INSTITUTE OF PHYSICS PUBLISHING PHYSICS IN MEDICINE AND BIOLOGY

Phys. Med. Biol. **48** (2003) 3543–3553 PII: S0031-9155(03)65955-5

A revision of the γ-evaluation concept for the comparison of dose distributions

Annemarie Bakai, Markus Alber and Fridtjof Nüsslin

Abteilung für Medizinische Physik, Universitätsklinik für Radioonkologie Tübingen, Hoppe-Seyler-Str. 3, 72076 Tübingen, Germany

E-mail: aebakai@med.uni-tuebingen.de

Received 11 July 2003
Published 10 October 2003
Online at stacks.iop.org/PMB/48/3543

Abstract

A method for the quantitative four-dimensional (4D) evaluation of discrete dose data based on gradient-dependent local acceptance thresholds is presented. The method takes into account the local dose gradients of a reference distribution for critical appraisal of misalignment and collimation errors. These contribute to the maximum tolerable dose error at each evaluation point to which the local dose differences between comparison and reference data are compared. As shown, the presented concept is analogous to the γ-concept of Low *et al* (1998a *Med. Phys.* **25** 656–61) if extended to (3+1) dimensions. The pointwise dose comparisons of the reformulated concept are easier to perform and speed up the evaluation process considerably, especially for fine-grid evaluations of 3D dose distributions. The occurrences of false negative indications due to the discrete nature of the data are reduced with the method. The presented method was applied to film-measured, clinical data and compared with γ-evaluations. 4D and 3D evaluations were performed. Comparisons prove that 4D evaluations have to be given priority, especially if complex treatment situations are verified, e.g., non-coplanar beam configurations.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

With the clinical implementation of new radiation therapy techniques, dosimetric verification of the dose computation algorithms involved becomes a key issue. This is definitely the case for intensity modulated radiotherapy (IMRT).

An advanced approach for IMRT treatment planning is to use Monte Carlo based dose calculations (Ma *et al* 2000, Leal *et al* 2003, Fippel *et al* 2003). To ensure high accuracy of the calculations, the employed head models in particular need explicit verification. Radiographic films, ionization chambers and arrays or portal imaging devices produce

measured data suitable for comparisons (Low *et al* 1998b, Van Esch *et al* 2002, Depuydt *et al* 2002, Pasma *et al* 1999). Even if the calculation methods prove to be accurate, measurements cannot be abandoned because data transfer and the delivery process still need to be checked.

If the correspondence between two IMRT dose distributions, a measured comparison and a reference distribution, is to be tested, a global dose difference acceptance criterion is not sufficient. Often, no point in the target volume (TV) exists which is irradiated by all beams or individual segments of an IMRT step-and-shoot treatment and high dose gradients even inside the TV will be observed for individual beams. For a homogeneous total dose within the TV, the correct matching of these gradients is essential.

Van Dyk *et al* (1993) described the large effects of small spatial shifts on observed dose differences for high dose gradients and introduced the concept of a distance-to-agreement (DTA) based on them. The DTA for a given reference dose point defines the maximum spatial distance within which a point in the comparison data set has to be found that receives the same dose. Van Dyk suggested to use a fixed dose criterion for the analysis of low-gradient regions and, complementary, perform a DTA-based analysis for high-gradient regions.

Instead of employing a DTA criterion, alternatively, one could enlarge the dose tolerances for points in such regions. Venselaar *et al* (2001) tabulated in their publication fixed tolerance values for application in different test configurations. For example, they proposed to use a 10% local dose criterion in high-gradient regions if a simple geometry is tested, while for a more complex geometry a 15% local dose criterion should be used.

An analysis based on such multiple criteria allows us to identify pass and fail regions within a data set but gives no quantitative index of the accuracy.

This paper takes up the theoretical concept of the γ-evaluation method of Low *et al* (1998a). Based on Van Dyk's proposal of complementary acceptance criteria for high and low dose gradients, Low *et al* developed a technique for the quantitative comparison of two-dimensional dose distributions in (2+1)-dimensional dose-distance space with proper Euclidean metric. Straightforward modification allows us to apply this technique also on one- and three-dimensional cases.

