *et al.* [27] van trobar en un estudi posterior que l'efecte de la pel·lícula circumdant, que ells anomenaren *efecte de diafonia (cross talk effect)*, només és rellevant per a regions petites (ço és, radi <5 mm) en combinació amb elevats gradients de dosi.

2.2.2 Homogeneïtat de resposta de la pel·lícula EBT2

Els lots inicials d'EBT2 presentaven heterogeneïtats de resposta excessives [4, 39], que suscitaven dubtes sobre la idoneïtat d'aquestes pel·lícules per a dosimetria absoluta precisa. Aquest treball va examinar l'homogeneïtat de resposta de la pel·lícula, concloent-hi que diferents regions presentaven respostes significativament diferents, tant abans com després de la irradiació. L'heterogeneïtat de la pel·lícula encara estava present fins i tot quan es va utilitzar com a resposta la densitat òptica neta (NOD), la qual cosa significa que, per corregir l'heterogeneïtat de la pel·lícula, calen mètodes més avançats que l'ús de NOD.

2.2.3 Corbes sensitomètriques i correccions laterals

A l'Article 1 es van comparar diferents models per a corbes sensitomètriques i correccions laterals. Es van seleccionar els millors models d'acord amb l'estimació de màxima versemblança. Per a la corba sensitomètrica, es va triar un polinomi de quart ordre. Pel que fa a la correcció lateral, dues funcions diferents van obtenir els menors valors de l'arrel de l'error quadràtic mig (RMSE):

$$v = a_1(x - x_c) + a_2(x - x_c)^2 + \hat{v} \tag{2.1}$$

$$v = a_1(x - x_c) + a_2(x - x_c)^2 + \hat{v}(1 + a_3(x - x_c) + a_4(x - x_c)^2) \tag{2.2}$$

On $\hat{v}$ és el PV abans de la correcció, x és la coordenada del píxel en l'eix paral·lel a la làmpada, x_c és la coordenada x del centre de l'escàner, v indica el PV corregit, i a_i són paràmetres de l'ajust.

L'eq.(2.1) va ser seleccionada d'acord amb el criteri d'informació d'Akaike (AIC) [40]. Curiosament, la correcció lateral de l'equació (2.2) va ser recomanada per Lewis *et al.* [41] en un article posterior. Per al rang de dosis estudiades a l'Article 1 (inferior a 4 Gy), l'equació (2.2) no va produir cap millora respecte a la correcció absoluta de l'equació (2.1). A més, Saur *et al.* [42] també van proposar correccions absolutes per a l'artefacte lateral. Tanmateix, l'artefacte lateral és més pronunciat per a dosis més altes [26], que podrien requerir el model examinat per Lewis *et al.* [41].

2.2.4 Calibratge basat en un pla

Per calibrar un lot de pel·lícules, s'han d'associar dosis de referència amb respostes del sistema de dosimetria. El mètode de calibratge més freqüent en la literatura [10, 36, 43, 44] fa servir fragments irradiats amb dosis conegudes. La resposta mesurada en una regió d'interès (ROI) d'un fragment s'associa amb la dosi del fragment. L'Article 1 va introduir un nou mètode de calibratge basat en un pla (*plan-based method*). En aquest mètode, una pel·lícula s'irradia amb una distribució de dosi coneguda. La resposta a cada píxel s'associa amb la dosi d'aquest píxel. Atès que la distribució de dosi de referència 2D es pot planificar amb un sistema de planificació de tractaments (TPS) o es pot mesurar amb un altre dosímetre, el mètode *plan-based* es va reanomenar com a mètode *plane-based* en l'Article 3.

El mètode basat en un pla és una alternativa al calibratge amb fragments. Un inconvenient del mètode basat en un pla és que la incertesa de les dosis de referència és més alta que amb fragments. Tanmateix, la mostra és més representativa, el temps necessari per al calibratge és més curt, i permet l'ajust simultani de corbes sensitomètriques i correccions laterals.

En aquest treball, la distribució de dosi de referència va ser obtinguda amb una falca dinàmica (EDW: Enhanced Dynamic Wedge) de 60° amb dimensions 20×20 cm^2 que cobria l'interval de dosis d'interès (de 75 cGy a 400 cGy). Es va calcular amb un TPS Eclipse v10.0 (Varian, Palo Alto, CA, EUA) utilitzant l'algoritme analític anisotròpic (AAA), i es va exportar com una matriu de dosis DICOM RT amb una resolució de 0.59 mm/px, que va ser interpolada bilinealment per enregistrar-la amb la pel·lícula. D'aquesta manera, es va obtenir un conjunt de dades de 1100×1300 punts. Cada punt consistia en les coordenades del píxel, la dosi de referència calculada amb el TPS, i els PV per als tres canals de color abans i després de la irradiació. Es van calcular les corbes sensitomètriques i les correccions laterals de cada canal, minimitzant les RMSE de les diferències entre les dosis de referència i les dosis de la pel·lícula. La minimització es va realitzar mitjançant un algoritme genètic. Les dosis de la pel·lícula es van calcular com la dosi mitjana ponderada dels tres canals de color. Els errors quadràtics mitjos dels calibratges de canal van ser emprats com a pesos.

Les dosis de la pel·lícula es van comparar amb les dosis del TPS utilitzant l'anàlisi gamma global amb toleràncies de 4% 3 mm i excloent-hi els punts amb dosi inferior al 20% de la dosi màxima. Per seleccionar entre calibratge amb fragments i calibratge basat en un pla, i entre dosimetria de tres canals i dosimetria de canal roig, es va utilitzar un conjunt de set casos diferents basats en les proves IAEA TECDOC-1583 [45] per a la posada en funcionament dels TPS. Es va concloure que la dosimetria de tres canals era superior a la dosimetria de canal roig i que el calibratge basat en un pla era una alternativa viable al calibratge amb fragments.

2.3 Article 2. On multichannel film dosimetry with channel-independent perturbations

2.3.1 Pertorbacions independents del canal

El model multicanal proposat a l'Article 1 era la mitjana ponderada dels canals de color. Aquest model suposa que no hi ha correlació entre les dosis dels diferents canals. Micke *et al.* [46] consideraven que les variacions en el gruix de la capa activa, els artefactes, les heterogeneïtats en la resposta de l'escàner i altres pertorbacions produïen correlacions entre els canals que podrien modelar-se amb pertorbacions independents del canal (CHIP). Mayer *et al.* [47] van derivar una solució analítica per obtenir la dosi assumint pertorbacions independents del canal. En ambdós casos, romania implícit que la funció de densitat de probabilitat (PDF) de la pertorbació era una distribució uniforme, ja que totes les possibles pertorbacions eren equiprobables. L'Article 2 va formular un marc on desenvolupar i examinar models CHIP amb diferents PDF de la pertorbació.

Si les pertorbacions són petites, podem aplicar una expansió de Taylor de primer ordre sobre la dosi al voltant de la pertorbació. Per tant, un model general de pertorbació independent del canal pot ser descrit per:

$$D(r) = D_k(r) + \dot{D}_k(r)\Delta(r) + \epsilon_k(r) \tag{2.3}$$

on k representa el canal de color (ço és, roig (R), verd (G) o blau (B)), $\dot{D}_k$ és la primera derivada de D_k respecte de la resposta, i $\epsilon_k(r)$ és un terme d'error, que representa la diferència entre la dosi absorbida real $D(r)$ al punt r i la dosi mesurada una vegada corregida la pertorbació $\Delta(r)$.

En aquest estudi, es van examinar tres PDF diferents de $\Delta(r)$: pertorbació amb distribució normal, amb distribució normal truncada i amb distribució uniforme. A banda de descriure la pertorbació, també cal conèixer la PDF de $\epsilon_k(r)$. Es van assumir distribucions normals dels termes d'error amb variància σ_k^2.

El valor estimat de la dosi absorbida d (és a dir, el valor més probable de D) i σ_D, la seva incertesa tipus B, van ser calculats:

$$d = \frac{A\beta - \gamma\delta}{A\alpha - \gamma^2} \tag{2.4}$$

i

$$\sigma_D = \sqrt{\frac{A}{A\alpha - \gamma^2}} \tag{2.5}$$

on

$$A = \frac{1}{\sigma_\Delta^2} + \sum_k \left(\frac{\dot{D}_k}{\sigma_k}\right)^2 \tag{2.6}$$

$$\alpha = \sum_k \frac{1}{\sigma_k^2} \tag{2.7}$$

$$\beta = \sum_k \frac{D_k}{\sigma_k^2} \tag{2.8}$$

$$\gamma = \sum_k \frac{\dot{D}_k}{\sigma_k^2} \tag{2.9}$$

$$\delta = \sum_k \frac{D_k \dot{D}_k}{\sigma_k^2} \tag{2.10}$$

i σ_Δ^2 és la variància de la pertorbació.

2.3.2 Selecció del protocol i del model CHIP

Es van comparar quatre models CHIP: el model de Mitjana Ponderada, el model proposat per Micke *et al.* [46] i Mayer *et al.* [47], un altre model amb Distribució Uniforme de les pertorbacions $\Delta(r)$, i un model amb una distribució Normal Truncada de $\Delta(r)$. El model de Mitjana Ponderada suposa que $\Delta(r)$ és zero. El model Micke-Mayer implícitament suposa que la PDF de $\Delta(r)$ és una distribució uniforme i les variàncies dels termes d'error (σ_k) són iguals. Tant en la Distribució Uniforme com en el model de Normal Truncada, els termes σ_k poden ser diferents.

També vam comparar les pel·lícules EBT2 amb les EBT3, l'escaneig en mode transmissió i en mode reflexió, temps d'espera postirradiació de 75 ± 5 min i de 20 ± 1 h, i emprar en el calibratge una distribució de dosi de referència planificada o una mesurada.

El mètode de calibratge seguit en l'estudi va ser el mètode basat en un pla. Per tal d'ampliar el rang de dosis i de reduir les variacions intralot, es van utilitzar tres pel·lícules irradiades amb falques EDW de 60°. Les RMSE de les diferències entre dosis de referència ($D(r)$) i dosis a la pel·lícula per a cada canal ($D_k(r)$) es van minimitzar amb un algoritme genètic. Aquesta optimització va ajustar les corbes sensitomètriques i els paràmetres de les correccions laterals. La Distribució Uniforme i el model de Normal Truncada requereixen també de l'optimització dels paràmetres que descriuen el PDF de la pertorbació. Aquests van ser ajustats optimitzant les RMSE de les diferències entre les dosis multicanal de les pel·lícules (d) i les dosis de referència.

Protocols i models CHIP van ser seleccionats comparant, mitjançant anàlisi gamma global, les distribucions de dosis de les pel·lícules amb les distribucions de dosis planificades i amb les mesurades. Es va avaluar una mostra de 14 casos considerats representatius de les distribucions clíniques de dosi.

No es van trobar diferències significatives entre l'escaneig en mode transmissió o en mode reflexió, ni entre utilitzar pel·lícules EBT2 o EBT3. Les finestres de

temps d'espera postirradiació curtes o llargues es van trobar igualment exactes, resultat que coincideix amb les conclusions de Lewis *et al.* [19]. El model de dosimetria multicanal de distribució Normal Truncada va proporcionar el millor acord entre les distribucions de dosi de referència i les pel·lícules, com és previsible tenint en compte que els altres tres models són casos particulars o límits del model de Normal Truncada. Aquest model es pot considerar un metamodel que minimitza la incertesa inherent a la dosi dels models CHIP objectes de l'estudi. Negligir les correlacions entre canals de color o, de manera equivalent, suposar que $\Delta = 0$, com fa el model de Mitjana Ponderada, va obtenir pitjors resultats. Per altra banda, considerar que totes les pertorbacions són igualment probables pot produir valors inacceptables de la dosi amb pel·lícula. Això és degut a que la incertesa de la dosi amb pel·lícula depèn de la forma de les corbes sensitomètriques. Per exemple, la incertesa de la dosi amb el model proposat per Micke *et al.* [46] i Mayer *et al.* [47] es pot calcular com

$$\sigma_D = \sigma_k \sqrt{\frac{\sum_{k=1}^{n} \dot{D}_k^2}{n \sum_{k=1}^{n} \dot{D}_k^2 - (\sum_{k=1}^{n} \dot{D}_k)^2}} \tag{2.11}$$

La relació entre la dosi absorbida per la pel·lícula i la resposta del sistema de dosimetria es modelitza en el calibratge. Es considera que el model amb menor RMSE de les diferències entre la pel·lícula i les dosis de referència és el més exacte. En conseqüència, es preveu que una menor RMSE en el calibratge correlaciona amb valors inferiors de $\overline{\gamma}$ i superiors de $\gamma_{<1}$ en l'anàlisi gamma dels casos de prova. Es va trobar que aquestes correlacions eren significatives.

2.4 Article 3. Model selection for radiochromic film dosimetry

2.4.1 Un Model de Pertorbació General

Els Models de Pertorbació Independent del Canal són casos particulars del Model de Pertorbació General presentat en l'Article 3. Segons aquest model, la dosi absorbida per la pel·lícula al punt r, $D(r)$, es pot expressar com:

$$D(r) = D_k(z_k(r) + \Psi_k(r, z_k)) + \Sigma_k(r) \tag{2.12}$$

on k representa el canal de color, D_k és la dosi obtinguda directament de la corba sensitomètrica (per al canal k), $z_k(r)$ és la resposta del conjunt escàner-pel·lícula, $\Psi_k(r, z_k)$ corregeix pertorbacions locals sistemàtiques i $\Sigma_k(r)$ representa el soroll.

El soroll consisteix en pertorbacions aleatòries i danys en el conjunt escàner-pel·lícula (*p.ex.*, pols, ratllades). Les pertorbacions aleatòries canvien o desapareixen entre escanejos repetits o entre els escanejos de les pel·lícules abans i després de la irradiació. Els danys en el conjunt escàner-pel·lícula provoquen greus

alteracions de la resposta dosimètrica. Les pertorbacions locals sistemàtiques persisteixen entre repeticions de l'escaneig (*p.ex.*, variacions de la capa activa, artefacte lateral).

Si les variacions locals sistemàtiques són petites en comparació amb la resposta, podem aplicar un desenvolupament de Taylor de primer ordre per a $D(r)$ al voltant de $\Psi_k(r, z_k)$ en eq.(2.12):

$$D(r) = D_k(z_k(r)) + \dot{D}_k(z_k(r))\Psi_k(r, z_k) + \epsilon_k(r) \tag{2.13}$$

on $\dot{D}_k$ és la primera derivada de D_k respecte a z_k, i $\epsilon_k(r)$ és un terme d'error que representa la diferència entre la dosi absorbida real, $D(r)$, i la dosi mesurada després de corregir la pertorbació. Els models CHIP són casos particulars d'aquest Model de Pertorbació General, els quals suposen que la pertorbació $\Psi_k(r, z_k)$ és igual en els tres canals de color.

2.4.2 Selecció del model

L'Article 3 exposa la base matemàtica del mètode seguit al llarg d'aquesta tesi per tal d'optimitzar tant el protocol com el model dosimètric amb pel·lícula radiocròmica. L'objectiu de la selecció dels models era maximitzar la probabilitat de que la dosi de la pel·lícula i la dosi real foren iguals.

Segons la teoria de la probabilitat bayesiana, la probabilitat d'un esdeveniment és proporcional a la probabilitat condicional de l'evidència donat l'esdeveniment per la probabilitat a priori de l'esdeveniment: $P(M|D) \propto P(D|M)\,P(M)$. En el nostre cas, l'esdeveniment era el model dosimètric, i l'evidència consistia en les dosis de referència i les dosis mesurades amb pel·lícula, utilitzades totes elles per al calibratge dels lots. L'opció Pareto-òptima per als priors és el pes universal basat en la complexitat de Kolmogorov del model [48, 49]. Tanmateix, atès que la complexitat de Kolmogorov no és computable finitament, el AIC [40] va ser escollit per a la selecció com a criteri pràctic i basat en la complexitat. Segons el AIC, el model més probable és aquell que té el valor AIC més baix: $AIC = 2\,c_M - 2\ln(P(D|M))$, on c_M és el nombre de paràmetres del model. Per tal de calcular $P(D|M)$, vam assumir que les diferències (ϵ_d) entre les dosis de referència ($D(r)$) i les dosis amb pel·lícula ($d(r)$) presentaven una distribució normal. El nombre de paràmetres dels models era insignificant en comparació amb la mida de la mostra de l'evidència. Llavors, el model de dosimetria més probable era el que tenia menor RMSE entre dosis de referència i dosis amb pel·lícula.

La mostra de dades a l'Article 3 incloïa diferents lots de pel·lícules amb diferents intervals de dosi i nombre de dades. Per donar el mateix pes a lots diferents, es van utilitzar diferències relatives de dosis en lloc de diferències absolutes.

Es va suposar que les diferències relatives entre dosis de referència i dosis amb pel·lícula ($d(r)$) també es distribuïen normalment:

$$\frac{\epsilon_d(r)}{d(r)} \sim \mathcal{N}(0, \sigma^2) \tag{2.14}$$

on σ es va anomenar *incertesa (relativa) del calibratge.*

L'entropia de la informació es defineix com la log versemblança negativa esperada d'una variable aleatòria.

L'entropia (diferencial) de la informació d'un model seguint l'equació (2.14) es pot expressar com:

$$h(M) = \frac{1}{2}\ln(2\pi e\sigma^2) \tag{2.15}$$

L'entropia d'un model en un conjunt de lots igualment ponderats és la mitjana aritmètica de l'entropia del model. Maximitzar la versemblança d'un model equival a minimitzar l'entropia de la informació. Llavors, es va seleccionar el model dosimètric amb la menor mitjana geomètrica, calculada sobre tots els lots estudiats, de la incertesa del calibratge.

2.4.3 Validació

Els models de dosimetria són aproximacions de la relació entre la dosi absorbida per la pel·lícula i la resposta del sistema de dosimetria. A més, la mostra de calibratge pot ser insuficientment representativa. Per fer una validació addicional de la selecció del model, es van comparar, mitjançant l'anàlisi de l'índex gamma global [50], les distribucions de dosi de les pel·lícules amb les distribucions de dosi planificades. Les toleràncies eren 3%, 3 mm i 2%, 2 mm, i es van excloure punts amb una dosi menor al 30 % de la dosi màxima planificada. Per tal d'evitar els artefactes de soroll [51] en els resultats de la gamma, les dosis amb pel·lícula van ser seleccionades com a distribucions de referència i les dosis planificades com a distribucions d'avaluació.

Es va trobar que menors incerteses del calibratge correlacionaven significativament amb majors valors mitjos de $\gamma_{<1}$ en l'anàlisi gamma. A més, tots dos mètodes per a la selecció del protocol i el model de dosimetria van arribar a les mateixes conclusions.

Com a validació addicional, també es va estudiar la significació estadística dels resultats de l'anàlisi gamma.

2.4.4 Selecció del model i del protocol de dosimetria

L'Article 3 continuava amb l'optimització del protocol de dosimetria. En aquest treball, el model de correcció lateral seleccionat en l'Article 1 es va comparar amb el proposat per Poppinga *et al.* [52], i es va trobar que tots dos models eren

igualment exactes. Es va examinar si calien correccions laterals quan s'utilitzaven models multicanal, concloent-hi que les correccions laterals milloraven l'exactitud dels resultats, fins i tot per a models multicanal. Es van comparar els models que utilitzaven la pertorbació normal truncada amb models que utilitzaven una distribució uniforme de les pertorbacions amb igual variància en els termes d'error (ço és, la pertorbació de Micke-Mayer); els resultats de la comparació recolzaven la recomanació d'utilitzar pertorbacions normals truncades. També es va trobar que l'ús de NOD com a resposta del sistema de dosimetria donava millors resultats que l'ús de PV. Tanmateix, aquesta millora no produïa valors significativament majors de $\gamma_{<1}$ (3% 3mm) en models de tres canals. Finalment, es va trobar que, amb els models multicanal estudiats, augmentar la quantitat de canals de color no necessàriament donava dosis més exactes amb pel·lícula radiocròmica.

En total, 42 casos de prova es van convertir a dosis, aplicant, a cadascun d'ells, 44 models de dosimetria diferents. Les distribucions de dosi més exactes es van trobar amb el model que aplicava correccions laterals, utilitzava NOD com a resposta, i combinava tots tres canals de color d'acord amb el CHIP amb pertorbacions normals truncades.

Cal assenyalar que l'aplicabilitat de les conclusions està constreta pels límits dels casos de prova (*p.ex.*, el rang de dosi de 20-600 cGy) i el disseny específic dels models (*p.ex.*, corbes sensitomètriques polinomials).

2.5 Article 4. Grid patterns, spatial inter-scan variations and scanning reading repeatability in radiochromic film dosimetry

2.5.1 Protocol d'escaneig

Diversos elements del protocol de dosimetria van ser estudiats i seleccionats en els Articles 1-3. Tanmateix, el protocol d'escaneig no es va modificar pràcticament des de l'Article 1. És a dir, les pel·lícules es van escanejar amb un escàner pla Epson Expression 10000XL mitjançant el programari Epson Scan v.3.x associat. S'escanejaren amb orientació vertical (ço és, el costat curt de la pel·lícula era paral·lel a la làmpada de l'escàner). Abans de les adquisicions, l'escàner es va escalfar durant almenys 30 minuts. Després d'això, i després de llargues pauses, es van realitzar cinc escanejos buits per estabilitzar la temperatura del llum de l'escàner. Cada pel·lícula es va escanejar cinc vegades seguides, es va descartar el primer escaneig, i la imatge resultant es va calcular amb la mitjana dels altres quatre. Les pel·lícules se centraren a l'escàner amb un marc opac. Es desactivaren les eines de processament d'imatges. Les imatges es van adquirir en mode RGB de 48 bits (16 bit per canal) i es van guardar com a fitxers TIFF. Les úniques desviacions de l'Article 4 respecte del protocol d'escaneig de l'Article 1, a part

d'utilitzar una versió actualitzada del programari Epson Scan, van consistir en col·locar una xapa de vidre de 3 mm de gruix a sobre de les pel·lícules per mantenir constant la distància entre pel·lícula i llum, i en utilitzar el mode transmissió en lloc del mode reflexió, tenint en compte que en l'Article 2 no es van trobar diferències significatives d'exactitud entre ambdós modes.

L'objectiu de l'Article 4 era examinar les incerteses relacionades amb la repetibilitat de l'escàner per tal de millorar el protocol d'escaneig.

2.5.2 Patrons de quadrícula

Calcular la mitjana de diversos escanejos redueix el soroll de la imatge i proporciona distribucions de dosi amb pel·lícula més exactes [53]. No obstant això, les variacions entre escanejos donen lloc a altres pertorbacions. Una d'aquestes pertorbacions són els patrons de quadrícula. Els patrons de quadrícula són artefactes espuris en forma de quadrícula. En general no són detectats, però de vegades poden aparèixer en distribucions de dosi o en anàlisis de l'índex gamma realitzats amb pel·lícules radiocròmiques. En aquest treball, es van trobar patrons de quadrícula amb resolucions de 50, 72 i 96 ppp, però no amb 150 ppp. Fins i tot van aparèixer en absència de llum transmesa. L'origen dels patrons de quadrícula és la variació periòdica del soroll de l'escàner al llarg d'ambdós eixos. Les diferents resolucions presenten diferents variacions periòdiques. La incertesa mitjana de la dosi deguda al soroll de l'escàner va ser almenys dues vegades superior a la diferència entre la màxima i la mínima incertesa de la dosi produïda pels patrons de quadrícula. És per això que, normalment, aquest artefacte no es percep. Tanmateix, és fàcilment visible quan s'escaneja amb una resolució de 72 dpi. Amb aquesta resolució, el patró de quadrícula té una forma sinusoïdal amb un període de 8.5 mm.

2.5.3 Variabilitat espacial entre escanejos

Lewis *et al* [30] van detectar que la resposta de l'escàner varia entre escanejos, generant canvis en la relació entre la dosi absorbida per la pel·lícula i la resposta del sistema de dosimetria. Si no es corregeixen, aquests canvis poden produir errors rellevants en la dosi mesurada amb pel·lícula. Lewis *et al* [30] van proposar corregir la resposta mitjançant un fragment no exposat escanejat juntament amb la pel·lícula analitzada. Suggerien calcular, per a cada escaneig, el PV mig en una ROI de referència al fragment no exposat. Aquest PV mig en la ROI de referència (Ref ROI) s'havia de comparar amb un Ref ROI PV de referència, com ara l'obtés durant el calibratge. La correcció consistia a escalar la resposta de l'escàner, en PV, pel factor necessari per obtenir en el Ref ROI el PV del calibratge. En l'Article 4, aquesta correcció es va anomenar *correcció mitjana*, i es va expressar com:

$$M(i,j) = v(i,j)\left\langle \frac{M(i_{Ref}, j_{Ref})}{v(i_{Ref}, j_{Ref})} \right\rangle \tag{2.16}$$

on (i,j) representa la posició del píxel a la imatge (i és la fila i j la columna), M simbolitza el valor de referència del píxel, v és el PV mesurat, i (i_{Ref}, j_{Ref}) és un píxel de la Ref ROI.

La correcció mitjana és idèntica per a cada píxel de l'escaneig, no varia en l'espai. Vam comparar-la amb una correcció que varia en l'espai: la *correcció de columna*. La correcció de columna calcula una correcció diferent per a cada posició en l'eix paral·lel al llum de l'escàner. Considera que les variacions entre escanejos depenen dels dispositius de càrrega acoblada (CCD), per la qual cosa originen patrons lineals perpendiculars a la làmpada de l'escàner en les distribucions de dosi amb pel·lícula. Les correccions de columna comparen el PV mig i el PV mig de referència per a cada columna de la Ref ROI:

$$M(i,j) = v(i,j)\left\langle \frac{M(i_{Ref}, j)}{v(i_{Ref}, j)} \right\rangle \tag{2.17}$$

Aplicant, bé la correcció mitjana de la resposta, bé la de columna, es van reduir les diferències de dosis entre escanejos repetits, en major mesura amb la correcció de columna. Sense aplicar cap correcció de resposta deguda a la variació entre escanejos, es van produir desviacions sistemàtiques respecte a la distribució de dosi de referència superiors a l'1% en moltes exploracions. Després d'aplicar correccions, no es va trobar cap desviació sistemàtica superior a l'1%. En l'eix perpendicular a la làmpada de l'escàner, les variacions espacials entre escanejos van ser insignificants, excepte en els escanejos inicials d'escalfament.

2.5.4 Repetibilitat de lectura de l'escàner

El posicionament inicial de la làmpada en l'eix paral·lel al seu moviment, i la velocitat del moviment, varien entre escanejos. Les diferències en el posicionament inicial es van trobar inferiors a 0.1 mm. No obstant això, aquestes diferències van créixer amb la distància des de la posició inicial a causa de les variacions de la velocitat. En els nostres experiments, vam trobar diferències de 0.7 mm a una distància de 20 cm des de la posició inicial. En conseqüència, es pot afirmar que els escanejos mitjans perden exactitud conforme s'allunyen de la posició inicial de la làmpada. L'impacte dosimètric d'aquestes variacions sol ser insignificant. Pot ser rellevant, però, i s'han d'adoptar les mesures adients, en alguns estudis, com ara mesures de penombres o calibratges basats en un pla amb falca.

2.6 Discussió, conclusions i treball futur

L'objectiu d'aquesta tesi era millorar l'exactitud de la dosimetria amb pel·lícula radiocròmica. El principal focus de recerca va ser l'optimització de protocols i models de dosimetria. Els treballs recollits ací van fer diverses contribucions al cos de coneixement en el camp de la dosimetria amb pel·lícula radiocròmica, com ara:

Es va introduir un Model de Pertorbació General per a la dosimetria multicanal amb pel·lícula radiocròmica i escàner pla. Es van desenvolupar i analitzar Models de Pertorbació Independent del Canal, que són casos particulars del Model de Pertorbació General. Es va obtenir la dosi absorbida estimada i la seva incertesa inherent per als models CHIP amb funcions de densitat de probabilitat de la pertorbació uniforme i normal truncada, inclòs el model CHIP proposat per Micke *et al.* [46] i Mayer *et al.* [47]. Es van explicar les assumpcions implícites del model Micke-Mayer. Es va detectar que totes les possibles pertorbacions són equiprobables en aquest model, la qual cosa el fa sensible a les propietats de les corbes sensitomètriques i pot provocar incerteses inacceptables. El model Micke-Mayer es va comparar amb el model de pertorbació normal truncada. Aquest últim va proporcionar resultats més exactes. També es van trobar millors resultats quan es van aplicar correccions laterals i usant NOD en lloc de PV com a resposta del sistema de dosimetria. Les correccions laterals van proporcionar una millor exactitud fins i tot per a models multicanal. Tanmateix, l'ús de NOD va ser significativament millor només per a la dosimetria monocanal. L'ús de NOD no va ser suficient per corregir les heterogeneïtats de la pel·lícula. Els modes d'escaneig de reflexió i de transmissió no van donar diferències significatives. Tampoc no va donar cap diferència significativa l'ús de pel·lícules EBT2 o EBT3, ni emprar temps d'espera postirradiació curts o llargs. Generalment, a la literatura s'han seleccionat models i protocols de dosimetria amb pel·lícula mitjançant comparacions de l'índex gamma. En aquesta tesi, es va proposar, justificar matemàticament i validar amb comparacions gamma, la selecció basada en la menor RMSE del calibratge. Cal tenir en compte que les conclusions relacionades amb els models i protocols seleccionats depenen dels límits dels casos de prova (*p.ex.*, el rang de dosis) i del disseny específic dels models (*p.ex.*, en aquest treball, corbes sensitomètriques polinomials). Per exemple, el model de correcció lateral absoluta seleccionat en aquesta tesi es va trobar tan exacte com el model recomanat per Lewis *et al.* [41]. Tanmateix, aquest resultat es va limitar a dosis inferiors a 4 Gy. Aquesta tesi també va presentar una alternativa al calibratge amb fragments: el calibratge basat en un pla. Aquest mètode és més ràpid i alhora permet la computació de la correcció lateral. A més, la mostra de calibratge és més representativa, la qual cosa el fa més robust contra pertorbacions. No obstant això, les dosis de referència tenen una major incertesa, és un mètode computacionalment més intensiu, i es pot veure afectat per la repetibilitat de lectura de l'escàner. El

concepte de repetibilitat de lectura de l'escàner es va introduir en aquest treball. Es va trobar que el posicionament inicial (en l'eix paral·lel al moviment del llum) i la velocitat de l'escàner varien entre escanejos. També es van descobrir en aquesta tesi els patrons de quadrícula, que són artefactes causats per la variació periòdica del soroll de l'escàner al llarg d'ambdós eixos. Una altra font d'incertesa produïda per la repetibilitat de l'escàner són les variacions entre escanejos. En aquesta tesi, es va proposar un nou mètode de correcció, la correcció de columna, per tenir en compte les desviacions dels diferents sensors CCD.

