

## Research Article

# A Probability Approach to the Study on Uncertainty Effects on Gamma Index Evaluations in Radiation Therapy

Francisco Cutanda Henríquez<sup>1</sup> and Silvia Vargas Castrillón<sup>2</sup>

<sup>1</sup> Servicio de Medicina Nuclear, Hospital General Universitario Gregorio Marañón, Calle Doctor Esquerdo, 46, 28007 Madrid, Spain

<sup>2</sup> Laboratorio de Metrología de Radiaciones Ionizantes, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas, Avenida Complutense, 22, 28040 Madrid, Spain

Correspondence should be addressed to Francisco Cutanda Henríquez, francisco.cutanda@salud.madrid.org

Received 21 March 2010; Accepted 4 November 2010

Academic Editor: Sivabal Sivaloganathan

Copyright © 2011 F. Cutanda Henríquez and S. Vargas Castrillón. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Two datasets of points of known spatial positions and an associated absorbed dose value are often compared for quality assurance purposes in External Beam Radiation Therapy (EBRT). Some problems usually arise regarding the pass fail criterion to accept both datasets as close enough for practical purposes. Instances of this kind of comparisons are fluence or dose checks for intensity modulated radiation therapy, modelling of a treatment unit in a treatment planning system, and so forth. The gamma index is a figure of merit that can be obtained from both datasets; it is widely used, as well as other indices, as part of a comparison procedure. However, it is recognized that false negatives may take place (there are acceptable cases where a certain number of points do not pass the test) due in part to computation and experimental uncertainty. This work utilizes mathematical methods to analyse comparisons, so that uncertainty can be taken into account. Therefore, false rejections due to uncertainty do not take place and there is no need to expand tolerances to take uncertainty into account. The methods provided are based on the rules of uncertainty propagation and help obtain rigorous pass/fail criteria, based on experimental information.

## 1. Introduction

Modern radiation therapy aims at a high level of accuracy and, as a consequence, becomes more demanding regarding quality assurance checks (even patient-specific checks) and measurement and computation performance. The use of comparisons of two datasets consisting of a sample of measured or computed absorbed dose points covering the treatment field or a patient tomographic slice is frequently performed on a routine basis. Therefore, the acceptance method should be both straightforward and reliable.

Traditionally, treatment goals in radiation therapy were achieved by choosing several directions around the patient so that the dose from all the beams was conformed to the target volume, sparing healthy tissues. Nowadays, it is possible to improve the homogeneity of absorbed dose in the planning target volume (PTV) and reduce the absorbed dose to healthy organs using several fields of non-uniform intensity (IMRT)

designed to combine in an optimised dose distribution inside the patient [1, 2].

The process is more complex than the one involved in conventional radiation therapy. The way the different beam orientations are combined could lead to practical problems, due to several issues: small and elongated beams are used, there are high dose gradients inside the fields, some features of the linear accelerator could have a noticeable effect, and treatment planning computation could not be accurate enough. These issues can make a particular plan unsuitable for treatment, and this is the reason why a comprehensive quality control of the technique and checks for each plan are often recommended [3, 4].

Two main types of patient specific checks have been recommended in the literature [3, 5].

(1) The first one consists on recomputing the plan substituting the representation of a suitable phantom for the patient representation and obtaining the 2D dose

distribution on several planes inside the phantom. Radiographic or radiochromic film is inserted in the phantom in the same positions where the 2D doses were computed, and it is irradiated with the whole treatment. These films are scanned with an appropriate device and compared with the computed dose planes; this is a way to check the combined dose distribution. Some 3D measurement devices are also available [6–9].

(2) Irradiations are carried out for each beam, with the film or 2D detector placed perpendicular to the beam direction. The dose distributions have been previously computed with the treatment planning system, and a corresponding set of 2D computed dose distributions (or fluence maps) has to be compared with the measured ones. A check of each fluence map is obtained with this technique.

In either case, a comparison of two datasets with a great number of points has to be performed. Similar situations arise when commissioning a treatment planning system, since computation results have to be checked against measurement results. This leads to the following problem.

Given two arrays of values (absorbed dose), maybe with different spacing, find a convenient criterion to decide whether or not they can be considered as coincident for practical purposes. The dose distributions to be expected in radiation therapy can have sharp gradients in the field boundary, and possibly also inside the field, where the dose is not homogeneous. Wherever a sharp gradient is present, the result could be affected by geometrical errors (i.e., error in the position of a collimator leaf, error in the computation of the dose on the edge of the collimator leaf) and the check method should be able to cope with this. A small geometrical error is considered acceptable, but a direct comparison of the measured and reference dose in this area could result in a value out of dose tolerance. This is the reason why acceptance criteria based on distance to agreement (DTA) were developed [3–5]. DTA is the distance from the measured point to the nearest point in the reference dose distribution with the same dose. DTA tolerances are usually set for penumbra regions (field edges), and tolerances based on absorbed dose differences are used for homogeneous regions inside or outside the field. Unfortunately, there is no reasonable criterion as to whether dose difference or DTA should be used for points inside a modulated field, because there could be gradients of very different magnitude.

A solution was proposed by Low et al. [10, 11], involving the computation of a single figure of merit for the quality of the match. It has become the method of choice for acceptance of IMRT plans. It involves an artificial distance in a 3D dose space. If the dose difference tolerance is  $\Delta D$  and the spatial tolerance is  $\Delta R$ , then the gamma index is

$$\gamma = \min \sqrt{\frac{D^2}{\Delta D^2} + \frac{R^2}{\Delta R^2}}, \quad (1)$$

where  $D$  and  $R$  are the dose difference and distance to the point in the reference dataset where the square root would reach a minimum. This minimum could be an interpolated point.

A point passes the check if this index is less or equal than to 1.  $\Delta D$  and  $\Delta R$  are no longer strict tolerances: dose difference could be greater than  $\Delta D$  for a point passing the gamma test; DTA could also be greater than  $\Delta R$  for a point with a gamma less than 1; although if dose difference is greater than  $\Delta D$  and DTA greater than  $\Delta R$  at the same point, the gamma test fails [10]. At the same time, the absolute value of  $\gamma$  at a point where the test is not passed is a measure of the severity of the failure.

The gamma index can be easily generalized to a 3D comparison, if DTA is computed with a 3D search [12]. A gamma filter method developed by Depuydt et al. [13] helps improve computation efficiency at the expense of not obtaining gamma values, but just checking whether or not every point is within tolerance.