To perform quantitative dose comparisons, Low suggested drawing acceptance ellipsoids around each dose point $D_r(\vec{r}_r)$ of a two-dimensional reference data set with the major axes of the ellipsoids scaled by the chosen, fixed acceptance criteria. $\Delta D_{\max}$, the maximum acceptable dose difference if two dose points are directly compared, defines the scale of the dose axis; the major axes in the two Cartesian directions, x and y, are both defined by the maximum acceptable DTA, $\Delta d_{\max}$. For each point $D_r(\vec{r}_r)$, the dose points within the planar comparison data set $D_c(\vec{r}_c)$ have to be searched for the point with the minimum distance. The distance of this point, intrinsically scaled to a fraction of the acceptance criteria, gives the local γ-value, the accuracy index.

Expressed in formulae, this concept reads

$$\gamma(x_r, y_r) = \min_{x_c, y_c}\{\Gamma_r(x_c, y_c, D_c)\} \tag{1}$$

with

$$\Gamma_r(x_c, y_c, D_c) = \sqrt{\frac{\Delta r^2}{\Delta d_{\max}{}^2} + \frac{\Delta D^2}{\Delta D_{\max}{}^2}}$$

where

$$\Delta r = |\vec{r}_c - \vec{r}_r| = (x_c - x_r)^2 + (y_c - y_r)^2 \tag{2}$$

and

$$\Delta D = D_c(\vec{r}_c) - D_r(\vec{r}_r).$$

ΔD is an absolute dose difference if $\Delta D_{\max}$ is also expressed in Gray. Otherwise, ΔD is a relative dose difference if D_c, D_r are scaled to the maximum of D_r.

If for a certain reference point $D_r(\vec{r}_r)$, γ is found to be $\leqslant 1$, the nearest $D_c(\vec{r}_c)$ is located in the ellipsoid of acceptance. If the nearest $D_c(\vec{r}_c)$ is outside the ellipsoid, the acceptance criteria cannot be met and the corresponding $\gamma > 1$ is a quantitative measure for the observed disagreement.

With the γ-evaluation method two main problems exist. First, if evaluations are extended to (3+1)D, many points need to be searched in the comparison data set, see equation (1). An extension to 4D is desirable for the critical appraisal of misalignment and collimation errors with realistic beam configurations and also for non-coplanar cases. Second, if the reference or measurement data are coarsely spaced, interpolations are required to give a better resolution and to avoid misinterpretations of the agreement. This causes even more evaluation points which might become critical. Alternatively, with a coarser comparison grid a multiple-level strategy could be used to reduce the occurrence of false negative indications allowing correct evaluation. This has been addressed by Depuydt *et al* (2002).

These problems give motivation to find an equivalent criterion that avoids the search as in equation (1).

Looking at Low's concept from a different perspective, an alternative quantitative measure, the acceptable local dose error, can be derived which can easily be interpreted from the laws of error propagation. With this new measure, reliable dose comparisons, even for fine-grid data sets, become feasible.

In this paper, the presented methodology is applied to film-measured, clinical IMRT data and compared to the results of γ-calculations.

2. Materials and methods

For simplicity, the new method is first developed in (1+1)D dose-distance space. Imagine a continuous profile of a reference dose distribution with acceptance ellipses drawn around each dose point. The envelope of all these ellipses forms a tube around the reference profile. Imagine further a second data set, a comparison dose profile. For those reference points where γ-values $\leqslant 1$ are observed, the comparison profile will run inside the tube, otherwise it will extend outside.

This situation is sketched in figure 1. The shaded area in this graph corresponds to the region where γ-values larger than 1 are found. Note that in the figure the dose distributions D_c and D_r are scaled by

$$c = \frac{\Delta d_{\max}}{\Delta D_{\max}} \tag{3}$$

to yield the functions x_{D_c} (labelled as 'comp') and x_{D_r} (labelled as 'ref'). This simple transformation deforms the ellipses to circles. The circles' radius is $\Delta d_{\max}$.

