A fast algorithm for gamma evaluation in 3D

Markus Wendling, Lambert J. Zijp, Leah N. McDermott, Ewoud J. Smit, Jan-Jakob Sonke, Ben J. Mijnheer, and Marcel van Herk[a)]
Department of Radiation Oncology, The Netherlands Cancer Institute - Antoni van Leeuwenhoek Hospital, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands

(Received 6 October 2006; revised 28 February 2007; accepted for publication 9 March 2007; published 19 April 2007)

The γ-evaluation method is a tool by which dose distributions can be compared in a quantitative manner combining dose-difference and distance-to-agreement criteria. Since its introduction, the γ evaluation has been used in many studies and is on the verge of becoming the preferred dose distribution comparison method, particularly for intensity-modulated radiation therapy (IMRT) verification. One major disadvantage, however, is its long computation time, which especially applies to the comparison of three-dimensional (3D) dose distributions. We present a fast algorithm for a *full* 3D γ evaluation at high resolution. Both the reference and evaluated dose distributions are first resampled on the same grid. For each point of the reference dose distribution, the algorithm searches for the best point of agreement according to the γ method in the evaluated dose distribution, which can be done at a subvoxel resolution. Speed, computer memory efficiency, and high spatial resolution are achieved by searching around each reference point with increasing distance in a sphere, which has a radius of a chosen maximum search distance and is interpolated "on-the-fly" at a chosen sample step size. The smaller the sample step size and the larger the differences between the dose distributions, the longer the γ evaluation takes. With decreasing sample step size, statistical measures of the 3D γ distribution converge. Two clinical examples were investigated using 3% of the prescribed dose as dose-difference and 0.3 cm as distance-to-agreement criteria. For 0.2 cm grid spacing, the change in γ indices was negligible below a sample step size of 0.02 cm. Comparing the full 3D γ evaluation and slice-by-slice 2D γ evaluations ("2.5D") for these clinical examples, the γ indices improved by searching in full 3D space, with the average γ index decreasing by at least 8%. © *2007 American Association of Physicists in Medicine*. [DOI: 10.1118/1.2721657]

I. INTRODUCTION

The high complexity of modern radiotherapy may result in unexpected deviations between the planned and delivered dose distribution. Many dose verification studies have been presented in the literature, in which delivered dose distributions measured with film, two-dimensional (2D) electronic devices, or gel are compared with calculated dose distributions. Alternatively, one might want to compare two dose distributions calculated by a treatment planning system (TPS) obtained in two different ways, for instance by using different dose calculation algorithms, beam models, or computed tomography (CT) data sets.

In 1998 Low *et al.*[1] introduced the γ-evaluation tool by which dose distributions can be compared in a quantitative manner. Since then, the original γ-evaluation method has been refined and modified by several authors.[2–7] This tool is fast becoming the reference dose distribution comparison method.

Already in 2D but especially in 3D, a major disadvantage of the γ-evaluation concept is its long computation time.[2,3] Bakai *et al.* revised the original γ concept by introducing gradient-dependent local acceptance thresholds, allowing fast 3D dose comparison.[4] In a recent study, Spezi and Lewis[7] compared 3D dose distributions, by performing an axial "slice-by-slice" 2D γ comparison, i.e., by using a stack of axial 2D γ maps. Although this approach might be useful in some cases, the comparison of 3D dose distributions using the γ tool should, in principle, be performed using a *full* 3D γ evaluation, i.e., including all three spatial dimensions of the dose distributions in the evaluation.[8] We present a fast algorithm in which the computation time is drastically reduced without any modification of the theoretical concept of the γ evaluation. The new algorithm is applicable to the 3D γ evaluation of large data sets. We illustrate our method with a comparison of a few simulated dose distributions and two clinical examples.

II. MATERIALS AND METHODS

A. Algorithm

The γ evaluation combines a dose-difference criterion and a distance-to-agreement (DTA) criterion for the comparison of two dose distributions.[1,2] Briefly, an evaluated dose distribution $D_e(\vec{r}_e)$ is compared to a reference dose distribution $D_r(\vec{r}_r)$, given at grid points $\vec{r}_e$ and $\vec{r}_r$, respectively. The dose difference is defined as

$$\delta(\vec{r}_e,\vec{r}_r) = D_e(\vec{r}_e) - D_r(\vec{r}_r), \tag{1}$$

and the spatial distance as

$$r(\vec{r}_e,\vec{r}_r) = |\vec{r}_e - \vec{r}_r|. \tag{2}$$

The generalized Γ function is defined as

$$\Gamma(\vec{r}_e,\vec{r}_r)=\sqrt{\frac{\delta^2(\vec{r}_e,\vec{r}_r)}{\Delta D^2}+\frac{r^2(\vec{r}_e,\vec{r}_r)}{\Delta d^2}}, \tag{3}$$

where ΔD and Δd are the acceptance criteria for the dose difference and DTA, respectively. As acceptance criterion ΔD for the dose difference, a local criterion can be used, e.g., 3% of the local dose, or a global criterion, e.g., 3% of a particular dose value, such as the prescribed dose. The evaluated point with the smallest $\Gamma(\vec{r}_e,\vec{r}_r)$ corresponds best to the reference point and is called the γ index of that reference point:

$$\gamma(\vec{r}_r)=\min\{\Gamma(\vec{r}_e,\vec{r}_r)\}\ \forall\ \{\vec{r}_e\}. \tag{4}$$