It is widely acknowledged that in few occasions measured and computed datasets pass this gamma test for every measurement point, and it is customary to allow for some percentage of points failing the test [3–5, 14]. In practice, the pass rate is checked, the percentage of points in the reference dataset passes the test, and the tolerance for this rate is set according to previous experience. Therefore, the occurrence of failing points does not mean that the plan has to be rejected. This is the reason to accept a pass rate that could be less than 100%. However, there are no other grounds to accept this tolerance in pass rate, but empirical evidence, unless experimental uncertainty for the check procedure is somehow taken into account and propagated to the test indices. Basran and Woo [14] show their method to set the acceptance pass rate. They check their history of previous checks, their pass rates and whether they have been accepted or not in order to find the pass rate value corresponding to a 95% confidence. This is a purely empirical method, that ensured self-consistency, but it does not address the causes of the failing points.

Palta et al. [15] proposed a method to set tolerances in the process of commissioning, according to the observed variability. This recommendation was included in the recent report by AAPM Task Group 119 [3], to account indirectly for uncertainty in the tolerance levels. In this case, commissioning tests provide experience about variability of results that can be attributed to the experimental procedure. Analysis of the results can help set expanded tolerance levels of acceptance pass rates.

In other kinds of comparisons (like comparisons between computed datasets), statistical information is not available and a decision about the percentage of failing points that can be tolerated has to be based on other considerations.

In this work, a novel method is presented that modifies the gamma index check, introducing uncertainty features into its computation. This method has some interesting properties: first, it is a direct propagation of experimental uncertainty, allowing for uncertainty analysis. Second, tolerance levels are not modified because of uncertainty; using this method, tolerance levels can be set to values close to the accuracy actually sought for. And third, experimental devices, computations, and their uncertainties are characterized by simple and physically meaningful parameters. Therefore, the study of the check procedure is

reduced to the *a priori* study of the devices and algorithms involved.

## 2. Methods and Materials

**2.1. Theoretical Background.** Each of the datasets has been represented as several arrays of random variables. Their mean values are the values in the dataset, labelled as small-case letters with subscripts for their position and superscripts for the dataset:  $x_{ij}^r, d_{kl}^r, \dots$ . There is one array for each of the spatial coordinates and another for the dose. Test and reference datasets are allowed to have different values of uncertainty and array spacing, but spatial uncertainty within one of the two datasets is supposed to be isotropic. Therefore, spatial uncertainty is described by one parameter for each dataset:  $\sigma_s^t$  for the test dataset and  $\sigma_s^r$  for the reference dataset. Along this work, the symbol  $\sigma$  stands for standard uncertainty (one standard deviation). Similarly,  $\sigma_d^t$  and  $\sigma_d^r$  are dose uncertainties.  $\Delta R$  and  $\Delta D$  are check tolerances for the comparison between both datasets (not related to dataset uncertainty). In the next paragraphs, the computation algorithm for the comparison with uncertainty evaluation will be presented and the derivation of the algorithm can be found in the appendix.

**2.1.1. Probability Check for Gamma Index (2D Datasets).** A pass/fail test has to be performed for each possible pair of points, one from each dataset: point  $ij$  from the reference dataset and  $kl$  from the test dataset.

*Step 1.* Compute the following parameters:

$$\begin{aligned} c_{1ijkl} &= \frac{\sigma_d^t + \sigma_d^r}{\Delta D^2} + 2 \cdot \frac{\sigma_s^t + \sigma_s^r}{\Delta R^2} + \frac{(d_{kl}^t - d_{ij}^r)^2}{\Delta D^2} \\ &\quad + \frac{(x_{kl}^t - x_{ij}^r)^2 + (y_{kl}^t - y_{ij}^r)^2}{\Delta R^2} \\ &= \frac{\sigma_d^t + \sigma_d^r}{\Delta D^2} + 2 \cdot \frac{\sigma_s^t + \sigma_s^r}{\Delta R^2} + \gamma_{ijkl}^2, \\ c_{2ijkl} &= \left( \frac{\sigma_d^t + \sigma_d^r}{\Delta D^2} \right)^2 + 2 \cdot \left( \frac{\sigma_s^t + \sigma_s^r}{\Delta R^2} \right)^2 \\ &\quad + 2 \cdot (\sigma_d^t + \sigma_d^r) \cdot \frac{(d_{kl}^t - d_{ij}^r)^2}{\Delta D^4} \\ &\quad + 2 \cdot (\sigma_s^t + \sigma_s^r) \cdot \frac{(x_{kl}^t - x_{ij}^r)^2 + (y_{kl}^t - y_{ij}^r)^2}{\Delta R^4}, \\ c_{3ijkl} &= \left( \frac{\sigma_d^t + \sigma_d^r}{\Delta D^2} \right)^3 + 2 \cdot \left( \frac{\sigma_s^t + \sigma_s^r}{\Delta R^2} \right)^3 \\ &\quad + 3 \cdot (\sigma_d^t + \sigma_d^r)^2 \cdot \frac{(d_{kl}^t - d_{ij}^r)^2}{\Delta D^6} \\ &\quad + 3 \cdot (\sigma_s^t + \sigma_s^r)^2 \cdot \frac{(x_{kl}^t - x_{ij}^r)^2 + (y_{kl}^t - y_{ij}^r)^2}{\Delta R^6}, \end{aligned}$$

$$h'_{ijkl} = \frac{c_2^{ijkl^3}}{c_3^{ijkl^2}},$$

$$\begin{aligned} y_{ijkl} &= \left( g^2 - \frac{\sigma_d^t + \sigma_d^r}{\Delta D^2} - 2 \cdot \frac{\sigma_s^t + \sigma_s^r}{\Delta R^2} - \gamma_{ijkl}^2 \right) \\ &\quad \cdot \frac{c_2^{ijkl}}{c_3^{ijkl}} + \frac{c_2^{ijkl^3}}{c_3^{ijkl^2}}. \end{aligned} \quad (2)$$

*Step 2.* Compute  $P_{ijkl} = P[\chi^2_{h'_{ijkl}} > y_{ijkl}]$  from a noncentral chi-square distribution probability function or table.

For each point in the reference dataset, the value  $P_{ij} = P[\max_{kl}(\Gamma_{ijkl}) > 1] = 1 - \prod_{kl}(1 - P[\Gamma_{ijkl} > 1])$  is computed and the test is passed if it is less than a preset significance figure  $\alpha$ .

A global modified pass rate can be reported with the results of this test for every point in the reference dataset. A value of  $\alpha = 0.05$  is used in this study.

**2.1.2. Probability Distribution of Gamma (3D Datasets).** In a similar fashion, the test can be carried out for 3D datasets.