2.1. Acceptance intervals

Instead of determining the minimum distance of a point in $D_r(\vec{r}_r)$ to $D_c(\vec{r}_c)$, we are going to perform the dose data analysis with respect to the acceptance tube enveloping $D_r(\vec{r}_r)$. To be more precise, the intervals defined by the vertical thicknesses of the tube are interpreted as the maximum acceptable local dose differences. We propose that for each point $\vec{r}$ with reference value $x_{D_r}(\vec{r})$ effective and reliable comparisons are possible if it is tested whether the value $x_{D_c}(\vec{r})$ is located inside the acceptance interval for that point. If reference and

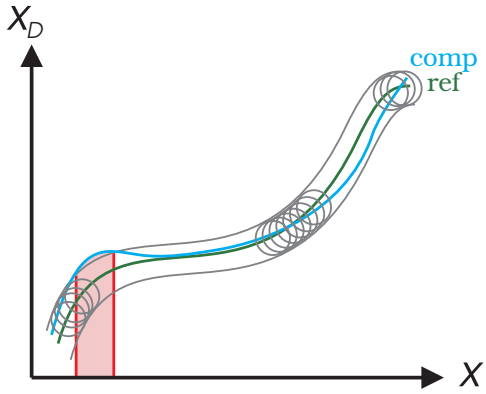

Figure 1. Two continuous dose profiles (ref: reference profile, comp: comparison profile). The vertical dose axis was scaled to $x_D = \frac{\Delta d_{\max}}{\Delta D_{\max}} \cdot D$ and has physical unit 'mm'. The acceptance ellipses of Low *et al* (1998a) that can be drawn around each reference point become circular due to the scaling and are enveloped by an acceptance tube.

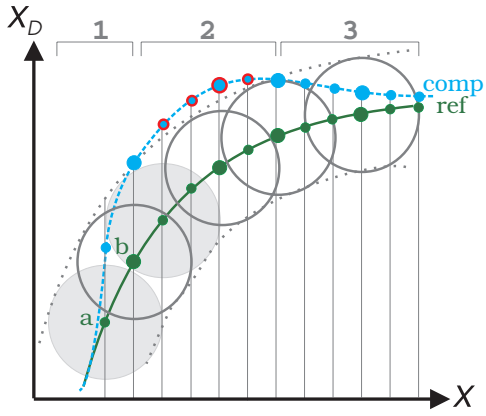

Figure 2. Reference and comparison x_D-profiles defined by discrete data points. The acceptance tube is indicated, as are a few acceptance circles. γ-analysis and dose-tube test give different results for reference points a and b in high-gradient region 1.

comparison data are defined on different grids, the comparison data set needs to be interpolated for this test.

Figure 2 shows a reference profile (ref) with its enveloping acceptance tube and a comparison profile (comp). In this case, both distributions are defined by discrete data points equally spaced along the x-axis as indicated by the point-marking dots and thin impulse lines. For illustration purposes, some acceptance circles are drawn. In region 3, all comparison points $x_{D_c}(\vec{r})$ are located within the acceptance tube and within the acceptance circle of at least one reference point $x_{D_r}(\vec{r})$ as indicated by two sample circles. Both methods, the γ-method and the acceptance interval test identify these points as being in agreement. The findings of both methods also correspond with the points in region 2, where disagreement between comparison and reference data is observed.

For region 1, being a typical example for a high-gradient region, the results of the two methods differ. While for both reference points in this region the acceptance tube-test detects agreement, the γ-method identifies agreement only for reference point b. Such local

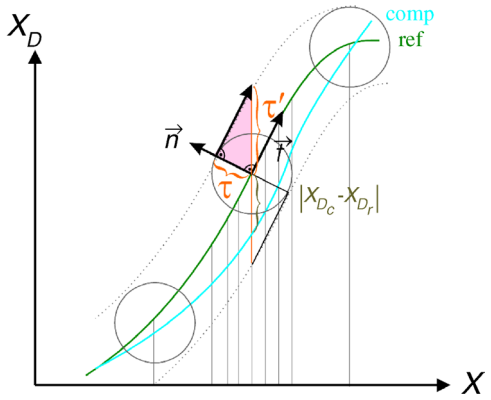

Figure 3. Schematic representation of the proposed method for the determination of the vertical acceptance tube thickness $2 \cdot \tau'$ at a certain point x. τ' can be calculated based on the local tangent and normal vectors, $\vec{t}$ and $\vec{n}$, and on the acceptance sphere radius τ. At point x, the comparison and reference value differ by $|x_{D_c} - x_{D_r}|$.

differences in the findings of the two evaluation methods strongly depend on the ratio of the grid spacing and $\Delta d_{\max}$, which defines the radius of the acceptance circles. The smaller the ratio $\frac{\text{grid spacing}}{\Delta d_{\max}}$, the better both methods will correspond. To avoid these discretization artefacts, the search for a minimum (see equation (1)) have to be performed on an interpolated profile.