Els models de pertorbació per a la dosimetria amb pel·lícula radiocròmica es basen en el supòsit que les pertorbacions són petites en comparació amb la resposta. Aquesta assumpció pot ser inadequada (*p.ex.*, l'artefacte lateral pot reduir considerablement la resposta lluny del centre de la làmpada de l'escàner) i provocar errors importants en el càlcul de la dosi. Un major coneixement de les fonts de pertorbació és clau per millorar la dosimetria amb pel·lícula. Correccions més precises dels artefactes laterals, de les variacions entre escanejos, de les variacions intralot, de les diferències amb el temps d'espera postirradiació, etc són objectius per futurs estudis. Una altra font d'incertesa són les heterogeneïtats en la capa activa de la pel·lícula. L'ús de NOD no les elimina. Una de les nostres línies de recerca actuals està centrada en aquestes heterogeneïtats. Hem trobat correlacions entre canals irradiats i no irradiats, que es poden modelar amb una funció de densitat de probabilitat Gaussiana Multivariada dels valors de píxel, és el que anomenem *model Multigaussian*. Hem comparat el model Multigaussian amb altres models dosimètrics amb resultats encoratjadors. Finalment, una altra línia de recerca a seguir és l'estudi de corbes sensitomètriques o correccions universals o genèriques, que podrien permetre mesures de dosi més senzilles, ràpides i exactes.

Bibliography

[1] A. Niroomand-Rad, C. R. Blackwell, B. M. Coursey, K. P. Gall, J. M. Galvin, W. L. McLaughlin, A. S. Meigooni, R. Nath, J. E. Rodgers, and C. G. Soares. "Radiochromic film dosimetry: Recommendations of AAPM Radiation Therapy Committee Task Group 55". *Medical Physics* 25.11 (1998), pp. 2093–2115.

[2] A. Rink, D. F. Lewis, S. Varma, I. A. Vitkin, and D. A. Jaffray. "Temperature and hydration effects on absorbance spectra and radiation sensitivity of a radiochromic medium". *Medical physics* 35.10 (2008), pp. 4545–4555.

[3] M. J. Butson, T. Cheung, P. K. Yu, and H. Alnawaf. "Dose and absorption spectra response of EBT2 Gafchromic film to high energy x-rays". *Australasian Physical & Engineering Science in Medicine* 32.4 (2009), pp. 196–202.

[4] B. Hartmann, M. Martišíková, and O. Jäkel. "Technical Note: Homogeneity of Gafchromic EBT2 film". *Medical Physics* 37.4 (2010), pp. 1753–1756.

[5] S. Reinhardt, M. Hillbrand, J. J. Wilkens, and W. Assmann. "Comparison of Gafchromic EBT2 and EBT3 films for clinical photon and proton beams". *Medical Physics* 39.8 (2012), pp. 5257–5262.

[6] J. Martín-Viera Cueto, V. Parra Osorio, C. Moreno Sáiz, F. Navarro Guirado, F. Casado Villalón, and P. Galán Montenegro. "A universal dose-response curve for radiochromic films". *Medical physics* 42.1 (2015), pp. 221–231.

[7] D. F. Lewis and M. F. Chan. "On GAFChromic EBT-XD film and the lateral response artifact". *Medical physics* 43.2 (2016), pp. 643–649.

[8] M. Callens, W. Crijns, V. Simons, I. De Wolf, T. Depuydt, F. Maes, K. Haustermans, J. D'hooge, E. D'Agostino, M. Wevers, et al. "A spectroscopic study of the chromatic properties of GafChromic EBT3 films". *Medical physics* 43.3 (2016), pp. 1156–1166.

[9] S. Devic, N. Tomic, and D. Lewis. "Reference radiochromic film dosimetry: review of technical aspects". *Physica Medica* 32.4 (2016), pp. 541–556.

[10] S. Devic, J. Seuntjens, E. Sham, E. B. Podgorsak, C. R. Schmidtlein, A. S. Kirov, and C. G. Soares. "Precise radiochromic film dosimetry using a flat-bed document scanner". *Medical Physics* 32.7 (2005), pp. 2245–2253.

[11] W. Crijns, F. Maes, U. A. van der Heide, and F. V. den Heuvel. "Calibrating page sized Gafchromic EBT3 films". *Medical Physics* 40.1 (2013), 012102 (13pp.)

[12] A. Rink, I. A. Vitkin, and D. A. Jaffray. "Energy dependence (75 kVp to 18 MV) of radiochromic films assessed using a real-time optical dosimeter". *Medical Physics* 34.2 (2007), pp. 458–463.

[13] C. Richter, J. Pawelke, L. Karsch, and J. Woithe. "Energy dependence of EBT-1 radiochromic film response for photon (10 kVp–15 MVp) and electron beams (6–18 MeV) readout by a flatbed scanner". *Medical Physics* 36.12 (2009), pp. 5506–5514.

[14] B. Arjomandy, R. Tailor, A. Anand, N. Sahoo, M. Gillin, K. Prado, and M. Vicic. "Energy dependence and dose response of Gafchromic EBT2 film over a wide range of photon, electron, and proton beam energies". *Medical Physics* 37.5 (2010), pp. 1942–1947.

[15] P. Lindsay, A. Rink, M. Ruschin, and D. Jaffray. "Investigation of energy dependence of EBT and EBT-2 Gafchromic film". *Medical Physics* 37.2 (2010), pp. 571–576.

[16] G. Massillon-JL, S. Chiu-Tsao, I. Domingo-Munoz, and M. Chan. "Energy Dependence of the New Gafchromic EBT3 Film:Dose Response Curves for 50 KV, 6 and 15 MV X-Ray Beams". *International Journal of Medical Physics, Clinical Engineering and Radiation Oncology* 1.2 (2012), pp. 60–65.

[17] H. Bekerat, S. Devic, F. DeBlois, K. Singh, A. Sarfehnia, J. Seuntjens, S. Shih, X. Yu, and D. Lewis. "Improving the energy response of external beam therapy (EBT) GafChromic dosimetry films at low energies (≤ 100 keV)". *Medical Physics* 41.2 (2014), 022101 (14pp.)

[18] S. Devic. "Radiochromic film dosimetry: past, present, and future". *Physica medica* 27 (2011), pp. 122–134.

[19] D. Lewis, A. Micke, X. Yu, and M. F. Chan. "An efficient protocol for radiochromic film dosimetry combining calibration and measurement in a single scan". *Medical Physics* 39.10 (2012), pp. 6339–6350.

[20] A. L. Palmer, A. Nisbet, and D. Bradley. "Verification of high dose rate brachytherapy dose distributions with EBT3 Gafchromic film quality control techniques". *Physics in medicine and biology* 58.3 (2013), p. 497.

[21] J. F. P. Azorín, L. I. R. García, and J. M. Martí-Climent. "A method for multichannel dosimetry with EBT3 radiochromic films". *Medical Physics* 41.6 (2014), 062101 (10pp.)

[22] I. Méndez. "Model selection for radiochromic film dosimetry". *Physics in medicine and biology* 60.10 (2015), p. 4089.

[23] C. Andrés, A. del Castillo, R. Tortosa, D. Alonso, and R. Barquero. "A comprehensive study of the Gafchromic EBT2 radiochromic film. A comparison with EBT". *Medical Physics* 37.12 (2010), pp. 6271–6278.

[24] L. Chang, S.-Y. Ho, H.-J. Ding, T.-F. Lee, and P.-Y. Chen. "Dependency of EBT2 film calibration curve on postirradiation time". *Medical physics* 41.2 (2014).

[25] F. Girard, H. Bouchard, and F. Lacroix. "Reference dosimetry using radiochromic film". *Journal of Applied Clinical Medical Physics* 13.6 (2012). ISSN: 15269914.

[26] A. A. Schoenfeld, D. Poppinga, D. Harder, K.-J. Doerner, and B. Poppe. "The artefacts of radiochromic film dosimetry with flatbed scanners and their causation by light scattering from radiation-induced polymers". *Physics in Medicine and Biology* 59.13 (2014), pp. 3575–3597.

[27] L. van Battum, H. Huizenga, R. Verdaasdonk, and S. Heukelom. "How flatbed scanners upset accurate film dosimetry". *Physics in medicine and biology* 61.2 (2015), p. 625.

[28] A. A. Schoenfeld, S. Wieker, D. Harder, and B. Poppe. "The origin of the flatbed scanner artifacts in radiochromic film dosimetry-key experiments and theoretical descriptions". *Physics in medicine and biology* 61.21 (2016), p. 7704.

[29] M. J. Butson, T. Cheung, and P. Yu. "Evaluation of the magnitude of EBT Gafchromic film polarization effects". *Australasian Physics & Engineering Sciences in Medicine* 32.1 (2009), pp. 21–25.

[30] D. Lewis and S. Devic. "Correcting scan-to-scan response variability for a radiochromic film-based reference dosimetry system". *Medical physics* 42.10 (2015), pp. 5692–5701.

[31] A. L. Palmer, D. A. Bradley, and A. Nisbet. "Evaluation and mitigation of potential errors in radiochromic film dosimetry due to film curvature at scanning". *Journal of Applied Clinical Medical Physics* 16.2 (2015).

[32] R. Dreindl, D. Georg, and M. Stock. "Radiochromic film dosimetry: Considerations on precision and accuracy for EBT2 and EBT3 type films". *Zeitschrift für Medizinische Physik* 24.2 (2014), pp. 153–163.

[33] L. Paelinck, W. D. Neve, and C. D. Wagter. "Precautions and strategies in using a commercial flatbed scanner for radiochromic film dosimetry". *Physics in Medicine and Biology* 52.1 (2007), pp. 231–242.

[34] B. Ferreira, M. Lopes, and M. Capela. "Evaluation of an Epson flatbed scanner to read Gafchromic EBT films for radiation dosimetry". *Physics in medicine and biology* 54.4 (2009), p. 1073.

[35] I. Méndez, Ž. Šljivić, R. Hudej, A. Jenko, and B. Casar. "Grid patterns, spatial inter-scan variations and scanning reading repeatability in radiochromic film dosimetry". *Physica Medica* 32.9 (2016), pp. 1072–1081.

[36] H. Bouchard, F. Lacroix, G. Beaudoin, J.-F. Carrier, and I. Kawrakow. "On the characterization and uncertainty analysis of radiochromic film dosimetry". *Medical Physics* 36.6 (2009), pp. 1931–1946.

[37] S. J. van Hoof, P. V. Granton, G. Landry, M. Podesta, and F. Verhaegen. "Evaluation of a novel triple-channel radiochromic film analysis procedure using EBT2". *Physics in Medicine and Biology* 57.13 (2012), pp. 4353–4368.

[38] L. Richley, A. C. John, H. Coomber, and S. Fletcher. "Evaluation and optimization of the new EBT2 radiochromic film dosimetry system for patient dose verification in radiotherapy". *Physics in Medicine and Biology* 55.9 (2010), p. 26012617.

[39] T. Kairn, T. Aland, and J. Kenny. "Local heterogeneities in early batches of EBT2 film: a suggested solution". *Physics in Medicine and Biology* 55.15 (2010), p. L37.

[40] K. P. Burnham and D. R. Anderson. *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Science & Business Media, 2002.

[41] D. Lewis and M. F. Chan. "Correcting lateral response artifacts from flatbed scanners for radiochromic film dosimetry". *Medical physics* 42.1 (2015), pp. 416–429.

[42] S. Saur and J. Frengen. "GafChromic EBT film dosimetry with flatbed CCD scanner: A novel background correction method and full dose uncertainty analysis". *Medical Physics* 35.7 (2008), pp. 3094–3101.

[43] M. Fuss, E. Sturtewagen, C. D. Wagter, and D. Georg. "Dosimetric characterization of GafChromic EBT film and its implication on film dosimetry quality assurance". *Physics in Medicine and Biology* 52.14 (2007), pp. 4211–4225.

[44] M. Martišíková, B. Ackermann, and O. Jäkel. "Analysis of uncertainties in Gafchromic EBT film dosimetry of photon beams". *Physics in Medicine and Biology* 53.24 (2008), pp. 7013–7027.

[45] I. A. E. Agency. *Commissioning of Radiotherapy Treatment Planning Systems: Testing for Typical External beam Treatment Techniques. IAEA-TECDOC-1583*. Tech. rep. Vienna, Austria: IAEA, 2008.

[46] A. Micke, D. F. Lewis, and X. Yu. "Multichannel film dosimetry with nonuniformity correction". *Medical Physics* 38.5 (2011), pp. 2523–2534.

[47] R. R. Mayer, F. Ma, Y. Chen, R. I. Miller, A. Belard, J. McDonough, and J. J. O'Connell. "Enhanced dosimetry procedures and assessment for EBT2 radiochromic film". *Medical Physics* 39.4 (2012), pp. 2147–2155.

[48] M. Hutter. *Universal Artificial Intelligence: Sequential Decisions based on Algorithmic Probability*. Berlin: Springer, 2005.

[49] M. Li and P. M. Vitányi. *An introduction to Kolmogorov complexity and its applications*. 3rd. New York: Springer-Verlag, 2009.

[50] D. A. Low, W. B. Harms, S. Mutic, and J. A. Purdy. "A technique for the quantitative evaluation of dose distributions". *Medical Physics* 25.5 (1998), pp. 656–661.

[51] B. M. Clasie, G. C. Sharp, J. Seco, J. B. Flanz, and H. M. Kooy. "Numerical solutions of the gamma-index in two and three dimensions". *Physics in Medicine and Biology* 57.21 (2012), pp. 6981–6997.

[52] D. Poppinga, A. A. Schoenfeld, K. J. Doerner, O. Blanck, D. Harder, and B. Poppe. "A new correction method serving to eliminate the parabola effect of flatbed scanners used in radiochromic film dosimetry". *Medical Physics* 41.2 (2014), 021707 (8pp.)

[53] J. A. V. Sánchez, C. R. Morales, and A. G. López. "Characterization of noise and digitizer response variability in radiochromic film dosimetry. Impact on treatment verification". *Physica Medica* 32.9 (2016), pp. 1167–1174.

Part II

Appended papers

Paper 1

Gafchromic EBT2 film dosimetry in reflection mode with a novel plan-based calibration method

I. Méndez, V. Hartman, R. Hudej, A. Strojnik and B. Casar

Medical physics 40.1 (2013), 011720

Paper 1. Gafchromic EBT2 film dosimetry in reflection mode with a novel plan-based calibration method

I. Méndez, V. Hartman, R. Hudej, A. Strojnik and B. Casar

Abstract

Purpose: A dosimetric system formed by Gafchromic EBT2 radiochromic film and Epson Expression 10000XL flatbed scanner was commissioned for dosimetry. In this paper, several open questions concerning the commissioning of radiochromic films for dosimetry were addressed: a) is it possible to employ this dosimetric system in reflection mode; b) if so, can the methods used in transmission mode also be used in reflection mode; c) is it possible to obtain accurate absolute dose measurements with Gafchromic EBT2 films; d) which calibration method should be followed; e) which calibration models should be used; f) does three-color channel dosimetry offer a significant improvement over single channel dosimetry. The purpose of this paper is to help clarify these questions.

Methods: In this study, films were scanned in reflection mode, the effect of surrounding film was evaluated and the feasibility of EBT2 film dosimetry in reflection mode was studied. EBT2's response homogeneity has been reported to lead to excessive dose uncertainties. To overcome this problem, a new plan-based calibration method was implemented. Plan-based calibration can use every pixel and each of the three color channels of the scanned film to obtain the parameters of the calibration model. A model selection analysis was conducted to select lateral correction and sensitometric curve models. The commonly used calibration with fragments was compared with red-channel plan-based calibration and with three-channel plan-based calibration.

Results: No effect of surrounding film was found in this study. The film response inhomogeneity in EBT2 films was found to be important not only due to differences in the fog, but also due to differences in sensitivity. The best results for lateral corrections were obtained using absolute corrections independent of the dose. With respect to the sensitometric curves, an empirical polynomial fit of order 4 was found to obtain results equivalent to a gamma-distributed single hit model based on physical assumptions. Three-channel dosimetry was found to be substantially superior to red-channel dosimetry.

Conclusions: Reflection mode with Gafchromic EBT2 radiochromic film was found to be a viable alternative to transmission mode. The same methods that are used in transmission mode can be followed in reflection mode. A novel plan-based method was developed for calibration and multichannel dosimetry. This novel method offers increased robustness against film response inhomogeneities and reduces considerably the time required for calibration.

1 Introduction

Radiochromic films present weak energy dependence, high spatial resolution, and near water equivalence. This makes them appropriate for measurements whenever nonequilibrium conditions exist, in fields with high dose gradients and in tissue heterogeneities: particularly for advanced radiotherapy techniques such as intensity modulated radiotherapy (IMRT) or stereotactic radiosurgery (SRS). Nevertheless, some questions remain open when radiochromic films are commissioned for dosimetry. The purpose of this paper is to help clarify these questions.

In this research, a dosimetric system formed by Gafchromic EBT2 radiochromic film (International Specialty Products, Wayne, NJ) and Epson Expression 10000XL flatbed scanner (Seiko Epson Corporation, Nagano, Japan) was commissioned for dosimetry. The Epson Expression 10000XL can scan either in reflection mode or in transmission mode, the later with a transparency adapter purchased separately. Then, the first question to answer is whether it is possible to employ this dosimetric system in reflection mode. If so, it should be investigated whether the methods used in transmission mode can also be used in reflection mode. Radiochromic film dosimetry has been previously developed in reflection mode[1, 2]. Nevertheless, to the authors' knowledge, the only analysis of dosimetry in reflection mode using Gafchromic EBT films (namely EBT2 films) was performed by Richley *et al.*[3], who reported an effect of surrounding film that makes it impossible to use the same protocols that are used in transmission mode dosimetry also in reflection mode. In this study films were scanned in reflection mode, the effect of surrounding film was evaluated and the feasibility of EBT2 film dosimetry in reflection mode was studied.

Gafchromic EBT film was replaced by EBT2 film in 2009. EBT film has been extensively commissioned and found to be reliable for dose measurements [4–9]. However, EBT2 properties have been studied in several papers[3, 10–12] and doubts have been cast on its response homogeneity [13, 14], which has been reported to lead to excessive dose uncertainties. A question arises whether it is possible to obtain accurate absolute dose measurements with Gafchromic EBT2 films. The answer to this question depends on the calibration method employed. In the literature, the most frequent calibration method uses fragments irradiated with different doses and scanned in different positions over the scanner[5, 8, 15, 16]. Another faster and accurate method was proposed by Menegotti *et al.* [9] who used single film exposure. A weakness of this method was the fact that only six levels of uniform dose placed in stripes were used to parametrize the sensitometric curve. This could reduce the accuracy of the sensitometric curve[15], especially when film response is inhomogeneous. To offer increased robustness against film response inhomogeneities, while being faster than the conventional fragment based method, a novel plan-based calibration method was developed in

this work. Independently of the calibration method, a suitable calibration model (consisting of lateral correction and sensitometric curve models) should be chosen. In this paper, a model selection analysis based on maximum likelihood estimation was performed to select the calibration model. The last question addressed in this paper is whether three-channel dosimetry offers a significant improvement over one channel dosimetry. Until recently[17–19], only one color channel has been commonly used for radiochromic film dosimetry. The red channel has been chosen because it has been found to provide the greatest sensitivity at lower doses [3]. Three-channel dosimetry using the weighted mean of the channels was developed both for calibration with fragments and plan-based calibration, and their results were compared.

2 Methods and materials

2.1 Dosimetric system

Gafchromic EBT2 films with dimensions 8 inch × 10 inch were used in this work. They were handled following recommendations outlined in the AAPM TG-55 report [20]. When smaller film pieces were required, films were divided into fragments with dimensions: 6.4 cm × 6.8 cm. The lot used was A04141003BB, except in the analysis of intralot and interlot variations, which included also lot A03171101A.

Films were digitized with an Epson Expression 10000XL flatbed scanner. This device is a 48-bit color scanner equipped with a linear charge-coupled device (CCD) as optical sensor, a xenon lamp as the light source and which can scan either in transmission mode (if a transparency adapter is acquired) or in reflection mode. In this work, each film was scanned in reflection mode.

2.2 Irradiation procedure

The films employed in preliminary tests and calibrations were irradiated in a 12×30×30 cm^3 Plastic Water phantom (Computerized Imaging Reference Systems Inc. Norfolk, VA, USA) with a 6 MV photon beam from a Novalis Tx accelerator (Varian, Palo Alto, CA, USA). Films were centered on the beam axis at a depth of 6 cm using SAD (source-axis distance) set-up.

The films employed in verification tests were irradiated in a CIRS Thorax phantom (Model 002LFC) with a 6 MV photon beam from a Varian Unique accelerator (Varian, Palo Alto, CA, USA). The CIRS Thorax phantom represents an average human torso, both in dimensions and structure. Its body is made of plastic water and it includes tissue heterogeneities corresponding to lung and bone. The films were placed with an offset of 1.5 cm from the beam axis to avoid the film and the beam axis being in the same plane [21]. The phantom was set-up at SAD.

2.3 Scanning protocol

Before acquisitions, the scanner was allowed to warm up for 30 min. Each film or film fragment was scanned in reflection mode and portrait orientation, 24 h after exposure (within a time window of less than 1 h)[11] and centered on the scanner with a black opaque cardboard frame.

Films were digitized using the associated software Epson Scan v.3.0. Images were acquired in "professional mode" with the image type set to 48-bit RGB (16 bit per channel), depending on the test the resolution was either 72 dpi or 150 dpi (0.35 mm/px or 0.17 mm/px) and the image processing tools were turned off. Data were saved as TIFF (tagged image file format) files.

Five consecutive scans were made for each film. The warm-up effect of the scanner lamp due to multiple scans [6, 8] was studied during the commissioning. The first scan was found to be markedly different from the last four scans and was therefore discarded. The resulting image was the average of the last four. Images were analyzed with the open-source software ImageJ v.1.44o (National Institutes of Health, USA).

2.4 Preliminary tests

Effect of surrounding film

Richley *et al.* [3], reported that the pixel value (PV) of a region of interest (ROI) was dependent on the PV of the surrounding film when scanned in reflection mode with Epson Expression 10000XL. This effect would implicate a serious disadvantage for dosimetry in reflection mode compared to transmission mode, requiring more tests and worsening the uncertainties of the dosimetric system.

To investigate this effect, seven fragments of a film were used. Six of them were irradiated with 1, 2, 3, 4, 5 and 6 Gy, respectively. The nonirradiated film was centered on the scanner and one of the irradiated pieces was positioned next to it along the x axis of the scanner (*i.e.*, parallel to the CCD array). Images were acquired with resolution of 72 dpi, and a 50×50 px ROI was measured at the center of the nonirradiated film. The process was repeated with each irradiated fragment.

Film response homogeneity

To examine the EBT2 film response homogeneity a film was cut into 12 fragments. Before irradiation, and 24h after being irradiated with 2 Gy, each fragment was centered on the scanner, and a 100×100 px ROI (3.5×3.5 cm^2) was measured at the center of the fragment.

2.5 Calibration

Subtracting the optical density (OD) of a film before irradiation from the OD after irradiation improves the accuracy of film dosimetry [6]. This is because this procedure partially accounts for the film response inhomogeneity. Following Ohuchi (2007) [22] the reflectance can be processed in the same way as the transmittance. Hence, net optical density (NOD) [16] was defined as:

$$NOD = z = \log_{10} \frac{v_{\text{nonirr}}}{v_{\text{irr}}} \tag{1.1}$$

Where v_{nonirr} and v_{irr} represent pixel values of nonirradiated and irradiated films, respectively, after correction according to lateral correction models.

Models for lateral correction

Lateral correction is necessary since the scanner's response over the scan field is not uniform [4–9, 23–25] . Deviation from the response in the center of the scanner is particularly important along the x axis (*i.e.*, parallel to the CCD array) and usually negligible along the y axis (*i.e.*, perpendicular to the CCD array). Besides, this correction could also be dependent on the pixel value (or equivalently on the dose). Then, considering that it is negligible along the y axis, lateral correction is a bidimensional function dependent on PV and pixel position along the x axis. Different approaches to this correction had been proposed in the literature: lateral correction function has been approximated by a matrix of correction factors [5, 8], it has been considered independent of the PV [25], the dependency on pixel position has been considered a parabola [9], etc. Based on the corrections proposed in the literature, four different bidimensional polynomial approximations to the lateral correction function were investigated. All of them are empirical, since the authors did not find in the literature any lateral correction function based on physical assumptions:

Type I:

$$v = a_1(x - x_c) + a_2(x - x_c)^2 + \hat{v} \tag{1.2}$$

Type II:

$$v = \hat{v}(1 + a_1(x - x_c) + a_2(x - x_c)^2) \tag{1.3}$$

Type III:

$$v = \hat{v}(1 + (a_1 + a_2\hat{v})(x - x_c) + (a_3 + a_4\hat{v})(x - x_c)^2) \tag{1.4}$$

Type IV:

$$v = a_1(x - x_c) + a_2(x - x_c)^2 + \hat{v}(1 + a_3(x - x_c) + a_4(x - x_c)^2) \tag{1.5}$$

Here $\hat{v}$ represents "raw" not corrected PV, x_c is the x coordinate of the center of the scanner, and v represents corrected PV.

Type I corresponds to an absolute correction independent of dose and second order in the distance from the center. Type II is a relative correction second order in the distance. Type III is a relative correction second order in the distance and in the PV. Type IV is a combination of types I and II.

Models for sensitometric curves

Throughout this work, calibration is considered a process that yields the dose measured in a point with pixel position and PVs before and after irradiation given as inputs. Then, it includes lateral correction, NOD calculation and sensitometric curve. Sensitometric curves convert NOD to absolute dose.

Two types of sensitometric curves with functional forms following the conditions stated by Bouchard *et al.* [15] were studied.

The first one is an empirical curve, a polynomial fit of order n:

$$D = \sum_{i=1}^{n} b_i z^i \tag{1.6}$$

where z represents NOD.

The second one is based on physical assumptions. A gamma-distributed single hit model derived from percolation theory [26]:

$$D = \left(\frac{b_1}{b_2 - z}\right)^{1/b_3} - \left(\frac{b_1}{b_2}\right)^{1/b_3} \tag{1.7}$$

2.6 Calibration with plan-based method

Matrix of data

Increasing the number of dose levels decreases the uncertainty in the sensitometric curve [15]. This is especially important when film response is inhomogeneous. Considering this, a film was irradiated with a 60° Enhanced Dynamic Wedge (EDW) field of dimensions 20×20 cm^2 with 438 MU (doses on the film ranging from approximately 75 cGy to approximately 400 cGy). The film was digitized before and after exposure at 150 dpi, obtaining a matrix of data with 1200×1500 px. A margin of 100×200 px was avoided during the computation.

A plan with the geometry of the irradiation was calculated using Eclipse v.10.0 (Varian Medical Systems) treatment planning system (TPS) with anisotropic analytical algorithm (AAA). The accuracy of the calculation of EDW by the TPS was commissioned previously with ionization chamber, linear diode array and 2D ion chamber array. The calculated absolute dose distribution on the plane of the film was exported to a matrix with resolution 0.59 mm/px. It was bilinearly interpolated to register to the film.

Using the described method, a matrix of 1100×1300 data points was obtained. Each data point included: x and y coordinates of the pixel, dose calculated by the TPS and pixel values before and after irradiation for all three color channels (R, G and B).

Model selection

To select the most appropriate lateral correction and sensitometric curve model, *i.e.*, the calibration model, least squares fitting was used as a maximum likelihood estimation. Therefore, the most probable calibration model is the one that minimizes the root-mean-square error (RMSE) of the differences between the doses measured with film and the doses calculated by the TPS.

To calculate the RMSE of a calibration model a program was depeloved in C++. A genetic algorithm [27] searched the parameters that minimize the RMSE of the calibration model. Using a PC with Intel Core 2 Duo 3.0 GHz Processor, the computation time was around 5 min per color channel. Optimized RMSEs for the red channel were obtained for different calibration models.

Once a calibration model was selected, all three color channels were calibrated with the genetic algorithm search. After that, film doses were calculated for each color channel. To combine the calibration of all three channels, the weighted mean dose was calculated. Channel doses were weighted with the variance estimated as the square of the RMSE of the channel.

2.7 Calibration with fragments

In the literature, the most frequent calibration method uses fragments irradiated with different doses and scanned in different positions over the scanner [5, 8, 15, 16]. This method was also followed in this work to compare it with the new plan-based method proposed.

Lateral correction

To evaluate the nonuniformity of the dosimetric system, five film fragments with different PV levels were digitized at different positions on the scanner. One of the fragments was nonirradiated, while the other four had been previously irradiated with different doses. Every fragment was scanned at 18 positions along the x axis. Two of the fragments were also scanned at 11 positions along the y axis. A 100×100 px ROI with resolution of 150 dpi was measured at the center of the fragment for every position.

Sensitometric curve

A film was divided in 12 fragments which were irradiated with 25, 50, 100, 150, 175, 200, 225, 250, 300, 350, 400 and 500 MU respectively (104 MU corresponded

to a dose of 1 Gy). Before irradiation and 24h after irradiation, each fragment was centered on the scanner and a 100×100 px ROI at 150 dpi was measured at the center of the fragment.

2.8 Comparison of calibration methods

Intralot and interlot variations

Three films from lot A03171101A (lot A) were calibrated using the plan-based method. Only the data from the red channel were used. Calibration parameters obtained for the three films from lot A and for the film previously used for plan-based calibration of lot A04141003BB (lot B) were employed to measure the dose on one of the three films from lot A. The RMSE of the differences between the doses measured with film and the doses calculated by the TPS was computed for each set of calibration parameters.

Verification tests

Seven different cases were tested. They were based on the IAEA TECDOC-1583 [28] tests for commissioning of TPS. They were planned using Eclipse TPS with AAA. Films were posteriorly irradiated in the phantom according to the plans. The geometry of the test cases is described in Table 1.1. The images were digitized before and 24 h after irradiation.

The calculated dose distribution on the plane of the film was exported with resolution 0.59 mm/px. The digitized films were converted to dose according to the previously selected calibration model. Four sets of images were created based on the parameters derived from: fragments using only the red channel, fragments using the three channels, plan-based red channel and plan-based three channels. 2D gamma analysis of the test cases was conducted. The selected criteria for the analysis were 4 % 3 mm excluding points with less than 20 % of the maximum dose. To automate the procedure of dose calculation and gamma analysis, a program was developed in C++.

3 Results and discussion

3.1 Preliminary tests

Effect of surrounding film

Contrary to Richley *et al.* [3], no effect of surrounding film was found in this study. Fig. 1.1 shows mean and standard deviation of the PV of the red channel measured in the nonirradiated fragment as a function of the dose of an abutting fragment, data are scaled so that the y-axis represents relative deviations of the PV with respect to the average of all six measurements. No statistically

Table 1.1: Geometry of the test cases. In all tests a CIRS Thorax phantom (Model 002LFC) was irradiated with an energy of 6 MV in SAD set-up.