*Step 1.* Compute the following parameters:

$$\begin{aligned} c_{1ijklmn} &= \frac{\sigma_d^t + \sigma_d^r}{\Delta D^2} + 3 \cdot \frac{\sigma_s^t + \sigma_s^r}{\Delta R^2} + \frac{(d_{lmn}^t - d_{ijk}^r)^2}{\Delta D^2} \\ &\quad + \frac{(x_{lmn}^t - x_{ijk}^r)^2 + (y_{lmn}^t - y_{ijk}^r)^2 + (z_{lmn}^t - z_{ijk}^r)^2}{\Delta R^2} \\ &= \frac{\sigma_d^t + \sigma_d^r}{\Delta D^2} + 3 \cdot \frac{\sigma_s^t + \sigma_s^r}{\Delta R^2} + \gamma_{ijklmn}^2, \\ c_{2ijklmn} &= \left( \frac{\sigma_d^t + \sigma_d^r}{\Delta D^2} \right)^2 + 3 \cdot \left( \frac{\sigma_s^t + \sigma_s^r}{\Delta R^2} \right)^2 \\ &\quad + 2 \cdot (\sigma_d^t + \sigma_d^r) \cdot \frac{(d_{lmn}^t - d_{ijk}^r)^2}{\Delta D^4} + 2 \cdot (\sigma_s^t + \sigma_s^r) \\ &\quad \cdot \frac{(x_{lmn}^t - x_{ijk}^r)^2 + (y_{lmn}^t - y_{ijk}^r)^2 + (z_{lmn}^t - z_{ijk}^r)^2}{\Delta R^4}, \\ c_{3ijklmn} &= \left( \frac{\sigma_d^t + \sigma_d^r}{\Delta D^2} \right)^3 + 3 \cdot \left( \frac{\sigma_s^t + \sigma_s^r}{\Delta R^2} \right)^3 \\ &\quad + 3 \cdot (\sigma_d^t + \sigma_d^r)^2 \cdot \frac{(d_{lmn}^t - d_{ijk}^r)^2}{\Delta D^6} + 3 \cdot (\sigma_s^t + \sigma_s^r)^2 \\ &\quad \cdot \frac{(x_{lmn}^t - x_{ijk}^r)^2 + (y_{lmn}^t - y_{ijk}^r)^2 + (z_{lmn}^t - z_{ijk}^r)^2}{\Delta R^6}, \\ h'_{ijklmn} &= \frac{c_2^{ijklmn^3}}{c_3^{ijklmn^2}}, \end{aligned}$$

$$y_{ijklmn} = \left( g^2 - \frac{\sigma_d^t + \sigma_d^r}{\Delta D^2} - 3 \cdot \frac{\sigma_s^t + \sigma_s^r}{\Delta R^2} - \gamma_{ijklmn}^2 \right) \cdot \frac{c_2^{ijklmn}}{c_3^{ijklmn}} + \frac{c_2^{ijklmn^3}}{c_3^{ijklmn^2}}. \quad (3)$$

*Step 2.* Compute  $P_{ijklmn} = P[\Gamma_{ijklmn} > 1] = P[\chi^2_{h'_{ijklmn}} > \gamma_{ijklmn}]$  from a noncentral chi-square distribution probability function or table.

For each point in the reference dataset, the value  $P_{ijk} = P[\max_{lmn}(\Gamma_{ijklmn}) > 1] = 1 - \prod_{lmn}[1 - P[\Gamma_{ijklmn} > 1]]$  is computed and the test is passed if it is less than a preset significance figure  $\alpha$ .

As in the 2D case, a global modified pass rate can be reported with the results of this test for every point in the reference dataset. A value of  $\alpha = 0.05$  is used in this study.

*2.2. Application.* A probabilistic method to check test datasets for coincidence with a reference dataset, taking uncertainty into account, was tested with an example. It has to be remarked that for the new test to be passed, every point has to pass the test, that is, the probability test has to be passed for each pair of points drawn from the reference and test datasets. Common practice when using classic gamma test is to allow a limited percentage of points to fail the test. For the application of the present method, the probability comparison will only be passed if all points pass the test.

A practical example with 5 segments was set up. Figure 1(a) shows the whole reference composite field on film and Figures 1(b), 1(c), 1(d), 1(e), and 1(f) the segments used to obtain the composite image. The composite irradiation was modified in order to introduce controlled defects.

*Case 1.* A 1 mm shift along  $X$  in the first segment, 1.5% more dose in the second, a 1 mm shift along  $Y$  in the third, and 0.5% less dose in the fourth.

*Case 2.* Same modifications, but the increment of dose in the second segment is 3% and the third is shifted 4 mm.

*Case 3.* The first segment has 2% less dose than the reference and is shifted 4 mm along  $X$ ; the second segment has been delivered with 5% more dose; the third segment shift along  $Y$  is 4 mm, and the fourth has 2% less dose.

*Case 4.* All segments but the smallest one were shifted 4 mm along the  $X$  axis.

*Case 5.* All segments but the smallest one were shifted 4 mm along the  $Y$  axis.

Therefore, each of the cases corresponds to a set of shifts and changes of intensities for every segment as exemplified in Figure 2. These controlled defects are simple enough as to make clear whether or not a test on coincidence with the reference unmodified image should pass or fail. However, the

algorithm was applied with the same rigour as it would have been done for a more complex fluence pattern.

The modified planar distributions (test datasets) were compared with the original one (reference dataset) with the following uncertainty parameters: 0.2% dose and 0.5 mm, 0.5% dose and 0.5 mm, and 0.2% dose and 1 mm. Dose uncertainty is relative, and this fact has been taken into account in the computation of the indices. Tests were performed for tolerances 2% dose and 2 mm and 3% dose and 3 mm.

A function in  $R$  statistical software [16] was used to perform all the computations. Graphs were obtained using the “rimage” package [17].

### 3. Results

Results for the gamma test are shown in Table 1. Pass rates for a classical test are presented along with the modified test. Shaded cells contain acceptable values: 100% pass rate for the modified test and more than 98% pass rate for classical tests.

Figure 3(a) shows a graph with points that fail the classic test for Case 2 and tolerances of 2 mm and 2%, Figure 3(b) shows the images of pass probability for the new test with 0.2% uncertainty in dose and 0.5 mm in position, the same tolerance values as in the previous case.

Figure 4 shows a sequence of pass probability images for Case 4, tolerance 3 mm and 3%, and different uncertainty values: 0.2 mm/0.2%, 0.5 mm/0.2%, 0.5 mm/0.5%, and 1.0 mm/0.2%. These uncertainty values have been chosen to illustrate the method.

### 4. Discussion and Conclusions

Case 1 is *a priori* an acceptable result, Case 2 is on the limit of acceptability, and the other ones are *a priori* unacceptable. It is clear that the classic test failed to discard the wrong irradiations even allowing for a percentage of failing points. The usual gamma index tests would have approved every case if a 97% pass rate would be allowed and 3%—3 mm tolerances would have been used. For tolerances of 2%—2 mm, only Case 4 would have been rejected. Case 4 is an extremely undesirable plan, with an unacceptable global shift, but, interestingly enough, Case 5, with the same shift along the other axis, would have been accepted with a passing rate greater than 98%. Comparison of pass rates as well as of images in Figure 3 shows that the novel test developed in this work would have rejected cases where the standard gamma index test would allow for a great number of points in gradient areas without rejecting the comparison. Therefore, the new test is less permissive than the classic one.

