Gamma Dose Distribution Evaluation Tool

Daniel A. Low, Ph.D.
Washington University, Department of Radiation Oncology, 4921 Parkview Place, St. Louis, MO 63110;
low@wustl.edu

Abstract. Quantitative comparisons of dose distributions are an integral component of a medical physicist’s responsibility of assuring high quality radiation therapy dose delivery. While overlays and other displays of multiple dose distributions are useful for such evaluations, quantitative evaluations require a mathematical comparison. The dose-difference is the most straightforward method for comparing two dose distributions, but it can show large differences in steep dose gradient regions, even for relatively small misalignments. A tool, termed $\gamma$, was developed to take both dose and spatial difference into account. It does this automatically, evaluating distributions for dose difference and spatial discrepancies in regions of shallow and steep dose gradients, respectively. The tool has been used extensively in commercial dose measurement and evaluation software. This chapter describes the tool, some alternative techniques, and limitations of the tool.

1. Dose Distribution Evaluations
Dosimeters provide critically important information about the characteristics of dose distributions. They allow us to directly measure a physical phenomenon that is capable of irradiation tumors such that they are sterilized while nearby normal organs retain their function. The dose accuracy required is typically assumed to be within a few percent in order to provide the potential benefit. This is quite challenging to achieve in a complex structure such as the human body, so radiation therapy relies on sophisticated computer dose calculations to predict the dose inside the patient and allow the treatment planner to visualize the impact of selecting different beam orientations, energies, and intensities. Computer-optimized dose distributions are also used, and these have very complex dose deliveries and consequently dose distribution characteristics, such as steep conformal dose gradients that wrap around the tumors and normal organs.

The complexity of the dose delivery and the resulting distributions, and the tight requirement for accuracy, result in the need for direct measurement-based verification of the calculated dose distributions. This has led to the main purpose of this conference; the 3-dimensional measurement of dose distributions. However, having a 3-dimensional calculation and measurements are only part of the dose distribution quality assurance (QA) process. The dose distributions need to be compared to one another using methods that are quantitative, efficient, and that take maximum advantage of the limited capability for visualizing the dose distributions, which are 3-dimensional scalar fields.

The first group to publish on applying quantitative tools for dose distribution comparisons was the group of Van Dyk et al. They showed that dose comparisons should consider the local dose gradient when determining the best way to compare doses. The most straightforward way to compare two dose distributions is to take the numerical difference. Presumably, if you want to know by how much the two dose distributions disagree, the difference makes it clear. However, there are practical limitations
to this method. First, the dose distributions may not be precisely aligned. If one of them is measured, there is a spatial tolerance on the ability to position the detector, align the phantom to the linear accelerator, and to conduct the readout process (e.g. optical densitometry). If there is a spatial error in the measurement, the dose distribution differences that lie within steep dose gradients will be artificially enhanced by the spatial offset of the two distributions.

While a spatial offset leads to sometimes large dose differences in steep dose gradient regions, one can ask the question; why are these differences not relevant to the patient’s treatment? They existed when a physicist made the dose distribution measurements, so why not consider them when evaluating the dose distribution quality? This is a valid argument, and the answer has to do with the purpose of the measurement. For most cases, the purpose of the measurement is to validate the output of the treatment planning system, or to validate the system-wide process of calculation, data transfer, and dose delivery. A reflection of the impact of measurement-induced errors is typically not desired by the person conducting or evaluating the QA results. Therefore, one of the desired features of dose distribution comparisons is to be somewhat immune to experimental error.

Van Dyk et al \(^1\) published on the importance of applying a particular dose distribution comparison test based on the local dose gradient. In regions of shallow dose gradient, the dose difference provided the physicist with a quantitative and straightforward method for comparing two dose distributions. However, in steep dose gradient regions, the dose difference was sensitive to the small spatial offsets commonly encountered in experiments, so even if the dose distributions agreed exactly, the dose differences were very large when even a small spatial error was encountered. As an example, IMRT dose distributions often dose gradients of close to 3% mm\(^{-1}\). Even a small position shift of one dose distribution relative to the other will cause large dose differences. In this example, a 1 mm shift will cause a 3% dose difference. Selecting a 3% dose difference criterion when comparing the two dose distributions will lead to false positive detections of regions where the calculation appears to fail to accurately model the dose distribution. Van Dyk \(^1\) developed the concept of distance to agreement (DTA), which Harms et al \(^2\) applied in a software tool to compare two-dimensional dose distributions.

