Radiotherapy and Oncology 79 (2006) 224–230
www.thegreenjournal.com

Gamma histograms for radiotherapy plan evaluation

Emiliano Spezi*, D. Geraint Lewis

Department of Medical Physics, Velindre Hospital, Cardiff, UK

Abstract

***Background and purpose*: The technique known as the 'γ evaluation method' incorporates pass-fail criteria for both distance-to-agreement and dose difference analysis of 3D dose distributions and provides a numerical index (γ) as a measure of the agreement between two datasets. As the γ evaluation index is being adopted in more centres as part of treatment plan verification procedures for 2D and 3D dose maps, the development of methods capable of encapsulating the information provided by this technique is recommended.**

***Patients and methods*: In this work the concept of γ index was extended to create gamma histograms (GH) in order to provide a measure of the agreement between two datasets in two or three dimensions. Gamma area histogram (GAH) and gamma volume histogram (GVH) graphs were produced using one or more 2D γ maps generated for each slice of the irradiated volume. GHs were calculated for IMRT plans, evaluating the 3D dose distribution from a commercial treatment planning system (TPS) compared to a Monte Carlo (MC) calculation used as reference dataset.**

***Results*: The extent of local anatomical inhomogenities in the plans under consideration was strongly correlated with the level of difference between reference and evaluated calculations. GHs provided an immediate visual representation of the proportion of the treated volume that fulfilled the γ criterion and offered a concise method for comparative numerical evaluation of dose distributions.**

***Conclusions*: We have introduced the concept of GHs and investigated its applications to the evaluation and verification of IMRT plans. The gamma histogram concept set out in this paper can provide a valuable technique for quantitative comparison of dose distributions and could be applied as a tool for the quality assurance of treatment planning systems.**

Keywords: Radiation therapy; Verification; Evaluation; Gamma index

As conformal radiotherapy and especially IMRT are complex techniques, which demand careful verification, the development and use of methods capable of assessing the quality of competing treatment plans and verifying their dosimetric accuracy has been strongly recommended [6,2]. The quantitative evaluation of dose distributions through a composite analysis of distance-to-agreement (DTA) and dose difference was presented by Low et al. [8]. This technique, known as the 'γ evaluation method', incorporates pass–fail criteria for both DTA and dose difference analysis of 3D dose distributions and provides a numerical index (γ) as a measure of the agreement between two datasets. A clinical assessment of the γ method has been reported by Depuydt et al. [3], who investigated the use of the γ algorithm as a routine verification quality control tool for IMRT dose distributions. Predicted portal images were compared to acquired images through a refined gamma filter developed by the authors. In this approach, adopted in most centres implementing IMRT, reference and evaluated dataset are visually examined for each treatment beam. In clinical practice it is realistic to expect regions where the γ criterion is not satisfied. The validation or rejection of the IMRT field is left to subjective observer interpretation of the obtained γ image. This is the limitation of this procedure. The visual examination is both time consuming and subject to the observer interpretation of the overall goodness of the agreement between evaluated and reference datasets. A tool capable of expressing quantitatively the confidence that can be associated with a certain dose calculation is thus needed.

In this work, the concept of γ index was extended to create gamma histograms (GH) in order to provide a measure of the agreement between two datasets in three dimensions. By analogy with dose volume histograms, DVHs [4], Gamma area histograms (GAHs) and Gamma Volume Histograms (GVHs) were produced on a basis of individual 2D γ maps

0167-8140/$ - see front matter © 2006 Elsevier Ireland Ltd. All rights reserved. doi:10.1016/j.radonc.2006.03.020

generated for each slice of the irradiated volume. GHs were calculated for two IMRT plans, evaluating the 3D dose distribution from a commercial treatment planning system (TPS) compared to a Monte Carlo (MC) calculation used as reference dataset.

We believe GHs are a useful tool for the graphical representation of area or volumetric gamma maps. A GH can also provide a method to extract an index characterizing the overall degree of acceptability of a 3D dose distribution with respect to a reference dataset. Our initial exposition of the use of gamma maps in histogram format was set out in previous work [14,15]. Recently, Stock et al. [16] proposed evaluation filters for IMRT hybrid plan verification based on the γ method and utilizing a similar concept to the one set out in this paper. This confirms the interest of the scientific community in decision guidelines for the practical implementation of gamma based tools for the verification of IMRT plans.