On the other hand, Figure 4 and their pass rates in Table 1 show the potential misleading effect of using measurement or computing methods not suitable for the task: as uncertainty grows larger, it is possible to accept an inadequate case (Case 4), if tolerances are also too large. It can be concluded that for tolerances of 3% and 3 mm, uncertainties of 0.2% and 1 mm are enough to make the test insensitive. Tolerances of 3 mm and 3% are currently used, but these results show that

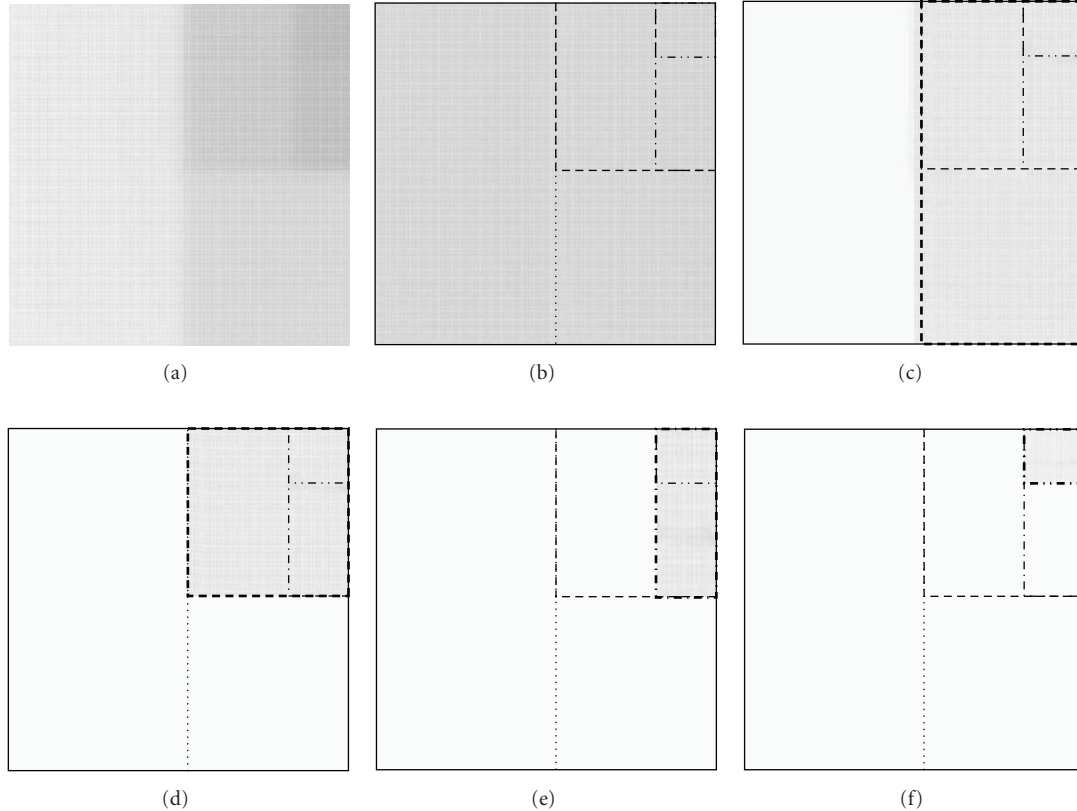


FIGURE 1: Reference dataset. (a) Composite irradiation. (b) Segment 1. (c) Segment 2. (d) Segment 3. (e) Segment 4. (f) Segment 5.

TABLE 1: Gamma results for Cases 1–5.

		Case 1	Case 2	Case 3	Case 4	Case 5
Tolerance 2%—2 mm	Dose Unc. 0.2%. Dist. Unc. 0.2 mm	<b>1.0000</b>	0.9929	0.9926	0.9696	0.9920
	Dose Unc. 0.2%. Dist. Unc. 0.5 mm	<b>1.0000</b>	0.9931	0.9930	0.9718	0.9927
	Dose Unc. 0.5%. Dist. Unc. 0.5 mm	<b>1.0000</b>	0.9932	0.9932	0.9726	0.9938
	Dose Unc. 0.2%. Dist. Unc. 1.0 mm	<b>1.0000</b>	0.9966	0.9966	0.9876	0.9972
	Classic test	<b>1.0000</b>	<b>0.9894</b>	<b>0.9861</b>	0.9615	<b>0.9893</b>
Tolerance 3%—3 mm	Dose Unc. 0.2%. Dist. Unc. 0.2 mm	<b>1.0000</b>	0.9957	0.9957	0.9847	0.9965
	Dose Unc. 0.2%. Dist. Unc. 0.5 mm	<b>1.0000</b>	0.9967	0.9968	0.9898	0.9979
	Dose Unc. 0.5%. Dist. Unc. 0.5 mm	<b>1.0000</b>	0.9968	0.9968	0.9909	0.9979
	Dose Unc. 0.2%. Dist. Unc. 1.0 mm	<b>1.0000</b>	<b>1.0000</b>	<b>1.0000</b>	<b>1.0000</b>	<b>1.0000</b>
	Classic test	<b>1.0000</b>	<b>0.9933</b>	<b>0.9932</b>	0.9794	<b>0.9946</b>

they are a compromise to account indirectly for uncertainty. The test developed in this paper would work with accuracy and sensitivity with 2%—2 mm tolerances that are closer to actual physical requirements.

The method presented in this work is potentially applicable to a broad set of comparisons: computer versus measured dose distributions for planning system commissioning, IMRT commissioning and patient checks, commissioning of measurement devices, and so forth. For any real experimental case, care should be taken to characterize its uncertainty. Furthermore, this method could be used to evaluate whether experimental uncertainties could deteriorate the sensitivity

of a test. Accuracy requirements in IMRT patient plan checks are very high, and it is useful to know if the checking device uncertainty could induce the checker to accept plans too easily.

Some alternative methods have been described in the literature in order to refine the standard gamma index test; but the result is a consensus about tolerances and pass rate criteria. It is interesting to look at some conclusions in the ESTRO Booklet no. 9 [5] in the sense that it is hard to decide if test failure is related to computer system, data transfer, linear accelerator, measurement, or data analysis. A document with a similar scope is the one published by the



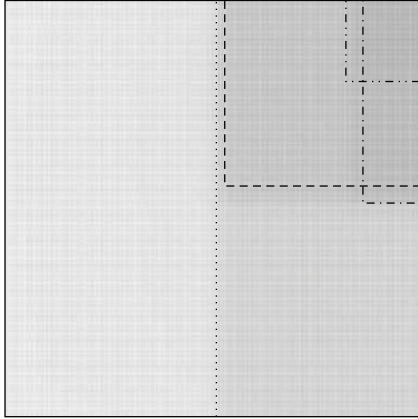


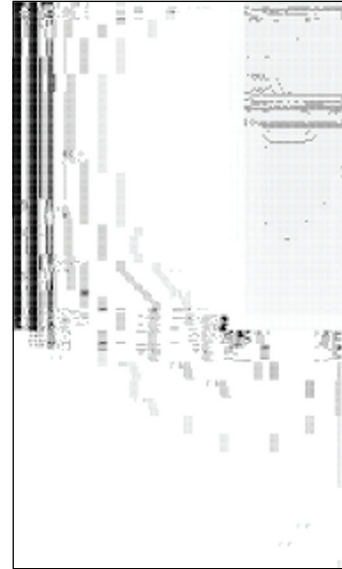
FIGURE 2: Example of setup of a test dataset by shifting and modifying segment intensity.