Test	Description	Field size (cm^2)	Gantry angle	Collimator angle	Beam modifiers
1	Square 10×10 cm^2	10×10	0	0	
2	Small field	4×4	0	0	
3	Lateral incidence	10×10	90	0	
4	Tangential field	10×15	90	90	
5	Four field box	10×15	0	0	
		8×15	90	0	
		10×15	180	0	
		8×15	270	0	
6	EDW and asymmetric fields A	10×10	300	90	EDW15IN
		20×20	60	90	EDW30OUT
		40×40	180	90	
		10×15	0	90	
7	EDW and asymmetric fields B	10×10	300	90	EDW45IN
		20×20	60	90	EDW60OUT
		5×5	180	90	
		15×15	0	90	

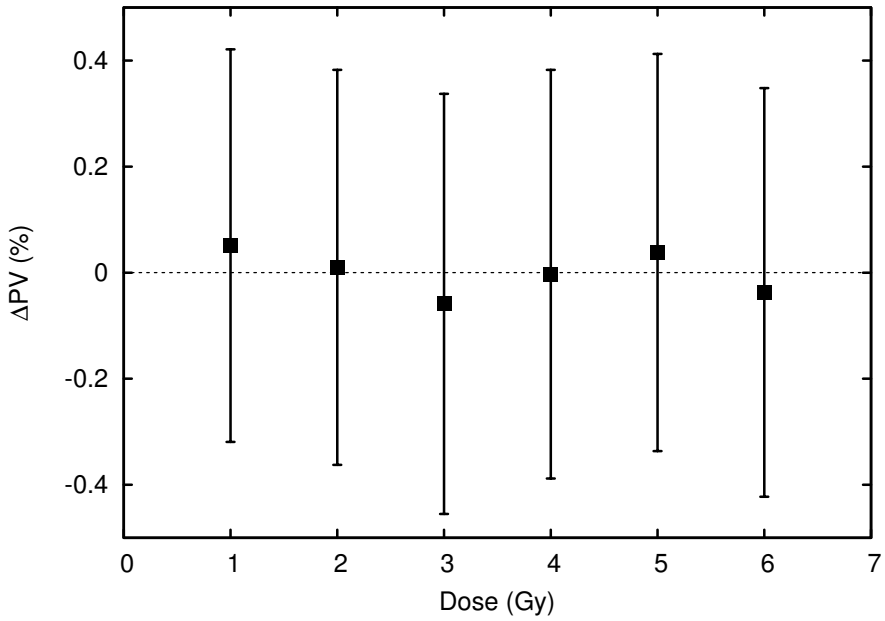


Figure 1.1: Mean and standard deviation of the PV measured in a nonirradiated fragment as a function of the dose of an abutting fragment, data are scaled so that the y-axis represents relative deviations of the PV with respect to the average of all six measurements.

significant (linear) correlation between PV of a fragment and dose of an abutting fragment was found ($p = 0.46$).

A slight effect of surrounding film is to be expected in the immediate vicinity of a stepwise change in dose, since the point spread function of the system cannot be a Dirac delta function. However, if the effect is significant some milimeters away from the step, the digitized image should be blurred. Possible explanations for the effect of surrounding film found by Richley *et al.* could be a problem with the optics of the scanner in reflection mode or a variation in the temperature of the scanner's bed.

Since no effect of surrounding film was found, this opens the possibility of using the same calibration methods in reflection mode as in transmission mode with the dosimetric system formed by Gafchromic EBT2 films and Epson Expression 10000XL scanner.

Film response homogeneity

The contribution of film response inhomogeneity to the uncertainty of the measured dose is known to be substantial [8]. It is reduced if the film is digitized before irradiation and NOD is calculated. However, this only accounts for the

background PV or fog, and not for differences in sensitivity (e.g. due to thickness variation of the active layer).

Fig. 1.2 presents mean PVs of the red channel measured on different fragments of a film nonirradiated and irradiated with 2 Gy. One-way analysis of variance (ANOVA) found statistically significant differences between PVs measured on different fragments for both nonirradiated ($p < 0.001$) and irradiated films ($p < 0.001$). Therefore, film response inhomogeneity is significant in Gafchromic EBT2 films. To check if this inhomogeneity is only due to fog differences, NOD was calculated for each pixel in the film. One-way ANOVA found statistically significant differences between fragments ($p < 0.001$). Maximum differences of 0.01 NOD between mean NODs of fragments irradiated with 2 Gy were observed with Tukey's HSD test.

Hence, film response inhomogeneity in Gafchromic EBT2 films increases the uncertainty of the dosimetry affecting the calibration and the final results. The use of NOD alone cannot correct this defect. More advanced procedures to correct film response inhomogeneity are necessary.

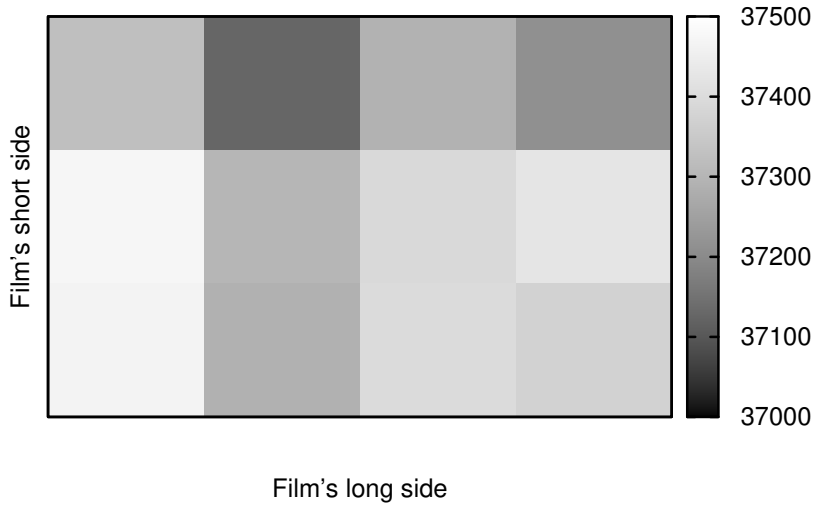
3.2 Calibration with plan-based method

Model selection:

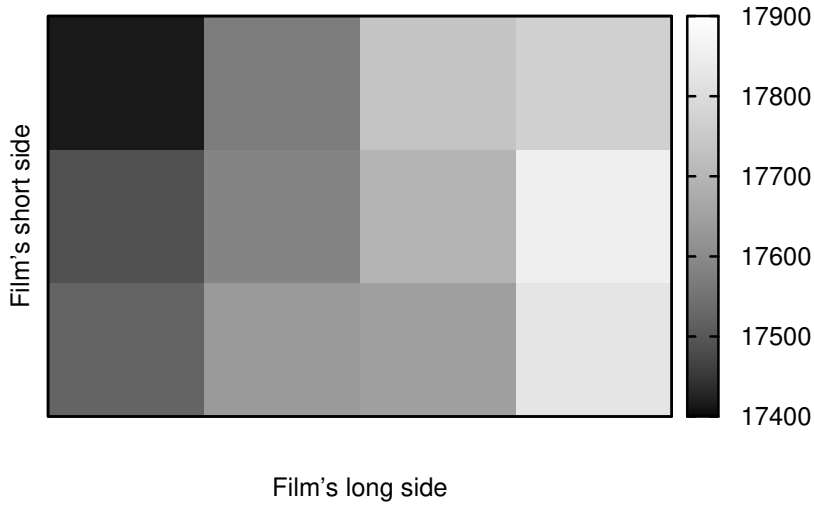
Optimized RMSEs for the red channel obtained with different calibration models are shown in Table 1.2. With respect to lateral correction, type I and type IV lateral correction functions obtained the lowest RMSE of 4.4 cGy, whereas type III obtained 4.5 cGy and type II 9.9 cGy. Considering that the sample size contains a matrix of 1100×1300 data points, these differences suggest substantial evidence for the superiority of models type I and type IV according to the Akaike information criterion (AIC). Since type IV includes type I functions, this supports the conclusions of Saur *et al.*, [25] who found a better agreement with absolute corrections independent of dose. Therefore, type I lateral correction functions were selected.

With respect to sensitometric curves, gamma-distributed single hit model and polynomial fits of orders 4 and 5 had the same RMSE. The complexity of the model was considered negligible comparing the number of parameters with the number of data points to fit. However less complexity facilitates the optimization. Gamma-distributed single hit curves exhibited less robustness during the optimization, *i.e.* small changes of the algorithm affected considerably the minimum RMSE found. Therefore, polynomial fit of order 4 was selected as the sensitometric curve.

It is important to note that even though genetic algorithm is a well established optimization method that effectively escapes from local minima, its results cannot be taken as global minima. As a consequence, it should not be concluded that



(a)



(b)

Figure 1.2: Mean PVs measured on different fragments of a film (a) nonirradiated and (b) irradiated with 2 Gy. Every fragment was centered in the scan and a 100×100 px ROI was measured at the center of the fragment.

Table 1.2: Optimized RMSE for the red channel obtained with different calibration models.

Lateral correction	Sensitometric curve	RMSE (cGy)
Type I	Polynomial order 4	4.4
Type II	Polynomial order 4	9.9
Type III	Polynomial order 4	4.5
Type IV	Polynomial order 4	4.4
Type I	Polynomial order 3	4.5
Type I	Polynomial order 5	4.4
Type I	Single hit model	4.4

the model selected is the best of the models analyzed, although this hypothesis becomes more plausible.

3.3 Calibration with fragments

Lateral correction

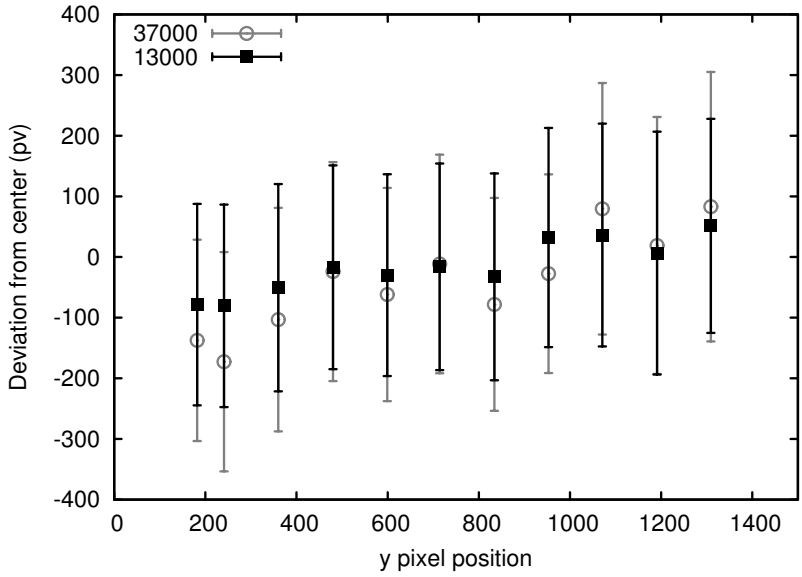
Deviation from the value in the center of the scanner for different PV levels in the red channel, as a function of the pixel position along the y axis and along the x axis is illustrated in Fig. 1.3. One-way analysis of variance (ANOVA) found statistically significant differences between PVs as a function of the pixel position along the y axis ($p < 0.001$) and along the x axis ($p < 0.001$). A linear regression of the measurements along the y axis obtained maximum differences of 0.002 NOD for doses around 4 Gy. Hence, along the y axis the lateral correction was considered to be negligible. Along the x axis the nonuniform response was fitted according to the model selected. The lateral correction function fitted from the fragments is shown.

Sensitometric curve

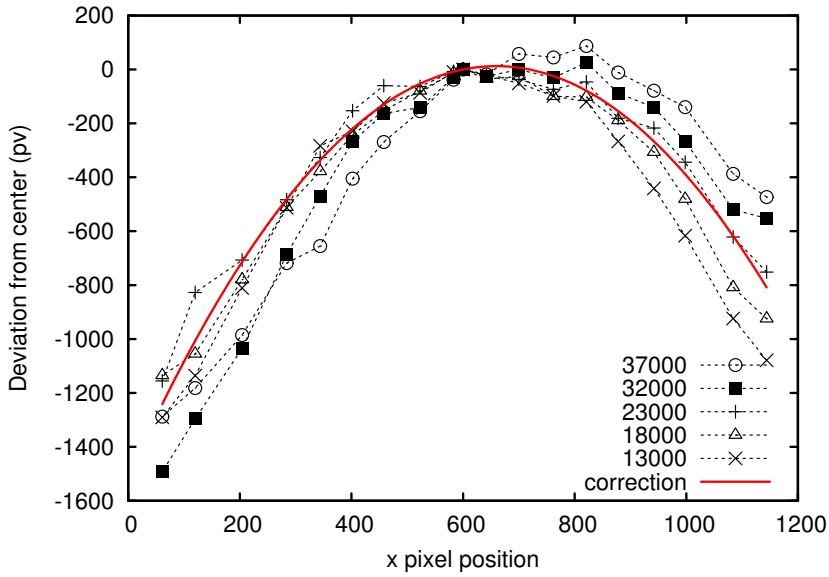
The sensitometric curve fitted from the fragments for the red channel, according to the model selected, is shown in Fig. 1.4. It is compared with the sensitometric curve obtained for the red channel with the plan-based method. The sensitometric curve obtained with the plan-based method is extrapolated for doses greater than 400 cGy.

3.4 Comparison of calibration methods

Table 1.3 compares RMSEs obtained with different calibration methods and calculated on different films. In the first part, plan-based calibration methods are compared. According to the AIC, the RMSEs present substantial evidence for



(a)



(b)

Figure 1.3: Deviation from the value in the center of the scanner for different pixel value levels, as a function of the pixel position (a) along the y axis (*i.e.*, perpendicular to the CCD array) (b) along the x axis (*i.e.*, parallel to the CCD array); the fixed line represents the lateral correction fitted from the fragments.

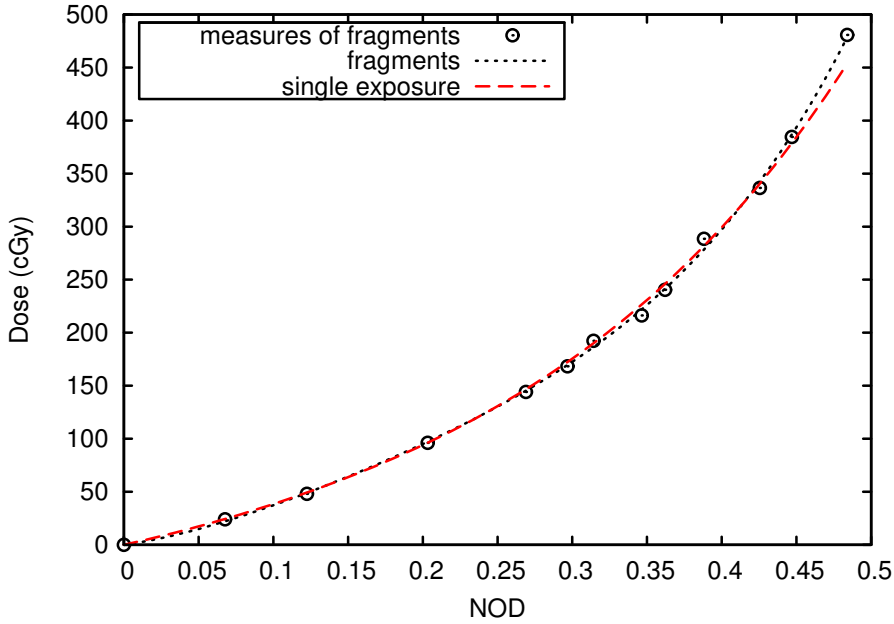


Figure 1.4: Sensitometric curves obtained with fragments (dotted) and with plan-based method (dashed). The sensitometric curve obtained with plan-based method is extrapolated for doses greater than 400 cGy.

the superiority of three-channel plan-based dosimetry compared to one-channel plan-based dosimetry. This is also the case for calibration with fragments, as it is shown in the second part. It has to be noted that pixel measures were aggregated in every ROI for the calculation of RMSEs on the sensitometric curve from fragments. These RMSEs would increase if pixel measures were disaggregated, and even more so if not only the sensitometric curve but also the residuals of the lateral correction's fit were considered. To calculate the weighted mean dose, channel doses were weighted with the variance estimated as the square of the RMSE of the channel. In the third part, calibration parameters obtained from calibration with fragments were employed to calculate RMSE on the film used for plan-based calibration. Calibration with fragments showed worse RMSEs than plan-based calibration. This outcome could be partially explained by the fact that the calibration with fragments is less robust to film inhomogeneities than plan-based calibration: plan-based calibration can use every pixel of the film, whereas calibration with fragments only uses a limited number of pixels which share coordinates. Calibration with fragments also needs a more complex measuring process. In addition, film-to-film variations are an important source of uncertainty too, as it is showed in Table 1.4.

Table 1.3: RMSEs obtained with different calibration methods and calculated on different films. Pixel measures were aggregated in every ROI for the calculation of RMSEs on the sensitometric curve from fragments.

Calibration method	Sample	Color channel	RMSE (cGy)
Plan-based	Plan-based film	Red	4.4
		Green	3.9
		Blue	10.3
		3 channel weighted mean	3.8
Fragments	Sensitometric curve from fragments	Red	5.0
		Green	4.5
		Blue	13.4
		3 channel weighted mean	4.4
Fragments	Plan-based film	3 channel weighted mean	7.3

Table 1.4: Intralot and interlot variations. RMSEs were calculated on film 1 (Lot A) with calibration parameters derived from different films of two different lots.

Test	Calibration on film	RMSE (cGy)
Intralot variation	Film 1 (Lot A)	5.8
	Film 2 (Lot A)	7.3
	Film 3 (Lot A)	6.3
Interlot variation	Film 1 (Lot B)	16.0

Intralot and interlot variations

Table 1.4 shows the influence of intralot (film-to-film) and interlot variations on RMSEs. The RMSEs were calculated on film 1 (Lot A) with calibration parameters derived from different films of two different lots (A and B). Film-to-film variations were found not negligible. This implies that to decrease the uncertainty of the sensitometric curve could be necessary not only to increase the number of dose levels, but also to calibrate several films simultaneously. Considering this, a possible improvement for the plan-based calibration method presented in this work would be to optimize simultaneously several films irradiated according to one or more reference plans.

Verification tests

In Fig. 1.5, histograms of gamma (4% 3mm) values obtained with different calibration methods are plotted. Gamma values were calculated for all the points in the test cases excluding points with less than 20 % of the maximum dose of the test. In Table 1.5, the percentage of points with $\gamma_{<1}$ and the γ_{mean} calculated with the compared calibration methods are presented for each test case, as well as the average values calculated for all the points in the test cases.

Three-channel calibration methods showed the best agreement with the TPS, followed by red channel plan-based; red channel calibration with fragments showed the worst agreement. The average number of points with $\gamma_{<1}$ was 90.7 % with red channel fragments, 93.3 % with red channel plan-based and 96.6 % with both three-channel plan-based calibration and three-channel calibration with fragments. The average γ_{mean} was 0.49 with red channel calibration with fragments, 0.46 with red channel plan-based, 0.39 with three-channel plan-based calibration and 0.38 with three-channel calibration with fragments.

The plan-based calibration method obtained comparable results to the well-established calibration method with fragments. It indicates that the plan-based calibration method is a feasible alternative to the calibration with fragments. However, possible film-to-film variations or systematic inaccuracies of the TPS

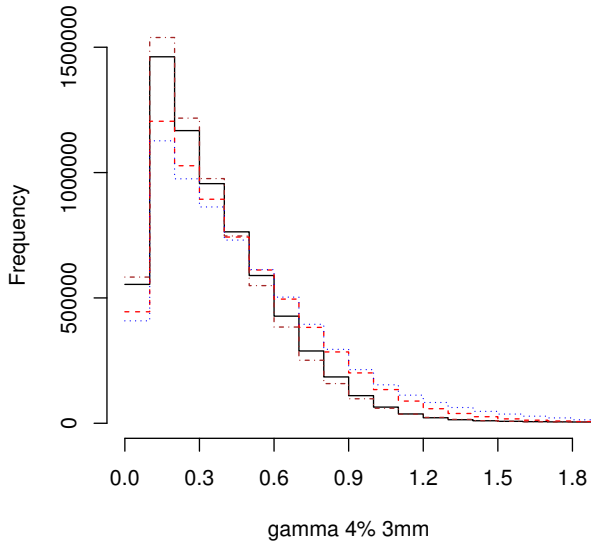


Figure 1.5: Histograms of gamma (4% 3mm) values obtained with different calibration methods: fragments red channel (dotted), fragments 3 channels (dotdash), plan-based red channel (dashed) and plan-based 3 channels (solid). Gamma values were calculated for all the points in the test cases excluding points with less than 20 % of the maximum dose of the test.

cannot be excluded. The plan-based calibration method offers increased robustness against film response inhomogeneities (since it can use every pixel of the film) and reduces considerably the time required for calibration (in this work, calibration time was reduced from several hours for calibration with fragments to minutes for plan-based calibration).

4 Conclusions

Radiochromic dosimetry in reflection mode using Gafchromic EBT2 films was found to be a viable alternative to transmission mode. In this study, no effect of surrounding film was found with the dosimetric system formed by Gafchromic EBT2 films and Epson Expression 10000XL scanner. This opens the possibility of using the same calibration methods in reflection mode as in transmission mode.

Film response inhomogeneity with EBT2 films was found to be important, not only due to differences in the fog but also to differences in sensitivity. The use of NOD alone cannot correct this defect. More advanced procedures to correct film response inhomogeneity are necessary.

Table 1.5: 2D gamma analysis (4% 3mm) of the test cases with different calibration methods.

	$\gamma_{<1}$ (%)				γ_{mean}			
	Fragments		Plan-based		Fragments		Plan-based	
Test	Red	Three channels	Red	Three channels	Red	Three channels	Red	Three channels
1	92.4	94.2	88.3	92.6	0.46	0.41	0.51	0.47
2	90.3	88.6	89.2	88.8	0.47	0.51	0.50	0.50
3	97.2	98.7	98.6	98.8	0.40	0.34	0.34	0.32
4	93.5	95.9	97.1	97.0	0.46	0.41	0.39	0.38
5	89.1	97.2	94.7	97.9	0.46	0.32	0.42	0.34
6	89.4	96.9	95.6	97.8	0.53	0.38	0.44	0.36
7	87.3	97.9	87.5	96.7	0.58	0.37	0.59	0.41
Average result	90.7	96.6	93.3	96.6	0.49	0.38	0.46	0.39

To offer increased robustness against film response inhomogeneities, a novel plan-based calibration method was developed. Plan-based calibration is a single exposure method that can use every pixel and each of the three color channels of the scanned film to obtain the parameters of the calibration model. Plan-based calibration uses a reference plan (in this study a field with 60° EDW) calculated by the TPS. The accuracy of the calculation was carefully commissioned. A film was irradiated following the reference plan. Least squares fitting was employed to find the parameters of the calibration model that minimized the differences between TPS and the doses measured with film. The complexity of the optimization made it necessary to use a genetic algorithm search. The calibration model (lateral correction and sensitometric curve models) was selected based on a maximum likelihood analysis. The best results for lateral corrections were obtained using absolute corrections independent of dose. With respect to sensitometric curves, an empirical polynomial fit of order 4 was found to obtain results equivalent to a gamma-distributed single hit model based on physical assumptions. Film-to-film variations were found to be not negligible, thus a possible improvement for the plan-based calibration method presented in this work would be to optimize simultaneously several films irradiated according to one or more reference plans.

Three-channel dosimetry was calculated using the weighted mean dose of the color channels. The variances of the calibration, estimated as the square of the RMSE for each channel, were used as weights. Three-channel dosimetry was found to be substantially superior to red-channel dosimetry.

Plan-based calibration method was found to be a feasible alternative to the well-established calibration method with fragments. This novel method offers increased robustness against film response inhomogeneities (since it can use every pixel of the film) and reduces considerably the time required for calibration (in this work, calibration time was reduced from several hours for calibration with fragments to minutes for plan-based calibration).

5 Acknowledgments

The authors would like to thank Sebastià Agramunt, Primož Peterlin, Juanjo Rovira and Attila Šarvari for many useful discussions and contributions to this work.

References

[1] H. Alva, H. Mercado-Uribe, M. Rodriguez-Villafuerte, and M. E. Brandan. “The use of a reflective scanner to study radiochromic film response”. *Physics in Medicine and Biology* 47.16 (2002), p. 2925.

[2] J. Kalef-Ezra and K. Karava. "Radiochromic film dosimetry: Reflection vs transmission scanning". *Medical Physics* 35.6 (2008), pp. 2308–2311.

[3] L. Richley, A. C. John, H. Coomber, and S. Fletcher. "Evaluation and optimization of the new EBT2 radiochromic film dosimetry system for patient dose verification in radiotherapy". *Physics in Medicine and Biology* 55.9 (2010), p. 26012617.

[4] C. Fiandra, U. Ricardi, R. Ragona, S. Anglesio, F. R. Giglioli, E. Calamia, and F. Lucio. "Clinical use of EBT model Gafchromic film in radiotherapy". *Medical Physics* 33.11 (2006), pp. 4314–4319.

[5] M. Fuss, E. Sturtewagen, C. D. Wagter, and D. Georg. "Dosimetric characterization of GafChromic EBT film and its implication on film dosimetry quality assurance". *Physics in Medicine and Biology* 52.14 (2007), pp. 4211–4225.

[6] L. Paelinck, W. D. Neve, and C. D. Wagter. "Precautions and strategies in using a commercial flatbed scanner for radiochromic film dosimetry". *Physics in Medicine and Biology* 52.1 (2007), pp. 231–242.

[7] L. J. van Battum, D. Hoffmans, H. Piersma, and S. Heukelom. "Accurate dosimetry with GafChromic EBT film of a 6 MV photon beam in water: What level is achievable?" *Medical Physics* 35.2 (2008), pp. 704–716.

[8] M. Martišíková, B. Ackermann, and O. Jäkel. "Analysis of uncertainties in Gafchromic EBT film dosimetry of photon beams". *Physics in Medicine and Biology* 53.24 (2008), pp. 7013–7027.

[9] L. Menegotti, A. Delana, and A. Martignano. "Radiochromic film dosimetry with flatbed scanners: A fast and accurate method for dose calibration and uniformity correction with single film exposure". *Medical Physics* 35.7 (2008), pp. 3078–3085.

[10] B. Arjomandy, R. Tailor, A. Anand, N. Sahoo, M. Gillin, K. Prado, and M. Vicic. "Energy dependence and dose response of Gafchromic EBT2 film over a wide range of photon, electron, and proton beam energies". *Medical Physics* 37.5 (2010), pp. 1942–1947.

[11] S. Devic, S. Aldelaijan, H. Mohammed, N. Tomic, L.-H. Liang, F. DeBlois, and J. Seuntjens. "Absorption spectra time evolution of EBT-2 model Gafchromic film". *Medical Physics* 37.5 (2010), pp. 2207–2214.

[12] C. Andrés, A. del Castillo, R. Tortosa, D. Alonso, and R. Barquero. "A comprehensive study of the Gafchromic EBT2 radiochromic film. A comparison with EBT". *Medical Physics* 37.12 (2010), pp. 6271–6278.

[13] B. Hartmann, M. Martišíková, and O. Jäkel. "Technical Note: Homogeneity of Gafchromic EBT2 film". *Medical Physics* 37.4 (2010), pp. 1753–1756.

[14] T. Kairn, T. Aland, and J. Kenny. “Local heterogeneities in early batches of EBT2 film: a suggested solution”. *Physics in Medicine and Biology* 55.15 (2010), p. L37.

[15] H. Bouchard, F. Lacroix, G. Beaudoin, J.-F. Carrier, and I. Kawrakow. “On the characterization and uncertainty analysis of radiochromic film dosimetry”. *Medical Physics* 36.6 (2009), pp. 1931–1946.

[16] S. Devic, J. Seuntjens, E. Sham, E. B. Podgorsak, C. R. Schmidtlein, A. S. Kirov, and C. G. Soares. “Precise radiochromic film dosimetry using a flat-bed document scanner”. *Medical Physics* 32.7 (2005), pp. 2245–2253.

[17] A. Micke, D. F. Lewis, and X. Yu. “Multichannel film dosimetry with nonuniformity correction”. *Medical Physics* 38.5 (2011), pp. 2523–2534.

[18] T. J. McCaw, J. A. Micka, and L. A. DeWerd. “Characterizing the marker-dye correction for Gafchromic® EBT2 film: a comparison of three analysis methods”. *Medical Physics* 38.10 (2011), pp. 5771–5777.

[19] R. R. Mayer, F. Ma, Y. Chen, R. I. Miller, A. Belard, J. McDonough, and J. J. O’Connell. “Enhanced dosimetry procedures and assessment for EBT2 radiochromic film”. *Medical Physics* 39.4 (2012), pp. 2147–2155.

[20] A. Niroomand-Rad, C. R. Blackwell, B. M. Coursey, K. P. Gall, J. M. Galvin, W. L. McLaughlin, A. S. Meigooni, R. Nath, J. E. Rodgers, and C. G. Soares. “Radiochromic film dosimetry: Recommendations of AAPM Radiation Therapy Committee Task Group 55”. *Medical Physics* 25.11 (1998), pp. 2093–2115.

[21] T. Künzler, I. Fotina, M. Stock, and D. Georg. “Experimental verification of a commercial Monte Carlo-based dose calculation module for high-energy photon beams”. *Physics in Medicine and Biology* 54.24 (2009), pp. 7363–7377.

[22] H. Ohuchi. “High sensitivity radiochromic film dosimetry using an optical common-mode rejection and a reflective-mode flatbed color scanner”. *Medical Physics* 34.11 (2007), pp. 4207–4212.

[23] S. Devic, Y.-Z. Wang, N. Tomic, and E. B. Podgorsak. “Sensitivity of linear CCD array based film scanners used for film dosimetry”. *Medical Physics* 33.11 (2006), pp. 3993–3996.

[24] B. D. Lynch, J. Kozelka, M. K. Ranade, J. G. Li, W. E. Simon, and J. F. Dempsey. “Important considerations for radiochromic film dosimetry with flatbed CCD scanners and EBT GAFCHROMIC film”. *Medical Physics* 33.12 (2006), pp. 4551–4556.

[25] S. Saur and J. Frengen. “GafChromic EBT film dosimetry with flatbed CCD scanner: A novel background correction method and full dose uncertainty analysis”. *Medical Physics* 35.7 (2008), pp. 3094–3101.

[26] F. del Moral, J. A. Vázquez, J. J. Ferrero, P. Willisch, R. D. Ramírez, A. Teijeiro, A. L. Medina, B. Andrade, J. Vázquez, F. Salvador, D. Medal, M. Salgado, and V. Muñoz. "From the limits of the classical model of sensitometric curves to a realistic model based on the percolation theory for GafChromic EBT films". *Medical Physics* 36.9 (2009), pp. 4015–4026.

[27] D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. 1st. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1989. ISBN: 0201157675.

[28] I. A. E. Agency. *Commissioning of Radiotherapy Treatment Planning Systems: Testing for Typical External beam Treatment Techniques. IAEA-TECDOC-1583*. Tech. rep. Vienna, Austria: IAEA, 2008.

Paper 2

On multichannel film dosimetry with channel-independent perturbations

I. Méndez, P. Peterlin, R. Hudej, A. Strojnik and B. Casar

Medical physics 41.10 (2014), 011705

Paper 2. On multichannel film dosimetry with channel-independent perturbations

I. Méndez, P. Peterlin, R. Hudej, A. Strojnik and B. Casar

Abstract

Purpose: Different multichannel methods for film dosimetry have been proposed in the literature. Two of them are the weighted mean method and the method put forth by Micke *et al* and Mayer *et al*. The purpose of this work was to compare their results and to develop a generalized channel-independent perturbations framework in which both methods enter as special cases.