Materials and methods

Evaluated and reference dataset

This investigation involves two retrospective IMRT treatment plans for tumours of the head and neck region, produced by our clinical TPS: Helax-TMS (Nucletron, The Netherlands). The pencil beam (PB) algorithm [1] was used for both beam fluence optimization and for the calculation of the final delivered dose with Helax-TMS. The beam setup for both treatment plans is shown in Fig. 1. Results have been evaluated and compared with calculations provided by the MC method. The BEAM MC code system [11] was used. A careful verification of the MC module definition for our radiotherapy linac (Varian Clinac 2100 CD incorporating a Millennium MLC-80 and operating at 6MV) has been presented elsewhere [12]. The CT dataset, TPS plan data (such as plan settings, etc.) and 3D dose distribution were exported in DICOM format [9] and processed within the Matlab environment (The MathWorks Inc., Natick USA) using a DICOM-RT Toolbox developed specifically for the evaluation and the verification of radiotherapy treatment plans [13]. MC calculations were performed in identical conditions on the basis of the optimized TPS dose plan and used as the reference dataset.

Gamma histograms

When evaluating the acceptability of a calculated dose distribution D_c with respect to a reference dataset D_m, for each reference point $\vec{r}_m$ the composite dose/distance value $\Gamma(\vec{r}_m,\vec{r}_c)$ is determined, with respect to each calculation point $\vec{r}_c$, as given by the following equation

$$\Gamma = \sqrt{\frac{r^2(\vec{r}_m,\vec{r}_c)}{\Delta d_M^2} + \frac{\delta^2(\vec{r}_m,\vec{r}_c)}{\Delta D_M^2}} \quad (1)$$

where

$$r(\vec{r}_m,\vec{r}_c) = |\vec{r}_c - \vec{r}_m| \quad (2)$$

and

$$\delta(\vec{r}_m,\vec{r}_c) = D_c(\vec{r}_c) - D_m(\vec{r}_m) \quad (3)$$

is the dose difference between calculated and reference dose distribution. These parameters are normally scaled to obtain dimensionless quantities and in this work they were set at 3% and 3 mm, respectively, based on the values selected by Low et al. [8] and Harms et al. [5]. For each reference point $\vec{r}_m$ there can then be defined a γ index at each point in the evaluation plane $\vec{r}_c - \vec{r}_m$ so that $\gamma(\vec{r}_m) = \min\{\Gamma(\vec{r}_c,\vec{r}_m)\}, \quad \forall\{\vec{r}_c\}$. The evaluation point passes the composite pass-fail criterion with respect to the reference dataset when $\gamma(\vec{r}_m) \leq 1$.

When evaluating a 3D dose distribution versus a reference dataset the γ calculation algorithm can be executed slice-by-slice, obtaining a 3D γ volume as a stack of 2D maps. This represents in most situations a very large amount of data, which needs to be classified.

In this investigation, we present in fuller form the concept and possible applications of gamma histograms. A Gamma histogram represents the relation between a certain γ value and the area or volume characterized by such a value. This is analogous to the definition of DVHs [4].

A frequency GAH (fGAH) or a cumulative GAH (cGAH) characterizes the information provided by a 2D γ map representing the number of pixels or the percentage of the area covered by a specific γ value. A frequency GVH