American Association of Physicists in Medicine [3]. Both et al. [18] performed a study of check results (dose difference and distance to agreement) in order to set reasonable acceptance values for the percentage of passing points (95% for prostate, 90% for other sites) and point dose error per field (3% for prostate and 5% elsewhere). Stock et al. [19] present a strategy of primary and secondary checks. They accept checks with  $\gamma$  pass rates of 5% and prescribe further evaluation ( $\gamma$  angle, e.g.) if  $\gamma$  pass rate is greater than 5% but less than 10%. Moran [20] designed a method to allow for small range failures in the test.

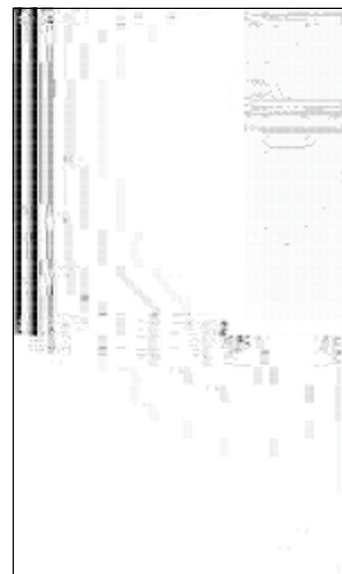
In the survey performed by Nelms and Simon [21], current practice (September 2007) in the USA is presented. It is far from clear that the consensus about how to accept results from a comparison check is actually used. From these sources, it seems there is no rigorous accepted method in the literature in order to consider measurement and computation uncertainty *a priori*.

This work shows a practical application of several results about the probability distribution of quadratic forms of normal random variables. Since no *a priori* relationship between dose and position uncertainty can be assumed, the expression for the gamma index cannot be reduced to a simple noncentral chi-square random variable. This is the reason why some more refined mathematics have been used. The use of a Monte Carlo method [22] would introduce more than a million iterations for each pair of points while the three moment approximation used in this work is fast and accurate enough. Computation does not involve more iterations than a classic gamma check.

A classical test (with  $\Delta D = 3\%$ ,  $\Delta R = 3$  mm, and a pass rate tolerance of 97%) accepts every case in this work, despite the fact that some of them were designed with controlled defects that should not be acceptable at all. Using tighter tolerances (presented,  $\Delta D = 2\%$ ,  $\Delta R = 2$  mm), only one case is rejected (Case 4) and, oddly enough, Case 5, with the same shift as Case 4 but along the central dose edge, passes this classical test. Figure 3(b) shows that this allowance in pass rate means that points in high gradient regions are allowed to fail the test. Unless further investigation is carried out it is



(a)



(b)

FIGURE 3: Image of pass probability for Case 2: (a) classic test with tolerances of 2 mm and 2%. (b) new test with tolerances of 2 mm and 2%.

not clear that those failing points are due to limitations in the measurement procedure.

When the new method is used, it becomes feasible to ensure whether or not points failing a classic test are a consequence of measurement limitations. If the new test does not yield a 100% pass rate it is possible to assert that the failing points cannot have been caused solely by the measurement procedure but there is also a problem with the irradiation. Therefore, no failing points are allowed.

As pointed out previously, this novel method relates experimental features (uncertainty) with test results. A well-defined answer in terms of probability, whether or not the

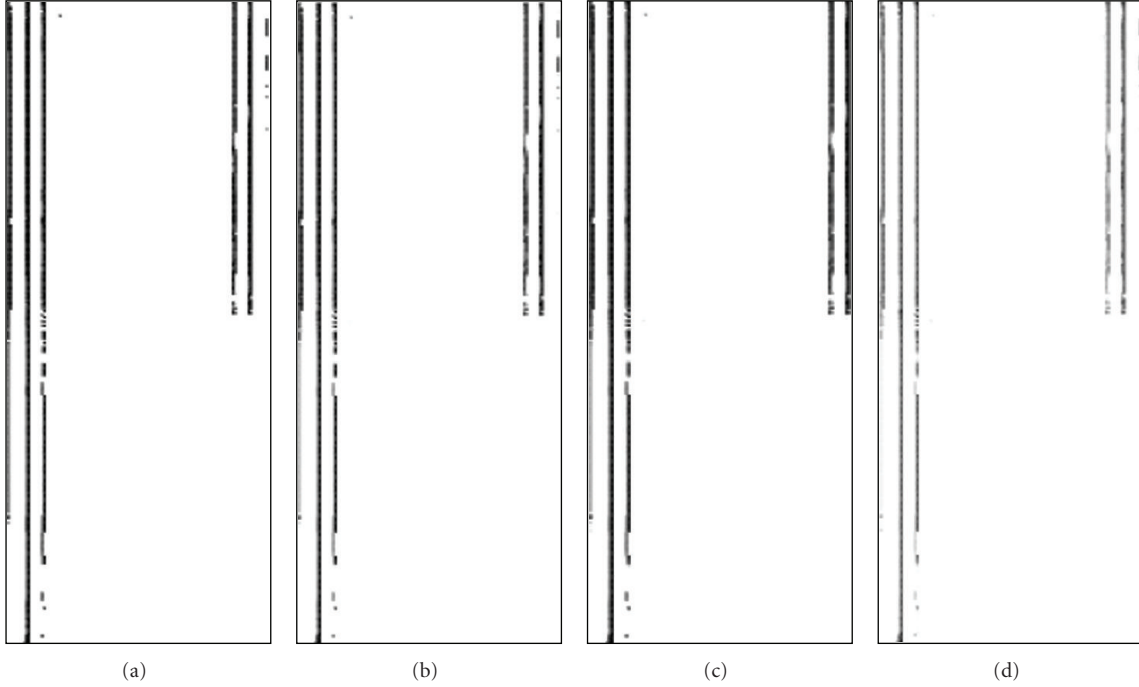


FIGURE 4: Image of pass probability for the new test for Case 4: (a) uncertainty 0.2%—0.2 mm and tolerances of 3 mm and 3%, (b) uncertainty 0.2%—0.5 mm and tolerances of 3 mm and 3%, (c) uncertainty 0.5 mm—0.5% and tolerances of 3 mm and 3%, (d) uncertainty 1.0 mm/0.2% and tolerances 3 mm and 3%.

probability of failing a gamma test at the point  $ij$  is larger than  $\alpha$ , is obtained. As long as the uncertainty properties of the experimental or computational procedure have been investigated, the user is provided with a method to obtain a definite answer. On the other hand, feasibility studies become possible and it is possible to evaluate whether or not a comparison procedure uncertainty features could affect sensitivity in the test results.