The γ index is the minimum distance in the renormalized dose-distance space. If $\gamma(\vec{r}_r)\leq 1$, the evaluated distribution is accepted at that point, otherwise it is rejected. As the γ index is defined for each reference point independently of the other points of the reference dose distribution, the γ-evaluation tool is in principle applicable to 1D, 2D, and 3D dose distributions. For an accurate dose comparison, the reference dose distribution should be contained within the space determined by the evaluated dose distribution, e.g., a 1D, 2D, or 3D reference dose distribution within a 3D evaluated dose distribution.[2]

The γ evaluation comes with two computational challenges. First, the γ evaluation should be performed quickly, in order to obtain verification results within a reasonable time. However, the computation is time consuming since the generalized Γ function needs to be calculated for every evaluated point $\vec{r}_e$ relative to each reference point $\vec{r}_r$. Secondly, the data sets have to fit into the available computer memory. For the latter, we consider a dose distribution data set of 256^3 points, corresponding to a cube of ~26 cm in each direction with a resolution of 0.1 cm. For 4-byte floating numbers, that data set would correspond to 64 MB of data. In general, the evaluated dose distribution should have a higher resolution than the reference distribution, e.g., 0.02 vs 0.1 cm, in order to avoid discretization artifacts. However, the memory required for only holding the evaluated distribution would then be 8000 MB. Since such a memory requirement exceeds modern PC specifications, we developed a concept, whereby both computation time and memory requirement can be solved at the same time.

First, the reference dose distribution $D_r(\vec{r}_r)$ and the evaluated dose distribution $D_e(\vec{r}_e)$ are resampled on the same uniform grid, i.e., a grid with a constant voxel spacing in each dimension (but not necessarily the same spacing for all dimensions). For every $\vec{r}_r$ we calculate $\gamma(\vec{r}_r)$ by searching through $D_e(\vec{r}_e)$, which can be done even at a subvoxel resolution. The search starts at $\vec{r}_e=\vec{r}_r$, i.e., at the same point in space, since the probability of finding the minimum Γ is generally highest around that point. We search within a sphere around this starting point. The sphere has a radius of a chosen maximum search distance and the dose values are interpolated linearly "on-the-fly" at a chosen sample step size; the maximum search distance should be (much) larger than the DTA criterion (e.g., 1.0 vs 0.3 cm) and the sample step size should be a fraction of the grid resolution, e.g., 0.02 vs 0.1 cm. Computation speed is gained by calculating (once) a sorted table of distances from the center in this sphere and then stepping through the distance space with increasing spatial distance from the starting point. The implementation of the algorithm is illustrated in 2D in Fig. 1.

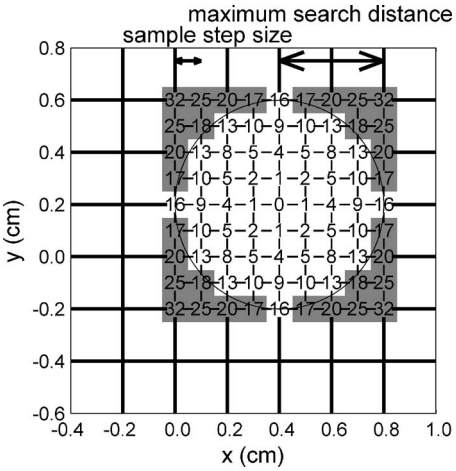

FIG. 1. Searching scheme for our fast γ-evaluation method, illustrated in 2D. The bold lines represent the grid of the reference and evaluated dose distributions. In this illustration the pixel size is 0.2×0.2 cm^2, the chosen sample step size 0.1 cm, and the chosen maximum search distance 0.4 cm (note that for the calculations in this study a maximum search distance of 1.0 cm was used). Around the reference point (here at $x=0.4$ cm and $y=0.2$ cm) a square with sides of twice the maximum search distance is created and interpolation factors for the given sample step size are determined. Points outside the circle with a radius of the maximum search distance are disregarded (gray background). The squared distances (in pixels) from the reference point are listed. By sorting this list, a search in a circle with increasing distance from the reference point is possible.

In order to interpolate the evaluated dose distribution at a certain point within the grid, the dose values at the surrounding grid points are weighted with *interpolation factors* and then summed. Because of identical uniform grids, these interpolation factors depend only on the spatial location of the "point-to-interpolate" relative to the surrounding grid points and are therefore independent of the reference point. Therefore, all interpolation factors necessary for the linear interpolation of the dose values of the evaluated distribution are calculated only once, which makes the interpolation fast.