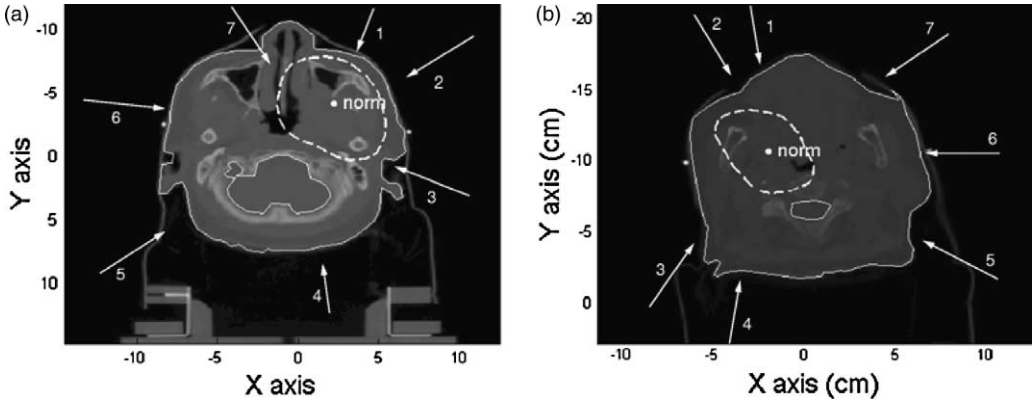

Fig. 1. Beam configuration for two IMRT plans, (a) p1 and (b) p2. Beam settings are displayed on the transverse plane through the isocentre (norm). The PTV outline is also shown (dashed line).

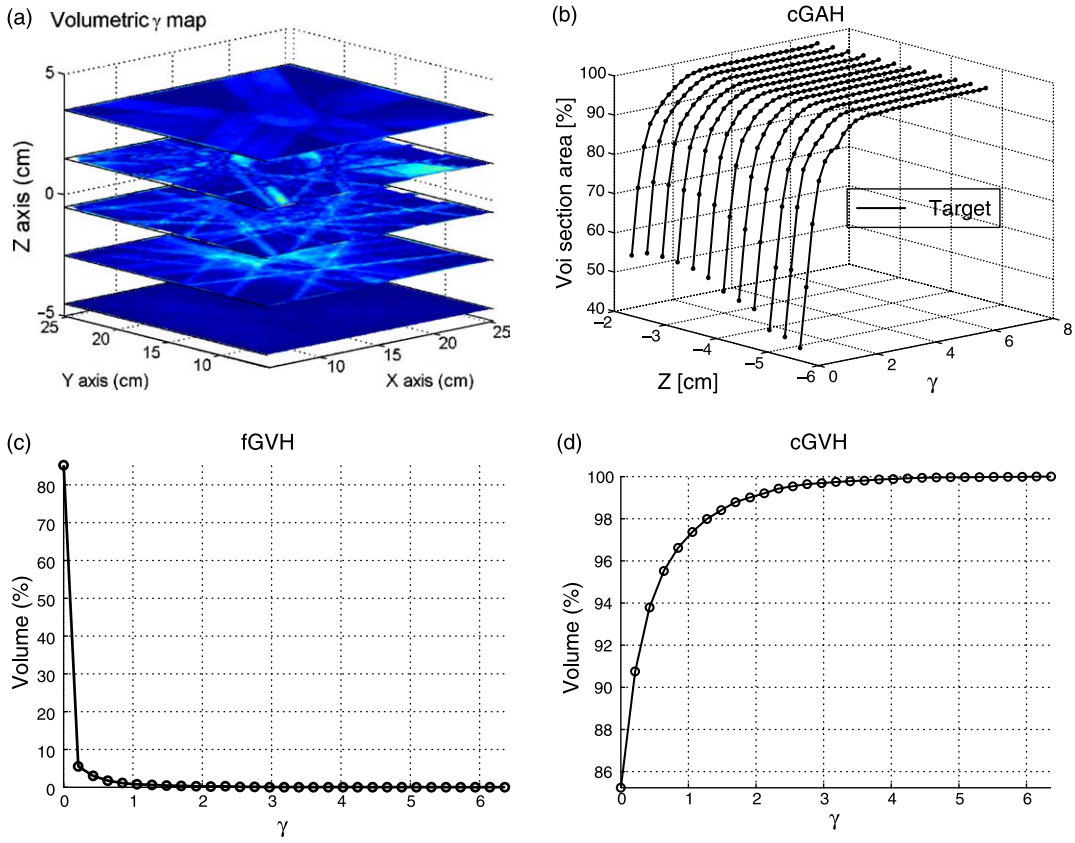

Fig. 2. Gamma histograms: the information provided by individual 2D γ maps generated for each slice of a treatment volume (a) is represented as stacked cumulative GAHs (b) or in frequency or cumulative GVHs (c and d). For clarity only a subset of gamma maps is shown in (a).