The DTA is typically computed by measuring the closest distance between one dose distribution to where the second distribution has the same dose level. This is equivalent to determining the closest approach to a point in one distribution of the isodose line that in the other distribution that has the same dose as the point. The DTA is a non-local function, meaning that when it is evaluated at a point, the other distribution is queried at a distance. The DTA function also does not commute. While the dose difference merely changes sign when the two dose distributions are swapped, the DTA is not the same value of the roles of the two distributions are swapped. Therefore it is useful to label the two distributions as the reference and evaluated distributions. The reference distribution is the one whereby the DTA is computed point by point, while the evaluated distribution is the one queried for the closest approach of the specific isodose.

One of the nice features of the dose difference distribution is its relatively straightforward interpretation. Most physicists can glance at a dose difference distribution and see and interpret the data. Because of the impact of steep dose gradients, the dose differences may not be clinically relevant, but the meaning behind the numbers is clear. While the dose difference is overly sensitive in steep dose gradient regions, the DTA is overly sensitive in shallow dose gradient regions. As an example, if the reference dose point has dose D, lying in a shallow dose gradient region, and the evaluated dose at that point is D+\(\delta D\), the distance in the evaluated distribution from the reference point that has the same dose D may be very far away. The DTA will therefore be very large, even if \(\delta D\) is small and has no clinical consequence. Because of the large values in shallow dose gradient regions, regions that typically dominate IMRT dose distributions, the DTA distribution is difficult, if not impossible to interpret by eye.

In steep dose gradient regions, the DTA can be interpreted as the distance between the two dose distributions. This interpretation is based on the assumption that the distance is caused primarily by a spatial offset between the two distributions. For distributions that differ by such an offset, the DTA distribution provides an effective and accurate measurement of the offset. However, as described
above, the discrepancy may not be due to an error of the dose calculation algorithm or delivery hardware, but simply due to experimental error. While as originally described, the DTA tool provides only the magnitude of the distance; the sign of the dose difference can be added to the DTA. This provides an indication of which of the two dose distributions is greater. It can also be used to quickly determine if the discrepancy is due to a spatial offset due to one distribution being “larger” than the other.

Because the dose difference and DTA are useful in regions of shallow and steep dose gradients, their use has been combined. A composite analysis was developed by Harms et al.\textsuperscript{2}, based on the concepts of Shiu et al.\textsuperscript{3} as applied by Cheng et al.\textsuperscript{4} used pass-fail criteria using both the dose difference and DTA. Dose-difference and DTA test criteria are selected to determine if each point passes or fails the dose distribution comparison test. The reference points are evaluated one by one to determine their dose difference and DTA and these are compared against the criteria. At points that fail both tests, the dose distribution comparison is said to have failed, otherwise it passes. This allows either the dose difference or DTA tests to be overly sensitive, so the other test determines if the dose comparison passes or fails.

While the composite test is capable of conducting a pass-fail analysis of the dose distribution comparison, it does not indicate by how much the test is passed or failed. In other words, 3\% dose difference, 3 mm DTA criteria can be exceeded by 1\% and 1 mm, or 10\% and 10 mm, and the result would be the same, although the clinical relevancy might differ.

In 2001, at the AAPM annual meeting, Low was explaining the dose difference, DTA, and composite tests to an employee of Computerized Medical Systems in the parking lot of one of the hotels used by the attendees. During the conversation, Dr. Low said he wished he could simultaneously test the dose difference and DTA, but the quantities, dose and distance, were not the same, so this could not be done. Instead, Dr. Low realized that if the dose difference and DTA were divided by their respective criteria, the results would be unitless, and one could evaluate both simultaneously. Because this was an analog to the composite criterion, which starts with the letter “c”, the third letter of the English alphabet, he named the quantity $\gamma$, the third letter of the Greek alphabet.