More speed is gained by not calculating $\Gamma(\vec{r}_e,\vec{r}_r)$ for evaluated points that "do not have a chance" to reach the minimal Γ: As soon as $r^2(\vec{r}_e,\vec{r}_r)/\Delta d^2$ becomes larger than the minimal Γ found so far, the search for that reference point can be stopped, since the distance itself prevents this point and all succeeding points from reaching an even smaller Γ value. The concept of searching by starting at the reference point, increasing the search distance, and terminating the search when Γ cannot be improved anymore has been

TABLE I. 3D γ evaluation, applied to examples A, B, and C of simulated dose distributions. The reference dose distribution was a cube of $(25.6\ \text{cm})^3$ with $(0.1\ \text{cm})^3$ voxels with a "central cube" of $(10\ \text{cm})^3$ at the center with dose values of 100 cGy (the "prescribed" dose); outside the central cube the dose values were set to 0 cGy. (A) For the evaluated dose distribution the central cube had dose values of 101 cGy and was shifted $0.1/\sqrt{3}$ cm in both x, y, and z direction relative to the reference central cube, i.e., the total length of the shift vector was 0.1 cm; (B) same as (A), but with dose values of 103 cGy and shifts of $0.3/\sqrt{3}$ cm; (C) same as (A), but with dose values of 105 cGy and shifts of $0.5/\sqrt{3}$ cm. For all cases, the dose-difference criterion was 3% of the prescribed dose, i.e., 0.03×100 cGy$=3$ cGy, and the DTA criterion was 0.3 cm. The maximum search distance was 1.0 cm. The grid resolution is the resolution of the (initially interpolated) reference and evaluated dose distributions before the on-the-fly interpolation of the evaluated distribution at the sample step size. In these examples, no on-the-fly interpolation of the evaluated dose distribution was necessary.

Example	Grid resolution (cm) [pixels]	Sample step size (cm)	γ_{avg}	γ_{max}	$\gamma_{1\%}$	$P_{\gamma\leq 1}$ (%)	Computation time (s)
A	0.1	0.10	0.34	0.67	0.47	100.0	2
B	$[256]^3$	0.10	1.01	1.53	1.20	94.1	15
C		0.10	1.68	2.40	1.94	0.0	52

presented for 2D dose distributions.[5] Additionally, effective programming is of course important, such as omitting superfluous calculations. For instance, only Γ^2 (and not Γ) is initially computed and minimized, so only one computationally expensive square root needs to be extracted per reference point.

In this work, the dose-difference criterion ΔD was 3% of the prescribed dose and the DTA criterion Δd was 0.3 cm. The maximum search distance was 1.0 cm. The sample step size was varied.

In addition to the full 3D γ evaluation, a slice-by-slice γ evaluation was performed by calculating 2D γ maps from corresponding slices of the two 3D dose distributions yielding a "pseudo" 3D γ distribution. This method will be called "2.5D γ evaluation." The 2.5D γ evaluation was done for coronal, axial, and sagittal slices. For the comparison of a 2.5D and the full 3D γ evaluation, the 3D γ-difference volume was calculated ("$\Delta\gamma=\gamma_{2.5D}-\gamma_{3D}$").

In order to display the 3D γ distribution, the maximum intensity projection (MIP) in the axial plane was used. To focus on the target volume, the 3D γ distribution was first cropped by taking axial slices between 7 cm superior and 7 cm inferior to the isocenter. The maximum γ index of all voxels projected in the superior-inferior direction was then determined to yield the γ^{MIP} image. The same technique was used for the 3D γ-difference volume, yielding the $\Delta\gamma^{MIP}$ image.

For the statistical evaluation of the γ distributions, we defined a volume of interest, bound by the 50% isodose surface (relative to the prescribed dose). Within this surface we determined the average γ (γ_{avg}), the maximum γ (γ_{max}), the 99th percentile ($\gamma_{1\%}$)[5] [symbol chosen in analogy with $D_{1\%}$ in dose-volume histograms, as 1% of points have an equal or higher γ value], and the percentage of points in agreement ($P_{\gamma\leq 1}$). The 50% isodose surface was chosen as it separates the "higher"-dose from the "lower"-dose regions, and we wanted to focus on the differences between the two dose distributions, which were more pronounced in the region of the target volume (see Sec. III). Including the lower-dose regions would lower the average γ and conceal differences. Note that the γ distributions themselves were calculated over the total volumes. For all computations we used a standard PC with a Pentium 4 processor (2.4 GHz, FSB 4×133 MHz).

B. Simulated and clinical examples

We first simulated an artificial reference dose distribution. The grid was a cube of $(25.6\ \text{cm})^3$ with $(0.1\ \text{cm})^3$ voxels. At the center there was a "central cube" of $(10\ \text{cm})^3$ with dose values of 100 cGy (the "prescribed" dose); outside the central cube the dose values were set to 0 cGy. For the evaluated distribution in example A, the central cube had dose values of 101 cGy and was shifted $0.1/\sqrt{3}$ cm in both x, y, and z direction, i.e., the total length of the shift vector was 0.1 cm. For the evaluated distributions in examples B and C, the dose values in the central cube were 103 and 105 cGy, respectively, and the total shifts were 0.3 and 0.5 cm, respectively, evenly distributed over all three spatial dimensions. For all cases, the dose-difference criterion was 3% of the prescribed dose, i.e., 0.03×100 cGy$=3$ cGy, and the DTA criterion was 0.3 cm.