(fGVH) or a cumulative GVH (cGVH) summarizes volumetric γ data in one plot providing the number of voxels or the percentage volume covered by a specific γ value. GHs can be also calculated for each volume of interest (VOI) defined in the segmentation process.

In this work, 20 γ maps were calculated for each slice of the 3D TPS dose datasets when compared with the reference 3D MC dose. A 3D γ map was then generated, as shown in Fig. 2a. GAHs were then calculated for each slice of the 3D volume. GAHs can be shown slice by slice or as a stacked histogram plot as depicted in Fig. 2b. Frequency GVHs and cGVHs are shown in Fig. 2c and d. In the treatment instances considered in this study GHs, similarly to DVHs, were calculated on the basis of the dose calculation grid set by the TPS, i.e. $(0.4\times0.4\times0.5)$ cm^3. Since the evaluation of γ can be a computationally intensive task, for large 3D matrices a 'search range' SR was implemented in our γ algorithm. If DTA or dose difference criteria are not satisfied within SR, the value $\gamma(r)$ is set to a specific character[1]. Therefore during the production of GHs, values of γ, which are set to this specific character are set to the highest γ number encountered.

[1] Correspond to the IEEE arithmetic representation for Not-a-Number (NaN).

Results

IMRT plan 1 (P1)

Fig. 3 shows the cGAH calculated for the PTV of P1. It can be noted that the percentage ($\hat{A}$) of each PTV section, which satisfies the acceptance criterion $\gamma(r)\geq1$ becomes gradually reduced towards the extremities of the tumour volume. This is also shown in Table 1 where $\hat{A}$ is reported for all the sections of the PTV. In particular from Table 1 one can observe that the agreement between TPS and MC is maximal on the central slices of the PTV (CT slice no. 6–8), with the lowest values of $\hat{A}$ scored in the inferior part of the target volume. This anatomical area, corresponding to the oral cavity (CT slice no. 1) and the patient's left mandible, is highly inhomogeneous and therefore more likely to show the limitations of the PB algorithm. This is shown in Fig. 4 where isodose contours are depicted for both TPS and MC calculations. This case shows the capabilities of GAHs in identifying regions of disagreement between reference and evaluated dose distributions.

IMRT plan 2 (P2)

The percentage area of each PTV section where the Helax-TMS dose distribution fulfils the acceptance criterion compared to MC is listed in Table 2. The inferior part of the PTV (towards CT slice no. 1) is characterised by a high level of agreement between TPS and MC. This is because in the region being considered the PTV section does not encompass any anatomical inhomogeneity. However, more superiorly