## Appendices

### A. Probability Check for Gamma Index (2D Datasets)

Gamma is often described as a distance in a  $N+1$  dimensional space. If the additional dimension were a spatial one and there would be the same tolerance and uncertainty value for every spatial direction, the problem of propagating uncertainty to gamma would have been much easier. But the additional dimension, absorbed dose, has independent tolerance and uncertainty values, and the problem, in its simpler formulation, is the following one: find the probability that the sum of two normal random variables, with zero mean, different variances, and different coefficients is less than  $g^2$ .

Two sets of 2D dose distributions are defined: the reference one and the test one. Both are regular arrays but their spacing could be different. The reference points are labelled with subscripts ( $ij$ ) and the test points with ( $kl$ ). For each of these positions, there will be three quantities: dose  $D_{ij}^r$ ,  $x$  coordinate  $X_{ij}^r$ , and  $y$  coordinate  $Y_{ij}^r$ . The notation for

the test 2D set will be a  $t$  superscript instead of  $r$ .  $X$  and  $Y$  axes are the same for both point sets. These quantities will be considered normally distributed random variables, with mean the measured or computed values. Capital letters will refer to random variables and small letters to their means ( $d_{ij}^r, x_{ij}^r, y_{ij}^r, d_{kl}^t, x_{kl}^t, y_{kl}^t$ ). Absorbed dose in the reference set has associated standard uncertainty  $\sigma_d^r$  and  $\sigma_d^t$  in the test set. Spatial uncertainty is isotropic in both datasets (same standard deviation for  $X$  and  $Y$ ) and will be referred to with the symbols  $\sigma_s^r$  and  $\sigma_s^t$ . Thus, we are assuming that

$$\begin{aligned} D_{ijkl} &= D_{kl}^t - D_{ij}^r \sim N\left(d_{kl}^t - d_{ij}^r, \sqrt{\sigma_d^{t2} + \sigma_d^{r2}}\right), \\ X_{ijkl} &= X_{kl}^t - X_{ij}^r \sim N\left(x_{kl}^t - x_{ij}^r, \sqrt{\sigma_s^{t2} + \sigma_s^{r2}}\right), \\ Y_{ijkl} &= Y_{kl}^t - Y_{ij}^r \sim N\left(y_{kl}^t - y_{ij}^r, \sqrt{\sigma_s^{t2} + \sigma_s^{r2}}\right). \end{aligned} \quad (\text{A.1})$$

The squared gamma random variable  $\Gamma_{ijkl}^2 = D_{ijkl}^2/\Delta D^2 + (X_{ijkl}^2 + Y_{ijkl}^2)/\Delta R^2$  is a weighted sum of normal random variables with different means and standard deviations. If this equation is written in the following way:

$$\begin{aligned} \Gamma_{ijkl}^2 &= \frac{\sigma_d^{t2} + \sigma_d^{r2}}{\Delta D^2} \cdot \frac{D_{ijkl}^2}{\sigma_d^{t2} + \sigma_d^{r2}} \\ &+ \frac{\sigma_s^{t2} + \sigma_s^{r2}}{\Delta R^2} \cdot \left[ \frac{X_{ijkl}^2}{\sigma_s^{t2} + \sigma_s^{r2}} + \frac{Y_{ijkl}^2}{\sigma_s^{t2} + \sigma_s^{r2}} \right], \end{aligned} \quad (\text{A.2})$$

TABLE 2

$\lambda_1 = (\sigma_d^t + \sigma_d^r)/\Delta D^2$	$h_1 = 1$	$\delta_1^2 = (d_{kl}^t - d_{ij}^r)/(\sigma_d^t + \sigma_d^r)$
$\lambda_2 = (\sigma_s^t + \sigma_s^r)/\Delta R^2$	$h_2 = 2$	$\delta_2^2 = ((x_{kl}^t - x_{ij}^r)^2 + (y_{kl}^t - y_{ij}^r)^2 + (z_{kl}^t - z_{ij}^r)^2)/(\sigma_s^t + \sigma_s^r)$

it becomes a weighted sum of noncentral chi-square random variables. The following random variables:

$$\begin{aligned}\tilde{D}_{ijkl} &= \frac{D_{ijkl}}{\sqrt{\sigma_d^t + \sigma_d^r}} \sim N\left(\frac{d_{kl}^t - d_{ij}^r}{\sqrt{\sigma_d^t + \sigma_d^r}}, 1\right), \\ \tilde{X}_{ijkl} &= \frac{X_{ijkl}}{\sqrt{\sigma_s^t + \sigma_s^r}} \sim N\left(\frac{x_{kl}^t - x_{ij}^r}{\sqrt{\sigma_s^t + \sigma_s^r}}, 1\right), \\ \tilde{Y}_{ijkl} &= \frac{Y_{ijkl}}{\sqrt{\sigma_s^t + \sigma_s^r}} \sim N\left(\frac{y_{kl}^t - y_{ij}^r}{\sqrt{\sigma_s^t + \sigma_s^r}}, 1\right)\end{aligned}\quad (\text{A.3})$$

have standard deviation 1. It is possible to use now some of the properties of the noncentral chi-squared distribution. Let  $U_n$  be a finite set of independent normally distributed random variables with means  $\mu_n$  and standard deviation 1. Then,  $W = \sum_n U_n^2$  will have a noncentral chi-square distribution  $\chi^2(n, \lambda)$  with  $n$  degrees of freedom and noncentrality parameter  $\lambda = \sum_n \mu_n^2$ . Thus,

$$\begin{aligned}\tilde{D}_{ijkl}^2 &= \frac{D_{ijkl}^2}{\sigma_d^t + \sigma_d^r} \sim \chi^2\left(1, \frac{(d_{kl}^t - d_{ij}^r)^2}{\sigma_d^t + \sigma_d^r}\right), \\ \tilde{R}_{ijkl}^2 &= \tilde{X}_{ijkl}^2 + \tilde{Y}_{ijkl}^2 = \frac{X_{ijkl}^2}{\sigma_s^t + \sigma_s^r} + \frac{Y_{ijkl}^2}{\sigma_s^t + \sigma_s^r} \\ &\sim \chi^2\left(2, \frac{(x_{kl}^t - x_{ij}^r)^2 + (y_{kl}^t - y_{ij}^r)^2}{\sigma_s^t + \sigma_s^r}\right).\end{aligned}\quad (\text{A.4})$$

The squared gamma index is

$$\Gamma_{ijkl}^2 = \frac{\sigma_d^t + \sigma_d^r}{\Delta D^2} \cdot \tilde{D}_{ijkl}^2 + \frac{\sigma_s^t + \sigma_s^r}{\Delta R^2} \cdot \tilde{R}_{ijkl}^2. \quad (\text{A.5})$$

Given a quadratic form of normally distributed variables, there always exists a transformation which reduces it to a weighted sum of noncentral chi-squared variables, corresponding to the orthogonal transformation that reduces the form to its canonical form. As a matter of fact, the previous derivation is a very simple particular case of this general result [23]. In the gamma test problem, it is necessary to evaluate the probability of the event  $\Gamma_{ijkl}^2 > 1$ .