Methods: Four models of channel-independent perturbations were compared: weighted mean, Micke-Mayer method, uniform distribution and truncated normal distribution. A closed-form formula to calculate film doses and the associated Type B uncertainty for all four models was deduced.

To evaluate the models, film dose distributions were compared with planned and measured dose distributions. At the same time, several elements of the dosimetry process were compared: film type EBT2 versus EBT3, different waiting-time windows, reflection mode versus transmission mode scanning, and planned versus measured dose distribution for film calibration and for γ-index analysis.

The methods and the models described in this study are publicly accessible through IRISEU. Alpha 1.1 (http://www.iriseu.com). IRISEU. is a cloud computing web application for calibration and dosimetry of radiochromic films.

Results: The truncated normal distribution model provided the best agreement between film and reference doses, both for calibration and γ-index verification, and proved itself superior to both the weighted mean model, which neglects correlations between the channels, and the Micke-Mayer model, whose accuracy depends on the properties of the sensitometric curves.

With respect to the selection of dosimetry protocol, no significant differences were found between transmission and reflection mode scanning, between 75 ± 5 min and 20 ± 1 h waiting-time windows or between employing EBT2 or EBT3 films. Significantly better results were obtained when a measured dose distribution was used instead of a planned one as reference for the calibration, and when a planned dose distribution was used instead of a measured one as evaluation for the γ-analysis.

Conclusions: The truncated normal distribution model of channel independent perturbations was found superior to the other three models under comparison and we propose its use for multichannel dosimetry.

1 Introduction

Radiochromic film dosimetry with flatbed scanners and Gafchromic films (Ashland Inc., Wayne, NJ) has been extensively studied in the literature [1–7]. High spatial resolution, near water equivalence[8, 9] and weak energy dependence [10–14] make radiochromic films convenient for measurements whenever sharp dose gradients, tissue heterogeneities or charged particle disequilibrium conditions exist. This opens up a wide range of applications for radiochromic films in the field of radiotherapy.

Recently, different multichannel dosimetry methods have been proposed to take into account the information conveyed by all three color channels delivered by the scanner. Micke *et al* [15] proposed the use of channel-independent perturbations to compensate for variations in the thickness of the active layer, artifacts, nonuniform response of the scanner or other disturbances. They found a substantial gain in dosimetric accuracy using this method. Van Hoof *et al* [16] found that this method performs at least as well as the conventional single-red-channel dosimetry. Mayer *et al* [17] derived a closed-form solution to obtain the dose employing channel-independent perturbations. They also compared different single, dual and triple channel methods, and found better agreement between planned and calculated dose distributions using the average dose of all three channels in comparison to using the channel-independent perturbations method. In an earlier article [18], our group suggested calculating the film dose as the weighted mean dose of all three channels. For each channel, the inverse of the mean square error obtained during the film calibration was used as weight. With this method, triple-channel dosimetry was found to be substantially superior to single-red-channel dosimetry.

The purpose of this work is to compare both weighted mean and Micke-Mayer methods, considering them as special cases of a more general channel-independent perturbations method. Deficiencies and important problems associated with both methods will be explained. To overcome these problems, an improved multichannel film dosimetry method will be introduced. Its performance against the other methods will be verified by comparing film dose distributions with planned as well as with measured dose distributions. In addition, other elements of the dosimetry process will be compared: film types [19], scanning modes [20], scanning waiting-time windows [21] and choices of reference dose distribution.

2 Methods and materials

2.1 Channel-independent perturbations

Channel-independent perturbations are obtained by applying a first order Taylor expansion to the dose due to a small perturbation:

$$\begin{cases} D(r) = D_R(r) + \dot{D}_R(r)\Delta(r) + \epsilon_R(r) \\ D(r) = D_G(r) + \dot{D}_G(r)\Delta(r) + \epsilon_G(r) \\ D(r) = D_B(r) + \dot{D}_B(r)\Delta(r) + \epsilon_B(r) \end{cases}, \tag{2.1}$$

Micke *et al* [15] derived the dose from the optical density (OD) of the irradiated film. Mayer *et al* [17] used pixel values directly. In this study, better results were found in preliminary tests using net optical density [1] (NOD) in comparison to using OD. Therefore, the channel-independent perturbation consists of a change in NOD and is represented by $\Delta(r)$. $D(r)$ represents the dose absorbed by the film at point r. D_k is the absolute dose measured by the channel k, *i.e.*, red (R), green (G) or blue (B) channel, when no disturbance is present, and it is calculated directly from the calibration model. In this study, the calibration model includes the lateral correction [2–5, 22–26] and the sensitometric curve. $\dot{D}_k(r)$ is the first derivative of the dose, with respect to the NOD, at point r. Finally, $\epsilon_k(r)$ is an error term accounting for the difference between the dose absorbed by the film and the dose measured in the channel k after correction by the perturbation.

Both for reflection [27] and transmission mode scanning, the NOD, denoted by z, was defined as

$$z \coloneqq \log_{10} \frac{v_{\text{nonirr}}}{v_{\text{irr}}}, \tag{2.2}$$

where v_{nonirr} and v_{irr} represent pixel values of nonirradiated and irradiated films, respectively, after applying lateral corrections. Our previous results [18] found better fit when lateral corrections are absolute corrections independent of dose, and sensitometric curves are polynomial fits of order four. Hence, lateral corrections were calculated as

$$v_k = a_{k_1}(x - x_c) + a_{k_2}(x - x_c)^2 + \hat{v_k}, \tag{2.3}$$

where $\hat{v_k}$ represents uncorrected pixel values, the x axis is parallel to the CCD array, x_c is the x coordinate of the center of the scanner, v_k represents corrected pixel values, and a_k are fitting parameters. Sensitometric curves followed

$$D_k = \sum_{j=1}^{4} b_{k_j} z_k^j, \tag{2.4}$$

and $\dot{D}_k$ was

$$\dot{D}_k = \sum_{j=1}^{4} j b_{k_j} z_k^{j-1}, \tag{2.5}$$

where b_k are fitting parameters.

2.2 Solving the equation system

The values of $\Delta(r)$ and $\epsilon_k(r)$ for $k = R, G, B$ in Eq.(2.1) are unknown. As a result, the absorbed dose $D(r)$ cannot be obtained directly. However, one can examine different probability density functions (pdf) for $\Delta(r)$ and $\epsilon_k(r)$ and, if $D(r)$ is known, analyze how well these models reproduce the absorbed dose distribution.

Probability density function of the dose

Given the pdfs of Δ, symbolized by $f(\Delta)$, and of each ϵ_k, symbolized by $g_k(\epsilon_k)$, the joint pdf of D, symbolized by $P(D)$, is:

$$P(D) = \int f(\Delta) \prod_k g_k(D - D_k - \dot{D}_k \Delta)\, \mathrm{d}\Delta, \tag{2.6}$$

taking into account that Δ and ϵ_k are not independent from each other:

$$\epsilon_k = D - D_k - \dot{D}_k \Delta. \tag{2.7}$$

Let us consider that the error terms are distributed normally with zero mean and σ_k^2 variance:

$$g_k(\epsilon_k) = \mathcal{N}(0, \sigma_k^2) \tag{2.8}$$

The joint pdf of D becomes:

$$P(D) = \int f(\Delta) \prod_k \frac{1}{\sigma_k \sqrt{2\pi}}\, \mathrm{e}^{-\frac{1}{2}\left(\frac{D - D_k - \dot{D}_k \Delta}{\sigma_k}\right)^2}\, \mathrm{d}\Delta. \tag{2.9}$$

Three different models for $f(\Delta)$ will be considered:
a) Normally distributed perturbation (*i.e.*, $f(\Delta) = \mathcal{N}(0, \sigma_\Delta^2)$):

$$P(D) = \frac{1}{(2\pi)^{\frac{n}{2}} \sigma_\Delta \prod_{k=1}^n \sigma_k}\, \frac{1}{\sqrt{A}}\, \mathrm{e}^{-\frac{1}{2}\left(C - \frac{B^2}{4A}\right)}, \tag{2.10}$$

where n represents the number of color channels (*i.e.*, $n = 3$) and

$$A = \frac{1}{\sigma_\Delta^2} + \sum_k \left(\frac{\dot{D}_k}{\sigma_k}\right)^2 \tag{2.11}$$

$$B = -2 \sum_k \frac{(D - D_k)\dot{D}_k}{\sigma_k^2} \tag{2.12}$$

$$C = \sum_k \left(\frac{D - D_k}{\sigma_k}\right)^2. \tag{2.13}$$

b) Truncated normal distribution with $\Delta \in (-\theta, \theta)$:

$$P(D) \propto \mathrm{e}^{-\frac{1}{2}\left(C-\frac{B^2}{4A}\right)} \left(\mathrm{erf}\left(\frac{\theta + \frac{B}{2A}}{\sqrt{\frac{2}{A}}} \right) - \mathrm{erf}\left(\frac{-\theta + \frac{B}{2A}}{\sqrt{\frac{2}{A}}} \right) \right), \tag{2.14}$$

excluding a normalizing term independent of D.

c) Uniform distribution with $\Delta \in (-\theta, \theta)$: is a special case of Eq.(2.14) where σ_Δ goes to infinity, therefore $A = \sum_k \left(\frac{\dot{D}_k}{\sigma_k}\right)^2$.

Dose calculation

The most likely value of the absorbed dose D, symbolized by d, is the one that maximizes $P(D)$. The exponential term in $P(D)$, $P(D) \propto \mathrm{e}^{-\frac{1}{2}\left(C-\frac{B^2}{4A}\right)}$, can be expressed in terms of D as a gaussian function:

$$P(D) \propto \mathrm{e}^{-\frac{1}{2}\left(\frac{D-\mu_D}{\sigma_D}\right)^2}, \tag{2.15}$$

where

$$\mu_D = d = \frac{A\beta - \gamma\delta}{A\alpha - \gamma^2} \tag{2.16}$$

and

$$\sigma_D = \sqrt{\frac{A}{A\alpha - \gamma^2}}, \tag{2.17}$$

A is defined in Eq.(2.11) and

$$\alpha = \sum_k \frac{1}{\sigma_k^2} \tag{2.18}$$

$$\beta = \sum_k \frac{D_k}{\sigma_k^2} \tag{2.19}$$

$$\gamma = \sum_k \frac{\dot{D}_k}{\sigma_k^2} \tag{2.20}$$

$$\delta = \sum_k \frac{D_k \dot{D}_k}{\sigma_k^2}. \tag{2.21}$$

Eq.(2.16) and Eq.(2.17) can be considered, respectively, as the estimated absolute dose and its type B uncertainty [28]. This result is exact for normally distributed perturbations and an approximation for truncated normal and uniform distributions.

Table 2.1: Models of channel-independent perturbations under comparison.

Model	Abbreviation	Assumptions
Weighted mean	WM	$\Delta(r) = 0$
Micke-Mayer method	MM	$f(\Delta)$ uniform distribution, σ_k are equal
Uniform distribution	UD	$f(\Delta)$ uniform distribution
Truncated normal distribution	TD	$f(\Delta)$ truncated normal distribution

2.3 Models of channel-independent perturbations under comparison

Four models of channel-independent perturbations were compared: weighted mean (WM), Micke-Mayer (MM) method, uniform distribution (UD) and truncated normal distribution (TN). They are summarized in Table 2.1.

The weighted mean method is a limit case of Eq.(2.1) in which $\Delta(r) = 0$. Thus, all three channels are independent of each other, which implies that correlations between channels are neglected.

The method employed by Micke *et al* [15] and Mayer *et al* [17] is a special case of Eq.(2.1) where all σ_k are equal and $f(\Delta)$ is uniformly distributed. Under these premises, Eq.(2.16) becomes:

$$d = \frac{A\beta - \gamma\delta}{A\alpha - \gamma^2} = \frac{(\sum_{k=1}^{n} \dot{D}_k)(\sum_{k=1}^{n} D_k \dot{D}_k) - (\sum_{k=1}^{n} \dot{D}_k^2)(\sum_{k=1}^{n} D_k)}{(\sum_{k=1}^{n} \dot{D}_k)^2 - n(\sum_{k=1}^{n} \dot{D}_k^2)}, \tag{2.22}$$

which coincides with the closed-form solution derived by Mayer *et al* [17]. The uncertainty in the dose associated to this model becomes:

$$\sigma_D = \sqrt{\frac{A}{A\alpha - \gamma^2}} = \sigma_k \sqrt{\frac{\sum_{k=1}^{n} \dot{D}_k^2}{n \sum_{k=1}^{n} \dot{D}_k^2 - (\sum_{k=1}^{n} \dot{D}_k)^2}} \tag{2.23}$$

The uniform distribution model is a more general and realistic model for the perturbation than the MM one. In this case, $f(\Delta)$ is uniformly distributed but the σ_k can differ.

Finally, the truncated normal distribution model considers that $f(\Delta)$ follows a truncated normal distribution. The WM model is a limit case and the UD and MM models are particular cases of this model.

2.4 Measurement protocol

Ten 8 inch × 10 inch EBT2 films from lot A03171101A and seventeen EBT3 films from lot A05151201 were employed. They were handled following recommendations from the AAPM TG-55 report [9].

Films were scanned with an Epson Expression 10000XL flatbed scanner (Seiko Epson Corporation, Nagano, Japan) using Epson Scan v.3.0 software. Images were acquired in 48-bit RGB mode (16 bit per channel), the resolution was 72 dpi (0.35 mm/px) and the image processing tools were turned off.

Before acquisitions, the scanner was warmed up for at least 30 min. After the warm-up, and whenever long interruptions occurred, five empty scans were taken to stabilize the temperature of the scanner lamp. Films were centered on the scanner with a black opaque cardboard frame and scanned in portrait orientation. Five consecutive scans were made for each film. To avoid the warm-up effect of the lamp due to multiple scans [3, 5] the first scan was discarded and the resulting image was the average of the remaining four.

Films were scanned before irradiation both in reflection and in transmission mode. After irradiation, two waiting-time windows were studied: films were first scanned after 75 ± 5 min in transmission mode, and again after 20 ± 1 h both in reflection and transmission mode.

Irradiation was delivered with a 6 MV photon beam from a Novalis Tx accelerator (Varian, Palo Alto, CA, USA). Three different phantoms were used: CIRS Thorax Phantom (Model 002LFC, Computerized Imaging Reference Systems Inc. Norfolk, VA, USA), CIRS Pelvic Phantom (Model 002PRA) and IBA MatriXX Evolution MULTICube (IBA Dosimetry GmbH, Germany). Source-axis distance (SAD) setup was used for all three phantoms. To avoid the films lying in the beam axis plane [29], the films were placed at an offset of 1.5 cm from the beam axis in the CIRS Thorax Phantom and of 1.3 cm in the CIRS Pelvic Phantom. The IBA MatriXX Evolution MULTICube was used jointly with the IBA MatriXX Evolution ionization chamber array, which measured the dose distribution delivered. The film was situated atop the detector.

The absolute dose distributions in the plane of the film were calculated with Eclipse v.10.0 (Varian Medical Systems, Palo Alto, CA, USA) treatment planning system (TPS) using the anisotropic analytical algorithm (AAA). The planned dose distributions were exported to dose matrices with a resolution of 0.49 mm/px. The dose values were scaled to correct for the daily output of the linac. Whenever MatriXX Evolution was used, the dose distribution was simultaneously measured. The dose values were scaled with a constant factor to correct for the distance (which was 3.5 mm) between the film and the plane at the effective depth of measurement. The MatriXX 2D array has a resolution of 7.62 mm/px. Planned and measured dose distributions were bicubically interpolated to the resolution of the scan and registered with the film.

Film scans, planned dose ditributions and measured dose distributions were uploaded and processed with IRISEU. Alpha 1.1 (http://www.iriseu.com). IRISEU. is a cloud computing web application for calibration and dosimetry of radiochromic films. It is developed by one of the authors (IM) and incorporates the methods and models described in this study. It was employed for the calibration, dosimetry and gamma index evaluation. Additional statistical analysis was performed with R statistical software [30].

2.5 Calibration

To fit the calibration parameters, the plan-based method [18] was chosen. Besides being faster than the calibration method with fragments, the plan-based method provides a more representative sample of perturbations (since it can use every pixel of the film). This method requires one or more 2D dose distributions as reference doses for the calibration. In order to obtain them, films were placed in the MatriXX Evolution phantom and irradiated with a 60° Enhanced Dynamic Wedge (EDW) field of dimensions 20×20 cm^2. To reduce intralot variations [18], three separate films from each lot were exposed. The range of doses relevant for this study and for posterior clinical use was estimated between 50 cGy and 400 cGy. To encompass the whole range, two different fields were used: the wedge dose spanned from approximately 75 cGy to approximately 400 cGy (535 MU) for two of the films from each lot and from approximately 50 cGy to approximately 300 cGy (401 MU) for the remaining one.

Following this procedure, one set with EBT2 and another with EBT3 films were irradiated. Posteriorly, the films were scanned following the three protocols previously mentioned: reflection mode with 20 ± 1 h time window, transmission mode with 20 ± 1 h time window and transmission mode with 75 ± 5 min time window. Each set of images (six sets in total) was employed to calibrate each of the four models of channel-independent perturbations. Each of the models was calibrated against planned dose distributions (calculated with the TPS) and against measured dose distributions (simultaneously measured with MatriXX during the irradiations). Altogether, a total of 48 calibrations were computed.

Pixel values of the films exposed were translated into doses, for each color channel independently, fitting the calibration parameters. A genetic algorithm was used to fit the parameters minimizing the root-mean-square error (RMSE) of the differences between film doses for each channel ($D_k(r)$) and reference doses ($D(r)$).

This optimization provides the parameters used in Eq.(3.7) and Eq.(3.4). This is enough for film dosimetry following WM or MM models. However, to obtain $d(r)$ using UD or TN models σ_k are necessary, and also σ_Δ if using the TN model. Knowing lateral correction, sensitometric curve parameters and the

standard deviation of $f(\Delta)$, which depends on σ_Δ and θ and will be symbolized by $\tilde{\sigma}_\Delta$, σ_k can be estimated with

$$\tilde{\sigma}_k^2 \simeq \left(E[\dot{D}_k]\,\tilde{\sigma}_\Delta\right)^2 + \sigma_k^2 \tag{2.24}$$

where $\tilde{\sigma}_k$ is the RMSE of the channel and $E[\dot{D}_k]$ is the expected value (*i.e.*, mean) of $\dot{D}_k$.

The values of $\tilde{\sigma}_\Delta$ for UD and TN models, and of σ_Δ for the TN model, were obtained optimizing the RMSE of the differences between film doses ($d(r)$) and reference doses ($D(r)$).

2.6 Verification

To evaluate the four models of channel-independent perturbations, film dose distributions were compared with planned and with measured dose distributions. Global gamma analysis was conducted. The tolerances were 4 %, 3 mm with 20% of the maximum dose as threshold. Fourteen different cases were tested (Table 2.2). The cases were chosen with the intention of compiling a representative sample of dose distributions: several simple geometries, tissue heterogeneities, three-dimensional conformal radiotherapy (3D-CRT) plans and intensity modulated radiation therapy (IMRT) plans, including volumetric modulated arc therapy (VMAT) plans, were selected. EBT3 films were irradiated with all the cases but only a subset (considered representative) of them was used with EBT2 films, as shown in Table 2.2. Appropriate phantoms were employed dependent on the test case.

At the same time, several elements of the dosimetry process were compared: film type EBT2 versus EBT3, different waiting-time windows (*i.e.*, 75 ± 5 min versus 20 ± 1 h), reflection versus transmission mode scanning and planned versus measured reference dose distribution for film calibration and for gamma index analysis.

As a result, seven EBT2 and fourteen EBT3 films were irradiated with the cases shown in Table 2.2. They were scanned following the three scanning protocols under study. Each image was translated into a dose distribution following each of the four models of channel-independent perturbations. The film dose distributions were compared with the planned dose distributions in the plane of the film. Whenever the test was irradiated in the MatriXX phantom, the film dose distributions were also compared with the measured dose distributions. When film dose distributions were compared with planned ones, the calibration parameters of the model had been fitted using planned reference dose distributions, and analogously with measured dose distributions. If both planned and measured reference dose distributions are accurate, they should provide similar sets of calibration parameters. Following this, and for the TN model only, film distributions obtained with calibration parameters fitted using measured

Table 2.2: Description of the test cases, including film type and phantom used in the measurements.

Test	Description	Film type	Phantom
A	Square 15×15 cm^2	EBT2,EBT3	MatriXX
B	Chair test [31]	EBT2,EBT3	MatriXX
C	Pyramid shaped in both axis [32]	EBT2,EBT3	MatriXX
D	EDW 30° field	EBT3	MatriXX
E	EDW 45° collimator 90 field	EBT3	MatriXX
F	Y-shaped 3D CRT field	EBT2,EBT3	MatriXX
G	Predominantly convex shaped 3D CRT field	EBT3	MatriXX
H	RapidArc prostate 1	EBT3	CIRS Pelvic
I	RapidArc prostate 2	EBT3	CIRS Pelvic
J	RapidArc prostate 3	EBT2,EBT3	CIRS Pelvic
K	Square 10×10 cm^2, lung inhomogeneity	EBT3	CIRS Thorax
L	Lateral incidence, lung inhomogeneity	EBT2,EBT3	CIRS Thorax
M	Four field box, lung inhomogeneity	EBT3	CIRS Thorax
N	EDW and asymmetric fields, lung inhomogeneity	EBT2,EBT3	CIRS Thorax

reference dose distribution were also compared with planned dose distributions, and vice versa (*i.e.*, film distributions obtained with calibration parameters fitted using planned reference dose distribution were compared with dose distributions measured with MatriXX).

3 Results and discussion

Twenty-four different dosimetry protocols were analyzed in this study. To represent each protocol in a clear and concise way, they will be named using four characters (Table 2.3). The characters stand for: gamma analysis with either planned (P) or measured (M) evaluation dose distributions, scanning in reflection mode with 20 ± 1 h time window (R), in transmission mode with 20 ± 1 h time window (T) or in transmission mode with 75 ± 5 min time window (t), film type either EBT2 (2) or EBT3 (3) and calibration with either planned (p) or measured (m) reference dose distributions.

3.1 Selection of model of channel-independent perturbations

Table 2.4 compares film doses ($d(r)$) with planned or measured doses ($D(r)$), data are aggregated by model of channel-independent perturbations and dosimetry

Table 2.3: Elements of the dosimetry protocol under comparison.

Element of the protocol	Alternative	Abbreviation
Evaluation dose distribution for the gamma analysis	Planned	P
	Measured	M
Scanning mode and time window	Reflection, 20 ± 1 h	R
	Transmission, 20 ± 1 h	T
	Transmission, 75 ± 5 min	t
Film type	EBT2	2
	EBT3	3
Reference dose distribution for the calibration	Planned	p
	Measured	m

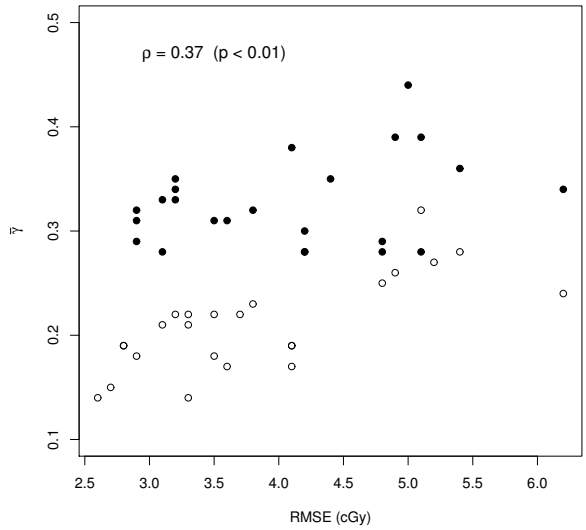
protocol. It contains RMSEs from the calibrations as well as gamma mean ($\overline{\gamma}$) and percentage of points with $\gamma_{<1}$ from the verification gamma analysis.

Considering the size of the sample and calculating likelihood from RMSE, Akaike Information Criterion (AIC) values are equivalent to calibration RMSEs. Since the TN model provided as good or better calibration RMSEs than the other models in all protocols, according to the AIC the TN model should be selected for multichannel dosimetry.

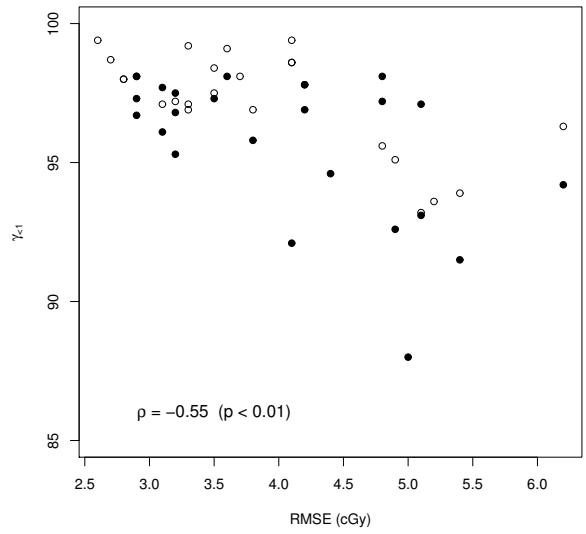
If the calibration data is well-suited to the problem of model selection, lower calibration RMSEs result in better agreement between film doses and planned or measured doses and, consequently, lower values of $\overline{\gamma}$ and higher $\gamma_{<1}$ in gamma analysis. In Fig. 3.1 it is shown that calibration RMSEs and $\overline{\gamma}$ or $\gamma_{<1}$ from Table 2.4 are significantly correlated. A consequence of this is shown in Table 2.5. It compares the models of channel-independent perturbations employing paired difference test. Models are paired for each dosimetry protocol. Differences are in $\overline{\gamma}$ values from Table 2.4. Mean differences between models are shown with dosimetry protocols grouped into protocols R and T. Protocols t were not used in this analysis in order to have both transmission mode and reflection mode scanning protocols equally weighted. The TN model provided significantly ($p < 0.05$) better results than the rest of models bringing together protocols R and T. Observing R protocols alone, the MM model was found significantly worse than WM and TN. Observing T protocols alone, the WM model was found significantly worse than the rest of models. Including in the analysis the rest of results, the TN model provided the best results both for R and T protocols, the UD model provided better results than the MM model, for R protocols the WM model provided better results than MM and UD models, however, for T protocols the WM model provided the worst results.

Table 2.4: Comparison of film doses ($d(r)$) with planned or measured doses ($D(r)$). Data are aggregated by model of channel-independent perturbations and dosimetry protocol. It contains RMSEs from the calibrations as well as gamma mean ($\overline{\gamma}$) and percentage of points with $\gamma_{<1}$ from the verification gamma analysis. The models of channel-independent perturbations include: weighted mean (WM), Micke-Mayer (MM) method, uniform distribution (UD) and truncated normal distribution (TN).

	WM			MM			UD			TN		
Protocol	RMSE (cGy)	$\overline{\gamma}$	$\gamma_{<1}$ (%)	RMSE (cGy)	$\overline{\gamma}$	$\gamma_{<1}$ (%)	RMSE (cGy)	$\overline{\gamma}$	$\gamma_{<1}$ (%)	RMSE (cGy)	$\overline{\gamma}$	$\gamma_{<1}$ (%)
PR2p	3.5	0.18	98.4	5.4	0.28	93.9	5.2	0.27	93.6	2.8	0.19	98.0
Pt2p	3.8	0.23	96.9	3.3	0.22	96.9	3.2	0.22	97.2	2.8	0.19	98.0
PT2p	5.1	0.32	93.2	3.7	0.22	98.1	3.5	0.22	97.5	3.1	0.21	97.1
PR3p	2.9	0.18	98.1	4.9	0.26	95.1	4.8	0.25	95.6	2.7	0.15	98.7
Pt3p	6.2	0.24	96.3	4.1	0.19	98.6	4.1	0.19	98.6	4.1	0.17	99.4
PT3p	3.3	0.21	97.1	3.6	0.17	99.1	3.3	0.14	99.2	2.6	0.14	99.4
MR2m	3.6	0.31	98.1	5.1	0.39	93.1	4.9	0.39	92.6	2.9	0.31	97.3
Mt2m	4.4	0.35	94.6	3.2	0.34	95.3	3.1	0.33	96.1	2.9	0.32	96.7
MT2m	5.4	0.36	91.5	3.5	0.31	97.3	3.2	0.35	96.8	3.2	0.33	97.5
MR3m	3.1	0.28	97.7	5.0	0.44	88.0	4.8	0.29	98.1	2.9	0.29	98.1
Mt3m	6.2	0.34	94.2	4.2	0.30	96.9	4.2	0.28	97.8	4.2	0.28	97.8
MT3m	4.1	0.38	92.1	5.1	0.28	97.1	4.8	0.28	97.2	3.8	0.32	95.8



(a)



(b)

Figure 2.1: Correlation between calibration RMSEs and (a) $\overline{\gamma}$ or (b) $\gamma_{<1}$, from Table 2.4. White dots correspond to planned dose distributions and black dots to measured dose distributions. Spearman's rank correlation coefficients and p-values of the correlations are shown.

Table 2.5: Comparison of models employing paired difference tests. Models are paired for each dosimetry protocol. Differences are in $\overline{\gamma}$ values from Table 2.4. Mean differences between models are shown. Negative values indicate that the first model obtained better results than the second one, and the opposite for positive values. Between parentheses, p-values of the t-tests are shown. The null hypothesis is that the mean difference between paired observations is zero. Protocols are grouped into protocols R and T.

Protocol group	WM - MM	WM - UD	WM - TN	MM - UD	MM - TN	UD - TN
R + T	-0.03 (0.40)	0.01 (0.61)	0.03 (<0.05)	0.02 (0.26)	0.06 (<0.05)	0.04 (<0.05)
R	-0.11 (<0.05)	-0.03 (0.51)	0.00 (0.81)	0.04 (0.32)	0.11 (<0.05)	0.07 (0.06)
T	0.05 (<0.05)	0.06 (<0.05)	0.05 (<0.05)	0.00 (0.73)	0.01 (0.32)	0.01 (0.43)

The WM model neglects correlations between channels. Poor performance of the WM model with T protocols can signify that these correlations (*e.g.*, due to variations in the thickness of the active layer) are important and can not be neglected. Still, this outcome does not mean that MM or UD models are preferable to WM in transmission mode scanning, it depends on the uncertainty σ_D (Eq.(2.17)), *i.e.*, on the properties of the dosimetry system under study. In fact, after analysing a sample of points from different test cases, σ_D was found to be the reason why MM and UD models provided worse results than the WM model with R protocols. To illustrate the importance of σ_D, the protocol PR2p was calibrated with MM using only red and green color channels. In our dosimetry system the sensitometric curves of both channels were very similar and this caused the RMSE of the calibration to be 3300 cGy. From Eq.(2.23), it follows that the accuracy of the MM model depends on the properties of the sensitometric curves, and can result in unacceptable uncertainties. Another flaw of the MM model is that all σ_k are considered equal. This hypothesis is usually false. As an example, it was found for protocol PR3p: σ_R = 3.5 cGy, σ_G = 2.8 cGy and σ_B = 6.8 cGy. The UD model corrects this deficiency, which could explain why it provided better results than the MM model. However, the accuracy of the UD model still depends on the properties of the sensitometric curves and of σ_k. With respect to the TN model, even though it is also submitted to Eq.(2.17), it can be considered as a metamodel that minimizes σ_D and can derive (as a limit) into the WM model, or into an intermediate case between models WM and UD. As a conclusion, we believe the superior performance of the TN model of channel-independent perturbations makes it the best choice for multichannel dosimetry.