As a practical illustration, two clinical patient plans for lung cancer treatment were used. Each plan was calculated on two different CT data sets. The first plan (example D, taken from Ref. 9) was a 3D conformal treatment consisting of four coplanar 8 MV photon beams with wedges. The dose prescription to the planning target volume was 67.5 Gy given over 30 fractions. The first CT scan (the "planning CT scan") was taken two weeks before treatment, whereas the second CT scan (the "repeat CT scan") was taken in the beginning of week 5.

The second clinical example (E) was an intensity-modulated radiation therapy (IMRT) treatment plan consisting of six coplanar 6 MV photon beams. The dose prescription to the planning target volume was 66.0 Gy, delivered over 24 fractions. A 4D respiration-correlated CT scan was acquired from which a maximum inhale (0% phase) and a

maximum exhale CT data set (50% phase) were constructed.[10] The CT scans representing these two respiration phases were only used for this study; the clinical plan was optimized for the time-averaged mean tumor position.

Note that in both examples (D,E) no plan parameters were changed, i.e., the same collimator and gantry angles, multileaf collimator settings, beam energies, and numbers of monitor units were used for both CT data sets in each example. The TPS Pinnacle (Philips Medical Systems, Eindhoven, The Netherlands) was used for planning and dose calculation. The TPS dose calculations were performed on a dose grid with $(0.4\ \text{cm})^3$ voxels. The dose-difference criterion was for both examples 3% of the prescribed dose and the DTA criterion was 0.3 cm.

III. RESULTS

A. Simulated examples

The results of the 3D γ evaluations for the three simulated examples A, B, and C are presented in Table I. The values for γ_{avg}, γ_{max}, $\gamma_{1\%}$, and $P_{\gamma\leq 1}$ are listed and agree with the values that can be calculated analytically using Eqs. (1) to (4), taking the finite resolution into account, as well as the fact that these statistical parameters were determined within the 50% isodose surface (relative to the prescribed dose). Let us, for instance, consider γ_{max} in example B. γ_{max} is the γ index of one corner of the central cube of the reference dose distribution corresponding to the same corner of the evaluated dose distribution. The dose difference δ was 3 cGy. The shift of the central dose cube was chosen as $0.3/\sqrt{3}$ cm in each dimension. On the 0.1 cm grid used, this corresponded effectively to a spatial distance r of 0.2 cm in each dimension. Therefore,

$$\gamma_{\text{max}} = \sqrt{\frac{(3\ \text{cGy})^2}{(0.03 \times 100\ \text{cGy})^2} + 3 \times \frac{(0.2\ \text{cm})^2}{(0.3\ \text{cm})^2}} = 1.53.$$

Furthermore, because the grid resolution and sample step size were equal, we can calculate $P_{\gamma\leq 1}=[(10.0\ \text{cm} - 0.2\ \text{cm})/(10.0\ \text{cm})]^3 = 94.1\%$ for example B. For example C, $P_{\gamma\leq 1}$ equals zero, because the introduced dose difference was larger than the dose-difference criterion ΔD.

With respect to computation time, the larger the introduced differences are, the longer the computation takes. For the largest differences (5%/0.5 cm) the computation time was approximately 52 s.

B. Clinical examples

In the first clinical example (D), a dose distribution based on the planning CT scan was compared with a dose distribution based on a repeat CT scan after several weeks of treatment. Axial slices through both CT scans are depicted in Figs. 2(a) and 2(b), where the shrunken tumor can be discerned. As a consequence of the changed anatomy, the dose distributions were different, which resulted in high γ values in that region. This can be seen in the axial γ^{MIP} image in Fig. 2(c).

The initial grid resolutions and the sample step size were varied for the γ evaluation. The numerical results are listed in Table II. In general, the higher the resolution, the longer the computation took. The values for γ_{avg}, γ_{max}, $\gamma_{1\%}$, and $P_{\gamma\leq 1}$ improved with decreasing sample step size, and eventually converged. Below a sample step size of 0.02 cm, the changes in γ_{avg}, γ_{max}, $\gamma_{1\%}$, and $P_{\gamma\leq 1}$ were negligible for 0.4, 0.2, and 0.1 cm grid resolution. The computation time required was 34, 246, and 2024 s, respectively.

The influence of the sample step size on the γ distribution is illustrated in Fig. 3, where we show two γ histograms. In both cases the dose distributions had an initial grid resolution of 0.2 cm, but one γ distribution was calculated with a sample step size of 0.20 cm, whereas for the other a sample step size of 0.02 cm was used. Like the statistical parameters, the histograms were determined within the 50% isodose surface (relative to the prescribed dose). In the histograms several distinct peaks can be discerned, the locations of which are indicated in the figure.