2.2. The evaluation factor χ

For computing the thickness of the acceptance tube in dose direction, the situation is sketched in figure 3. One side of the highlighted triangle in the figure runs parallel to the local tangent vector

$$\vec{t} = \begin{pmatrix} 1 \\ \frac{\mathrm{d}x_{D_r}}{\mathrm{d}x} \end{pmatrix}$$

the second side extends along the reference distribution normal vector

$$\vec{n} = \frac{1}{\sqrt{1 + \left(\frac{\mathrm{d}x_{D_r}}{\mathrm{d}x}\right)^2}} \begin{pmatrix} -\frac{\mathrm{d}x_{D_r}}{\mathrm{d}x} \\ 1 \end{pmatrix}$$

and has length τ. The length of the vertical side of the triangle, τ', follows from the vector equation

$$\tau \cdot \vec{n} + a \cdot \vec{t} = \tau' \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

where a is a multiplicative factor. Then

$$\tau' = \tau \cdot \sqrt{1 + \left(\frac{\mathrm{d}x_{D_r}}{\mathrm{d}x}\right)^2}.$$

With the transformation factor c, as introduced in (3), τ' is obtained as

$$\tau' = \tau \cdot \sqrt{1 + c^2 \cdot \left(\frac{\mathrm{d}D_r}{\mathrm{d}x}\right)^2} = \tau \cdot \sqrt{1 + c^2 \cdot \nabla {D_r}^2}$$

or, expressed in doses instead of Cartesian coordinates,

$$\tau' = \sqrt{\Delta D_{\max}^2 + \Delta d_{\max}^2 \cdot \nabla D_r{}^2}. \tag{4}$$

This equation can readily be extended to 3D:

$$\tau' = \sqrt{\Delta D_{\max}^2 + \Delta d_{\max}^2 \cdot \|\vec{\nabla} D_r\|^2}. \tag{5}$$

In case the second derivative of the reference dose distribution is relatively small, $2 \cdot \tau'$ is a good measure for the vertical thickness of the acceptance tube or the acceptance interval at any profile point. Depending on the absolute value and the signature of the curvature, the real value will be slightly over- or underestimated.

Equations (4) and (5) can be easily interpreted with respect to error propagation. The equations show that in addition to the measurement error $\Delta D_{\max}$ the misalignment error(s) $|\Delta d_{\max} \cdot \nabla_x D_r|$, $|\Delta d_{\max} \cdot \nabla_y D_r|$ and $|\Delta d_{\max} \cdot \nabla_z D_r|$ contribute(s) to the total local dose error or, respectively, to the maximum acceptable local dose difference τ'.

We introduce the new evaluation factor χ as

$$\begin{aligned} \chi &\equiv \frac{D_c(\vec{r}) - D_r(\vec{r})}{\tau'} \\ &= \frac{D_c(\vec{r}) - D_r(\vec{r})}{\sqrt{\Delta D_{\max}^2 + \Delta d_{\max}^2 \cdot \|\vec{\nabla} D_r\|^2}}. \end{aligned} \tag{6}$$

The value of $|\chi|$ ($\simeq\gamma$) provides a quantitative measure for the agreement of two dose points. If $|\chi| \leqslant 1$, their difference in dose is acceptable. Note that the evaluation of each pair of measurement and reference data points takes place at the same grid point $\vec{r}$, no search is necessary.

2.3. Clinical test case

At the University Hospital Tübingen, IMRT planning is performed with the inverse treatment planning system HYPERION which utilizes the fast Monte Carlo code XVMC for the dose calculations (Alber *et al* 2000, Fippel *et al* 2003).