Once the TN model was selected, gamma analysis was conducted for Pm and Mp protocols. Values of $\overline{\gamma}$ and $\gamma_{<1}$ for this protocols are shown in Table 2.6. RMSEs from the calibrations are not included since they are already present in Table 2.4 (*i.e.*, the RMSE from the calibration is independent of the evaluation dose distribution used for gamma analysis).

3.2 Selection of dosimetry protocol

Comparisons of elements of the dosimetry process were made employing paired difference tests for the TN model. For each point of each test case the difference in γ values between two protocols was calculated. Between both protocols, only one element of the dosimetry process was modified. Since some test cases were not present in some protocols (*e.g.*, Test H in protocols M), the numbers of test cases differ between the comparisons. Results of the comparisons are shown in Table 2.7.

Table 2.7a and Table 2.7b compare transmission mode scanning with 75 ± 5 min time window (t) versus transmission mode scanning with 20 ± 1 h time

Table 2.6: Comparison of film doses ($d(r)$) with planned or measured doses ($D(r)$). Data are aggregated by model of channel-independent perturbations and dosimetry protocol. It completes Table 2.4 for the truncated normal distribution (TN) model.

	TN	
	$\overline{\gamma}$	$\gamma_{<1}$
PR2m	0.15	99.4
Pt2m	0.17	99.2
PT2m	0.20	98.5
PR3m	0.15	98.8
Pt3m	0.19	98.3
PT3m	0.18	98.4
MR2p	0.47	88.6
Mt2p	0.35	95.5
MT2p	0.43	91.2
MR3p	0.41	89.6
Mt3p	0.39	91.1
MT3p	0.33	95.2

window (T), and reflection mode scanning with 20 ± 1 h time window (R) versus transmission mode scanning with 20 ± 1 h time window, respectively. Protocols t provided better results than T, and T better than R. However, the differences are not significant. In Table 2.7c, there is almost no difference between employing film type EBT2 (2) or EBT3 (3). Table 2.7d shows significant ($p < 0.05$) differences between calibration with planned (p) or measured (m) reference dose distributions. This result could be explained assuming that, for the EDW plan used in the calibration, the dose distribution measured with MatriXX has less uncertainty than the dose planned with Eclipse 10. Table 2.7e shows significantly ($p < 0.05$) better results for the gamma analysis with planned (P) than with measured (M) dose distributions. This is a consequence of the resolution of the evaluation dose distribution which is much worse for MatriXX. The resolution of the array affects negatively the value of the γ-index in spite of using bicubic interpolation. Swapping reference and evaluation dose distributions was discarded since it would induce noise artifacts [33].

Taking into account these comparisons, we selected the following dosimetry protocol: calibration with measured reference dose distributions, using film type EBT3, scanning in transmission mode with 75 ± 5 min time window and comparing the results with gamma analysis using planned evaluation dose distributions (*i.e.*, protocol Pt3m). Following this protocol allowed us to improve our previous γ-index tolerances from 4 % 3 mm to 3 % 3 mm or even 2.5 % 2.5 mm, results

Table 2.7: Comparing dosimetry protocols. When protocols with trasmission-mode scanning were grouped together they are symbolized with: (t+T). Differences are in γ values. Mean differences between protocols as well as the mean of the means and its standard deviation are shown. Negative values indicate that the first model obtained better results than the second one, and the opposite for positive values. Between parentheses, the p-value of the t-test is included. The null hypothesis is that the mean difference between paired observations is zero.

(a) Protocols t versus T.

P2p	P3p	M2m	M3m	P2m	P3m	M2p	M3p	Mean (p-value)
-0.034	0.013	-0.002	-0.046	0.032	-0.007	-0.132	0.023	-0.02 ± 0.06 (p = 0.33)

(b) Protocols R versus T.

P2p	P3p	M2m	M3m	P2m	P3m	M2p	M3p	Mean (p-value)
0.005	0.046	0.020	0.021	-0.003	-0.019	0.071	0.168	0.04 ± 0.06 (p = 0.11)

(c) Protocols 2 versus 3.

PRp	P(t+T)p	MRm	M(t+T)m	PRm	P(t+T)m	MRp	M(t+T)p	Mean (p-value)
0.038	0.061	-0.066	-0.003	0.008	0.024	-0.064	-0.001	-0.001 ± 0.050 (p = 0.97)

(d) Protocols p versus m.

PR2	P(t+T)2	PR3	P(t+T)3	MR2	M(t+T)2	MR3	M(t+T)3	Mean (p-value)
0.084	0.027	0.049	0.002	0.232	0.102	0.179	0.085	0.10 ± 0.08 ($p < 0.05$)

(e) Protocols P versus M.

R2p	(t+T)2p	R3p	(t+T)3p	R2m	(t+T)2m	R3m	(t+T)3m	Mean (p-value)
-0.370	-0.251	-0.370	-0.239	-0.193	-0.149	-0.226	-0.151	-0.24 ± 0.09 ($p < 0.05$)

Table 2.8: Gamma analisys of the test cases with dosimetry model TN, protocol Pt3m and different tolerances: 4 % 3 mm (with 20% of the dose maximum (D_{max}) as threshold), 3 % 3 mm (threshold 10% of D_{max}) and 2.5 % 2.5 mm (threshold 10% of D_{max}).

	γ (4 %, 3 mm)		γ (3 %, 3 mm)		γ (2.5 %, 2.5 mm)	
Test	$\overline{\gamma}$	$\gamma_{<1}$ (%)	$\overline{\gamma}$	$\gamma_{<1}$ (%)	$\overline{\gamma}$	$\gamma_{<1}$ (%)
A	0.15	99.6	0.17	99.5	0.22	97.9
B	0.23	97.7	0.26	96.0	0.34	93.6
C	0.13	100	0.22	99.4	0.30	97.1
D	0.19	96.7	0.13	99.9	0.18	99.3
E	0.15	97.7	0.22	96.8	0.31	95.8
F	0.14	100	0.14	99.8	0.18	99.3
G	0.14	99.9	0.18	99.2	0.23	97.9
H	0.18	98.6	0.22	97.7	0.31	94.8
I	0.10	99.8	0.17	98.8	0.23	97.0
J	0.15	99.5	0.20	98.3	0.27	96.1
K	0.15	99.7	0.24	98.4	0.30	95.9
L	0.17	99.7	0.24	98.0	0.33	93.5
M	0.21	98.4	0.32	94.0	0.42	88.6
N	0.39	93.5	0.52	83.5	0.66	71.9

are presented in Table 2.8.

3.3 Summary and recommendations

With respect to the model of channel-independent perturbations:

1. We recommend using the truncated normal distribution model because it can be considered as a metamodel which minimizes the uncertainty in the dose inherent in the method of channel-independent perturbations. The weighted mean model neglects correlations between the channels, which can be important, and the accuracy of the Micke-Mayer model depends on the properties of the sensitometric curves, which can result in unacceptable uncertainties for particular dosimetry systems. Since the other models are either limit cases or particular cases of the TN model, the latter should provide at least as good results as them.

2. For film calibration using the TN model, it is recommended to calibrate each color channel first. After that, two parameters: $\tilde{\sigma}_{\Delta}$ and of σ_{Δ}, are obtained optimizing the RMSE of the differences between film doses ($d(r)$) and reference doses ($D(r)$), according to Eq.(2.16) and Eq.(2.24).

3. Film doses can be calculated following a closed-form formula (Eq.(2.16)). In addition, the type B uncertainty in the dose implicit in the method can be calculated (Eq.(2.17)).

With respect to the dosimetry protocol, and excluding the comparisons between the particular TPS and array dosimeter used in this study:

1. No significant differences were found between transmission and reflection mode scanning.

2. Short waiting-time windows can be employed without losing accuracy, as pointed out by Lewis *et al* [34].

3. No significant differences were found between using EBT2 or EBT3 films.

4 Conclusions

Four models of channel-independent perturbations for multichannel film dosimetry were examined. Two of them based on the literature: a model which employs channel-independent perturbations as proposed by Micke *et al* [15] and further developed by Mayer *et al* [17], and another one which uses the weighted mean of all three channels to obtain the dose [18]. In addition to these, two novel models were proposed, a more realistic extension to the Micke-Mayer model which uses uniform distributed perturbations but allows the error terms to differ from one channel to another, and a truncated normal distribution, which comprises the other models as particular or limit cases.

A closed-form formula for dose calculation was derived for all four models, and it coincides with the published one [17] in the case of the Micke-Mayer model. In addition, Type B uncertainties in film dose due to the channel-independent perturbations method were obtained.

In order to assess the performance of the models, a set of tests was devised in which the dose distributions obtained from films were compared to either planned, or measured dose distributions. In these tests, the truncated normal distribution model provided the best agreement between film and reference doses, both for calibration and γ-index verification, and proved itself superior to both the weighted mean model, which neglects correlations between the channels, and the Micke-Mayer and the uniform distribution models, whose accuracy depends on the properties of the sensitometric curves. As a conclusion, we feel confident to recommend the truncated normal distribution model of channel-independent perturbations for multichannel dosimetry.

Along with the models, other factors which could influence the dosimetry process were also evaluated. No significant differences were found between transmission mode scanning and reflection mode scanning, between 75 ± 5 min

versus 20 ± 1 h waiting-time window or between employing EBT2 or EBT3 films. However, significantly better results were obtained when a measured dose distribution was used instead of a planned one as reference for the calibration, and when a planned dose distribution was used instead of a measured one as evaluation for the γ-analysis.

5 Acknowledgments

The authors would like to thank Denis Brojan, Víctor Hernández and Sašo Pulko for their contributions to this work.

References

[1] S. Devic, J. Seuntjens, E. Sham, E. B. Podgorsak, C. R. Schmidtlein, A. S. Kirov, and C. G. Soares. "Precise radiochromic film dosimetry using a flat-bed document scanner". *Medical Physics* 32.7 (2005), pp. 2245–2253.

[2] M. Fuss, E. Sturtewagen, C. D. Wagter, and D. Georg. "Dosimetric characterization of GafChromic EBT film and its implication on film dosimetry quality assurance". *Physics in Medicine and Biology* 52.14 (2007), pp. 4211–4225.

[3] L. Paelinck, W. D. Neve, and C. D. Wagter. "Precautions and strategies in using a commercial flatbed scanner for radiochromic film dosimetry". *Physics in Medicine and Biology* 52.1 (2007), pp. 231–242.

[4] L. J. van Battum, D. Hoffmans, H. Piersma, and S. Heukelom. "Accurate dosimetry with GafChromic EBT film of a 6 MV photon beam in water: What level is achievable?" *Medical Physics* 35.2 (2008), pp. 704–716.

[5] M. Martišíková, B. Ackermann, and O. Jäkel. "Analysis of uncertainties in Gafchromic EBT film dosimetry of photon beams". *Physics in Medicine and Biology* 53.24 (2008), pp. 7013–7027.

[6] L. Richley, A. C. John, H. Coomber, and S. Fletcher. "Evaluation and optimization of the new EBT2 radiochromic film dosimetry system for patient dose verification in radiotherapy". *Physics in Medicine and Biology* 55.9 (2010), p. 26012617.

[7] S. Devic. "Radiochromic film dosimetry: past, present, and future". *Physica medica* 27 (2011), pp. 122–134.

[8] W. Crijns, F. Maes, U. A. van der Heide, and F. V. den Heuvel. "Calibrating page sized Gafchromic EBT3 films". *Medical Physics* 40.1 (2013), 012102 (13pp.)

[9] A. Niroomand-Rad, C. R. Blackwell, B. M. Coursey, K. P. Gall, J. M. Galvin, W. L. McLaughlin, A. S. Meigooni, R. Nath, J. E. Rodgers, and C. G. Soares. "Radiochromic film dosimetry: Recommendations of AAPM Radiation Therapy Committee Task Group 55". *Medical Physics* 25.11 (1998), pp. 2093–2115.

[10] A. Rink, I. A. Vitkin, and D. A. Jaffray. "Energy dependence (75 kVp to 18 MV) of radiochromic films assessed using a real-time optical dosimeter". *Medical Physics* 34.2 (2007), pp. 458–463.

[11] C. Richter, J. Pawelke, L. Karsch, and J. Woithe. "Energy dependence of EBT-1 radiochromic film response for photon (10 kVp–15 MVp) and electron beams (6–18 MeV) readout by a flatbed scanner". *Medical Physics* 36.12 (2009), pp. 5506–5514.

[12] B. Arjomandy, R. Tailor, A. Anand, N. Sahoo, M. Gillin, K. Prado, and M. Vicic. "Energy dependence and dose response of Gafchromic EBT2 film over a wide range of photon, electron, and proton beam energies". *Medical Physics* 37.5 (2010), pp. 1942–1947.

[13] P. Lindsay, A. Rink, M. Ruschin, and D. Jaffray. "Investigation of energy dependence of EBT and EBT-2 Gafchromic film". *Medical Physics* 37.2 (2010), pp. 571–576.

[14] G. Massillon-JL, S. Chiu-Tsao, I. Domingo-Munoz, and M. Chan. "Energy Dependence of the New Gafchromic EBT3 Film:Dose Response Curves for 50 KV, 6 and 15 MV X-Ray Beams". *International Journal of Medical Physics, Clinical Engineering and Radiation Oncology* 1.2 (2012), pp. 60–65.

[15] A. Micke, D. F. Lewis, and X. Yu. "Multichannel film dosimetry with nonuniformity correction". *Medical Physics* 38.5 (2011), pp. 2523–2534.

[16] S. J. van Hoof, P. V. Granton, G. Landry, M. Podesta, and F. Verhaegen. "Evaluation of a novel triple-channel radiochromic film analysis procedure using EBT2". *Physics in Medicine and Biology* 57.13 (2012), pp. 4353–4368.

[17] R. R. Mayer, F. Ma, Y. Chen, R. I. Miller, A. Belard, J. McDonough, and J. J. O'Connell. "Enhanced dosimetry procedures and assessment for EBT2 radiochromic film". *Medical Physics* 39.4 (2012), pp. 2147–2155.

[18] I. Méndez, V. Hartman, R. Hudej, A. Strojnik, and B. Casar. "Gafchromic EBT2 film dosimetry in reflection mode with a novel plan-based calibration method". *Medical Physics* 40.1 (2013), 011720 (9pp.)

[19] S. Reinhardt, M. Hillbrand, J. J. Wilkens, and W. Assmann. "Comparison of Gafchromic EBT2 and EBT3 films for clinical photon and proton beams". *Medical Physics* 39.8 (2012), pp. 5257–5262.

[20] S. Park, S.-K. Kang, K.-H. Cheong, T. Hwang, H. Kim, T. Han, M.-Y. Lee, K. Kim, H. Bae, H. S. Kim, J. H. Kim, S. J. Oh, and J.-S. Suh. "Variations in dose distribution and optical properties of Gafchromic EBT2 film according to scanning mode". *Medical Physics* 39.5 (2012), pp. 2524–2535.

[21] S. Devic, S. Aldelaijan, H. Mohammed, N. Tomic, L.-H. Liang, F. DeBlois, and J. Seuntjens. "Absorption spectra time evolution of EBT-2 model Gafchromic film". *Medical Physics* 37.5 (2010), pp. 2207–2214.

[22] C. Fiandra, U. Ricardi, R. Ragona, S. Anglesio, F. R. Giglioli, E. Calamia, and F. Lucio. "Clinical use of EBT model Gafchromic film in radiotherapy". *Medical Physics* 33.11 (2006), pp. 4314–4319.

[23] S. Devic, Y.-Z. Wang, N. Tomic, and E. B. Podgorsak. "Sensitivity of linear CCD array based film scanners used for film dosimetry". *Medical Physics* 33.11 (2006), pp. 3993–3996.

[24] B. D. Lynch, J. Kozelka, M. K. Ranade, J. G. Li, W. E. Simon, and J. F. Dempsey. "Important considerations for radiochromic film dosimetry with flatbed CCD scanners and EBT GAFCHROMIC film". *Medical Physics* 33.12 (2006), pp. 4551–4556.

[25] L. Menegotti, A. Delana, and A. Martignano. "Radiochromic film dosimetry with flatbed scanners: A fast and accurate method for dose calibration and uniformity correction with single film exposure". *Medical Physics* 35.7 (2008), pp. 3078–3085.

[26] S. Saur and J. Frengen. "GafChromic EBT film dosimetry with flatbed CCD scanner: A novel background correction method and full dose uncertainty analysis". *Medical Physics* 35.7 (2008), pp. 3094–3101.

[27] H. Ohuchi. "High sensitivity radiochromic film dosimetry using an optical common-mode rejection and a reflective-mode flatbed color scanner". *Medical Physics* 34.11 (2007), pp. 4207–4212.

[28] JCGM. *JCGM 100:2008. Evaluation of measurement data - Guide to the expression of uncertainty in measurement.* 1st. Paris: JCGM, 2008.

[29] T. Künzler, I. Fotina, M. Stock, and D. Georg. "Experimental verification of a commercial Monte Carlo-based dose calculation module for high-energy photon beams". *Physics in Medicine and Biology* 54.24 (2009), pp. 7363–7377.

[30] R Core Team. *R: A Language and Environment for Statistical Computing.* ISBN 3-900051-07-0. R Foundation for Statistical Computing. Vienna, Austria, 2012. URL: `http://www.R-project.org/`.

[31] A. Van Esch, J. Bohsung, P. Sorvari, M. Tenhunen, M. Paiusco, M. Iori, D. P. Huyskens, et al. "Acceptance tests and quality control (QC) procedures for the clinical implementation of intensity modulated radiotherapy (IMRT) using inverse planning and the sliding window technique: experience from five radiotherapy departments". *Radiotherapy and Oncology* 65.1 (2002), pp. 53–70.

[32] S. G. Ju, Y. C. Ahn, S. J. Huh, and I. J. Yeo. "Film dosimetry for intensity modulated radiation therapy: Dosimetric evaluation". *Medical Physics* 29.3 (2002), pp. 351–355.

[33] B. M. Clasie, G. C. Sharp, J. Seco, J. B. Flanz, and H. M. Kooy. "Numerical solutions of the gamma-index in two and three dimensions". *Physics in Medicine and Biology* 57.21 (2012), pp. 6981–6997.

[34] D. Lewis, A. Micke, X. Yu, and M. F. Chan. "An efficient protocol for radiochromic film dosimetry combining calibration and measurement in a single scan". *Medical Physics* 39.10 (2012), pp. 6339–6350.

Paper 3

Model selection for radiochromic film dosimetry

I. Méndez

Physics in medicine and biology 60.10 (2015), 4089

Paper 3. Model selection for radiochromic film dosimetry

I. Méndez

Abstract

Abstract:

The purpose of this study was to find the most accurate model for radiochromic film dosimetry by comparing different channel independent perturbation models. A model selection approach based on (algorithmic) information theory was followed, and the results were validated using gamma-index analysis on a set of benchmark test cases. Several questions were addressed: (a) whether incorporating the information of the non-irradiated film, by scanning prior to irradiation, improves the results; (b) whether lateral corrections are necessary when using multichannel models; (c) whether multichannel dosimetry produces better results than single-channel dosimetry; (d) which multichannel perturbation model provides more accurate film doses. It was found that scanning prior to irradiation and applying lateral corrections improved the accuracy of the results. For some perturbation models, increasing the number of color channels did not result in more accurate film doses. Employing Truncated Normal perturbations was found to provide better results than using Micke-Mayer perturbation models. Among the models being compared, the triple-channel model with Truncated Normal perturbations, net optical density as the response and subject to the application of lateral corrections was found to be the most accurate model. The scope of this study was circumscribed by the limits under which the models were tested. In this study, the films were irradiated with megavoltage radiotherapy beams, with doses from about 20 cGy to 600 cGy, entire (8 inch × 10 inch) films were scanned, the functional form of the sensitometric curves was a polynomial and the different lots were calibrated using the plane-based method.

1 Introduction

Near water-equivalence[1, 2], high spatial resolution and weak energy dependence across a broad range of energies [3–8] make radiochromic film dosimetry with Gafchromic films (Ashland Inc., Wayne, NJ) and flatbed scanners the dosimetry system of choice for many applications in radiation therapy. Radiochromic films darken upon irradiation, which makes it possible to measure the absorbed dose using a scanner. However, when they are digitized with a color scanner, three different dose distributions - one for each color channel - are obtained. To combine

the information provided by all three channels into a single and more accurate dose distribution, multichannel dosimetry methods have been proposed[9–16].

An emerging field of research in multichannel radiochromic film dosimetry are perturbation models [11, 13, 15, 16]. These models consider that, for each channel, the measured dose distribution deviates from the true absorbed dose distribution because of small perturbations in the film-scanner response. To combine all three color channels, certain assumptions on the characteristics of the perturbations are necessary. This approach has shown promising results [11, 13, 15–17].

The purpose of this study is to select the most accurate model for radiochromic film dosimetry using Gafchromic EBT2 and EBT3 films and a flatbed scanner. In order to do so, different channel independent perturbation (CHIP) models are compared. This work examines whether incorporating the information of the non-irradiated film, by scanning prior to irradiation, improves the dosimetry. To what extent perturbation models correct the deviation from the response at the center of the scanner along the axis parallel to the scanner lamp is analyzed. Finally, single-, dual- and triple-channel models are compared to determine if by increasing the number of combined color channels the accuracy of radiochromic film dosimetry is also increased.

The results of this study should be considered applicable within the limits under which the models were tested. Through this work, the films were irradiated with megavoltage radiotherapy beams, in the dose range from about 20 cGy to 600 cGy, and whole (8 inch × 10 inch) films were scanned. The functional form of the sensitometric curves was a polynomial and the different lots were calibrated using the plan-based method [14].

2 Methods and materials

2.1 A general perturbation model

Many disturbances affect the film-scanner response [2, 11, 18, 19]: thickness variations in the active layer coated on the film, electronic noise, scanner instability, lateral artifact, local variations produced by systematic problems of the scanner, Newton rings, dust, scratches or other damage, etc. According to their persistence, disturbances can be classified into two groups: systematic local variations and random perturbations. Random perturbations (*i.e.*, noise) are not consistent in time: they change or disappear between scan repetitions or between non-irradiated and irradiated film scans. Electronic noise, scanner instability and Newton rings can be included within this category. Systematic variations are consistent between film scans. They include thickness variations in the active layer, lateral artifact, local variations produced by systematic problems of the scanner, etc. Some other disturbances such as dust, scratches or other damage

can be random or systematic, but they generally produce large alterations in the response, cannot be treated as perturbations, and should be considered as another source of noise.

The dose absorbed by the film at point r, $D(r)$, can be described by:

$$D(r) = D_k(z_k(r) + \Psi_k(r, z_k)) + \Sigma_k(r) \tag{3.1}$$

where k represents the color channel (*i.e.*, red (R), green (G) or blue (B)), D_k denotes the sensitometric curve (for channel k), $z_k(r)$ is the film-scanner response (either pixel value, optical density (OD) or net optical density (NOD)) at point r, $\Psi_k(r, z_k)$ corrects the systematic local variations and $\Sigma_k(r)$ represents noise disturbances.

If $\Psi_k(r, z_k)$ is small, we can apply a first-order Taylor expansion of $D(r)$ in terms of $\Psi_k(r, z_k)$. Hence, a general perturbation model for $D(r)$ can be expressed as:

$$D(r) = D_k(z_k(r)) + \dot{D}_k(z_k(r))\Psi_k(r, z_k) + \epsilon_k(r) \tag{3.2}$$

where $\dot{D}_k$ is the first derivative of D_k with respect to z_k and $\epsilon_k(r)$ is an error term that accounts for the difference between the true absorbed dose, $D(r)$, and the measured dose after correction of the perturbation. The noise is included in the error term.

2.2 Channel independent perturbation models in the literature

The Micke and Mayer models

Micke *et al* [11] were the first to suggest a perturbation model. Mayer *et al* [13] found a closed-form solution to that model. They proposed a CHIP model:

$$D(r) = D_k(z_k(r)) + \dot{D}_k(z_k(r))\Delta(r) + \epsilon_k(r) \tag{3.3}$$

which implicitly assumes that the probability density functions (pdf) of all the $\epsilon_k(r)$ terms are equal and the pdf of Δ is a uniform distribution [15].

The Truncated Normal model

In an earlier article [15], a generalization of the Micke-Mayer (MM) model was introduced. The pdf of Δ was considered to be a truncated normal (TN) distribution and the $\epsilon_k(r)$ terms could be different.

In this study, the MM and TN models were compared while different elements of the functional form that translates film-scanner responses into doses were varied. It should be pointed out that the MM model is not the same as the Micke or Mayer multichannel film dosimetry methods. For example, the Mayer method uses pixel value as film-scanner response, by contrast, in this study OD and NOD

were used. The Micke method uses a rational form for the sensitometric curves, by contrast, in this study the sensitometric curves followed

$$D_k = \sum_{n=0}^{4} b_{k_n} z_k^n \tag{3.4}$$

where b_k are fitting parameters.

However, both the Micke and Mayer methods share the same form of the perturbation, which diverges from the form defined by the TN model.

2.3 Scanning before and after irradiation

The response of the dosimetry system to irradiation is expressed in terms of pixel value, OD or NOD. OD is defined as:

$$z := \log_{10} \frac{v_{\max}}{v_{\text{irr}}} \tag{3.5}$$

where v_{irr} denotes the pixel values of the irradiated film and $v_{\max}$ is the maximum possible pixel value of the scanner. Between pixel value and OD there is only a change of coordinates. For this reason, in this study pixel value was not included in the comparison.

The information of the non-irradiated film can be incorporated in the response. NOD is defined as:

$$z := \log_{10} \frac{v_{\text{nonirr}}}{v_{\text{irr}}} \tag{3.6}$$

where v_{nonirr} denotes the pixel values of the non-irradiated film.

2.4 The lateral artifact

The lateral artifact is the deviation from the response at the center of the scanner along the axis parallel to the scanner lamp [14, 18, 20–29]. It is caused by the interplay between the light scattering from the polymers created in the active layer of the film and the properties of the scanner [30]. It is approximately parabolic in shape, with lower pixel values along the edges than at the center of the scan. It is dependent on the color channel and OD. The lateral correction is modeled empirically. In this paper, lateral corrections are calculated as [14, 27]:

$$v_k = a_{k_1}(x - x_c) + a_{k_2}(x - x_c)^2 + \hat{v_k} \tag{3.7}$$

where $\hat{v_k}$ is the pixel value before correction, x is the coordinate of the pixel in the axis parallel to the CCD array, x_c is the x coordinate of the center of the scanner, v_k represents the corrected pixel value, and a_k are fitting parameters.

Multichannel correction methods have been found to substantially mitigate the lateral artifact [11, 31]. This raises the question of whether additional

steps to correct the lateral artifact are necessary. If the lateral correction is considered included in the perturbation term, additional steps are unnecessary. An alternative approach is to explicitly apply the lateral correction following Eq.(3.7). In this study, both approaches were compared.

2.5 Single-channel dosimetry vs. multichannel dosimetry

Multichannel dosimetry combines the doses measured with different color channels ($D_k(r)$) into a single dose value ($d(r)$). Considering $D_k(r)$ as different measurements of the dose, if they were uncorrelated the uncertainty of the dose $d(r)$ could not be higher than the uncertainties of each of the channels separately. However, they are correlated [15], which implies that the uncertainty of $d(r)$ can be higher than the uncertainty of $D_k(r)$. Thus, multichannel dosimetry is not necessarily better than single-channel dosimetry. It depends, first, on the functional form of the particular multichannel model. According to the functional form, it can be derived that the uncertainty of the CHIP models compared in this work cannot worsen when the number of combined channels increases. Second, this is only valid if the functional form correctly describes the physics of radiochromic film dosimetry.

In this study, single-, dual- and triple-color-channel dosimetry models were compared.

2.6 Model selection

The most accurate model for radiochromic film dosimetry will be the one with the highest degree of coincidence between the film dose ($d(r)$) and the true dose ($D(r)$). The degree of coincidence is a qualitative magnitude, in this work it was quantified as the probability of both being equal.

According to Bayesian probability theory, given the measured data, the probability that a model is the one where the data come from is proportional to the probability of obtaining these data from the model times the prior probability of the model: $P(M|D) \propto P(D|M)\, P(M)$. The ideal choice for the prior is the universal weight based on the Kolmogorov complexity of the model [32, 33]. Unfortunately, the Kolmogorov complexity is not finitely computable. A practical complexity-based approach consists of selecting the model using the Akaike information criterion (AIC) [34]. Thus, the most probable model is the one with the lowest AIC value: $AIC = 2\, c_M - 2 \ln(P(D|M))$, where c_M is the number of parameters of the model.

Usually, when selecting a model for film dosimetry, the functional form is predefined and the only parameters to select fit the lateral corrections and sensitometric curves. This is done for each lot of films and is referred to as 'calibration'. In this study, the functional form (MM or TN, OD or NOD, etc.) had to be selected as well. Each lot of films was calibrated with each functional

form. For each calibration, to calculate $P(D|M)$, it was considered that the differences (ϵ_d) between the reference doses ($D(r)$) and measured doses ($d(r)$) were normally distributed. Thus, and considering that the number of parameters was negligible compared to the size of the data sample, the model with the lowest root mean square error (RMSE) comparing the reference and measured doses was regarded as the most likely model. Therefore, given a functional form, the parameters of the model were selected minimizing the RMSE. In addition, for each lot, the most probable functional form was given by the model with the lowest RMSE.

Henceforth, a given use of the term 'model' will refer to all those models with the same functional form. A calibration will select the parameters that maximize the likelihood of that model.

To exclude the possible dependence on a particular lot, a sample with four different lots was employed. Calibrations of different lots had different dose ranges and quantities of data. To balance their weights, relative rather than absolute differences were compared. For each calibration, it was assumed that the relative differences between reference doses ($D(r)$) and measured doses ($d(r)$) were also normally distributed:

$$D(r) = d(r) + \epsilon_d(r) \tag{3.8}$$

with:

$$\frac{\epsilon_d(r)}{d(r)} \sim \mathcal{N}(0, \sigma^2) \tag{3.9}$$

where σ is referred to as the relative 'uncertainty of the calibration'. To estimate the uncertainty of the measured dose, σ should be combined with the uncertainty of $D(r)$.