The comparison between the axial 2.5D γ evaluation and the full 3D γ evaluation is illustrated in Fig. 2(d) by the axial $\Delta\gamma^{\text{MIP}}$ image. Note that all differences of the $\Delta\gamma$ volume were non-negative, i.e., 2.5D γ indices were never smaller than the corresponding 3D γ indices. There were large deviations in the region of the tumor, where—by including the third spatial dimension (superior-inferior)—smaller γ indices could be found. This can be derived quantitatively from Table II by comparing the 2.5D and 3D γ approaches for a grid resolution of 0.2 cm and a sample step size of 0.02 cm, e.g., γ_{avg} decreased from 0.41 for the axial 2.5D γ evaluation to 0.37 for the full 3D γ evaluation.

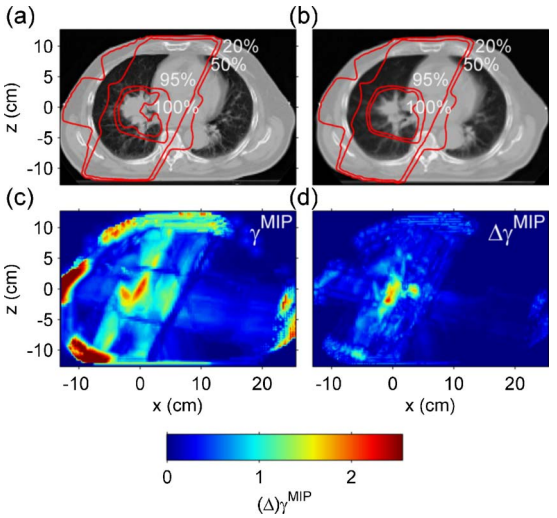

FIG. 2. Axial CT slices (through the isocenter) from (a) the original planning CT scan and (b) a repeat CT scan of a lung cancer patient (example D). The 3D conformal treatment plan was calculated on every CT data set. Isodose lines, relative to the prescribed dose, are indicated. (c) Axial γ^{MIP} image (MIP, maximum intensity projection) from the 3D γ evaluation, 0.2 cm grid resolution and 0.02 cm sample step size. (d) Axial $\Delta\gamma^{\text{MIP}}$ image showing the difference between the axial 2.5D γ evaluation and the full 3D γ evaluation. Both MIP images were calculated from axial slices between 7 cm superiorly and 7 cm inferiorly around the isocenter.

TABLE II. 3D γ evaluation, applied to examples D and E of clinical dose distributions. (D) Clinical example of a 3D conformal lung cancer treatment calculated on both the planning and a repeat CT scan after four weeks of treatment. (E) Clinical example of an IMRT lung cancer treatment calculated on both the maximum inhale and the maximum exhale CT scan. The rows with "2.5D" refer to the slice-by-slice γ evaluations as described in the text. Computation times were not listed for the 2.5D γ evaluations as a fast implementation was irrelevant for this study. The dose-difference criterion was for both examples 3% of the prescribed dose and the DTA criterion was 0.3 cm. The maximum search distance was 1.0 cm. The grid resolution is the resolution of the (initially interpolated) reference and evaluated dose distributions before the on-the-fly interpolation of the evaluated distribution at the sample step size.

Example	Grid resolution (cm) [pixels]	Sample step size (cm)	γ_{avg}	γ_{max}	$\gamma_{1\%}$	$P_{\gamma\leq1}$ (%)	Computation time (s)
D	0.4	0.40	0.53	3.19	2.00	87.5	<1
	[97×89×64]	0.20	0.49	3.06	1.61	93.4	<1
		0.10	0.43	3.03	1.55	94.3	<1
		0.05	0.39	3.03	1.53	94.7	3
		0.02	0.37	3.02	1.53	94.8	34
		0.01	0.36	3.02	1.53	94.8	253
	0.2	0.20	0.49	3.11	1.59	93.5	1
	[193×177×127]	0.10	0.43	3.11	1.54	94.3	3
		0.05	0.40	3.10	1.52	94.8	19
		0.02	0.37	3.10	1.52	94.8	246
		0.01	0.37	3.10	1.52	94.9	1851
	0.1	0.10	0.43	3.15	1.52	94.4	28
	[385×353×253]	0.05	0.40	3.14	1.51	94.8	157
		0.02	0.38	3.14	1.51	94.9	2024
		0.01	0.37	3.14	1.51	94.9	15163
2.5D, coronal	0.2	0.02	0.44	15.44	2.52	92.7	–
2.5D, axial	[193×177×127]	0.02	0.41	3.24	1.67	93.6	–
2.5D, sagittal		0.02	0.41	7.17	1.62	93.8	–
E	0.2	0.02	0.29	3.09	1.88	93.4	106
2.5D, axial	[205×169×203]	0.02	0.38	5.25	2.71	89.2	–

When we compare the values for the different 2.5D approaches, i.e., coronal, axial, and sagittal, some variations could be found (see Table II). For instance, γ_{avg} varied between 0.41 and 0.44 depending on the selected slice direction for the 2.5D γ evaluation. The minimum γ index of all three 2.5D γ evaluations at each voxel was never smaller than the corresponding γ index of the full 3D γ evaluation.