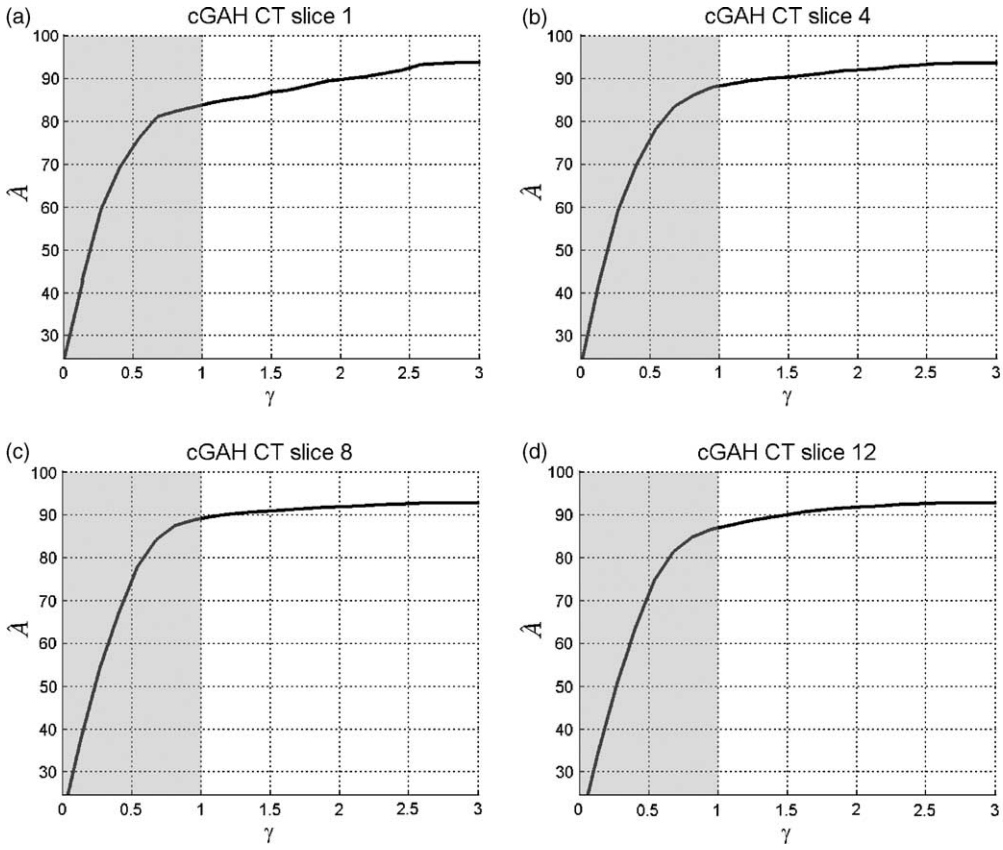

Fig. 3. Cumulative gamma area histograms for P1 target volume. cGAH is shown for CT slice no. 1 (a), no. 4 (b), no. 8 (c) and no. 12 (d) (cf. Table 1). Â values are lowest for the inferior part of the CT volume (a), where inhomogeneities are most prevalent. The area where the γ criterion is satisfied is also shown.

Table 1
Percentage are Â for each PTV section of P1, which satisfies the acceptance criterion $\gamma \leq 1$

CT slice no.	1	2	3	4	5	6	7	8	9	10	11	12
Z (cm)	−5.64	−5.34	−5.04	−4.74	−4.44	−4.14	−3.84	−3.54	−3.24	−2.94	−2.64	−2.34
Â	83.7	83.6	87.1	88.1	88.4	88.9	89.1	89.0	88.2	87.6	87.2	86.7

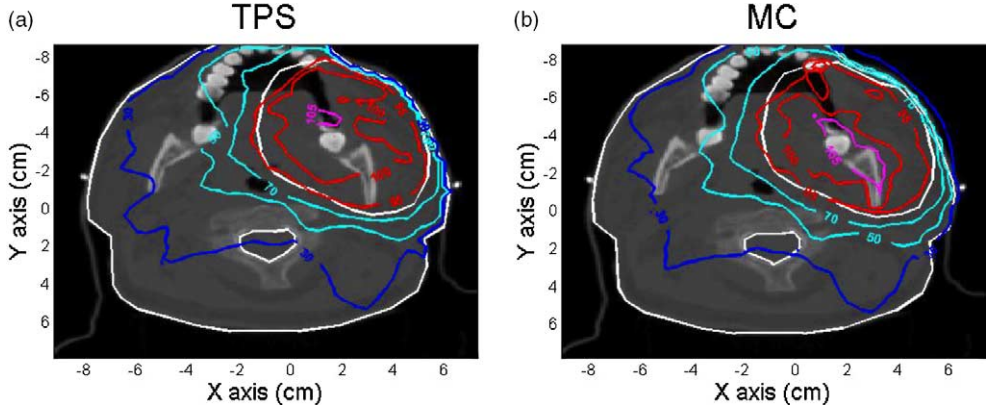

Fig. 4. P1 isodose distribution in the transverse plane through the inferior target slice for TPS (a) and MC (b) generated data. PTV, cord and patient outlines are shown in white.