2. What $\gamma$ is

2.1. Definition
As in many new endeavors, there are multiple changes that take place as the ideas begin to gel. Such was the case with $\gamma$ and the labeling of the two dose distributions, as well as the nomenclature used to describe the criteria and quantities. For simplicity, the nomenclature and symbology used here reflect the second $\gamma$ paper by Low and Dempsey\textsuperscript{5}. In that paper, the two dose distributions are the reference and evaluated distributions. The reference distribution is intended to be the one relative to which the other is compared. It is typically the measurement in a measurement versus calculation, or it can be a Monte Carlo calculation when calculations are compared against each other. The reference distribution is queried point by point, and $\gamma$ computed individually for each. Therefore, the reference distribution can be a single point, for example an ionization chamber measurement point. The evaluated distribution needs to be a distribution, and in early $\gamma$ papers, it was assumed that it had to have a relatively high spatial resolution. This is because the $\gamma$ calculation can have relatively large errors in steep dose gradient regions when the spacing between evaluated distribution points is close to the DTA criterion.

The $\gamma$ distribution calculation is typically conducted as a search. As in the DTA, one needs to search the local evaluated distribution. In the DTA, the value being calculated at each evaluated point is simply the Euclidean distance to the reference point. However, in the $\gamma$ calculation, the distance is generalized in the dose and distance space. The generalized distance is labeled $\Gamma$. $\Gamma$ is the generalized Euclidean distance in the renormalized dose and distance space, where distance and dose axes have been divided by the DTA and dose difference criteria, respectively.
\[ \Gamma(\vec{r}_e, \vec{r}_r) = \sqrt{\frac{\left| \vec{r}_e - \vec{r}_r \right|^2}{\Delta d^2} + \frac{\left| D_e(\vec{r}_e) - D_r(\vec{r}_r) \right|^2}{\Delta D^2}} \]  

(1)

where \( \vec{r}_e \) and \( \vec{r}_r \) are the vector positions of the evaluated and reference points, respectively, \( D_e(\vec{r}_e) \) and \( D_r(\vec{r}_r) \) are the evaluated and reference doses, respectively, and \( \Delta d \) and \( \Delta D \) are the DTA and dose difference criteria, respectively. The generalized \( \Gamma \) function can be computed for any pair \( \vec{r}_e \) and \( \vec{r}_r \), so for each reference point, there are as many values of \( \Gamma \) as there are evaluated points (infinite number with interpolation). The minimum value of \( \Gamma \) is the value of \( \gamma \).

\[ \gamma(\vec{r}_r) = \min \{ \Gamma(\vec{r}_e, \vec{r}_r) \} \forall \vec{r}_e \]  

(2)

Equation 2 states that \( \gamma \) is simply the minimum value in all of the evaluated distribution search space of \( \Gamma \). While equations 1 and 2 provide the factual definition of \( \Gamma \), they do not impart any intuition for what \( \gamma \) means and its utility.

2.2. Interpretation

The \( \gamma \) function is the minimum distance between two dose distributions, but in an unusual space. The distance includes not only space, but dose as well. The spatial axes are the normal ones, and there are as many as necessary to describe the evaluated dose distribution. Recall that the reference dose distribution can be a single point, so the number of dimensions is determined by the evaluated distribution. While the distance axes are the same as normal axes, they are renormalized by the DTA criterion \( \Delta D \), so they have no units. The distance corresponding to \( \Delta d = 1 \). There is an extra axis, which is the dose axis, and this is made possible by the renormalization of the dose axis by the dose difference criterion, \( \Delta D \). The origin is placed at the reference dose point (its position and dose), so travelling along a distance axis corresponds to moving away from the reference point and travelling along the dose axis corresponds to changing the dose difference. The distance from the origin along the dose axis is 1 at the dose difference corresponding to \( \Delta D \).