Multi-leaf collimated step-and-shoot segments for a 6 MV IMRT head-and-neck (H&N) treatment with gantry angles of 0°, 30°, 75°, 160°, 200°, 225° and 324° were optimized based on patient CT data. For verification purposes, the resulting treatment plan was applied without modifications to a cubic 30 cm solid water phantom. The iso-centre was transferred to the midpoint of the phantom whose outer contours were aligned with the right-left, anterior-posterior and superior-inferior directions. Monte Carlo doses were calculated on a 4 mm × 4 mm × 4 mm dose grid covering the whole phantom.

Kodak EDR2 verification film (Eastman Kodak Co., Rochester, NY, USA) was inserted into the middle of a corresponding slab phantom and the 2D dose was recorded for the coronal iso-centric plane. The verification measurement was performed at an Elekta SLi15 linear accelerator (Elekta Ltd, Crawley, UK) according to the simulation calculations. Absolute film doses were calculated from the digitized film transmission values employing a transmission value-to-dose calibration curve which was derived from film measurements immediately before the actual measurements. The films were digitized using a Vidar VXR-12 scanner (Vidar Systems Corporation, VA, USA). The scanned data had a resolution of ≈0.8 mm per pixel.

To allow the comparison of the two data sets based on the presented χ-methodology, both distributions were interpolated onto a 1 mm grid. This grid size is small compared to the maximum accepted misalignment errors of 3 mm and is a compromise between the original

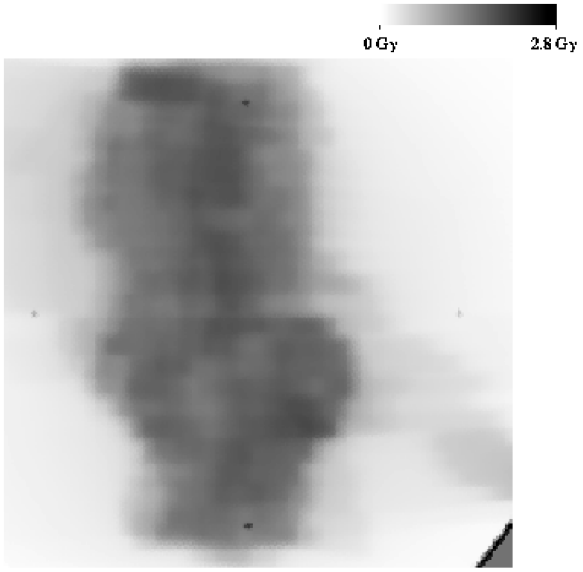

Figure 4. Film-measured dose distribution of a H&N treatment plan consisting of seven beams. The measurement was performed in the coronal, iso-centric plane of a cubic phantom.

resolutions of the two data sets. The Monte Carlo doses, which were taken as reference values, were polynomially interpolated in 3D. The interpolation yielded 301 $\times$ 301 equidistant dose points for the coronal iso-centric plane and the same number of dose points at the same (x, y)-grid points for the neighbouring anterior and posterior dose planes.

Employing a value of 3% (compared to plane maximum dose $D_{\max}$) for the measurement error $\Delta D_{\max}$ and $\Delta d_{\max} = 3$ mm, χ-ratios were calculated for each grid point. For the same sets of discrete data and also based on 3%/3 mm acceptance criteria, γ-indices were calculated so as to be compared with the results of the χ-analyses. With both methods, (2+1)D and (3+1)D evaluations of the data were performed. For the (3+1)D evaluation cases, it was assumed that the comparison dose plane could have been measured anywhere within the phantom cube.

For 4D evaluations with the γ-method, $(x_c - x_r)^2 + (y_c - y_r)^2 + (z_c - z_r)^2$ has to be used for Δr instead of $(x_c - x_r)^2 + (y_c - y_r)^2$ in equation (2). To speed up the calculation process of the γ-indices, the search areas for the minimum Γ-values were restricted to 7 $\times$ 7 $\times$ 7 pixel large sub-grids around each dose point in question.

3. Results and discussion

Figure 4 shows the dose distribution obtained from film measurement for the investigated H&N IMRT plan.