In the general case of a quadratic form, the noncentrality parameters are linear combinations of the means. Thus, a quadratic form of central normal variables results in a linear combination of central chi-squared variables. The normal variables involved in  $\Gamma$  are noncentral ones; their means are the differences between doses or between spatial coordinates in the test and the reference datasets.

Different expansions of the distribution function of a weighted sum of noncentral chi-squared variables can

be found in the literature, and they could be used for this problem. Shah and Khatri [24] found a power series expansion, Ruben [25] developed series of distribution functions of central and noncentral chi-squared variables, with coefficients recursively defined, and Shah and other authors [26] proposed series involving Laguerre polynomials. There is also a simple approximation based on a study on relationships between chi-squared and Poisson variables first proposed by Patnaik [27], improved by Pearson [28], which gives accurate results especially in the tails [29]. Imhoff [23] rewrote this approximation for the weighted sum of noncentral chi-squared variables. It uses probability values for a single central chi-squared variable. This approach was used in the present work. The accuracy of this method leaves out of the question a Monte Carlo approach based on ISO recommendations [22], which would lead to a minimum of  $10^7$  iterations for each pair of points.

According to Imhoff, if  $Q = \sum_{r=1}^m \lambda_r \chi_{h_r, \delta_r^2}^2$  ( $\delta_r^2$  being noncentrality parameters in his notation), then

$$P[Q > x] \cong P[\chi_{h'}^2 > y], \quad (\text{A.6})$$

where

$$\begin{aligned}h' &= \frac{c_2^3}{c_3^2}, \\ y &= (x - c_1) \cdot \left(\frac{h'}{c_2}\right)^{1/2} + h', \\ c_j &= \sum_{r=1}^m \lambda_r^j \cdot (h_r + j\delta_r^2), \quad j = 1, 2, 3.\end{aligned}\quad (\text{A.7})$$

Applying this approximation to the problem of finding  $P[\Gamma_{ijkl}^2 > g^2]$ , the set of parameters (2) in the main text are obtained, taking into account that the chi squared parameters are those in Table 2, taking into account that for our case, and, therefore, (2) in the main text are obtained. With those parameters  $P_{ijkl} = P[\Gamma_{ijkl} > g] = P[\chi_{h'}^2 > y_{ijkl}]$ , and the probability of gamma being greater than  $g^2$  for the points  $(ij)$  in the reference dataset and  $(kl)$  in the test dataset, taking into account spatial and dosimetry uncertainties in both datasets has been obtained.

It is possible to modify the original gamma test for the  $ij$  point if the probabilities

$$P_{ij} = P\left[\max_{kl}(\Gamma_{ijkl}) > 1\right] = 1 - \prod_{kl} (1 - P[\Gamma_{ijkl} > 1]) \quad (\text{A.8})$$

are defined and the following criterion is set:  $ij$  passes the test if  $P_{ij} < \alpha$  being  $\alpha$  a significance figure, set by the user as the maximum probability to be accepted.



## B. Probability Distribution of Gamma (3D Datasets)

The squared gamma random variable is now  $\Gamma_{ijklmn}^2 = D_{ijklmn}^2/\Delta D^2 + (X_{ijklmn}^2 + Y_{ijklmn}^2 + Z_{ijklmn}^2)/\Delta R^2$ . As long as the spatial standard uncertainties are isotropic, the same rearrangement as in the 2D case can be done

$$\begin{aligned} \Gamma_{ijklmn}^2 &= \frac{\sigma_d^t + \sigma_d^r}{\Delta D^2} \cdot \frac{D_{ijklmn}^2}{\sigma_d^t + \sigma_d^r} \\ &+ \frac{\sigma_s^t + \sigma_s^r}{\Delta R^2} \cdot \left[ \frac{X_{ijklmn}^2}{\sigma_s^t + \sigma_s^r} + \frac{Y_{ijklmn}^2}{\sigma_s^t + \sigma_s^r} + \frac{Z_{ijklmn}^2}{\sigma_s^t + \sigma_s^r} \right], \end{aligned} \quad (\text{B.1})$$

and, defining the new random variable  $\tilde{Z}_{ijklmn} = Z_{ijklmn}/\sqrt{\sigma_s^t + \sigma_s^r} \sim N(z_{lmn}^t - z_{ijk}^r/\sqrt{\sigma_s^t + \sigma_s^r}, 1)$ , we obtain

$$\begin{aligned} \tilde{D}_{ijklmn}^2 &= \frac{D_{ijklmn}^2}{\sigma_d^t + \sigma_d^r} \sim \chi^2 \left( 1, \frac{(d_{lmn}^t - d_{ijk}^r)^2}{\sigma_d^t + \sigma_d^r} \right), \\ \tilde{R}_{ijklmn}^2 &= \tilde{X}_{ijklmn}^2 + \tilde{Y}_{ijklmn}^2 + \tilde{Z}_{ijklmn}^2 \\ &= \frac{X_{ijklmn}^2}{\sigma_s^t + \sigma_s^r} + \frac{Y_{ijklmn}^2}{\sigma_s^t + \sigma_s^r} + \frac{Z_{ijklmn}^2}{\sigma_s^t + \sigma_s^r} \\ &\sim \chi^2 \left( 3, \frac{(x_{lmn}^t - x_{ijk}^r)^2 + (y_{lmn}^t - y_{ijk}^r)^2 + (z_{lmn}^t - z_{ijk}^r)^2}{\sigma_s^t + \sigma_s^r} \right), \\ \Gamma_{ijklmn}^2 &= \frac{\sigma_d^t + \sigma_d^r}{\Delta D^2} \cdot \tilde{D}_{ijklmn}^2 + \frac{\sigma_s^t + \sigma_s^r}{\Delta R^2} \cdot \tilde{R}_{ijklmn}^2. \end{aligned} \quad (\text{B.2})$$

Using Imhoff's three moments approximation, the parameter values are the ones in equations (3) in the main text, and  $P_{ijklmn} = P[\Gamma_{ijklmn} > 1] = P[\chi^2_{h'_{ijklmn}} > y_{ijklmn}]$ . The test is  $P_{ijk} < \alpha$  with  $P_{ijk} = P[\max_{lmn}(\Gamma_{ijklmn}) > 1] = 1 - \prod_{lmn}[1 - P[\Gamma_{ijklmn} > 1]]$ .