Given that the dose-difference and DTA criteria are met at the distance of 1, we define the unit circle, sphere, or hypersphere, depending on the dimensionality of the dose distributions, as the limit within which a comparison between the reference and evaluated distribution passes. Figure 1a shows a schematic of the unit circle in a one-dimensional dose distribution evaluation. The horizontal and vertical axes are the distance and dose axes, respectively. The unit circle corresponds to the region within which the hyperdistance is less than 1.
The evaluated distribution can be plotted in the renormalized space, and consists of a line, sheet, or hypersheet if the reference distribution is one, two, or more dimensions. In figure 1b, an example reference distribution is drawn to explain the features of the calculation. The distance from the origin to the evaluated distribution is the $\Gamma$ distribution, and the minimum value is the closest that the evaluated distribution gets to the reference point. When that distance is less than (Figure 1b) or greater than (Figure 1c) 1, the evaluated distribution passed and failed the test, respectively.

The passing and failing of the $\gamma$ test is very similar to the composite test of Harms et al.\textsuperscript{2} The difference lies in the fact that, while the test does identify if the evaluated distribution passed or failed, it also identifies by how much. A value of $\gamma=1.1$ fails, but by only 10%. This corresponds to a distance of 0.3 mm, or a dose error of 0.3% for 3 mm and 3% DTA and dose difference criteria, a relatively minor infraction. On the other hand, a value of $\gamma=3$ fails by a factor of 3, corresponding to 9 mm or 9%, a large difference from the previous example.

The $\gamma$ tool automatically defaults to the dose-difference and DTA tests in regions of shallow and steep dose gradients, respectively. Figures 1d and 1e show the $\gamma$ evaluation under conditions of shallow and steep dose gradients, respectively. In shallow dose gradients (Figure 1d), the evaluated dose is nearly constant near the reference point, so the reference curve appears flat and the closest point lies near the dose axis. A comparison that takes place along the dose axis is essentially the dose difference, so the $\gamma$ tool defaults to the dose difference test at shallow gradients. In steep dose gradients (Figure 1e), the nearest point lies in the direction of one of the spatial axes. The distance between the evaluated distribution where it crosses the spatial axis and the reference point is the DTA, so as the gradient increases, the $\gamma$ direction points towards the spatial axes and automatically becomes the DTA test. The angle between the dose axis and the $\gamma$ vector can be used to determine if the difference between the evaluated and reference distribution is due to the dose difference, DTA, or an intermediate reason.

An interesting question is; for real clinical data and reasonable criteria, what is the maximum angle that the $\gamma$ vector makes to the dose axis? For a dose gradient of 3%/mm and DTA and dose difference criteria of 3% and 3mm, the angle the evaluated distribution makes to the dose axis is almost 72 degrees, or 18 degrees from the distance axis.

3. Other Approaches and Limitations

3.1. Other Approaches
The original $\gamma$ calculation required an exhaustive search in the evaluated dose distribution space to determine the minimum $\Gamma$ value. This caused a significant amount of calculation time, especially when the evaluated dose distribution was two or three dimensional. This limitation led to a number of groups to develop methods for modifying or improving the speed of the $\gamma$ calculation.

Soon after the publication of the initial $\gamma$ paper, Depuydt et al. $^6$ identified some of the limitations of the initial report: 1) the initial $\gamma$ paper did not deal with the two distributions having different spatial densities or locations, 2) the need for interpolation to reduce the artifacts caused by large spacing between evaluated points. In order to reduce the impact of these issues on the $\gamma$ values, they decided to reduce the problem to whether $\gamma$ was greater or less than 1 (fail or pass, respectively) rather than compute the value of $\gamma$. They employed a filter cascade process to determine which evaluated points lay within or outside the $\gamma = 1$ hypercircle. The restricted the search to regions nearby the reference point, and once finding a point for which $\gamma < 1$, they stopped the search. Barring finding points that passed, they evaluated whether a steep dose gradient caused the values of $\gamma$ to be greater than 1 even though an interpolated evaluated distribution would have passed within the $\gamma = 1$ hypercircle. The intersection of the evaluated dose and the spatial axis (axes) is checked to determine if it lies within the $\gamma = 1$ hypercircle. The advantage of this approach was that it was significantly faster than the original algorithm, but it did not have the benefit of providing the user with the magnitude of $\gamma$ to determine the magnitude of failure.