In information theory, entropy is defined as the expected negative log-likelihood of a random variable. Hence, maximizing the likelihood of a model is equivalent to minimizing the information entropy of the errors. The (differential) entropy of a model following Eq.(3.9) can be calculated as:

$$h(M) = \frac{1}{2}\ln(2\pi e\sigma^2) \tag{3.10}$$

Assigning the same weight to each lot, the entropy of a model in a sample of lots is the mean entropy of the model in the sample. As a result, it was considered that the most probable model for radiochromic film dosimetry was the one with the lowest geometric mean of the calibration uncertainty for all lots under study.

2.7 Scanning protocol

Gafchromic EBT2 and EBT3 films with dimensions 8 inch × 10 inch were used. They were handled in conformity to the recommendations of the AAPM TG-55

report [2]. The films were scanned with an Epson Expression 10000XL flatbed scanner (Seiko Epson Corporation, Nagano, Japan) prior to irradiation and within the time-window following irradiation. The scanner was warmed up for at least 30 min before use. Before acquisitions, and after long pauses, five empty scans were taken to stabilize the scanner lamp. The films were centered on the scanner with a black opaque cardboard frame. They were scanned in portrait orientation (*i.e.*, the short side of the film parallel to the scanner lamp and the long side parallel to the lamp movement axis). Scans were acquired with image-type set to 48-bit RGB (16 bit per channel), a resolution of 72 dpi and image processing tools turned off. They were saved as TIFF files. Five consecutive scans were taken for each film and the first scan was discarded to avoid the warm-up effect of the scanner lamp occurring with multiple scans [22, 25].

2.8 Calibration

To calibrate a lot according to the models under comparison, a set of reference doses should be associated with a representative sample of responses, lateral positions and perturbations.

To obtain a representative calibration sample, the plan-based method was chosen [14]. Films were placed in a Plastic Water (Computerized Imaging Reference Systems Inc. Norfolk, VA, USA) phantom at source-axis distance (SAD). They were irradiated with a 6 MV photon beam and a 60° Enhanced Dynamic Wedge field of dimensions 20×20 cm^2 at SAD. The range of doses delivered to the film encompassed the range of doses of interest. For some lots, the irradiation was repeated with several films and different monitor units (MU). In this way, the dose range was extended and the intralot variability mitigated. The dose distribution in the plane of the film was either calculated with the treatment planning system (TPS) or, preferably, measured simultaneously with an IBA MatriXX Evolution ionization chamber array (IBA Dosimetry GmbH, Germany). Since the reference dose plane can be either planned or measured, from now on this method will be referred to as the 'plane-based' method.

The plane-based method is an alternative to the well-established calibration method with fragments [14]. A disadvantage of the plane-based method is that the reference doses have higher uncertainty than with the calibration with fragments. However, better precision in the reference dose is useless if the calibration sample is biased. The plane-based method provides a more representative calibration sample, making it more robust against perturbations. Additionally, another advantage of the plane-based method is the efficiency, since the time required for calibration is considerably reduced.

Four lots were used, two of them with EBT2 films: lot A04141003BB (Lot A) and lot A03171101A (Lot B); and the other two with EBT3: lot A05151201 (Lot

C) and lot A03181301 (Lot D). They were calibrated following slightly different variations of the protocol:

The Lot A films were irradiated at a depth of 6 cm in a 12×30×30 cm^3 phantom in slab form with a Novalis Tx accelerator (Varian, Palo Alto, CA, USA). One film was irradiated with a wedge field with 660 MU (the reference doses extended from 100 cGy to 600 cGy, approximately). The reference dose plane was calculated with Eclipse v.10.0 (Varian Medical Systems) using the Anisotropic Analytical Algorithm (AAA). The films were scanned in reflection mode and the waiting time-window was 24 ± 1 h.

The Lot B and Lot C films were irradiated atop the IBA MatriXX detector inside the IBA MatriXX Evolution MULTICube with a Novalis Tx accelerator. Three films were used for the calibration, two of them were irradiated with a wedge field with 535 MU (dose range: 75-400 cGy, approximately) and the other one with a wedge with 401 MU (dose range: 50-300 cGy, approximately). The reference doses were measured simultaneously. The films were scanned in reflection and transmission mode. The time-window was 20 ± 1 h.

The Lot D films were also irradiated and measured with IBA MatriXX Evolution. They were irradiated with a Varian 2100 CD accelerator. Three films were used for the calibration using wedge fields with 130, 350 and 550 MU (approximate dose ranges: 20-100, 50-270 and 80-420 cGy, respectively). The films were scanned in transmission mode and the time-window was 24 ± 1 h.

Whenever the reference doses were measured with the detector array, the dose values were scaled with a factor of 1.015 in order to correct for the distance between the plane of the film and the plane of measurement of the detector. When the reference doses were calculated with the TPS, they were corrected with the daily output of the Linac.

The reference dose planes were exported (with a resolution of 0.49 mm/px for the planned doses and 7.62 mm/px for doses measured with MatriXX). They were uploaded together with the film scans to Radiochromic.com. All the dosimetry models analyzed in this study were incorporated in a research version based on Radiochromic.com v1.6 (http://radiochromic.com).

A total of six calibration samples (*i.e.*, one for each lot except for Lots B and C, which had two samples: one scanned in reflection and another in transmission mode) were uploaded.

Each was calibrated using MM and TN perturbations, using OD and NOD as film-scanner response, applying and not applying the lateral correction defined in Eq.(3.7) and employing each of the seven possible combinations of color channels (*i.e.*, R, G, B, RG, RB, GB, RGB). Taking into account that the MM and TN models only apply to multichannel combinations, 44 different models for radiochromic film dosimetry were compared for each of the calibration samples. Since Radiochromic.com optimizes the calibrations using an evolutionary

algorithm, each calibration was repeated three times and the one with the lowest uncertainty was selected.

2.9 Validation

Given the calibration sample, the model with the lowest calibration uncertainty was considered the most accurate one. This model is expected to correctly describe the relationship between the film-scanner response and the absolute dose absorbed by the film. However, this expectation can be disproved when studying new data, especially if the calibration sample is not representative. To validate the model selection, film dose distributions were compared with planned dose distributions by means of global gamma-index analyses [35]. Tolerances were set at 3%, 3 mm excluding points with less than 30% of the maximum planned dose. This threshold was chosen to prevent extrapolating the sensitometric curves of the calibrations for lower doses when using OD. To avoid noise artifacts [36] (*i.e.*, noise in the evaluation distribution which spuriously improves the gamma-index), the film and planned distributions were, respectively, the reference and evaluation distributions.

Benchmark test cases described in an earlier paper [15] were employed. To allow other researchers to reproduce and verify the results, or to improve the present models under analysis, these benchmark tests are publicly available at Radiochromic.com. They were considered a representative sample of the dose distributions, both for the EBT2 and EBT3 films. The EBT2 films belonged to Lot B and the EBT3 films to Lot C. The films were exposed with a Novalis Tx accelerator and scanned, in reflection and transmission modes, following the same protocol as for the calibrations. Three phantoms were used: IBA MatriXX Evolution MULTICube, CIRS Pelvic Phantom (Model 002PRA, Computerized Imaging Reference Systems Inc. Norfolk, VA, USA) and CIRS Thorax Phantom (Model 002LFC). They were set up at SAD. The films were placed atop the MatriXX Evolution detector (coronal plane) in the MatriXX phantom and with an offset from the beam axis in the CIRS phantoms (transversal plane), of 1.5 cm in the thorax phantom and 1.3 cm in the pelvic phantom, to diminish beam attenuation by the film [37, 38]. Calibration and verification tests which were simultaneously measured with MatriXX were compared with the TPS, obtaining a mean $\gamma_{<1}$ (3% 3mm) of 98.9%.

The film scans (42 tests) were converted to dose distributions applying each of the 44 dosimetry models of the calibration. Radiochromic.com was employed for the calculations.

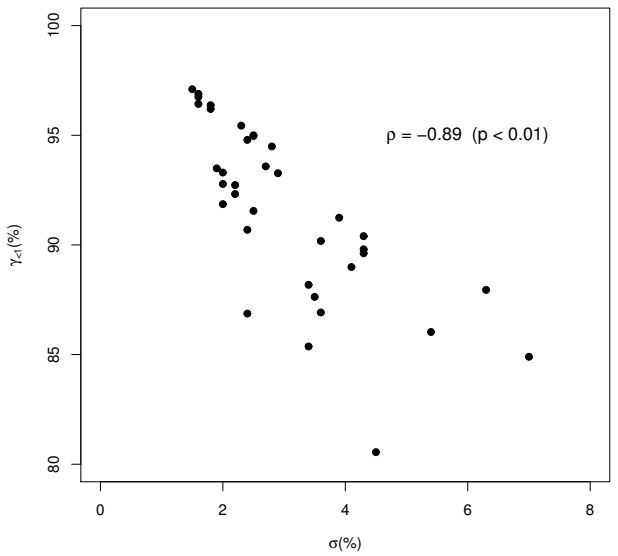


Figure 3.1: Correlation between the uncertainty of the calibration and the mean, aggregating all 42 test cases, $\gamma_{<1}$ (3% 3mm). The Spearman's rank correlation coefficient and p-value are included

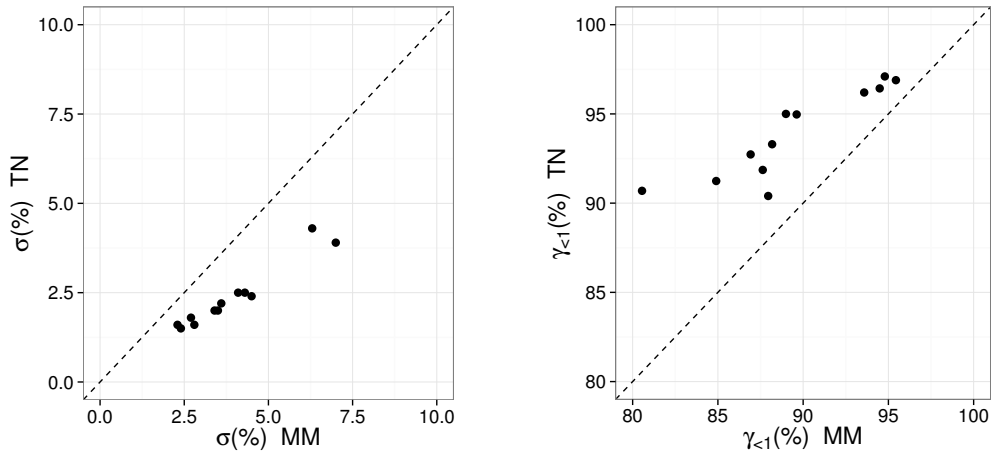
3 Results and discussion

If the model selection based on the calibration is correct, the uncertainty of the calibration (σ) should be correlated with the percentage of points with $\gamma_{<1}$ in the gamma analysis. In Figure 3.1, it is found that the correlation between σ and the mean, aggregating all 42 test cases, $\gamma_{<1}$ (3% 3mm) is significant.

In this section, the 44 models are compared. Each dosimetry model comprises: perturbation form for multichannel models (MM/TN), response (OD/NOD), lateral correction (applied/not applied) and the combination of color channels (R, G, B, RG, RB, GB, RGB). Models that only differ in one of these elements will be contrasted according to the σ and the mean $\gamma_{<1}$ (3% 3mm). For the sake of clarity of the presentation, points with σ larger than 10% or $\gamma_{<1}$ less than 80% are not displayed in the figures.

3.1 MM vs. TN perturbations

Figure 3.2 contrasts multichannel models that only differ in the form of the perturbation: they use either MM or TN perturbations. The TN models were found to be better than the MM models in every case. This confirms the recommendation of using TN perturbations as opposed to MM perturbations



(a) (b)

Figure 3.2: Contrast of multichannel models that only differ in the form of the perturbation: MM vs. TN perturbations. According to a) the uncertainty of the calibration and b) the mean $\gamma_{<1}$ (3% 3mm)

when using CHIP models. It is due to the fact that the TN perturbation generalizes MM and minimizes the uncertainty in the dose inherent to CHIP models [15].

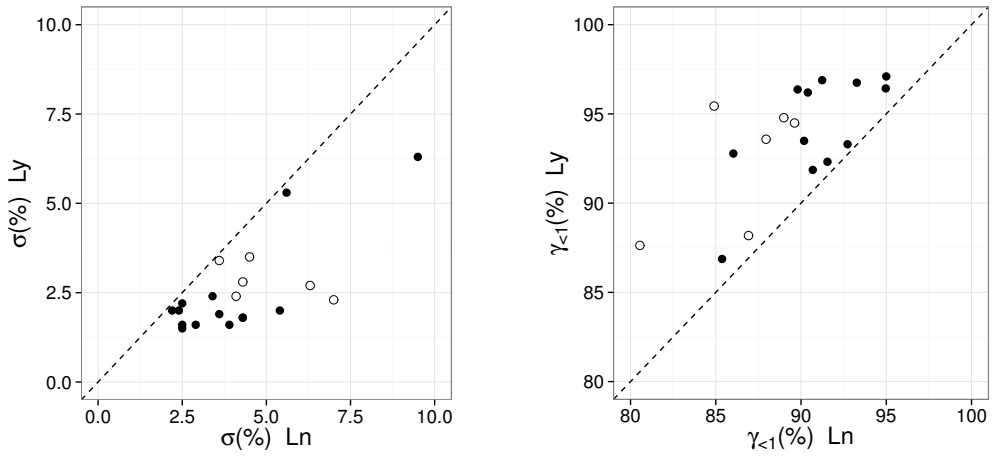
3.2 Lateral correction

Figure 3.3 contrasts the models that apply (Ly) with those that do not apply (Ln) lateral corrections. Applying lateral corrections improved the results for all the cases, even for multichannel models.

Recently, a new empirical formula to correct the lateral artifact has been published by Poppinga *et al* [29]. Figure 3.4 illustrates how this formula compares with the correction used in this work (Eq.(3.7)), which will be denoted as Absolute correction.

Four fragments from lot A irradiated with different doses were scanned at five different positions along the axis parallel to the scanner lamp. A 50 × 50 px ROI with resolution 72dpi was measured at the center of each fragment. The lateral artifacts for all three color channels were fitted following the Absolute correction and the Poppinga correction formulas. In Figure 3.4 deviations in pixel value from the value measured at the center of the scanner are plotted. The maximum percentage differences between OD obtained using both correction methods were 0.9%, 0.3% and 0.3% for the R, G and B channels, respectively.

To evaluate both corrections, residuals of the fits were calculated, which are shown in Table 3.1. The Poppinga formula was better for the R channel, and the Absolute formula for the G and B channels. It was considered that both lateral



(a) (b)

Figure 3.3: Contrast of models that apply (Ly)/do not apply (Ln) lateral corrections. According to a) the uncertainty of the calibration and b) the mean $\gamma_{<1}$ (3% 3mm). White points are models using MM perturbations.

Table 3.1: Residuals fitting lateral artifacts with the Absolute and Poppinga formulas for all three color channels.

	Residuals (pixel value)		
Lateral correction	R	G	B
Absolute	65.3	33.7	28.4
Poppinga	46.3	57.9	38.9

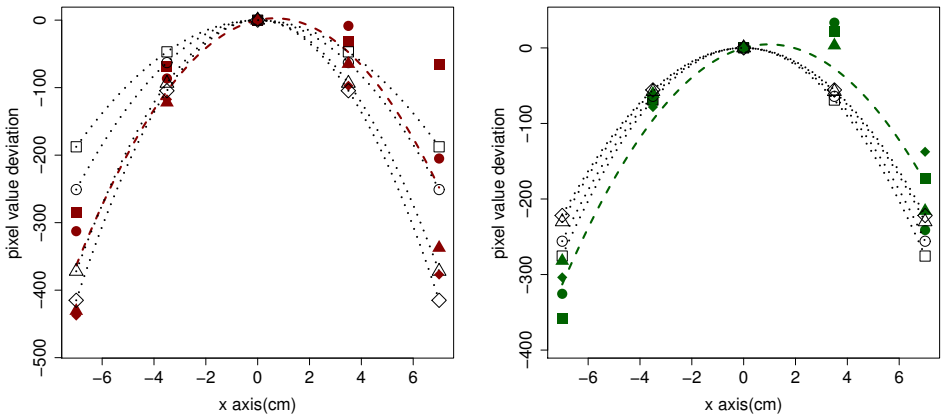
corrections were equally valid for empirically modeling the lateral artifact present in radiochromic film dosimetry.

3.3 Scanning before and after irradiation

Figure 3.5 contrasts models that only need the information of the irradiated scan (*i.e.*, use OD as response) with models that also need the information of the non-irradiated scan (*i.e.*, use NOD as response). White points are models that do not apply lateral corrections or use MM perturbations.

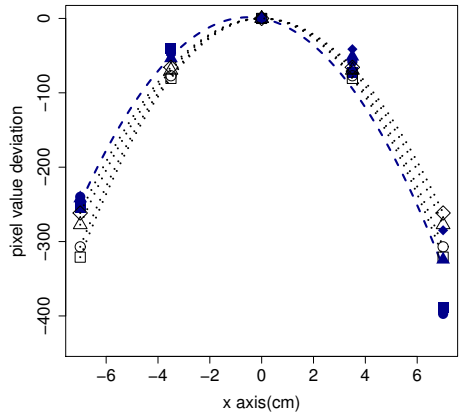
Incorporating the information of the non-irradiated scan was found to be correlated with lower σ and higher $\gamma_{<1}$ (3% 3mm).

When using NOD, dose-independent perturbations in the film-scanner response (e.g., thickness variations of non-active layers) are canceled out, dose-dependent perturbations, like the lateral artifact or film inhomogeneities, are reduced. However, this is valid if the non-irradiated and the irradiated film



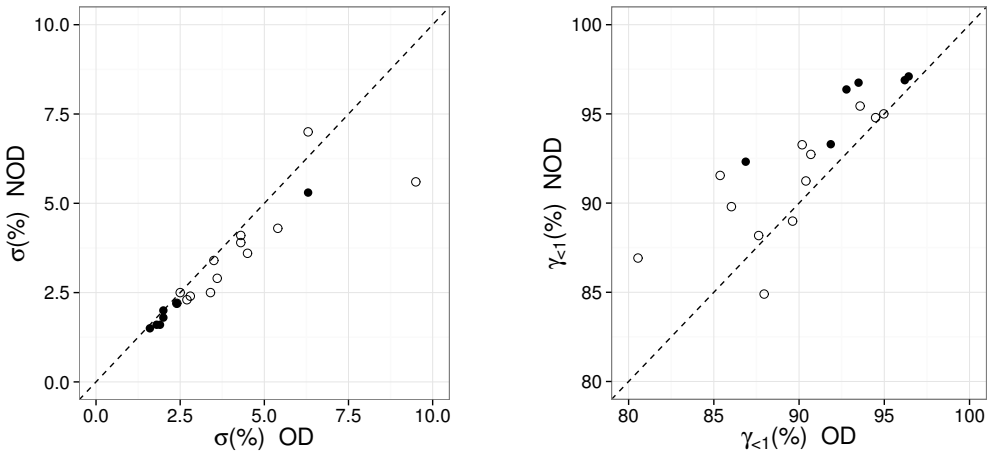
(a)

(b)



(c)

Figure 3.4: Film fragments scanned at different positions along the axis parallel to the scanner lamp for the three color channels. Deviations in pixel value from the value measured at the center of the scanner are shown. The colored points represent measurements, the dashed lines Absolute corrections, and the dotted lines and empty points Poppinga corrections. The shape of the point identifies the dose level of the fragment: 0 Gy (□), 1 Gy (○), 4 Gy (△), 6 Gy (◇).



(a) (b)

Figure 3.5: Contrast of models using OD vs. NOD as film-scanner response, according to the uncertainty of the calibration, and according to the mean $\gamma_{<1}$ (3% 3mm). White points are models that do not apply lateral corrections or use MM perturbations.

images are perfectly registered. Registration errors introduce another source of noise/uncertainty that will increase in importance the steeper the gradients of the perturbations are. Registration errors and pixel value noise could increase the noise in the film dose distributions. Even though higher noise in the reference dose distribution is not expected to improve the results of the gamma analysis, in contrast to higher noise in the evaluation distribution, the effect of increasing the noise in the film dose distribution was studied. The dose distribution of a RapidArc prostate plan (case J from the benchmark tests) [15] was calculated using the TN-NOD-Ly-RGB model (comprising TN perturbation, with NOD as response, applying lateral correction, and employing all three color channels). The film belonged to Lot C scanned in reflection mode. The noise in the film dose map was either smoothed using median filters or increased with Gaussian noise. The percentage of points with $\gamma_{<1}$ was calculated using tolerances of 3% 3mm and 2% 2mm. Data are shown in Table 3.2. It was found that higher noise in the dose distribution did not improve the gamma results.

3.4 Single-channel dosimetry vs. multichannel dosimetry

Table 3.3 and Table 3.4 include, respectively, the uncertainty of the calibration and the mean percentage of points with $\gamma_{<1}$ (3% 3mm), taking into account all 42 test cases, for the dosimetry models under comparison. It can be observed that increasing the number of color channels does not necessarily result in more accurate film doses. This is conspicuous for the MM-RG models, which have very

Table 3.2: Percentage of points with $\gamma_{<1}$ (2% 2mm) and $\gamma_{<1}$ (3% 3mm) for the test J of Lot C scanned in reflection mode after filtering the dose distribution with median filters or Gaussian noise.

Filter	$\gamma_{<1}$(2% 2mm)	$\gamma_{<1}$(3% 3mm)
Median filter 7x7 px	97.2	99.8
Median filter 5x5 px	97.1	99.8
Median filter 3x3 px	96.9	99.8
No filter	96.0	99.6
1% standard deviation noise	95.1	99.3
2% standard deviation noise	92.6	98.3
3% standard deviation noise	88.7	96.8

Table 3.3: Uncertainty of the calibration for the dosimetry models under comparison.

	Color channel combination						
Dosimetry model	R	G	RG	B	RB	GB	RGB
MM-OD-Ln	5.4	3.4	81.7	9.5	6.3	4.5	4.3
MM-OD-Ly	2.0	2.4	20.6	6.3	2.7	3.5	2.8
MM-NOD-Ln	4.3	2.5	62.8	5.6	7.0	3.6	4.1
MM-NOD-Ly	1.8	2.2	14.7	5.3	2.3	3.4	2.4
TN-OD-Ln	5.4	3.4	3.6	9.5	4.3	2.4	2.5
TN-OD-Ly	2.0	2.4	1.9	6.3	1.8	2.0	1.6
TN-NOD-Ln	4.3	2.5	2.9	5.6	3.9	2.2	2.5
TN-NOD-Ly	1.8	2.2	1.6	5.3	1.6	2.0	1.5

high uncertainties and lower mean $\gamma_{<1}$ (3% 3mm) values than the corresponding R single-channel models. These high uncertainties and lower mean $\gamma_{<1}$ values were caused by the sensitivity of the inherent dose uncertainty in the MM model with the properties of the sensitometric curves [15]; in particular, in these cases, poor outcomes occurred because the sensitometric curves of both the R and G channels were very similar.

From the data presented in Table 3.3 and Table 3.4, it is found that applying lateral corrections and scanning prior and following to irradiation had a larger effect on the results than using multichannel models. For instance, the OD-Ln-R model had 5.4% of calibration uncertainty and the mean percentage of points with $\gamma_{<1}$ 3% 3mm was 86%, the NOD-Ly-R model had 1.8% and 96.4% of calibration uncertainty and mean $\gamma_{<1}$ 3% 3mm, respectively, a greater improvement than the one obtained with the MM-OD-Ln-RGB (4.3%, 89.6%) or the TN-OD-Ln-RGB model (2.5%, 95.0%).

Table 3.4: Mean percentage of points, taking into account all 42 test cases, with $\gamma_{<1}$ (3% 3mm) for the dosimetry models under comparison.

	Color channel combination						
Dosimetry model	R	G	RG	B	RB	GB	RGB
MM-OD-Ln	86.0	85.4	59.5	59.2	88.0	80.6	89.6
MM-OD-Ly	92.8	86.9	72.1	64.6	93.6	87.6	94.5
MM-NOD-Ln	89.9	91.6	59.6	76.5	84.9	86.9	89.0
MM-NOD-Ly	96.4	92.3	75.3	76.2	95.4	88.2	94.8
TN-OD-Ln	86.0	85.4	90.2	59.2	90.4	90.7	95.0
TN-OD-Ly	92.8	86.9	93.5	64.6	96.2	91.9	96.4
TN-NOD-Ln	89.9	91.6	93.3	76.5	91.2	92.7	95.0
TN-NOD-Ly	96.4	92.3	96.8	76.2	96.9	93.3	97.1

Table 3.5: The best models according to: calibration uncertainty (σ), mean percentage of points with $\gamma_{<1}$ 2% 2mm (mean γ_{22}) and 3% 3mm (mean γ_{33}), both gamma dimensions taking into account the 42 test cases. They are sorted by mean γ_{33}.

Dosimetry model	σ (%)	mean γ_{22} (%)	mean γ_{33} (%)
TN-NOD-Ly-RGB	1.5	88.9	97.1
TN-NOD-Ly-RB	1.6	88.1	96.9
TN-NOD-Ly-RG	1.6	88.6	96.8
TN-OD-Ly-RGB	1.6	87.2	96.4
NOD-Ly-R	1.8	87.6	96.4
TN-OD-Ly-RB	1.8	86.4	96.2

Both taking into account the calibration uncertainty and the mean percentage of points with $\gamma_{<1}$ (3% 3mm), the dosimetry model comprising TN perturbation, with NOD as response, applying the lateral correction, and employing all three color channels (*i.e.*, the TN-NOD-Ly-RGB model) obtained the best results and was regarded as the most accurate model in the study.

The models which were among the ten best according to each of these dimensions: calibration uncertainty (σ), mean percentage of points with $\gamma_{<1}$ 2% 2mm (denoted as mean γ_{22}) and 3% 3mm (mean γ_{33}), both gamma dimensions taking into account the 42 test cases, are shown in Table 3.5. According to these criteria, there were six models. All of them use lateral corrections. Five of them are multichannel models, the other one is a single-channel (R) model. All the multichannel models use TN perturbations. Four out of six models incorporate the information of the non-irradiated film scan. All these models include the R color channel, which is the channel with the highest absorption [39].

Table 3.6: Comparison of models using paired differences of $\gamma_{<1}$(3% 3mm) values. The sample is composed of the 42 test cases. Median $\gamma_{<1}$ values are shown. Positive values indicate that model A obtained better results, and the opposite for negative values. *P*-values using the Wilcoxon signed-rank test are included.

Model A	Model B	$\gamma^A_{<1} - \gamma^B_{<1}$	*p*-value
OD-Ln-R	OD-Ly-R	-4.47	<0.01
OD-Ln-R	NOD-Ln-R	-3.05	<0.01
OD-Ln-R	NOD-Ly-R	-7.67	<0.01
OD-Ly-R	NOD-Ly-R	-1.27	<0.01
OD-Ly-R	NOD-Ln-R	1.58	0.05
NOD-Ln-R	NOD-Ly-R	-3.75	<0.01
NOD-Ly-R	MM-NOD-Ly-RGB	1.38	0.02
NOD-Ly-R	TN-NOD-Ly-RGB	-0.11	0.02
MM-NOD-Ln-RGB	MM-NOD-Ly-RGB	-2.47	<0.01
MM-NOD-Ly-RGB	MM-OD-Ly-RGB	0.22	0.47
MM-NOD-Ly-RGB	TN-NOD-Ly-RGB	-1.78	<0.01
TN-NOD-Ln-RGB	TN-NOD-Ly-RGB	-0.48	<0.01
TN-NOD-Ly-RGB	TN-OD-Ly-RGB	0.01	0.14

3.5 Statistical hypothesis testing

The model selection based on the uncertainty of the calibration using the plane-based method was confirmed as a convenient approach to compare film dosimetry models, in the sense that lower uncertainties are correlated with higher expected $\gamma_{<1}$ values in the gamma analysis. Both model selection and expected gamma approaches produced the same conclusions. An additional validation was conducted by testing the statistical significance of the gamma analysis results.

Models were compared using paired differences of $\gamma_{<1}$(3% 3mm) values. The sample was composed of the 42 benchmark cases. Since the probability density functions of $\gamma_{<1}$ differences were not normal in general, the Wilcoxon signed-rank test was used. Thus, the null hypothesis was that the median difference between paired observations was zero. *P*-values of less than 0.05 were considered as statistically significant, *i.e.*, indicative that one of the models provided higher $\gamma_{<1}$(3% 3mm) results. There are 990 possible pairwise comparisons with 44 models, for clarity only the 13 comparisons considered most relevant were studied. They are shown in Table 3.6. It was found that applying lateral corrections significantly improved the gamma analysis results both for single-channel (R) and triple-channel models. Using NOD as film-scanner response improved the results, significantly for single-channel (R) but not significantly for triple-channel models. The NOD-Ly-R model was found to be significantly better than the MM-NOD-Ly-RGB model, and the most accurate model according to the model

selection approach (*i.e.*, the TN-NOD-Ly-RGB model) also produced significantly higher $\gamma_{<1}$(3% 3mm) values than both the MM-NOD-Ly-RGB and the NOD-Ly-R models.

4 Conclusions

Under the scope of applicability defined by the limits under which the models were tested (*i.e.*, the films were irradiated with megavoltage radiotherapy beams, with doses in the dose range of 20-600 cGy, entire films were scanned, the functional form of the sensitometric curves was a polynomial and the calibration followed the plane-based method [14]), it was confirmed that using TN perturbations provided better results than using MM perturbation models. It was found that applying lateral corrections produced more accurate film doses. The same occurred if the information of the non-irradiated film was incorporated by scanning prior to irradiation, however this improvement did not produce significantly higher $\gamma_{<1}$(3% 3mm) values with triple-channel models. Scanning prior to irradiation and applying lateral corrections had a greater effect in terms of improving the accuracy of the results than the increase of the number of combined color channels, which, for some models, did not necessarily improve the results.

Among the models under comparison, the most accurate was found to be the dosimetry model comprising TN perturbation, with NOD as response, applying the lateral correction, and employing all three color channels (*i.e.*, the TN-NOD-Ly-RGB model).

5 Acknowledgments

The author would like to thank Domingo Granero, Primož Peterlin and Facundo Ballester for their very helpful comments during the preparation of this paper.

The author is co-founder of Radiochromic.com.

References

[1] W. Crijns, F. Maes, U. A. van der Heide, and F. V. den Heuvel. “Calibrating page sized Gafchromic EBT3 films”. *Medical Physics* 40.1 (2013), 012102 (13pp.)

[2] A. Niroomand-Rad, C. R. Blackwell, B. M. Coursey, K. P. Gall, J. M. Galvin, W. L. McLaughlin, A. S. Meigooni, R. Nath, J. E. Rodgers, and C. G. Soares. “Radiochromic film dosimetry: Recommendations of AAPM Radiation Therapy Committee Task Group 55”. *Medical Physics* 25.11 (1998), pp. 2093–2115.

[3] A. Rink, I. A. Vitkin, and D. A. Jaffray. "Energy dependence (75 kVp to 18 MV) of radiochromic films assessed using a real-time optical dosimeter". *Medical Physics* 34.2 (2007), pp. 458–463.