Table 2
Percentage are Â for each PTV section of P2, which satisfies the acceptance criterion $\gamma \leq 1$

CT slice no.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Z (cm)	−8.32	−8.02	−7.72	−7.42	−7.12	−6.82	−6.52	−6.22	−5.92	−5.62	−5.32	−5.02	−4.72	−4.42	−4.12	−3.82	−3.52	−3.22	−2.92	−2.62	−2.32	−2.02	−1.72
Â	100.0	100.0	96.9	98.0	99.0	97.0	95.5	95.1	93.5	92.4	90.5	90.5	90.5	90.2	90.1	90.1	89.8	89.3	88.7	87.9	86.8	86.5	86.6

(towards slice no. 23), the agreement between TPS and MC calculations becomes gradually poorer due to the presence of inhomogeneities such as air cavities and bony structures involved in the PTV. This is shown in Fig. 5 where the gamma map for PTV sections no. 2, and no. 20 (Ref. Table 2) are displayed with the corresponding CT slice and PTV outline at that coordinate and the calculated cGAH. It can be noted that the value of γ rises in correspondence with anatomical inhomogeneities, which increase in number and in extent as one moves superiorly through the volume. Fig. 5a shows a very good agreement between TPS and MC. This is confirmed by the cGAH in Fig. 5b where all the evaluation points satisfy the pass–fail criterion. However, Fig. 5c and d indicate how the γ criterion is not satisfied in slice no. 20 for $\approx$12% of the calculation points. The disagreements correspond to air gaps and bony structures encompassed by the PTV, as shown in the transverse section of Fig. 5c. Overall the percentage of PTV, which satisfies the acceptance criterion ($\gamma \leq 1$) for P2 is 86.6%. This is similar to what was found for the P1 case. However, the spectrum of values in 3D can be very different. In Fig. 6, the cGVHs for both evaluation instances are compared. Although, the overall level of acceptance is similar, there is a higher percentage of volume satisfying lower γ for P1 than P2. In the P1 case 70% of the volume satisfies $\gamma \leq 0.5$ compared to 60% of P2. This difference increases for lower values of γ. Although, P1 and P2 are two different radiotherapy cases, this comparative analysis of cGVHs gives greater confidence in the P1 evaluation case, as the level of agreement between the TPS and MC distributions is better than in the P2 evaluation case.

Discussion and conclusions

We believe the gamma histogram concept set out in this paper provides a valuable method for quantitative comparison of dose distributions. DVHs and GHs provide complementary functions. While DVHs describe how the calculated dose is delivered to anatomical structures such as target and organs at risk, GHs describe how good the performed calculation is with respect to a reference dataset, which can be measured or MC-generated. In particular GAHs yield an immediate visual representation of the proportion of the treated area that fulfils the agreement criteria. Moreover, a stack of GAH plots (cf. Fig. 2b) can be very useful in locating regions where unacceptable discrepancies between evaluated and reference dose distributions occur. The GVH provides a representative index of the agreement over a given volume. GH analysis results can be easily included in local guidelines as a criterion to accept or reject radiotherapy treatment plans. Mijnheer et al. [10] have recently suggested that GAHs could be used as a tool for the quality assurance of treatment planning systems. The possibility of defining levels of acceptance criteria of dose plans when compared to reference datasets using the proposed GH concept could be explored using this approach.

As with DVHs, GHs represent a surrogate dataset where spatial information about the γ map is lost. However, spatial information is already difficult to interpret in intensity modulated beams, which are characterized by very complex