Moran, et al. $^7$ defined a dose comparison tool that examined the local dose gradient. The user selected a distance tolerance to deal with such issues as measurement spatial registration errors. The dose difference and the local dose gradients are computed. The product of the gradient and the distance tolerance indicates the maximum dose difference that could be associated with a spatial shift of the distance tolerance magnitude. The product of the gradient and the dose distance tolerance is subtracted from the dose difference magnitude to determine a gradient-adjusted difference. It is this modified dose difference that is examined.

Stock et al. $^8$ in 2005 developed a tool for applying the $\gamma$ analysis for two-dimensional dose distributions (e.g. film and a two-dimensional cross-sectional sample of a 3-dimensional dose calculation). They were the first to describe the use of the $\gamma$ angle, the angle that the $\gamma$ vector makes to the dose axis. Figure 2 shows an example from Stock et al. $^8$ of a $\gamma$ distribution for a head-and-neck treatment plan. The $\gamma$ distribution shows discrepancy near the posterior head, where $\gamma$ values approach 2. The $\gamma$ angle distribution at the same location indicates that the discrepancy lies within a region of relatively steep dose gradient and that the $\gamma$ extends from approximately $\pi/4$ to $3 \pi /8$. Finally, Stock et al. $^8$ also introduced $\gamma$ histograms (also evaluated by Spezi and Lewis $^9$) as a method for distilling the complex distributions into a more easily interpreted form (Figure 2c). A histogram allows the reviewer a relatively straightforward method for determining how many points are failing the $\gamma$ test as well as the number of points with very large values of $\gamma$. 


Jiang et al.\textsuperscript{10} developed a method that translated the spatial tolerance into the dose domain. They introduced the concepts of equivalent dose tolerance, maximum allowed dose difference (MADD), and the normalized dose difference (NDD). They also used dose-difference and DTA criteria (termed predetermined dose tolerance and predetermined spatial tolerance, respectively). The equivalent dose tolerance is computed by shifting the evaluated distribution until the DTA is equal to the DTA criterion. The equivalent dose tolerance is then the dose-difference between the reference point and the shifted evaluated dose distribution. If the equivalent dose tolerance (in magnitude) is less than the dose-difference, then the DTA must have been less than the DTA criterion, so only a dose-difference test is required. Jiang et al.\textsuperscript{10} also defined an acceptance box, consisting of a rectangular region with the dose-difference and DTA criteria defining the box outlines. If the evaluated distribution intersected the box, the dose distribution passed the test. They also examined the pass-fail criteria in terms of mathematical equations. Figure 3 shows examples of the criteria graphed for the composite evaluation, $\gamma$, and the box.

The MADD defined an acceptance region in the dose-distance space such that it provided a pass/fail boundary such as shown in Figure 3 for other evaluations. The evaluated distribution is shifted until it is tangent to the MADD boundary. The MADD is the dose difference in this case, and if the dose difference is less than MADD (in magnitude), the distribution test passes. An interesting feature of the paper by Jiang et al.\textsuperscript{10} was the appearance of $\gamma$ artifacts caused by the finite evaluated dose distribution pixel spacing.

Figure 2: $\gamma$ distributions for a head-and-neck quality assurance procedure. The isodose distribution is shown as white contours. a) the $\gamma$ distribution. b) the $\gamma$ angle distribution. c) a histogram of $\gamma$ for three plane orientations. Reprinted with permission from Stock et al.\textsuperscript{8}
3.2. Calculation time and spatial resolution artifacts

Some of the concerns with the use of $\gamma$ is the calculation time. When first developed, the $\gamma$ distribution was calculated reference point-by-point and an exhaustive search conducted in the evaluated dose distribution for the minimum value of $\Gamma$. This did not pose a large computational burden for two-dimensional dose distributions, but was significant for three-dimensional dose distributions.