In figure 5, a graphical comparison of the χ- and γ-calculation results for that plan is given. The upper two graphs in the figure represent $|\chi|$-distributions and the lower two graphs represent γ-distributions. While the graphs in the left column are results of full (3+1)D evaluations, the graphs in the right column are results of (2+1)D evaluations and do not consider misalignment in the anterior-posterior z-direction. In all four graphs, function values between 0 and 3 are colour coded. Pixels with values $\leqslant 1$ have the same colour shade to assist

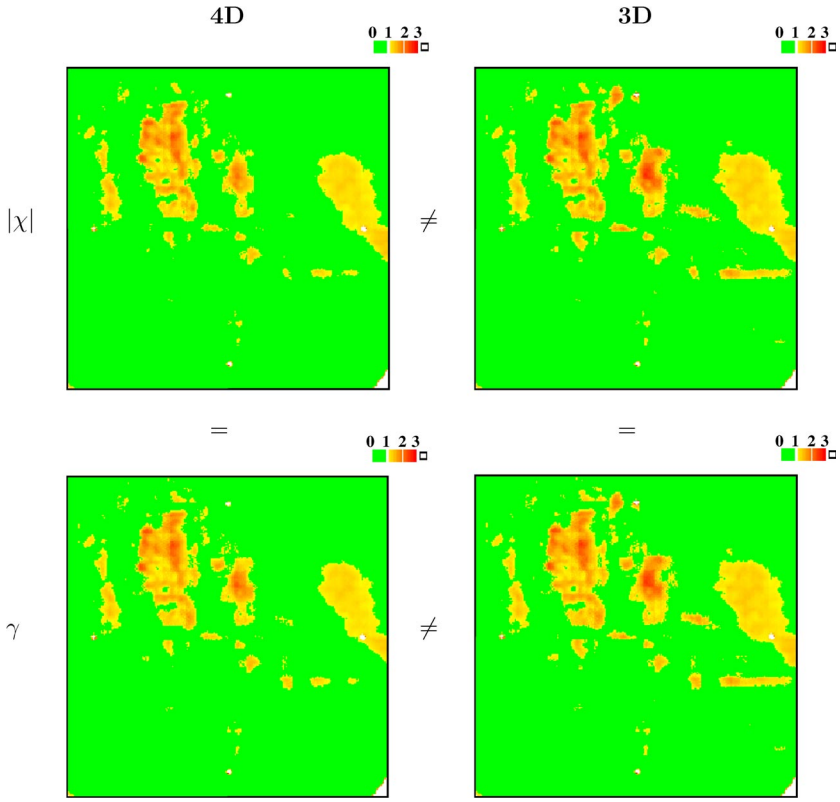

Figure 5. Results of the $|\chi|$- (upper row) and γ- (lower row) evaluations for the film-measured dose distribution in figure 4. For the comparisons, the Monte Carlo dose distribution was taken as reference. In the left column, the results of 4D evaluations are shown, in the right column, the results of the 3D analyses. All pixels with $|\chi|$- and γ-values $\leqslant 1$ are displayed in the same shade of green.

in the evaluation. Function values >3 are not resolved. A total of 241×241 points was analysed.

In all four graphs, function values $\leqslant 1$ dominate. $|\chi|$- and γ-values larger than 1 can be observed in the upper half of each graph. In these regions leaf transmission highly contributes to the film dose causing over-response of the film.

Both graphs of the 4D evaluations and, likewise, both of the 3D comparisons correspond with each other very well. This was expected because of the equivalent nature of the two evaluation methods. The few differences that can be observed are due to the discrete nature of the data.

While with the γ-method, each reference point was compared with $7^3 = 343$ points which, altogether, took around 25 s ($\approx 20 \times 10^6$ calculations), the χ-evaluation was approximately a factor 120 faster. Computations were performed using a computer with a 1 GHz processor.

The histograms in figure 6 provide an overview of the relative quantitative distribution of the indices $|\chi|$ and γ based on the 241×241 analysed points. The width of the histogram bins is 0.5. The similar histogram shapes in the left (4D evaluation) and right (3D evaluation) panel once more confirm the analogy of the evaluation strategies. The observed differences between the $|\chi|$ and γ data are mainly due to the coarse binning. The fact that the histogram