## Conflict of Interests

No conflict of interests exists for any of the authors.

## Acknowledgments

The authors are indebted to C. Weatherill from Boston University for some very valuable suggestions. They are also indebted to E. Overton and A. Baker for their careful revision of this paper and valuable suggestions.

## References

- [1] J. A. Purdy, "Intensity-modulated radiotherapy: current status and issues of interest," *International Journal of Radiation Oncology Biology Physics*, vol. 51, no. 4, pp. 880–914, 2001.
- [2] C. Burman, C. S. Chui, G. Kutcher et al., "Planning, delivery, and quality assurance of intensity-modulated radiotherapy using dynamic multileaf collimator: a strategy for large-scale implementation for the treatment of carcinoma of the prostate," *International Journal of Radiation Oncology Biology Physics*, vol. 39, no. 4, pp. 863–873, 1997.
- [3] G. A. Ezzell, J. W. Burmeister, N. Dogan et al., "IMRT commissioning: multiple institution planning and dosimetry comparisons, a report from AAPM Task Group 119," *Medical Physics*, vol. 36, no. 11, pp. 5359–5373, 2009.
- [4] J. R. Palta, C. Liu, and J. G. Li, "Quality assurance of intensity-modulated radiation therapy," *International Journal of Radiation Oncology Biology Physics*, vol. 71, no. 1, supplement, pp. S108–S112, 2008.
- [5] ESTRO Booklet no. 9, *Guidelines for the Verification of IMRT*, ESTRO, 2008.
- [6] J. G. Li, G. Yan, and C. Liu, "Comparison of two commercial detector arrays for IMRT quality assurance," *Journal of Applied Clinical Medical Physics*, vol. 10, no. 2, pp. 62–74, 2009.
- [7] R. Sadagopan, J. A. Bencomo, R. L. Martin, G. Nilsson, T. Matzen, and P. A. Balter, "Characterization and clinical evaluation of a novel IMRT quality assurance system," *Journal of Applied Clinical Medical Physics*, vol. 10, no. 2, pp. 104–119, 2009.
- [8] P. A. Jursinic and B. E. Nelms, "A 2-D diode array and analysis software for verification of intensity modulated radiation therapy delivery," *Medical Physics*, vol. 30, no. 5, pp. 870–879, 2003.
- [9] E. Spezi, A. L. Angelini, F. Romani, and A. Ferri, "Characterization of a 2D ion chamber array for the verification of radiotherapy treatments," *Physics in Medicine and Biology*, vol. 50, no. 14, pp. 3361–3373, 2005.
- [10] D. A. Low, W. B. Harms, S. Mutic, and J. A. Purdy, "A technique for the quantitative evaluation of dose distributions," *Medical Physics*, vol. 25, no. 5, pp. 656–661, 1998.
- [11] D. A. Low and J. F. Dempsey, "Evaluation of the gamma dose distribution comparison method," *Medical Physics*, vol. 30, no. 9, pp. 2455–2464, 2003.
- [12] M. Wendling, L. J. Zipp, L. N. McDermott et al., "A fast algorithm for gamma evaluation in 3D," *Medical Physics*, vol. 34, no. 5, pp. 1647–1654, 2007.
- [13] T. Depuydt, A. van Esch, and D. P. Huyskens, "A quantitative evaluation of IMRT dose distributions: refinement and clinical assessment of the gamma evaluation," *Radiotherapy and Oncology*, vol. 62, no. 3, pp. 309–319, 2002.
- [14] P. S. Basran and M. K. Woo, "An analysis of tolerance levels in IMRT quality assurance procedures," *Medical Physics*, vol. 35, no. 6, pp. 2300–2307, 2008.
- [15] J. Palta, S. Kim, J. Li, and C. Liu, "Tolerance limits and action levels for planning and delivery of IMRT," in *Tolerance limits and action levels for planning and delivery of IMRT*, J. R. Palta and T. R. Mackie, Eds., Medical Physics Monograph, pp. 593–612, Medical Physics Publishing, Madison, Wis, USA, 2003.
- [16] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2007.
- [17] Nikon Systems Inc., *rimage: Image Processing Module for R*, R package version 0.5-7, 2005.

- [18] S. Both, I. M. Alecu, A. R. Stan et al., “A study to establish reasonable action limits for patient-specific quality assurance in intensity-modulated radiation therapy,” *Journal of Applied Clinical Medical Physics*, vol. 8, no. 2, pp. 1–8, 2007.
- [19] M. Stock, B. Kroupa, and D. Georg, “Interpretation and evaluation of the  $\gamma$  index and the  $\gamma$  index angle for the verification of IMRT hybrid plans,” *Physics in Medicine and Biology*, vol. 50, no. 3, pp. 399–411, 2005.
- [20] J. M. Moran, J. Radawski, and B. A. Fraass, “A dose gradient analysis tool for IMRT QA,” *Journal of Applied Clinical Medical Physics*, vol. 6, no. 2, pp. 62–73, 2005.
- [21] B. E. Nelms and J. A. Simon, “A survey on planar IMRT QA analysis,” *Journal of Applied Clinical Medical Physics*, vol. 8, no. 3, pp. 76–90, 2007.
- [22] JCGM 101:2008, Evaluation of measurement data—Supplement 1 to the “Guide to the expression of uncertainty in measurement”—Propagation of distributions using a Monte Carlo method, JCGM, 2008.
- [23] J. P. Imhoff, “Computing the distribution of quadratic forms in normal variables,” *Biometrika*, vol. 48, pp. 419–426, 1961.
- [24] B. K. Shah and C. G. Khatri, “Distribution of a definite quadratic form for non-central normal variates,” *The Annals of Mathematical Statistics*, vol. 32, pp. 883–887, 1961.
- [25] H. Ruben, “Probability content of regions under spherical normal distributions, IV: the distribution of homogeneous and non-homogeneous quadratic functions of normal variables,” *The Annals of Mathematical Statistics*, vol. 33, pp. 542–570, 1962.
- [26] B. K. Shah, “Distribution of definite and of indefinite quadratic forms from a non-central normal distribution,” *The Annals of Mathematical Statistics*, vol. 34, pp. 186–190, 1963.
- [27] P. B. Patnaik, “The non-central  $X^2$ - and F-distribution and their applications,” *Biometrika*, vol. 36, pp. 202–232, 1949.
- [28] E. S. Pearson, “Note on an approximation to the distribution of non-central  $X^2$ ,” *Biometrika*, vol. 46, p. 364, 1959.
- [29] N. L. Johnson, “On an extension of the connexion between Poisson and  $X^2$  distributions,” *Biometrika*, vol. 46, pp. 352–363, 1959.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