[4] C. Richter, J. Pawelke, L. Karsch, and J. Woithe. "Energy dependence of EBT-1 radiochromic film response for photon (10 kVp–15 MVp) and electron beams (6–18 MeV) readout by a flatbed scanner". *Medical Physics* 36.12 (2009), pp. 5506–5514.

[5] B. Arjomandy, R. Tailor, A. Anand, N. Sahoo, M. Gillin, K. Prado, and M. Vicic. "Energy dependence and dose response of Gafchromic EBT2 film over a wide range of photon, electron, and proton beam energies". *Medical Physics* 37.5 (2010), pp. 1942–1947.

[6] P. Lindsay, A. Rink, M. Ruschin, and D. Jaffray. "Investigation of energy dependence of EBT and EBT-2 Gafchromic film". *Medical Physics* 37.2 (2010), pp. 571–576.

[7] G. Massillon-JL, S. Chiu-Tsao, I. Domingo-Munoz, and M. Chan. "Energy Dependence of the New Gafchromic EBT3 Film:Dose Response Curves for 50 KV, 6 and 15 MV X-Ray Beams". *International Journal of Medical Physics, Clinical Engineering and Radiation Oncology* 1.2 (2012), pp. 60–65.

[8] H. Bekerat, S. Devic, F. DeBlois, K. Singh, A. Sarfehnia, J. Seuntjens, S. Shih, X. Yu, and D. Lewis. "Improving the energy response of external beam therapy (EBT) GafChromic dosimetry films at low energies (≤ 100 keV)". *Medical Physics* 41.2 (2014), 022101 (14pp.)

[9] O. Hupe and J. Brunzendorf. "A novel method of radiochromic film dosimetry using a color scanner". *Medical Physics* 33.11 (2006), pp. 4085–4094.

[10] S. Devic, N. Tomic, C. G. Soares, and E. B. Podgorsak. "Optimizing the dynamic range extension of a radiochromic film dosimetry system". *Medical Physics* 36.2 (2009), pp. 429–437.

[11] A. Micke, D. F. Lewis, and X. Yu. "Multichannel film dosimetry with nonuniformity correction". *Medical Physics* 38.5 (2011), pp. 2523–2534.

[12] T. J. McCaw, J. A. Micka, and L. A. DeWerd. "Characterizing the marker-dye correction for Gafchromic® EBT2 film: a comparison of three analysis methods". *Medical Physics* 38.10 (2011), pp. 5771–5777.

[13] R. R. Mayer, F. Ma, Y. Chen, R. I. Miller, A. Belard, J. McDonough, and J. J. O'Connell. "Enhanced dosimetry procedures and assessment for EBT2 radiochromic film". *Medical Physics* 39.4 (2012), pp. 2147–2155.

[14] I. Méndez, V. Hartman, R. Hudej, A. Strojnik, and B. Casar. "Gafchromic EBT2 film dosimetry in reflection mode with a novel plan-based calibration method". *Medical Physics* 40.1 (2013), 011720 (9pp.)

[15] I. Méndez, P. Peterlin, R. Hudej, A. Strojnik, and B. Casar. "On multichannel film dosimetry with channel-independent perturbations". *Medical Physics* 41.1 (2014), 011705 (10pp.)

[16] J. F. P. Azorín, L. I. R. García, and J. M. Martí-Climent. "A method for multichannel dosimetry with EBT3 radiochromic films". *Medical Physics* 41.6 (2014), 062101 (10pp.)

[17] S. J. van Hoof, P. V. Granton, G. Landry, M. Podesta, and F. Verhaegen. "Evaluation of a novel triple-channel radiochromic film analysis procedure using EBT2". *Physics in Medicine and Biology* 57.13 (2012), pp. 4353–4368.

[18] S. Devic, Y.-Z. Wang, N. Tomic, and E. B. Podgorsak. "Sensitivity of linear CCD array based film scanners used for film dosimetry". *Medical Physics* 33.11 (2006), pp. 3993–3996.

[19] H. Bouchard, F. Lacroix, G. Beaudoin, J.-F. Carrier, and I. Kawrakow. "On the characterization and uncertainty analysis of radiochromic film dosimetry". *Medical Physics* 36.6 (2009), pp. 1931–1946.

[20] C. Fiandra, U. Ricardi, R. Ragona, S. Anglesio, F. R. Giglioli, E. Calamia, and F. Lucio. "Clinical use of EBT model Gafchromic film in radiotherapy". *Medical Physics* 33.11 (2006), pp. 4314–4319.

[21] B. D. Lynch, J. Kozelka, M. K. Ranade, J. G. Li, W. E. Simon, and J. F. Dempsey. "Important considerations for radiochromic film dosimetry with flatbed CCD scanners and EBT GAFCHROMIC film". *Medical Physics* 33.12 (2006), pp. 4551–4556.

[22] L. Paelinck, W. D. Neve, and C. D. Wagter. "Precautions and strategies in using a commercial flatbed scanner for radiochromic film dosimetry". *Physics in Medicine and Biology* 52.1 (2007), pp. 231–242.

[23] M. Fuss, E. Sturtewagen, C. D. Wagter, and D. Georg. "Dosimetric characterization of GafChromic EBT film and its implication on film dosimetry quality assurance". *Physics in Medicine and Biology* 52.14 (2007), pp. 4211–4225.

[24] L. J. van Battum, D. Hoffmans, H. Piersma, and S. Heukelom. "Accurate dosimetry with GafChromic EBT film of a 6 MV photon beam in water: What level is achievable?" *Medical Physics* 35.2 (2008), pp. 704–716.

[25] M. Martišíková, B. Ackermann, and O. Jäkel. "Analysis of uncertainties in Gafchromic EBT film dosimetry of photon beams". *Physics in Medicine and Biology* 53.24 (2008), pp. 7013–7027.

[26] L. Menegotti, A. Delana, and A. Martignano. "Radiochromic film dosimetry with flatbed scanners: A fast and accurate method for dose calibration and uniformity correction with single film exposure". *Medical Physics* 35.7 (2008), pp. 3078–3085.

[27] S. Saur and J. Frengen. "GafChromic EBT film dosimetry with flatbed CCD scanner: A novel background correction method and full dose uncertainty analysis". *Medical Physics* 35.7 (2008), pp. 3094–3101.

[28] F. Girard, H. Bouchard, and F. Lacroix. "Reference dosimetry using radiochromic film". *Journal of Applied Clinical Medical Physics* 13.6 (2012). ISSN: 15269914.

[29] D. Poppinga, A. A. Schoenfeld, K. J. Doerner, O. Blanck, D. Harder, and B. Poppe. "A new correction method serving to eliminate the parabola effect of flatbed scanners used in radiochromic film dosimetry". *Medical Physics* 41.2 (2014), 021707 (8pp.)

[30] A. A. Schoenfeld, D. Poppinga, D. Harder, K.-J. Doerner, and B. Poppe. "The artefacts of radiochromic film dosimetry with flatbed scanners and their causation by light scattering from radiation-induced polymers". *Physics in Medicine and Biology* 59.13 (2014), pp. 3575–3597.

[31] D. Lewis, A. Micke, X. Yu, and M. F. Chan. "An efficient protocol for radiochromic film dosimetry combining calibration and measurement in a single scan". *Medical Physics* 39.10 (2012), pp. 6339–6350.

[32] M. Hutter. *Universal Artificial Intelligence: Sequential Decisions based on Algorithmic Probability*. Berlin: Springer, 2005.

[33] M. Li and P. M. Vitányi. *An introduction to Kolmogorov complexity and its applications*. 3rd. New York: Springer-Verlag, 2009.

[34] K. P. Burnham and D. R. Anderson. *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Science & Business Media, 2002.

[35] D. A. Low, W. B. Harms, S. Mutic, and J. A. Purdy. "A technique for the quantitative evaluation of dose distributions". *Medical Physics* 25.5 (1998), pp. 656–661.

[36] B. M. Clasie, G. C. Sharp, J. Seco, J. B. Flanz, and H. M. Kooy. "Numerical solutions of the gamma-index in two and three dimensions". *Physics in Medicine and Biology* 57.21 (2012), pp. 6981–6997.

[37] T. Künzler, I. Fotina, M. Stock, and D. Georg. "Experimental verification of a commercial Monte Carlo-based dose calculation module for high-energy photon beams". *Physics in Medicine and Biology* 54.24 (2009), pp. 7363–7377.

[38] B. Arjomandy, R. Tailor, L. Zhao, and S. Devic. "EBT2 film as a depth-dose measurement tool for radiotherapy beams over a wide range of energies and modalities". *Medical Physics* 39.2 (2012), pp. 912–921.

[39] S. Devic, N. Tomic, Z. Pang, J. Seuntjens, E. B. Podgorsak, and C. G. Soares. "Absorption spectroscopy of EBT model Gafchromic film". *Medical Physics* 34.1 (2007), pp. 112–118.

Paper 4

Grid patterns, spatial inter-scan variations and scanning reading repeatability in radiochromic film dosimetry

I. Méndez, Ž. Šljivić, R. Hudej, A. Jenko and B. Casar

Physica Medica 32.9 (2016), 1072-1081

Paper 4. Grid patterns, spatial inter-scan variations and scanning reading repeatability in radiochromic film dosimetry

I. Méndez, Ž. Šljivić, R. Hudej, A. Jenko and B. Casar

Abstract

Purpose: When comparing different scans of the same radiochromic film, several patterns can be observed. These patterns are caused by different sources of uncertainty, which affect the repeatability of the scanner. The purpose of this work was to study these uncertainties.
Methods: The variance of the scanner noise, as a function of the pixel position, was studied for different resolutions. The inter-scan variability of the scanner response was analyzed taking into account spatial discrepancies. Finally, the distance between the position of the same point in different scans was examined.
Results: The variance of noise follows periodical patterns in both axes, causing the grid patterns. These patterns were identified for resolutions of 50, 72 and 96 dpi, but not for 150 dpi. Specially recognizable is the sinusoidal shape with a period of 8.5 mm that is produced with 72 dpi. Inter-scan variations of the response caused systematic relative dose deviations larger than 1% in 5% of the red channel images, 9% of the green and 51% of the blue. No systematic deviation larger than 1% was found after applying response corrections. The initial positioning and the speed of the scanner lamp vary between scans.
Conclusions: Three new sources of uncertainty, which influence radiochromic film dosimetry with flatbed scanners, have been identified and analyzed in this work: grid patterns, spatial inter-scan variations and scanning reading repeatability. A novel correction method is proposed, which mitigates spatial inter-scan variations caused by deviations in the autocalibration of the individual Charge Coupled Device detectors.

1 Introduction

The system composed of radiochromic films and a flatbed scanner is the dosimeter of choice for many applications in radiology and radiation therapy[1]. This dosimetry system is affected by several sources of uncertainty. Some of them involve only the film: for example, the thickness variations of the active layer[2], the change in film darkening as a function of post-irradiation time [3], the influence of humidity and temperature [4, 5], the UV-induced polymerization [6], etc. Some other uncertainties are a consequence of the interaction of the characteristics of

both the film and scanner: for example, the lateral artifact [7, 8], Newton rings [9], the dependency with the orientation of the film on the scanner bed [10], the cross talk effect [8], the dependency on film-to-light source distance [11, 12], etc. Finally, other uncertainties are intrinsic to the scanner: for example, noise [13, 14], the inter-scan variability of the scanner response [11], warming-up of the lamp [15, 16], differences between color channels [17–20], etc.

Despite all those perturbations, GAFChromic films (Ashland Inc., Wayne, NJ) have been repeatedly found to be capable of delivering accurate dose measurements [20–23]. Still, to further improve the accuracy of the dosimetry system, thorough knowledge of its uncertainties is necessary.

GAFChromic EBT3 films were used in this study, in combination with the Epson Expression 10000XL scanner (Seiko Epson Corporation, Nagano, Japan). In the literature, the Epson Expression 10000XL scanner has been selected numerous times [3, 7, 8, 11, 24, 25] for radiochromic film dosimetry. In this work, the repeatability of this scanner has been examined. As a result, three new artifacts have been identified and analyzed: grid patterns, spatial inter-scan variations and scanning reading repeatability.

2 Methods and materials

GAFChromic EBT3 films from lot 06061401 were employed. They were irradiated with a Novalis Tx accelerator (Varian, Palo Alto, CA, USA). The darkening of the films was measured with an Epson Expression 10000XL scanner. The scanner was powered on 30 min before readings and five scans were taken to warm up its lamp. The films were placed on the center of the scanner with an opaque frame. To avoid the Callier effect [12, 24], a glass sheet, with a thickness of 3 mm, was placed on top of the films. They were scanned in portrait orientation (*i.e.*, the short side of the film parallel to the scanner lamp) and transmission mode. Images were acquired using the Epson Scan v3.49a software, in 48-bit RGB (16 bit per channel) format, while processing tools were turned off. Images were saved as TIFF files. Data analysis was performed with the R programming language [26].

2.1 Preliminary test

A film was placed at a depth of 11 cm in an IBA MULTICube phantom (IBA Dosimetry GmbH, Schwarzenbruck, Germany). Source-to-film distance was 100 cm. The film was irradiated with a step pattern composed of six stripes with doses of 0.25, 1, 4, 8, 2 and, again, 0.25 Gy. It was scanned ten consecutive times 24 h after irradiation, with a resolution of 72 dpi. The mean of the ten scans was calculated. For each scan and color channel, the difference image between the scan and the mean scan was also computed.

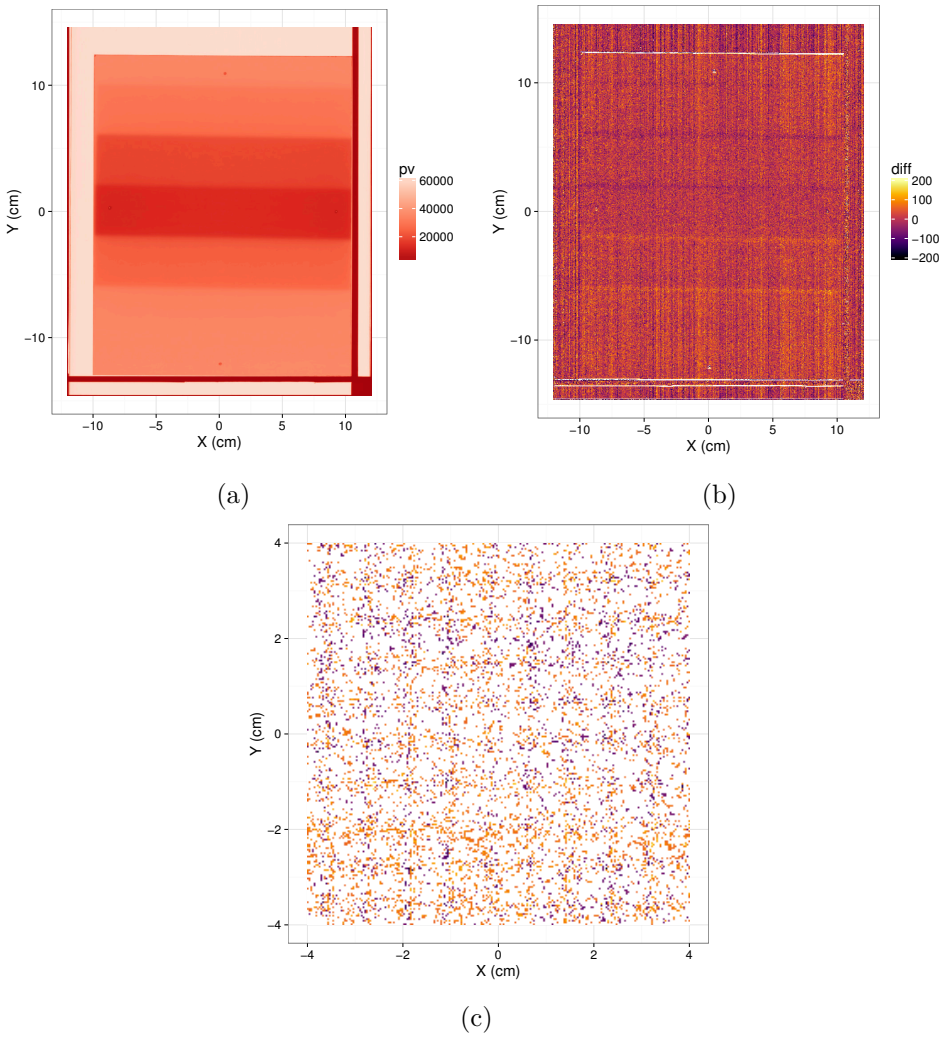


Figure 4.1: Pixel values (pv) in the red channel for a) one of the scans of the step pattern, b) the difference between this scan and the mean scan, excluding absolute differences larger than 200 pv, and c) a zoom of the difference image, excluding absolute differences larger than 200 pv or smaller than 60 pv.

Figure 4.1 shows pixel values in the red channel of one scan image as well as the difference between this scan and the mean scan. Several patterns can be observed in the difference image: for example, the edges of the steps generate thick lines approximately parallel to the scanner lamp (axis X). There are many thin linear patterns perpendicular to the scanner lamp (axis Y). In addition, there is a grid pattern, which can be better perceived in Figure 4.1c. These artifacts were present in most of the difference images. They were also found using the Epson Scan v3.41 software, as well as with another Epson Expression 10000XL scanner. The following tests were developed to analyze them.

2.2 Grid pattern

Four different setups were studied. In the first one, an unexposed film was scanned. In the second one, without the presence of the film, the light transmitted through the flattening glass sheet was measured, with the image referred to as white background. In the third one, the bed of the scanner, except for the calibration area, was covered with a black opaque plastic in order to avoid the transmission of light to the detectors; this was called the black background. In the last one, three previously irradiated film stripes were scanned; their dimensions were 20.3 cm × 4 cm and had received homogeneous doses of 100, 200, and 400 cGy, respectively. Each setup was scanned with resolutions of 50, 72, 96 and 150 dpi. While the Epson Expression 10000XL scanner has an optical resolution of 2400 dpi, these resolutions were regarded as the most commonly used for film dosimetry. For each resolution, 20 scans were taken.

For each of the four setups, resolution and color channel, the mean scan image was calculated. The difference between each scan and the corresponding mean image was obtained. Pixel value differences were grouped by column (X axis) and row (Y axis), while the mean absolute deviations (MADs) of the differences were computed. The MAD is a measure of statistical dispersion which is more robust to outliers than the standard deviation. If the sample is normally distributed, as was generally the case, the MAD is an estimator proportional to the standard deviation of the population. Hence, the objective of this test was to obtain the dispersion of the measures of the scanner (*i.e.*, the noise) as a function of the pixel position.

For the three irradiated stripes, relative dose uncertainties resulting from repeated scans were calculated. Dose uncertainties for each pixel were obtained as the product of the standard deviation of the pixel value, which can be determined from the MAD, times the derivative of the dose with respect to the pixel value.

Figure 4.2: Spatial inter-scan variability: setup of film fragments.

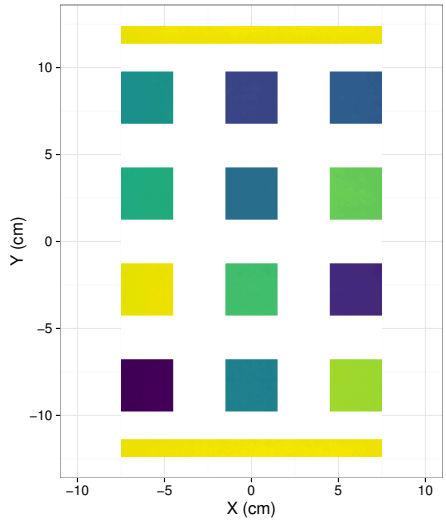


Figure 4.3: Spatial inter-scan variability: ROIs analyzed.

2.3 Spatial inter-scan variability

Measurements

Inter-scan variations of the scanner response produce discrepancies in the dose-response relationship between the calibration and subsequent scans, which, if not corrected, can cause important errors in film dosimetry [11]. One of the objectives of this test was to investigate spatial variations of the inter-scan repeatability. In order to do so, a film was divided in 16 fragments. Twelve of them, the central fragments, were 6.0 cm × 5.5 cm. The superior and inferior margins were 20.3 cm × 1.7 cm. The other two fragments, the lateral margins, were 1.2 cm × 22.0 cm.

Each central fragment was separately placed at a depth of 7 cm in a 14×30×30 cm^3 Plastic Water™phantom (Computerized Imaging Reference Systems Inc. Norfolk, VA, USA) at 100 cm SSD (source-to-surface distance). They were irradiated with a a 10 cm × 10 cm field, at 150, 400, 300, 100, 250, 50, 0, 75, 500, 750, 200 and 25 cGy (fragments 1-12). Doses were randomly distributed to prevent misleading patterns arising.

The film was reassembled, as shown in Figure 4.2, and scanned with resolutions of 50, 72, 96 and 150 dpi, four months after irradiation. Each resolution was scanned 20 consecutive times. For each resolution and color channel, the mean scan image was calculated.

Regions of interest (ROIs), with dimensions of 3 cm × 3 cm centered on each of the central fragments were selected, while two ROIs were centered on the superior and inferior margins with dimensions of 15 cm × 1 cm. They are shown in Figure 4.3. Only pixels contained in the ROIs were analyzed to avoid the edges of the fragments.

Corrections

Another objective of this test was to find the most accurate model to correct the inter-scan variability, taking into account spatial differences.

Even though, in clinical dosimetry, the reference dose-response relationship should be the sensitometric curve obtained during the calibration, in analysis of the inter-scan variations we can select any image or combination of images as reference. In this study, the reference image was considered to be the mean scan. Applying the correction to a scan image should reduce the differences between it and the reference. Several corrections were examined and two of them were finally chosen: the mean correction and the column correction. The superior and inferior margin ROIs, which were unexposed, were used as the reference material (Ref ROI) to derive the corrections.

The mean correction was calculated as follows:

$$M(i,j) = v(i,j)\left\langle \frac{M(i_{Ref}, j_{Ref})}{v(i_{Ref}, j_{Ref})} \right\rangle \tag{4.1}$$

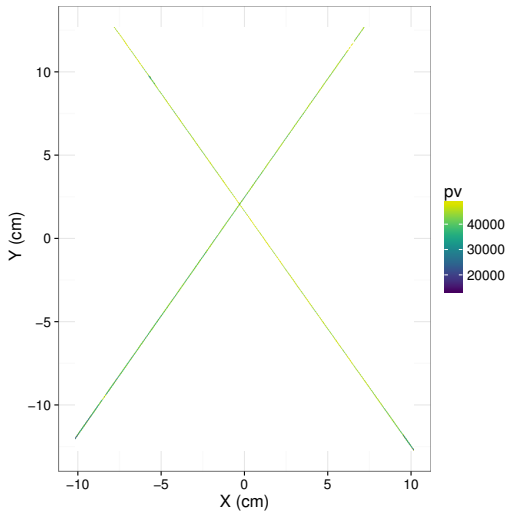


Figure 4.4: Scanning reading repeatability : the cross shape.

where (i, j) symbolizes the pixel position in the image (i is the row and j the column), M is the value of the pixel in the mean scan, v is the value in the scan being corrected, and (i_{Ref}, j_{Ref}) is a pixel in the Ref ROI. Therefore, the mean correction is the average of the factors applied to each of the pixels in the Ref ROI to obtain the values of the mean image. The mean correction is constant for every pixel of the scan, it is spatially invariant.

The column correction can be described as follows:

$$M(i,j) = v(i,j)\left\langle \frac{M(i_{Ref}, j)}{v(i_{Ref}, j)} \right\rangle \tag{4.2}$$

Thus, the column correction only averages the factors of the pixels in the Ref ROI which are in the column of the pixel being corrected. In this way, the deviations of the individual charge-coupled device (CCD) detectors are rectified. The column correction depends on the scan and on the position of the pixel in the scan, it is a spatial correction.

2.4 Scanning reading repeatability

A cross shape was drawn on a transparency sheet. It was placed on the center of the scanner under the flattening glass. Fifty scans were taken for each resolution, employing resolutions of 50, 72, 96 and 150 dpi. To select only the pixels of the cross shape, pixel values higher than 50000 were removed. All three color channels were combined in a single image. One of the scans can be seen in Figure 4.4.

The arms of the cross, which can be several pixels thick, were transformed into lines. To do this, the weighted mean column position of the pixels was

calculated for each arm separately and each row of the scan. Thus, each row of the scan was associated with the most likely positions of the cross shape, namely, two positions, one for each arm, with the exception of the point where both arms cross.

Although the inverse of the pixel value was employed as weight to compute the most likely positions, different weights were tested with negligible influence in the results. In an analogous fashion, row positions of the cross shape were associated with the columns of the scan. Additionally, the mean or reference cross for each resolution was computed by combining all the pixels of each scan, and calculating the weighted mean positions of the arms.

For each row and each column of the scan (or, equivalently, for each X and Y position), the distance, in each axis, between the most likely positions of the reference cross shape and of the cross shape of each scan was calculated.

3 Results

3.1 Grid pattern

Figure 4.5 plots the MAD of the differences in pixel value with respect to the mean image as a function of the column, resolution and color channel for the unexposed film, white and black backgrounds. For the sake of clarity, only 100 columns are included. Nevertheless, the same patterns with the same periodicity appear in the rows and in the rest of the columns.

To discard that the patterns found in the black background were caused by scattered light, measurements were repeated covering the scanner with opaque plastics, as well as preventing the transmission of light to the detectors with different opaque materials. Similar results were obtained in every case.

Figure 4.6 presents the dosimetric impact of the grid patterns. The dosimetric impact depends on the color channel, the dose, the scanner repeatability and the scanner resolution. For instance: for 100 cGy, 96 dpi and the blue channel, the relative dose uncertainty varied from approximately 6% to 8% as a function of the position, while for 400 cGy, 72 dpi and the red channel, it varied from approximately 0.4% to less than 0.6%. Even though the dosimetric impact was greater in the blue channel, tipically this channel is not used alone to convert pixel value to dose. Therefore, this impact will be reduced employing multichannel film dosimetry.

3.2 Spatial inter-scan variability

Table 4.1 and Table 4.2 contain the standard deviations of the relative differences between the mean scan image and the scan images with and without corrections, for each resolution and each color channel. Table 4.1 displays pixel value

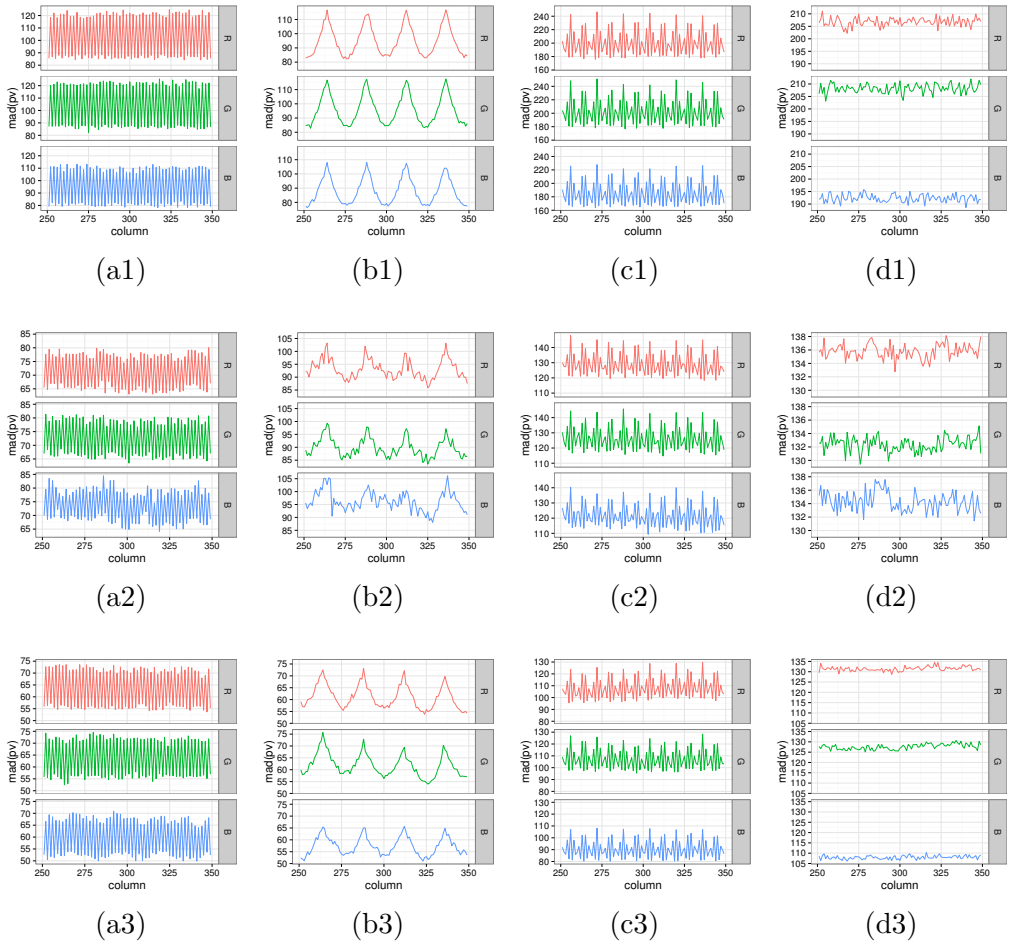


Figure 4.5: Mean absolute deviations (MADs) of the differences in pixel value with respect to the mean image as a function of the column for each setup (black background: 1, white background: 2, unexposed film: 3), resolution (50: a, 72: b, 96: c, 150: d) and color channel (red: R, green: G, blue: B).

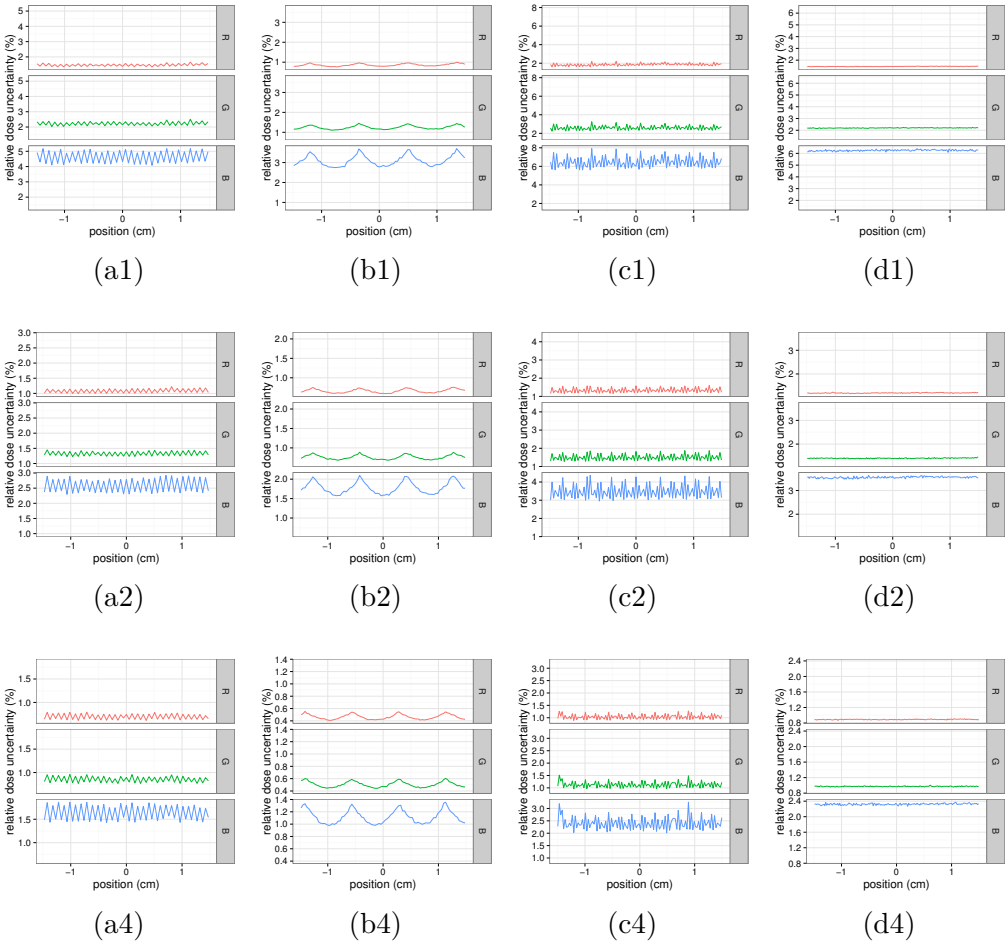


Figure 4.6: Relative dose uncertainties as a function of the position for each resolution (50: a, 72: b, 96: c, 150: d), dose (100 cGy: 1, 200 cGy: 2, 400 cGy: 4), and color channel (red: R, green: G, blue: B).