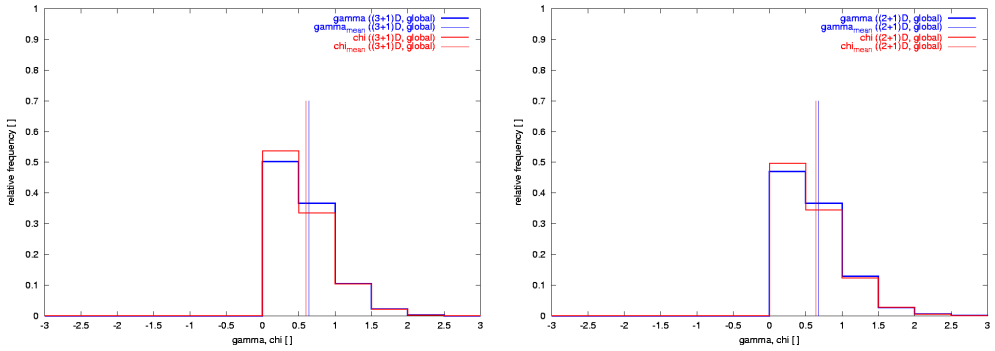

Figure 6. Histograms showing the $|\chi|$- and γ-value relative distributions for the graphs in figure 5 (left: 4D evaluation, right: 3D evaluation). Histogram bin width is 0.5. The vertical lines indicate the mean $|\chi|$ and γ.

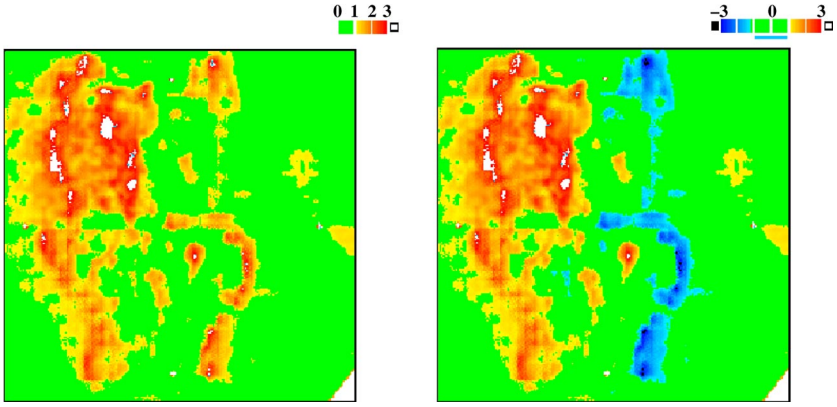

Figure 7. Results of 4D evaluations if film and Monte Carlo distributions are shifted 6 mm against each other in right–left direction. The right graph follows from the γ-analysis, the left is the result of a χ-analysis. The χ-analysis considers regions of under- and overdosage separately.

bars and the mean values in the left panel are shifted to smaller $|\chi|$- and γ-values proves that it is important for a correct data analysis not only to consider possible 2D misalignments but also 3D misalignment errors. This is especially important if oblique or even non-coplanar beams are applied or if antropomorphic phantoms which are difficult to position are used for the measurements.

In table 1 the mean values of $|\chi|$ and γ for the presented 3D and 4D cases and the percentages of the identified points with $|\chi| \leqslant 1$ or, respectively, $\gamma \leqslant 1$ are collected. As the data show, the full 4D analyses result in lower mean quantitative measures, while at the same time also the number of accepted data points grows. Such data and histograms (see figure 6) might assist dose evaluations.

If χ-ratios are calculated according to equation (6), positive and negative results may result depending on the dose values D_c and D_r. A graph with χ-values ranging from $[-3 \cdots 3]$ is presented in figure 7. The graph is the result of a 4D analysis of the sample case if film and Monte Carlo data are shifted by 6 mm against each other. This graph allows easy interpretation

Table 1. Resulting mean $|\chi|$ and γ values and percentages of the analysed points that have a $|\chi|$ or $\gamma \leqslant 1$ for the different evaluations presented in figures 5 and 6.

Evaluation method	$\lvert\chi\rvert_{\text{mean}}$, γ_{mean}	Percentage of analysed points with $\lvert\chi\rvert$ or $\gamma \leqslant 1$ (%)
$\lvert\chi\rvert$ (3D)	0.64	83.6
γ (3D)	0.68	84.2
$\lvert\chi\rvert$ (4D)	0.60	87.1
γ (4D)	0.64	86.8

concerning the shifts of the data sets. For comparison, the result of the corresponding 4D γ-analysis is plotted in the left panel of the figure.