Another concern was related to $\gamma$ distribution artifacts that were seen in steep dose gradient regions. Figure 4 shows an example as published by Jiang et al.\textsuperscript{10} where two square-field dose distributions are being compared, one rotated with respect to the other. Near the field boundary, the $\gamma$ distribution appears to have undulations. These are caused by the finite evaluated dose distribution spacing in steep dose gradient regions. If the $\gamma$ distribution is computed only at each evaluated dose point, the actual minimum distance between the two dose distributions may be missed. This is shown in Figure 5, where the computed and interpolated values of $\gamma$ are presented. When the dose distributions align such that the evaluated point sits exactly at the closest approach to the reference point, $\gamma$ is accurately computed. When the dose distributions are slightly shifted, the value of $\gamma$ is overestimated (Figure 5). For conditions where the dose distributions are slowly shifting (such as in Figure 4), this presents $\gamma$ as a wavy distribution.
Figure 5: Example of finite dose spacing error in $\gamma$ calculation errors. a) Case where interpolation is not needed, the determination of $\gamma$ is accurate. b) Case where the actual value of $\gamma$ is smaller than the computed value (the value as computed at the closest evaluated point).

The solution to this problem was first by interpolation, making the interpolated evaluated distribution sufficiently fine that errors in $\gamma$ caused by the finite spacing were relatively small. Wendling et al.\textsuperscript{11} dealt with the long times for computing $\gamma$ accurately by limiting the search space and reducing 3D $\gamma$ calculations to multiple 2D planes. Ju et al.\textsuperscript{12} looked at $\gamma$ calculations as a geometric problem. The interpolation was essentially the question of the closest approach of the linearly interpolated evaluated distribution, described as a set of line segments, to the reference position. Because this is essentially the question of the closest each line segment (interpolated evaluation distribution) gets to the reference point, Ju et al.\textsuperscript{12} interpreted the problem geometrically.

In higher dimensions, the interpolation was more challenging. For example, a two-dimensional dose distribution would be represented by a series of four evaluation points projecting to a rectangle on the distance axes. The interpolated surface between these points is not well defined, so a fifth point, lying in the middle of the four in the spatial axis and being the mean of the four doses on the dose axis was defined and the rectangular space segmented into four triangles, each with two vertices being the evaluated points and one vertex the interpolated point. This method made the evaluated distribution into a series of triangles, which are considered simplexes, as are line segments in the one-dimensional case. The algorithm was able to rapidly evaluate the distance from the reference point to the simplexes and recursively compute the distance from the surface, edge, or points, whichever was closest. Applications of the algorithm to 3D cases showed that, not only were the artifacts caused by finite reference dose spacing gone, but the calculation was much faster, a 3D calculation requiring a couple of minutes.

3.3. Noise artifacts
One of the uses for $\gamma$ was to evaluate Monte Carlo dose distributions, but these contain pixel-to-pixel noise. Because the $\gamma$ calculation can be interpreted geometrically, the presence of noise would impact the closest approach determination and therefore the value of $\gamma$. This was evaluated by Low and Dempsey\textsuperscript{5} by examining $\gamma$ histograms for dose distributions with and without noise. The greatest impact of noise occurred in regions of shallow dose gradient. The impact of noise in the evaluated distribution always led to a reduction in the value of $\gamma$ because the closest approach for random noise was not likely to get farther away, but closer. Figure 6 shows an example of the impact of noise on homogeneous dose distributions with differences of 6% and a dose difference criterion of 3% ($\gamma = 2$ for no noise). A histogram of $\gamma$ showed that the value was 2. As pseudo-random noise was added to the reference distribution, the histogram of $\gamma$ showed a reduction in the values of $\gamma$ such that $\gamma$ was approximately 1 for 3% noise. This implied that the impact of noise in the evaluated distribution led to
A reduction in the measured value of $\gamma$ that was linearly proportional to the noise (normalized by the dose-difference criterion) until the noise approached the dose difference, whereafter the values of $\gamma$ were very small. A different picture appeared when noise was added to the reference distribution. In this case, because $\gamma$ is computed point-by-point for the reference distribution, the impact of noise was essentially to change the dose difference at each point and the histogram was simply smeared by a width corresponding to the added noise.

Figure 6. The impact of pseudorandom noise on $\gamma$ determinations for dose distributions with no gradient and 6% dose differences. The dose-difference criterion was 3%. Shown are $\gamma$ histograms stacked one per row, with the first row having 0 noise and subsequent rows having added noise. a) Noise added to the evaluated distribution, b) Noise added to the reference distribution. Figure from Low and Dempsey. 