Table 4.1: Standard deviations of the relative pixel value differences (%) between the mean scan image and the scans with and without corrections. For ease of analysis, the uncertainties are not included; they were equal to or lower than 2.10^{-4}%.

Resolution (dpi)	50			72			96			150		
Color channel	R	G	B	R	G	B	R	G	B	R	G	B
No correction	0.237	0.201	0.263	0.256	0.215	0.231	0.406	0.339	0.406	0.441	0.373	0.444
Mean correction	0.226	0.192	0.228	0.220	0.188	0.218	0.397	0.333	0.398	0.425	0.359	0.425
Column correction	0.218	0.183	0.221	0.213	0.181	0.213	0.390	0.324	0.392	0.414	0.347	0.415

Table 4.2: Standard deviations of the relative dose differences (%) between the mean scan image and the scans with and without corrections. For ease of analysis, the uncertainties are not included; they were lower than 1.10^{-3}%.

Resolution (dpi)	50			72			96			150		
Color channel	R	G	B	R	G	B	R	G	B	R	G	B
No correction	1.0	1.9	7.8	1.1	1.9	6.9	1.6	2.8	10.5	1.8	3.1	10.8
Mean correction	0.9	1.6	6.2	0.9	1.5	6.2	1.6	2.7	10.2	1.7	3.0	10.2
Column correction	0.8	1.5	6.0	0.8	1.4	6.0	1.5	2.6	10.0	1.6	2.8	10.0

differences and Table 4.2 dose differences. No image darkening or trend in the inter-scan variability was noticed.

Figure 4.7 presents the density of the relative differences (in pixel value and dose) between the mean image and one of the images, both in the green channel and scanned with a resolution of 72 dpi. The map of the differences for this same scan is plotted in Figure 4.8. In this case, there was a bias or systematic deviation when no inter-scan correction was applied: the deviation with respect to zero for the mean relative pixel value difference could not be explained by the variance of pixel value differences. Systematic dose deviations were found in many other scans also, independently of the resolution. In 5% of the red channel images, 9% of the green and a 51% of the blue the mean relative dose difference from the reference image was greater than 1%. No systematic deviation larger than 1% was found among the corrected images.

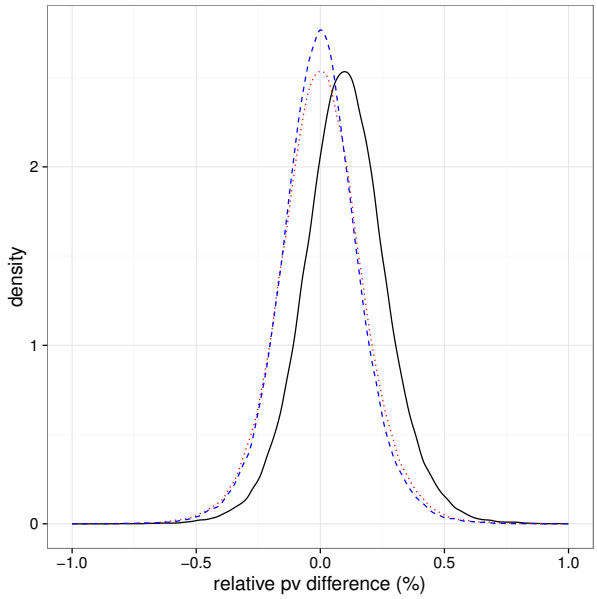
3.3 Scanning reading repeatability

The mean distance between pixels in the reference cross shape and the same pixels in each of the scans is presented in Figure 4.9. The variations were considered negligible and were, presumably, caused by noise in the X axis. They were not negligible in the Y axis. The distance in the Y axis, as a function of the Y position of the pixel in the reference image, is shown in Figure 4.10. The signals were noisy, and local polynomial regression fitting was applied to smooth them. Even though 50 scans are represented in this figure, many lines are overlapped. The initial distance in the Y axis between the reference cross and the cross in each scan is neither zero nor unique; it seems to have a set of possible discrete values. Furthermore, this distance does not remain constant, but approximately increases linearly with the lamp movement. Meanwhile, rather than a continuous of possible slopes, a discrete set was found. Both Figure 4.9 and Figure 4.10 were obtained from the scans with a resolution of 72 dpi. Still, all the other resolutions produced similar results.

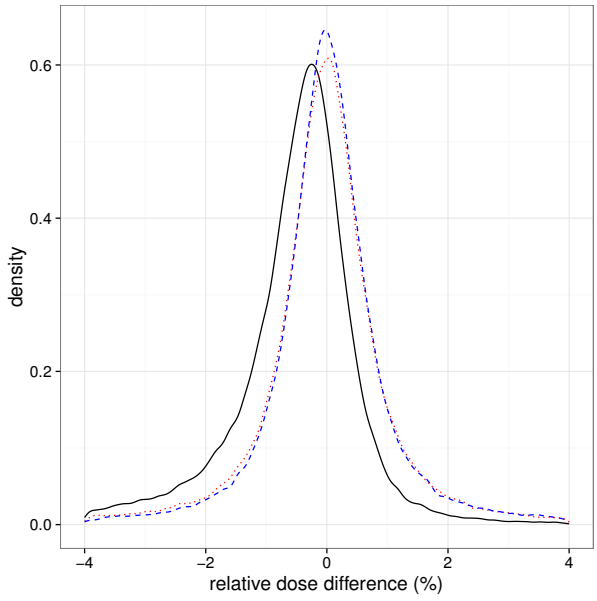
4 Discussion

4.1 Grid pattern

Measurements of the scanner are affected by noise. It is well known [16] that the variance of the noise depends on the resolution of the scanner: the larger the resolution, then the larger the variance. However, this variance is not constant throughout the entire scanner bed. For the scanner and scanning software being studied, periodical patterns in both axes have been found using resolutions of 50, 72 and 96 dpi. These patterns are independent of films: they even appear in the absence of transmitted light.



(a)



(b)

Figure 4.7: Density of the relative differences , a) in pixel value and b) in dose, between the mean image and one of the images, both in the green channel and scanned with a resolution of 72 dpi. The solid line represents the differences without any correction, while the dotted line applies to the mean correction and the dashed line applies to the column correction.

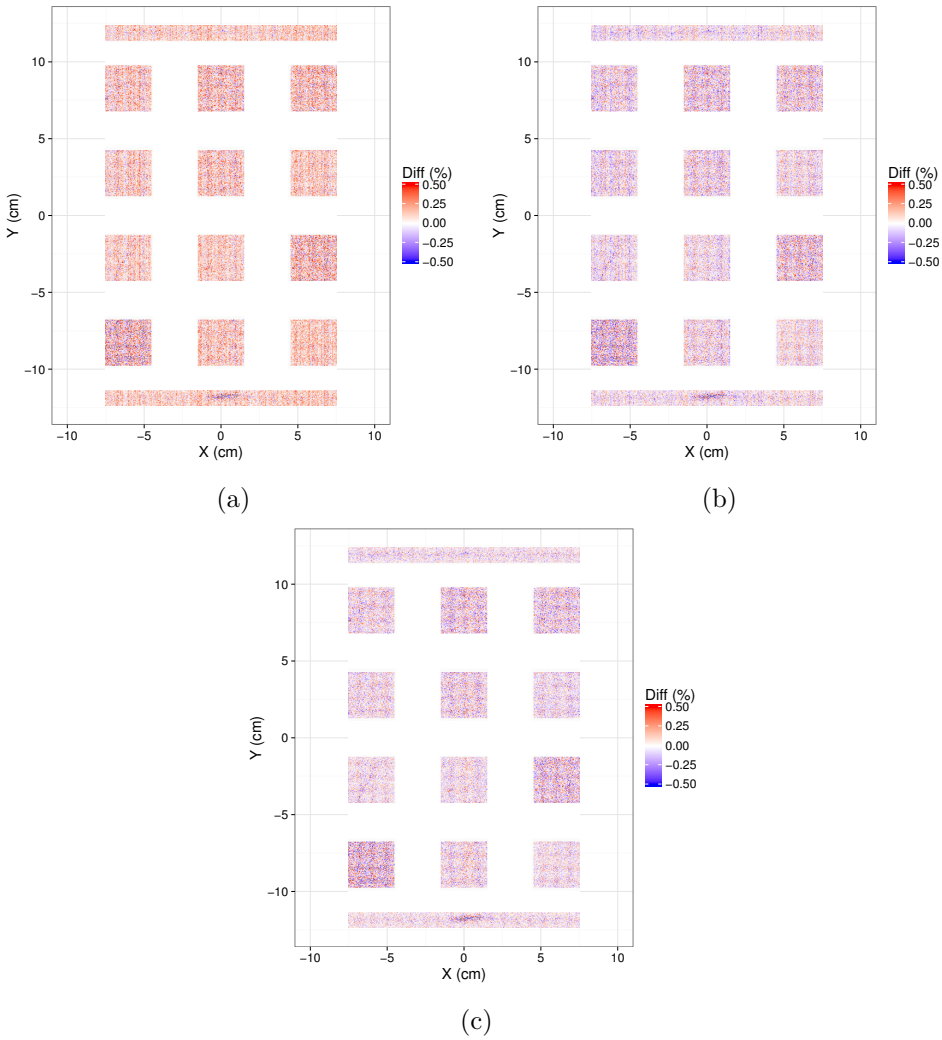
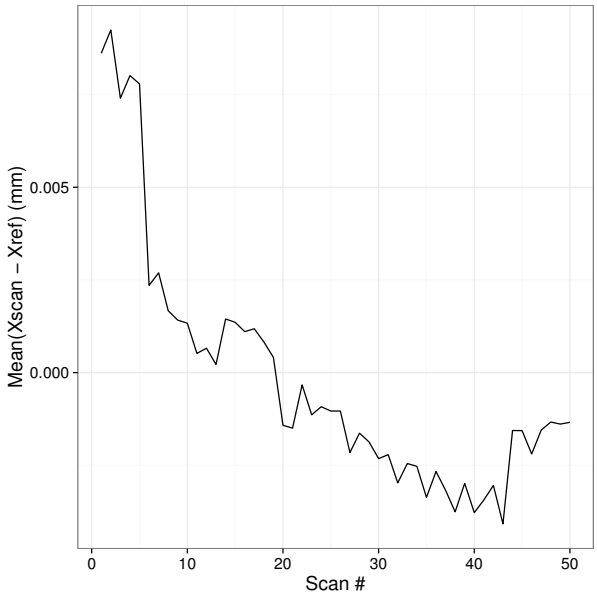
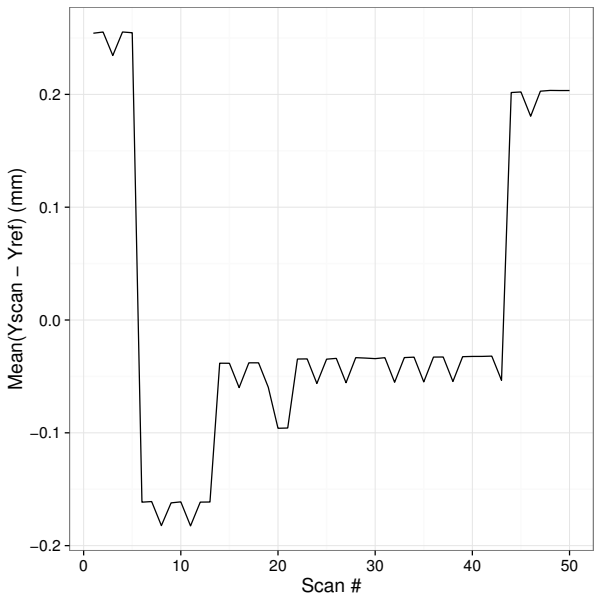


Figure 4.8: Distribution of the relative differences (%) between the mean image and one of the images, both in the green channel and scanned with a resolution of 72 dpi: a) without any correction, b) with mean correction, and c) with column correction.

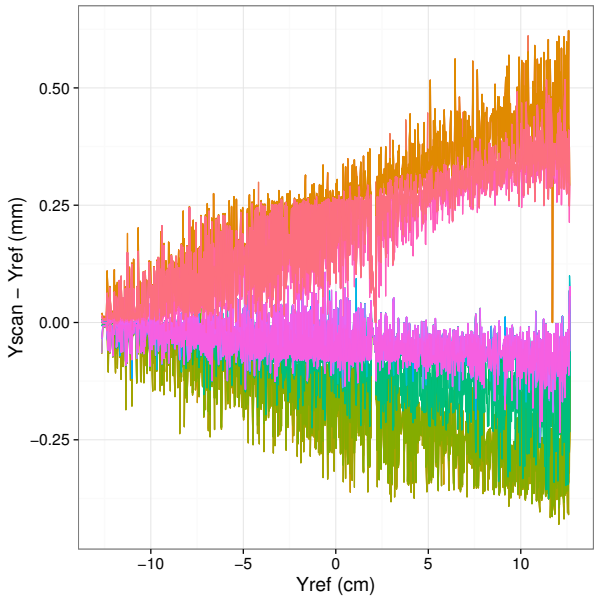


(a)

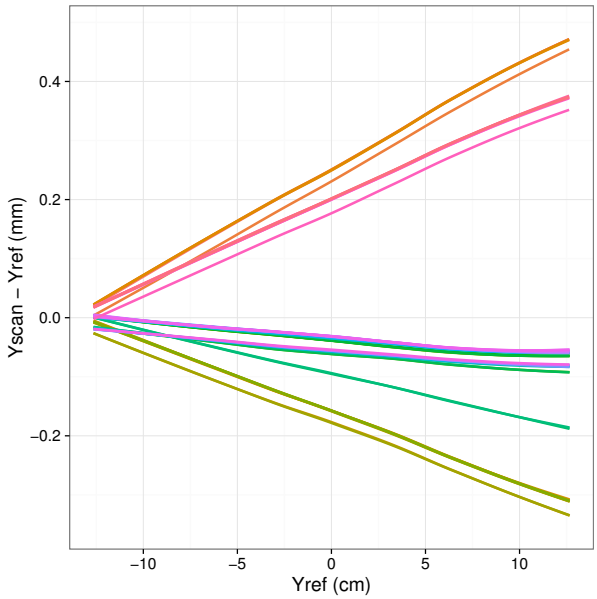


(b)

Figure 4.9: Mean distance between the position of the cross shape in the reference and in each of the scans: a) X axis, b) Y axis.



(a)



(b)

Figure 4.10: Distance in the Y axis between pixels of the cross shape in the reference and in each of the scans, as a function of the reference Y position: a) raw differences, b) smoothed differences. Different scans are displayed with different colors.

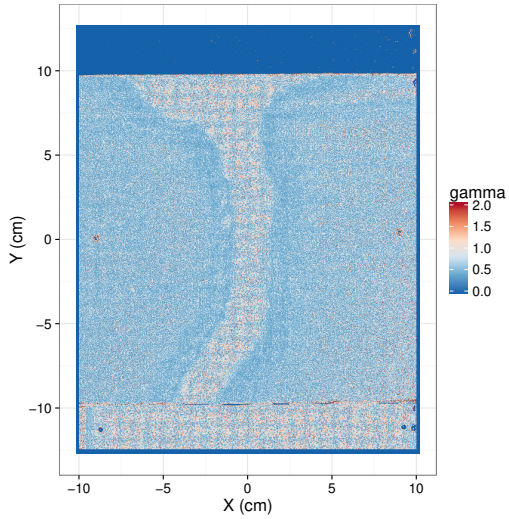


Figure 4.11: Grid pattern in gamma comparison.

The dosimetric impact of the grid patterns depends on the slope of the sensitometric curve (and, consequently, on the dose and the color channel), the scanner repeatability and the noise variance, which in turn depends on the scanner resolution. Still, the mean dose uncertainty was at least two times greater than the difference between the maximum and the minimum dose uncertainty, for each case under study. This can explain why grid patterns are rarely detected. Nevertheless, more important than the amplitude of the differences is their periodicity, which can occasionally produce misleading grid artifacts in film dose distributions or gamma index comparisons which could have clinical implications. An example of a gamma analysis, which is affected by the grid pattern, is presented in Figure 4.11. It is a gamma 1% 1 mm comparing the dose distributions calculated with two different multichannel dosimetry models, the only difference being the shape of the probability density function (pdf) of the perturbation term [19]. The film was scanned with 72 dpi. With this resolution, the pattern has a sinusoidal shape with a period of 8.5 mm, which makes it particularly apparent.

Devic *et al* [27] proposed scanning at a high resolution (*e.g.*, with 150 dpi) and downscale to obtain the resolution of interest. In this way, the standard deviations associated with the average pixel values can be computed. In light of the results of the present research, this approach offers the additional benefit of preventing grid artifacts.

The spatial variation of the pdf of pixel value differences between repeated scan images determines both grid patterns and spatial inter-scan variability. The pdf variance causes grid patterns, and the pdf mean causes the spatial inter-scan variability.

4.2 Spatial inter-scan variability

Applying the mean correction reduced the differences between the scan images and the mean scan image for each color channel and resolution. A larger reduction was achieved by applying the column correction.

The mean correction is equivalent to the correction proposed by Lewis and Devic [11], who recommended the use of an unexposed film piece as reference for the scanner response in each scan image. We support this recommendation, as neglecting this correction can cause systematic errors in the determination of the dose with radiochromic films. Additionally, the response correction can be enhanced including one or several pieces irradiated with known doses to rescale the sensitometric curves (*e.g.*, using the efficient protocol for radiochromic film dosimetry proposed by Lewis *et al* [21]).

Even though these methods mitigate the inter-scan variability of the scanner, they neglect spatial discrepancies in the repeatability. The column correction method presented in this study mitigates the spatial inter-scan variations caused by deviations in the autocalibration of the individual CCD detectors with respect to their reference state. This method is superior to the mean correction method reducing response inter-scan variations while also removing the systematic errors caused by these variations.

Even though this work employed the mean scan image as reference, as long as the Ref ROI stays in the same position between scans, any other scan or scan average could be used as reference for the correction. If the reference is the average of the scans taken for the calibration, employing the average of repeated scans in subsequent cases should reduce discrepancies in the dose-response relationship. Still, dosimetrically relevant errors caused by scans with large systematic deviations cannot be excluded. Thus, any average of scans should also be corrected using either the mean or the column correction.

Several other correction methods were tested and discarded in the preparation of this work. Some of them were aimed at reducing possible spatial inter-scan variations present in the axis perpendicular to the scanner lamp. None of them improved the results achieved with the column correction. As a consequence, spatial inter-scan variations in this axis were considered to be negligible. Nevertheless, it was observed that they were not negligible in the initial five scans, which were employed for warming up the scan lamp, as can be seen in Figure 4.12. This image corresponds to the red channel of one of the warm up scans after applying column corrections; the resolution of the image was 72 dpi. Spatial inter-scan variability in both axes was frequent in the warm up scans, which confirms the importance of warming up the scanner lamp before using it or after long pauses [16].

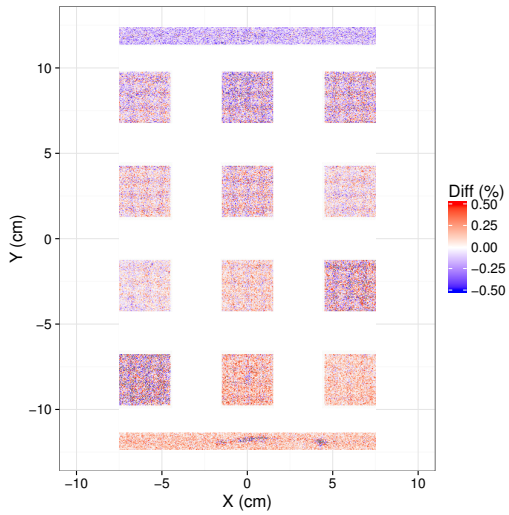


Figure 4.12: Spatial inter-scan variation in the axis perpendicular to the scanner lamp.

4.3 Scanning reading repeatability

The initial reading positioning (in the Y axis) and the reading speed of the scanner lamp vary between scans. Differences in the initial positioning were found to be less than 0.1 mm. However, 20 cm away from the initial position, the variations in the speed produced differences of 0.7 mm. Calculating the average result of several scans reduces the noise; nevertheless, the scanning reading repeatability blurs the resulting image. The blurring increases with the distance from the initial position of the scanning.

The distribution of reading positioning differences can be conservatively estimated as a uniform distribution with a support of length $\Delta_Y = 0.1 + 0.003y$, where y is the distance from the initial reading positioning in mm. The dosimetric impact of this distribution depends on the dose gradient. For instance, let us consider a film irradiated with a 60° Enhanced Dynamic Wedge field of dimensions 20×20 cm^2, which has been scanned several times. Excluding penumbras and out of field areas, the maximum relative dose difference between two scans would be 2.0%, which corresponds to a point 25 cm away from the initial reading positioning in the extreme of the wedge with the lowest dose. The relative dose uncertainty associated with this point according to the uniform distribution would be 0.6%, and the uncertainty of the reading positioning would be 0.2 mm. This conservative estimation of the maximum relative dose uncertainty would be substantially reduced simply by placing the lowest dose of the wedge at the beginning of the scanning path. The scanning reading repeatability should also be considered when films are used to measure penumbras. For instance, let us

consider a beam profile that is an ideal step function with zero dose out of the field. The penumbra, which is defined as the distance between the points with 20% and 80% of the field dose, measures 0 mm for the step function. However, if it is calculated employing the mean of several scans of this field, and is situated 25 cm away from the initial reading positioning, it will measure 0.5 mm. While it is true that this value is a conservative estimation of an ideal worst case scenario, it is a value comparable with the maximum broadening of the penumbra observed by Agostinelli *et al* [28] for a type 31014 PinPoint ion chamber (PTW, Freiburg, Germany), which was 0.6 mm. Nevertheless, the broadening of the penumbra when using radiochromic films can be prevented by employing single scans for the measurements.

5 Conclusions

For the scanner and scanning software under study, three new sources of uncertainty in radiochromic film dosimetry have been identified and analyzed: the grid pattern, spatial inter-scan variations and scanning reading repeatability.

The grid patterns appear because the variance of noise is not constant throughout the entire scanner bed: it follows periodical patterns in both axes. These patterns have been identified using resolutions of 50, 72 and 96 dpi. The mean dose uncertainty due to noise and scanner repeatability was found to be at least two times greater than the difference between the maximum and the minimum dose uncertainty caused by the grid patterns, which explains why grid patterns are usually undetected. However, they can produce misleading grid artifacts in film dose distributions or gamma comparisons, with potential clinical implications.

Inter-scan variations produce discrepancies in the dose-response relationship between the calibration and subsequent scans. Response correction methods mitigate these variations and eliminate the systematic errors. In this work, a novel correction method has been proposed to reduce inter-scan variations addressing the deviations in the response of individual CCD detectors with respect to their reference state.

The initial positioning (in the Y axis) and the speed of the scanner lamp vary between scans. The differences in the initial positioning were found negligible; however, they increase with the distance from the initial position due to the variations in reading speed. As a consequence, average scans are less accurate at the end of the scanning reading than at the beginning. Given the submillimetric scale of the positioning uncertainty, the dosimetric impact is usually negligible. Still, in some measurements this uncertainty can be relevant and actions should be taken to reduce it.

6 Acknowledgments

The authors would like to thank Primož Peterlin and Juan José Rovira for their contributions to this work.

One of the authors (I.M.) is co-founder of Radiochromic.com.

References

[1] S. Devic. "Radiochromic film dosimetry: past, present, and future". *Physica medica* 27 (2011), pp. 122–134.

[2] B. Hartmann, M. Martišíková, and O. Jäkel. "Technical Note: Homogeneity of Gafchromic EBT2 film". *Medical Physics* 37.4 (2010), pp. 1753–1756.

[3] C. Andrés, A. del Castillo, R. Tortosa, D. Alonso, and R. Barquero. "A comprehensive study of the Gafchromic EBT2 radiochromic film. A comparison with EBT". *Medical Physics* 37.12 (2010), pp. 6271–6278.

[4] A. Rink, D. F. Lewis, S. Varma, I. A. Vitkin, and D. A. Jaffray. "Temperature and hydration effects on absorbance spectra and radiation sensitivity of a radiochromic medium". *Medical physics* 35.10 (2008), pp. 4545–4555.

[5] F. Girard, H. Bouchard, and F. Lacroix. "Reference dosimetry using radiochromic film". *Journal of Applied Clinical Medical Physics* 13.6 (2012). ISSN: 15269914.

[6] A. Niroomand-Rad, C. R. Blackwell, B. M. Coursey, K. P. Gall, J. M. Galvin, W. L. McLaughlin, A. S. Meigooni, R. Nath, J. E. Rodgers, and C. G. Soares. "Radiochromic film dosimetry: Recommendations of AAPM Radiation Therapy Committee Task Group 55". *Medical Physics* 25.11 (1998), pp. 2093–2115.

[7] A. A. Schoenfeld, D. Poppinga, D. Harder, K.-J. Doerner, and B. Poppe. "The artefacts of radiochromic film dosimetry with flatbed scanners and their causation by light scattering from radiation-induced polymers". *Physics in Medicine and Biology* 59.13 (2014), pp. 3575–3597.

[8] L. van Battum, H. Huizenga, R. Verdaasdonk, and S. Heukelom. "How flatbed scanners upset accurate film dosimetry". *Physics in medicine and biology* 61.2 (2015), p. 625.

[9] R. Dreindl, D. Georg, and M. Stock. "Radiochromic film dosimetry: Considerations on precision and accuracy for EBT2 and EBT3 type films". *Zeitschrift für Medizinische Physik* 24.2 (2014), pp. 153–163.

[10] M. J. Butson, T. Cheung, and P. Yu. "Evaluation of the magnitude of EBT Gafchromic film polarization effects". *Australasian Physics & Engineering Sciences in Medicine* 32.1 (2009), pp. 21–25.

[11] D. Lewis and S. Devic. "Correcting scan-to-scan response variability for a radiochromic film-based reference dosimetry system". *Medical physics* 42.10 (2015), pp. 5692–5701.

[12] A. L. Palmer, D. A. Bradley, and A. Nisbet. "Evaluation and mitigation of potential errors in radiochromic film dosimetry due to film curvature at scanning". *Journal of Applied Clinical Medical Physics* 16.2 (2015).

[13] H. Bouchard, F. Lacroix, G. Beaudoin, J.-F. Carrier, and I. Kawrakow. "On the characterization and uncertainty analysis of radiochromic film dosimetry". *Medical Physics* 36.6 (2009), pp. 1931–1946.

[14] S. J. van Hoof, P. V. Granton, G. Landry, M. Podesta, and F. Verhaegen. "Evaluation of a novel triple-channel radiochromic film analysis procedure using EBT2". *Physics in Medicine and Biology* 57.13 (2012), pp. 4353–4368.

[15] L. Paelinck, W. D. Neve, and C. D. Wagter. "Precautions and strategies in using a commercial flatbed scanner for radiochromic film dosimetry". *Physics in Medicine and Biology* 52.1 (2007), pp. 231–242.

[16] B. Ferreira, M. Lopes, and M. Capela. "Evaluation of an Epson flatbed scanner to read Gafchromic EBT films for radiation dosimetry". *Physics in medicine and biology* 54.4 (2009), p. 1073.

[17] A. Micke, D. F. Lewis, and X. Yu. "Multichannel film dosimetry with nonuniformity correction". *Medical Physics* 38.5 (2011), pp. 2523–2534.

[18] R. R. Mayer, F. Ma, Y. Chen, R. I. Miller, A. Belard, J. McDonough, and J. J. O'Connell. "Enhanced dosimetry procedures and assessment for EBT2 radiochromic film". *Medical Physics* 39.4 (2012), pp. 2147–2155.

[19] I. Méndez, P. Peterlin, R. Hudej, A. Strojnik, and B. Casar. "On multi-channel film dosimetry with channel-independent perturbations". *Medical Physics* 41.1 (2014), 011705 (10pp.)

[20] J. F. P. Azorín, L. I. R. García, and J. M. Martí-Climent. "A method for multichannel dosimetry with EBT3 radiochromic films". *Medical Physics* 41.6 (2014), 062101 (10pp.)

[21] D. Lewis, A. Micke, X. Yu, and M. F. Chan. "An efficient protocol for radiochromic film dosimetry combining calibration and measurement in a single scan". *Medical Physics* 39.10 (2012), pp. 6339–6350.

[22] A. L. Palmer, A. Nisbet, and D. Bradley. "Verification of high dose rate brachytherapy dose distributions with EBT3 Gafchromic film quality control techniques". *Physics in medicine and biology* 58.3 (2013), p. 497.

[23] I. Méndez. "Model selection for radiochromic film dosimetry". *Physics in medicine and biology* 60.10 (2015), p. 4089.

[24] D. Lewis and M. F. Chan. "Correcting lateral response artifacts from flatbed scanners for radiochromic film dosimetry". *Medical physics* 42.1 (2015), pp. 416–429.

[25] M. Martišíková, B. Ackermann, and O. Jäkel. "Analysis of uncertainties in Gafchromic EBT film dosimetry of photon beams". *Physics in Medicine and Biology* 53.24 (2008), pp. 7013–7027.

[26] R Core Team. *R: A Language and Environment for Statistical Computing*. ISBN 3-900051-07-0. R Foundation for Statistical Computing. Vienna, Austria, 2012. URL: `http://www.R-project.org/`.

[27] S. Devic, N. Tomic, and D. Lewis. "Reference radiochromic film dosimetry: review of technical aspects". *Physica Medica* 32.4 (2016), pp. 541–556.

[28] S. Agostinelli, S. Garelli, M. Piergentili, and F. Foppiano. "Response to high-energy photons of PTW31014 PinPoint ion chamber with a central aluminum electrode". *Medical physics* 35.7 (2008), pp. 3293–3301.