The comparisons in this work were based on a global $\Delta D_{\max}$ value (3% of $D_{\max}$) since film measurements were performed. If equipment with reduced systematic measurement errors is used, such as an ion chamber, array evaluations can be based on local $D_{\max}$ values. This may have a higher clinical relevance.

4. Conclusions

In this paper, a local dose error based method for the quantitative evaluation of dose distributions is presented. The formalism of the new method can be derived from the γ-evaluation concept if the acceptance regions on which the γ-analysis is based are abstracted to an acceptance tube with spatially different vertical thicknesses. This different formalism allows quantitative evaluation of dose distributions simply by the comparison of the locally observed dose differences and the maximum accepted local dose error for a pair of dose points. This works not only well for discrete data sets as a sample analysis based on film and Monte Carlo dose data shows, but is much faster than searching the minimum Γ of a whole set of points. The accepted dose errors that are given by the sum of the squared measurement and misalignment errors can be calculated in advance since they are independent of the measured comparison data. In general, it is recommended to consider misalignment errors in all three spatial directions, because it corresponds to the three-dimensional character of irradiations.

A possible extension of the presented concept could be anisotropic acceptance criteria. This might be desirable if treatment plans are to be verified where highly dose-sensitive healthy structures are in the direct neighbourhood of the high dose regions of the TV and where an alignment error of $\Delta d_{\max}$ appears unacceptable. Considering spatially variant DTA acceptance criteria $\Delta\vec{d}(\vec{r}) = (\Delta d_x(\vec{r}), \Delta d_y(\vec{r}), \Delta d_z(\vec{r}))$, equation (5) changes to

$$\tau' = \sqrt{\Delta D_{\max}^2 + \|\Delta\vec{d}(\vec{r}) \cdot \vec{\nabla} D_r\|^2}. \tag{7}$$

An example of use of equation (7) could be the shielding of the myelon in H&N treatments. Verification should prove the exact position of the gradient regions. It might, therefore, be desirable to apply acceptance criteria $\Delta\vec{d}(\vec{r})$ where misalignment errors in the anterior-posterior z-direction could be more highly penalized for the myelon-facing parts of the dose distribution.

Finding proper mathematical expressions for $\Delta\vec{d}(\vec{r})$ will be the subject of future research.

Acknowledgments

This work was partly supported by the Deutscher Akademischer Austauschdienst eV, Bonn, and the *fortüne* program of the Tübingen University Hospital and by DFG grant no 33/7-1.

References

Alber M *et al* 2000 Hyperion—an integrated IMRT planning tool *The Use of Computers in Radiation Therapy* ed W Schlegel and T Bortfeld (Heidelberg: Springer) pp 46–8

Depuydt T *et al* 2002 A quantitative evaluation of IMRT dose distributions: refinement and clinical assessment of the gamma evaluation *Radiother. Oncol.* **62** 309–19

Fippel M *et al* 2003 A virtual photon energy fluence model for Monte Carlo dose calculation *Med. Phys.* **30** 301–11

Leal A *et al* 2003 Routine IMRT verification by means of an automated Monte Carlo simulation system *Int. J. Radiat. Oncol. Biol. Phys.* **56** 58–68

Low D A *et al* 1998a A technique for the quantitative evaluation of dose distributions *Med. Phys.* **25** 656–61

Low D A *et al* 1998b Quantitative dosimetric verification of an IMRT planning and delivery system *Radiother. Oncol.* **49** 305–16

Ma C M *et al* 2000 Monte Carlo verification of IMRT dose distributions from a commercial treatment planning optimization system *Phys. Med. Biol.* **45** 2483–95

Pasma K L *et al* 1999 Dosimetric verification of intensity modulated beams produced with dynamic multileaf collimation using an electronic portal imaging device *Med. Phys.* **26** 2373–78

Van Dyk J *et al* 1993 Commissioning and quality assurance of treatment planning computers *Int. J. Radiat. Oncol. Biol. Phys.* **26** 261–73

Van Esch A *et al* 2002 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 *Radiother. Oncol.* **65** 53–70

Venselaar J *et al* 2001 Tolerances for the accuracy of photon beam dose calculations of treatment planning systems *Radiother. Oncol.* **60** 191–201