In the second example (E), two dose distributions were compared based on CT data sets for (maximum) inhale and exhale positions of the diaphragm (Fig. 4). The different position of the tumor can be seen in the coronal slices through the CT scans [Figs. 4(a) and 4(b)]. The 3D γ evaluation showed differences larger than $\gamma=2$, i.e., larger than 6% or 6 mm, in the region of the (moving) tumor. Also, in this example, the comparison between the full 3D γ evaluation and the axial 2.5D slice-by-slice γ evaluation shows large differences [Fig. 4(d)]. Again, all differences in the $\Delta\gamma$ volume were non-negative. In the $\Delta\gamma^{MIP}$ image large deviations in the region of the tumor can be seen, illustrating that inclusion of the third spatial dimension (superior-inferior) led to smaller γ indices in that region. For instance, from Table II it can be seen that γ_{avg} decreased from 0.38 for the axial 2.5D γ evaluation to 0.29 for the full 3D γ evaluation.

IV. DISCUSSION

3D dose distributions can be compared using the 3D γ-evaluation method. As an illustration, we showed that tumor regression and respiration-induced tumor motion can have considerable impact on the dose distributions and that the differences can be rapidly quantified using our 3D γ-evaluation algorithm. Moreover, a full 3D approach is necessary for comprehensive quantitative results.

Comparison of computation speed for the γ evaluation is difficult since the speed of the algorithm depends on the general correspondence between the two dose distributions relative to the acceptance criteria. Therefore, γ evaluations will be faster if the differences are smaller for the same acceptance criteria. In addition, the computation speed depends on the shape of the dose distributions. For steeper dose gradient regions the Γ function has to be determined over a larger region potentially, since computation of the Γ function

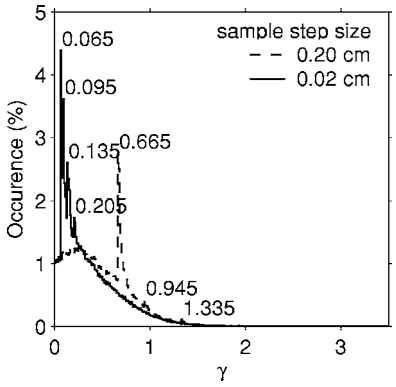

FIG. 3. Histograms of two γ distributions comparing a dose distribution based on the planning CT scan with a dose distribution based on a repeat CT scan of a lung cancer patient (example D). The γ distributions were calculated for 0.2 cm grid resolution with sample step sizes of 0.20 cm (dashed line) and 0.02 cm (solid line). The histograms were made for the γ values within the 50% isodose surface (relative to the prescribed dose) using a binning width of 0.01. The location of several distinct peaks in the histogram, specified by the center of the bin, is indicated in the figure.

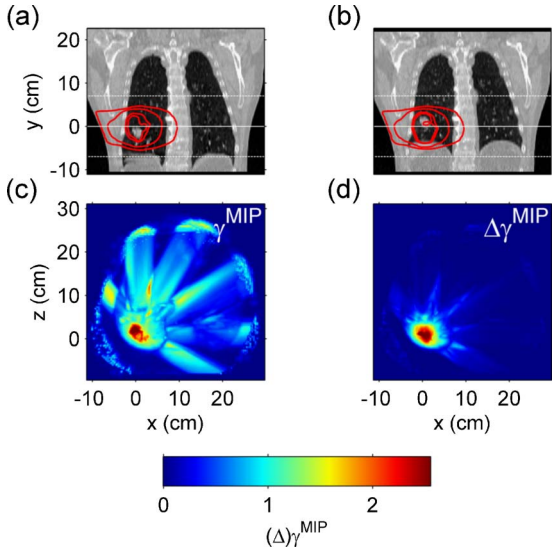

FIG. 4. Coronal CT slices (through the isocenter) from the CT scans for (a) maximum inhale and (b) maximum exhale of a lung cancer patient (example E). The IMRT treatment plan was calculated on every CT data set. Isodose lines (100, 95, 50, and 20% relative to the prescribed dose) are indicated. The horizontal solid white lines indicate the axial slice at the position of the isocenter; the horizontal dashed white lines indicate the axial slices between which the MIP images in (c) and (d) were calculated. (c) Axial γ^{MIP} image (MIP, maximum intensity projection) from the 3D γ evaluation, 0.2 cm grid resolution and 0.02 cm sample step size. (d) Axial $\Delta\gamma^{MIP}$ image showing the difference between the axial 2.5D γ evaluation and the full 3D γ evaluation. Both MIP images were calculated from axial slices between 7 cm superiorly and 7 cm inferiorly around the isocenter.

starts at the position of the reference point. We used several simulated dose distributions, for which all details have been given. In this way, interested readers can simulate the same dose distributions, and the speed of other γ-evaluation implementations can be tested against our implementation. For comparison, Low *et al.*[2] published a typical calculation time of 5 min for a 14.1×15.2 cm^2 radiographic film with a 1×1 mm^2 resolution on an 800 MHz Pentium 4. After correction for the difference in CPU frequencies by a factor of 1/3 (800 MHz versus 2.4 GHz in our study), we arrive at approximately 100 s for 141×152 pixels. The calculation time for example D at 0.1 cm grid resolution and sample step size was only 28 s (see Table II), despite the fact that we used a much larger 3D data set.