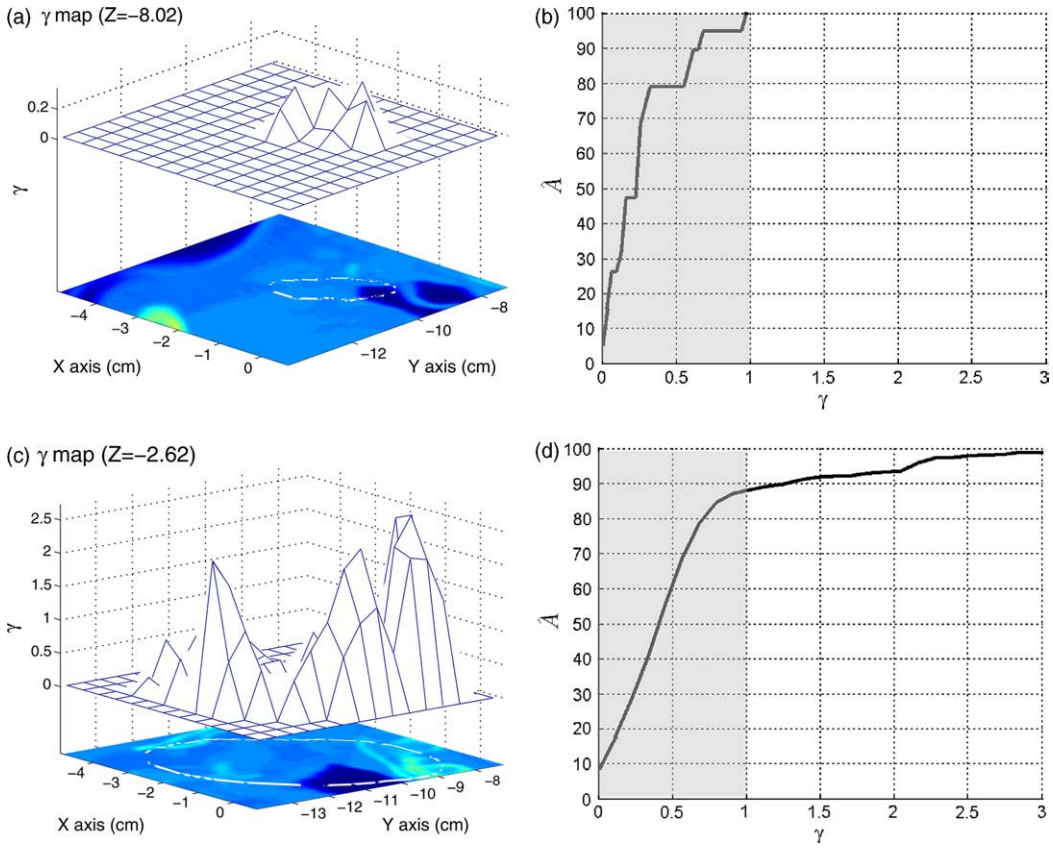

Fig. 5. Gamma maps for PTV sections no. 2 and no. 20 in P2 (ref. Table 2). Respectively, the corresponding CT slices and PTV outline are also shown (a, c), together with the calculated cGAHs (b, d). The shaded area where the γ criterion is satisfied is also shown.

and non intuitive fluence maps. Moreover, the clinical influence of the spatial information in current IMRT verification techniques, where radiation beams are delivered to a uniform water phantom still needs to be addressed.

In this paper, we set out a methodology and describe a tool that can be used in a simple and effective way to represent the degree of acceptability of radiotherapy plans. The extraction of indices characterizing the overall degree of acceptability of a single plan (or in other words the definition of GH-based criteria to be used for validating or rejecting a treatment plan) is a subsequent step, which undoubtedly needs to be addressed. The approach followed by Stock et al. [16] in defining gamma based filters for IMRT hybrid plan verification represents an example of how such tools could be used. However, the definition of such recommendations for reporting dose validation in radiotherapy treatment planning, similarly to published ICRU guidelines, is work, which would also involve extended and multi-centred practical clinical experience.

Finally, we are currently investigating the benefits of implementing a full 3D algorithm for the calculation of volumetric γ maps, rather than volumes built up as a summation of 2D slices, as is commonly the case. The inclusion of the third dimension is expected to improve the γ based evaluation of treatment plans, as non-transverse dose gradients will be taken into account in the analysis. The GH method could also be used successfully to compare the various classes of treatment planning algorithms, e.g. pencil beam v. 'collapsed cone' convolution. The increased availability of reference 3D dose distributions independently generated using MC technology or polymer gel dosimetry [7] is expected to make this application very useful in clinical practice.