These results showed that care needs to be taken when $\gamma$ is used to analyze noisy data.

4. Conclusions and Future
The $\gamma$ tool has proven useful in the quantitative evaluation of dose distributions. While artifacts caused by the finite evaluated dose distribution spatial resolution appeared early on, the geometric interpretation of $\gamma$ allowed a more rapid and accurate computation. Most of the dose measurement and evaluation software has incorporated $\gamma$ into the evaluation tool suite and the evaluation is used by most medical physicists when attempting to interpret differences between multiple dose distributions.

The original description of $\gamma$ used constant criteria, often 3% and 3 mm dose-difference and DTA criteria, respectively. However, this is entirely arbitrary and should be expanded to consider clinical relevance. For example, the dose-difference at low doses could exceed 3% with no clinical consequence, or the DTA could be considered more important near a critical structure than in the periphery of the dose distribution. The criteria should be modified according to clinical needs, and ideally according to dose, and tumor and normal organ dose tolerances and spatial uncertainty. This would imply that there is a one-to-one relationship between dose and tumor and organ geometry, which is not true for measurement-based patient-specific QA. However the field is going to move towards independent, 3D calculations conducted in the patient geometry, so the $\gamma$ tolerances could be spatially and dosimetrically defined.

References
[1] J. Van Dyk, R. B. Barnett, J. E. Cygler, and P. C. Shragge, "Commissioning and quality assurance of treatment planning computers," Int J Radiat Oncol Biol Phys. 26, 261-273 (1993).
[2] W. B. Harms, Sr., D. A. Low, J. W. Wong, and J. A. Purdy, "A software tool for the quantitative evaluation of 3D dose calculation algorithms," Med Phys. 25, 1830-1836 (1998).
[3] A. S. Shiu, S. Tung, K. R. Hogstrom, J. W. Wong, R. L. Gerber, W. B. Harms, J. A. Purdy, R. K. Ten Haken, D. L. McShan, and B. A. Fraass, "Verification data for electron beam dose algorithms," Med Phys. 19, 623-636 (1992).
[4] A. Cheng, W. B. Harms, Sr., R. L. Gerber, J. W. Wong, and J. A. Purdy, "Systematic verification of a three-dimensional electron beam dose calculation algorithm," Med Phys. 23, 685-693 (1996).

[5] D. A. Low and J. F. Dempsey, "Evaluation of the gamma dose distribution comparison method," Medical Physics. 30, 2455-2464 (2003).

[6] T. Depuydt, A. Van Esch, and D. P. Huyskens, "A quantitative evaluation of IMRT dose distributions: refinement and clinical assessment of the gamma evaluation," Radiotherapy and Oncology. 62, 309-319 (2002).

[7] J. M. Moran, J. Radawski, and B. A. Fraass, "A dose-gradient analysis tool for IMRT QA," Journal of Applied Clinical Medical Physics. 6, 62-73 (2005).

[8] M. Stock, B. Kroupa, and D. Georg, "Interpretation and evaluation of the gamma index and the gamma index angle for the verification of IMRT hybrid plans," Physics in Medicine and Biology. 50, 399-411 (2005).

[9] E. Spezi and D. G. Lewis, "Gamma histograms for radiotherapy plan evaluation," Radiotherapy and Oncology. 79, 224-230 (2006).

[10] S. B. Jiang, G. C. Sharp, T. Neicu, R. I. Berbeco, S. Flampouri, and T. Bortfeld, "On dose distribution comparison," Physics in Medicine and Biology. 51, 759-776 (2006).

[11] M. Wendling, L. J. Zijp, L. N. McDermott, E. J. Smit, J. J. Sonke, B. J. Mijnheer, and M. Van Herk, "A fast algorithm for gamma evaluation in 3D," Medical Physics. 34, 1647-1654 (2007).

[12] T. Ju, T. Simpson, J. O. Deasy, and D. A. Low, "Geometric interpretation of the gamma dose distribution comparison technique: Interpolation-free calculation," Medical Physics. 35, 879-887 (2008).