Bakai *et al.* revised the γ evaluation by introducing gradient-dependent local acceptance thresholds.[4] Although gradient estimation is a potential problem in the presence of noise in the data sets and although slightly different results are obtained compared with the original γ concept, their algorithm is fast in 3D because no search in the 3D space is necessary. The calculation time for both the γ evaluation and the gradient method is approximately proportional to the number of points in the reference dose distribution. By using the information given in Ref. 4, i.e., both dose distributions with 0.1 cm grid resolution and evaluation of 241^2 pixels in 25/120 s, we estimate that a full 3D evaluation of our data set D with $385\times353\times253$ voxels at 0.1 cm grid resolution and sample step size, would have taken approximately 120 s using that method. If we correct this by a factor of 1/2.4 for the difference in CPU frequencies (1 versus 2.4 GHz in our case), we arrive at 50 s, which should be compared to 28 s for our implementation. As this is a very rough estimate based on limited information, the computation speed can be considered to be similar.

For our fast 3D algorithm, both dose distributions have first to be resampled on the same uniform grid. The uniform grid is necessary for the fast on-the-fly interpolation of the evaluated dose distribution using only-once-calculated linear-interpolation factors. This uniform grid is not a serious problem, because dose distributions are usually generated on such a grid, like in the TPS, or can easily be resampled.

Our algorithm is faster than a straightforward implementation of the γ-evaluation tool since searching is done in an efficient way. Efficient searching by starting at the reference point, then increasing the search distance, and finally stopping the search when no improvement is possible anymore, has been presented by Stock *et al.* for 2D dose distributions.[5] Note that in their method, the γ evaluation needed to be performed with the evaluated dose distribution at the same resolution as the reference dose distribution. The interpolation of the evaluated dose distribution *on-the-fly* in our method is fast because the linear-interpolation factors are calculated only once and the problem of vast computer memory requirements for high-resolution evaluations is solved. Note that in our algorithm no compromises or approximations are made to the γ-evaluation concept, except for an optional maximum search distance. Restriction of the search distance will not affect γ values below the maximum search distance divided by the distance acceptance criterion. A maximum search distance should therefore be set in such a way that it does not influence γ values within the range of interest. In our case of a maximum search distance of 1.0 cm and a DTA criterion of 0.3 cm, only γ indices above 3.33 would be affected. In our examples, all maximum γ values inside the 50% isodose surfaces were below this value for the

3D γ evaluations. The high values of γ_{max} in example D for the coronal and sagittal 2.5D and in example E for the axial 2.5D γ evaluations might be improved by increasing the maximum search distance, however, this effect was not considered relevant; in 3D, these γ_{max} values were smaller due to the inclusion of the third dimension.

The sample step size should be a (small) fraction of the grid resolution. Decreasing the sample step size improved the results. The values for γ_{avg}, γ_{max}, $\gamma_{1\%}$, and $P_{\gamma\leq 1}$ converge. This can be understood by considering a certain point of the reference dose distribution. When changing from a large to a small sample step size, more points of the evaluated distribution are evaluated, and therefore the γ index, being the minimum of the Γ function over the evaluated points, is likely not to increase; this is certainly the case if the larger sample step size is a multiple of the smaller one. The γ index will therefore approach a certain minimum end value for decreasing sample step size. Theoretically, this minimum γ index can be found for an infinitesimally small sample step size. Numerically, it can be calculated by solving a nonlinear minimization problem, which, however, takes a long time.

The influence of the sample step size on the γ distribution was illustrated in Fig. 3, where the shift of the histogram with decreasing sample step size towards lower γ values can be seen. Several peaks can be discerned in these histograms. The peaks are related to the finite sample step size. For example, with a sample step size of 0.20 cm and a DTA criterion of 0.3 cm, γ indices below $\sqrt{(0.2\text{ cm})^2/(0.3\text{ cm})^2} \approx 0.667$ are determined by dose difference only (note that with a binning width of 0.01, this γ value falls into the γ bin centered at 0.665). For larger γ values, dose difference *and* distance can play a role, and therefore *suddenly* more points can contribute to this γ bin. This is a "discontinuity" in the γ method due to the finite sample step size. The next peak occurs at $\sqrt{2\times(0.2\text{ cm})^2/(0.3\text{ cm})^2} \approx 0.943$ (γ bin centered at 0.945), etc. All peaks in the γ histograms—also for the other γ distribution calculated with 0.02 cm sample step size—can be explained in this way.