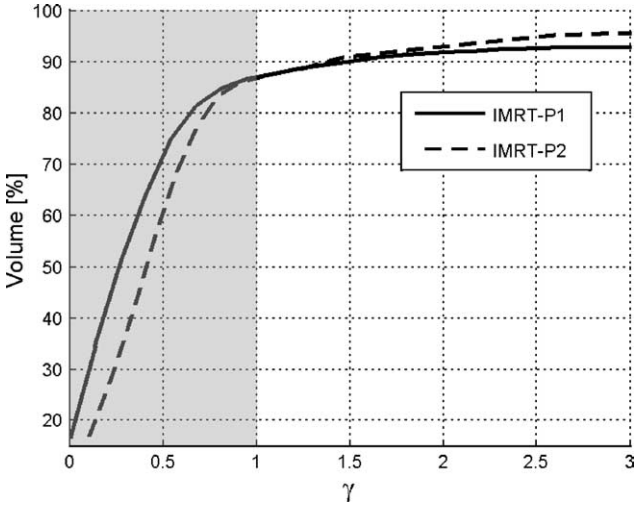

Fig. 6. Comparison of cumulative gamma volume histogram (cGVH) for P1 and P2. The shaded area where the γ criterion is satisfied is also shown.

Acknowledgements

ES is grateful for financial support from Cancer Research Wales and Yr Ysgol Uwchradd Tregaron.

* **Corresponding author. E. Spezi,** Present address: Servizio di Fisica Sanitaria, Policlinico S. Orsola-Mapighi, via Massarenti 9, 40138 Bologna (Italy). *E-mail address:* espezi@aosp.bo.it

Received 16 June 2005; received in revised form 12 January 2006; accepted 21 March 2006; Available online 11 May 2006

References

[1] Ahnesjö A, Saxner M, Trepp A. A pencil beam model for photon dose calculation. Med Phys 1992;19:263–73.
[2] Deasy JO, Niemierko A, Herbert D, et al. Methodological issues in radiation dose-volume outcome analysis: summary of a joint AAPM/NIH workshop. Med Phys 2002;29:2109–27.
[3] Depuydt T, van Esch A, Huyskens DP. A quantitative evaluation of IMRT dose distributions: refinement and clinical assessment of the gamma evaluation. Radiat Oncol 2002;62:309–19.
[4] Drzymala RE, Mohan R, Brewster L, et al. Dose—volume histograms. Int J Radiat Oncol Biol Phys 1991;21:71–8.
[5] Harms W, Low DA, Wong JW, Purdy JA. A software tool for the quantitative evaluation of 3D dose calculation algorithms. Med Phys 1998;25:1830–6.
[6] IMRTCWG Intensity-modulated radiation therapy: current status and issues of interest. Int J Radiat Oncol Biol Phys 2005;51(4):880–914.
[7] Love PA, Lewis DG, Smith GW, et al. A three-way comparison of radiotherapy dose distributions using polymer gel dosimetry, Monte Carlo simulation and treatment planning system. Med Phys 1999;26:1085.
[8] Low DA, Harms WB, Mutic S, Purdy JA. A technique for the quantitative evaluation of dose distributions. Med Phys 1998; 25:656–61.
[9] NEMA Digital Imaging and Communications in Medicine (DICOM). Rosslyn: National Electrical Manufacturers Association; 2001.
[10] Mijnheer B, Olszewska A, Fiorino C, et al. Quality assurance of treatment planning systems: practical examples for non-IMRT photon beams. Brussels: ESTRO Booklet; 2004.
[11] Rogers DWO, Faddegon BA, Ding GX, Ma CM, Wei J, Mackie TR. BEAM: a Monte Carlo code to simulate radiotherapy treatment units. Med Phys 1995;22:503–24.
[12] Spezi E, Lewis DG. Full forward Monte Carlo calculation of portal dose from MLC collimated treatment beams. Phys Med Biol 2002;47:377–90.
[13] Spezi E, Lewis DG, Smith CW. DICOM-RT-based toolbox for the evaluation and verification of radiotherapy plans. Phys Med Biol 2002;47:4223–32.
[14] Spezi E. A Monte Carlo investigation of the accuracy of intensity modulated radiotherapy. PhD dissertation, University of Wales (UK); 2003.
[15] Spezi E, Lewis DG, Millin AE, Cuffin R. MC based QA of IMRT. Radiat Oncol 2003;68:S46.
[16] Stock M, Kroupa B, Georg D. Interpretation and evaluation of the γ index and the γ index angle for the verification of IMRT hybrid plans. Phys Med Biol 2005;50:399–411.