Low *et al.* recommended 1/3 of the DTA criterion Δd for the pixel spacing of the evaluated dose distribution (corresponding to the sample step size in our implementation).[2] Nevertheless, we still see changes in γ_{avg}, for example, for smaller sample step sizes (see Table II). If one needs precise quantitative γ results, 1/10 Δd appears to be a value more appropriate for the pixel or voxel spacing (sample step size) of the evaluated distribution.

Jiang *et al.* developed a general framework for the comparison of dose distributions.[6] Their concept, called "maximum allowed dose difference," includes the γ evaluation as one special case of comparison methods. For the numerical implementation of the γ evaluation itself, they proposed an iterative multiresolution algorithm to deal with the resolution issue. In their approach, the sample step size of the evaluated distribution is adapted (made finer) during the search for the smallest Γ for each reference point. This means that search and interpolation are also necessary in that implementation.

The aim of this study was to implement a fast full 3D γ-evaluation method. Pseudo 3D γ evaluations have been presented as stacks of 2D γ maps, i.e., in a slice-by-slice fashion.[7] For a 2D stack of γ maps, the γ evaluation depends on the selected slice direction in which the 2D maps are taken (see Table II). Including the third spatial dimension correctly in a full 3D evaluation is therefore important. Note that a full 3D γ analysis will always yield lower γ indices than a 2.5D stack analysis.[7,8] Moreover, the minimum of the three 2.5D approaches will usually still be larger than the full 3D approach, because in the former situation only dose values on the three intersecting planes instead of the whole volume are evaluated for every reference point.

The ability to perform (3D) γ evaluations much faster will possibly increase the various application types of the γ concept, for example, in online treatment verification or in optimization routines for treatment planning. As most treatment plans are based on 3D dose calculations, a fast and accurate full 3D γ evaluation is essential for comparing and verifying dose distributions of high-precision treatment techniques.

V. CONCLUSIONS

A full 3D γ evaluation can be performed in a reasonable amount of time by searching in the evaluated dose distribution with increasing distance in a sphere around each point of the reference dose distribution. Small sample step sizes in the evaluated distribution can be obtained without large computer memory requirements by calculating the linear-interpolation factors only once and interpolating the evaluated dose distribution within the search sphere on-the-fly.

ACKNOWLEDGMENTS

This work was funded by the Dutch Cancer Society (Grant No. NKI 2000-2255) and the Netherlands Foundation for Technology and Science (Stichting Technologie en Wetenschappen, Grant BKG.7184). Ewoud Smit was supported by the Department of Biomedical Engineering, University of Groningen, Groningen, The Netherlands.

[a]Author to whom correspondence should be addressed: The Netherlands Cancer Institute - Antoni van Leeuwenhoek Hospital, Department of Radiation Oncology, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands. Electronic mail: m.v.herk@nki.nl

[1]D. A. Low, W. B. Harms, S. Mutic, and J. A. Purdy, "A technique for the quantitative evaluation of dose distributions," Med. Phys. **25**, 656–661 (1998).

[2]D. A. Low and J. F. Dempsey, "Evaluation of the gamma dose distribution comparison method," Med. Phys. **30**, 2455–2464 (2003).

[3]T. Depuydt, A. Van Esch, and D. P. Huyskens, "A quantitative evaluation of IMRT dose distributions: Refinement and clinical assessment of the gamma evaluation," Radiother. Oncol. **62**, 309–319 (2002).

[4]A. Bakai, M. Alber, and F. Nüsslin, "A revision of the γ-evaluation concept for the comparison of dose distributions," Phys. Med. Biol. **48**, 3543–3553 (2003).

[5]M. Stock, B. Kroupa, and D. Georg, "Interpretation and evaluation of the γ index and the γ index angle for the verification of IMRT hybrid plans," Phys. Med. Biol. **50**, 399–411 (2005).

[6]S. B. Jiang, G. C. Sharp, T. Neicu, R. I. Berbeco, S. Flampouri, and T. Bortfeld, "On dose distribution comparison," Phys. Med. Biol. **51**, 759–776 (2006).

[7]E. Spezi and D. G. Lewis, "Gamma histograms for radiotherapy plan evaluation," Radiother. Oncol. **79**, 224–230 (2006).

[8]S. Gillis, C. De Wagter, J. Bohsung, B. Perrin, P. Williams, and B. J. Mijnheer, "An inter-centre quality assurance network for IMRT verification: Results of the ESTRO QUASIMODO project," Radiother. Oncol. **76**, 340–353 (2005).
[9]L. N. McDermott, M. Wendling, J.-J. Sonke, M. van Herk, and B. J. Mijnheer, "Anatomy changes in radiotherapy detected using portal imaging," Radiother. Oncol. **79**, 211–217 (2006).
[10]J. W. H. Wolthaus, C. Schneider, J.-J. Sonke, M. van Herk, J. S. A. Belderbos, M. M. G. Rossi, J. V. Lebesque, and E. M. F. Damen, "Mid-ventilation CT scan construction from four-dimensional respiration-correlated CT scans for radiotherapy planning of lung cancer patients," Int. J. Radiat. Oncol., Biol., Phys. **65**, 1560–1571 (2006).