Improving radiotherapy treatment for left-sided breast cancer

Cho, B-C. J.

Citation for published version (APA):
Chapter 2

THE EFFECT OF SETUP UNCERTAINTIES, CONTOUR CHANGES, AND TISSUE INHOMOGENEITIES ON TARGET DOSE-VOLUME HISTOGRAMS

B.C. John Cho¹,², M.D.
Coen W. Hurkmans¹, M.Sc.
Eugene M. F. Damen¹, Ph.D.
Lambert J. Ziip¹, M.Sc.
Ben J. Mijnheer¹, Ph.D.

¹The Netherlands Cancer Institute/
Antoni van Leeuwenhoek Hospital
Department of Radiotherapy
Plesmanlaan 121
1066 CX Amsterdam
The Netherlands

²The Cross Cancer Institute
Department of Radiation Oncology
11560 University Ave
Edmonton, Alberta
T6G 1Z2
Canada

2.1 ABSTRACT

Introduction: Understanding setup uncertainty effects on dose distributions is an important clinical problem but difficult to model accurately due to their dependence on tissue inhomogeneities and changes in the surface contour (i.e. variant effects).

Purpose: (1) to evaluate and quantify the invariant and variant effects of setup uncertainties, contour changes and tissue inhomogeneities on target dose-volume histograms (DVHs); (2) to propose a method to interpolate variant DVHs.

Materials and Methods: We present a lung cancer patient to estimate the significance of setup uncertainties, contour changes and tissue inhomogeneities in target DVHs. Differential DVHs are calculated for 15 displacement errors (with respect to the isocenter) using: 1) an invariant shift of the dose distribution at isocenter, 2) a full variant calculation, and 3) a B-spline interpolation applied to sparsely sampled variant DVHs. The collapsed cone algorithm was used for all dose calculations. Dosimetric differences are quantified with the root mean square (RMS) deviation and the equivalent uniform dose (EUD). To determine setup uncertainty effects, weighted mean EUDs, assuming normally distributed displacement errors, are used.

Results: The maximum absolute difference and RMS deviation in the integral DVHs' relative dose between: 1) the invariant and calculated curves are 65.2% and 5.8% and, 2) the interpolated and calculated curves are 16.9% and 2.5%. Similarly, the maximum absolute difference and RMS deviation in mean EUD as a function of the setup uncertainty's standard deviation between: 1) the invariant and calculated curves are 0.02 and 0.01 Gy and, 2) the interpolated and calculated curves are 0.01 and 0.006 Gy.

Conclusions: Since a “worst case” example is selected, we conclude that, in the majority of clinical cases, the variant effects of contour changes, tissue inhomogeneities and setup uncertainties on EUD are negligible. Interpolation is a valid, efficient method to approximate DVHs.

2.2 INTRODUCTION

2.2.1 Setup Uncertainties

2.2.1.1 Treatment Margins

Target volume definition helps to ensure the tumour receives an adequate dose. They were formally introduced by the International Commission on Radiation Units and Measurements (ICRU) and discussed in Report 50 [1] and its supplement [2]. The report defines several related conceptual target volumes such as the gross target volume (GTV), the clinical target volume (CTV) and the planning target volume (PTV). Therefore, if all uncertainties, movements and variations are well characterized and well understood, then one can, in principle, define the PTV from only the CTV.
Figure 2-1. Representative axial CT slices of the clinical example with a gray scale dose wash to represent the dose distribution for a displacement error ($\delta$) of 0 cm (A) and a $\delta=+3.6$ cm (from isocenter) for an invariant dose distribution, ignoring the effects of contour changes and tissue inhomogeneities (B), and a variant recalculated dose distribution with the same displacement error (C). The straight lines represent the beam and leaf edges. The multileaf collimator leaves are tightly conformed to the planning target volume with a 0 cm margin. The apparent discrepancy in the margins is due to the oblique view of the slice with respect to the beam.

The margin between the CTV and PTV is meant to account for all geometric errors such that there is a high likelihood the CTV receives a sufficient clinical dose. If the margin is too large, then this results in greater normal tissue complications (than necessary) due to excessive irradiation of the surrounding organs at risk. If the margin is too
small, then this results in poorer outcomes due to inadequate irradiation of the target from geographical misses.

Stroom et al [3] and van Herk et al [4] both have proposed a “margin recipe” where an adequate treatment margin is defined using known random and systematic errors for a specified minimum CTV dose for a given proportion of the population. One disadvantage of population based margins is that they may overestimate the volume of the required PTV if a relatively high population threshold used. The influence of margin width, systematic (preparation) and random (execution) setup uncertainties on biologic indices such as equivalent uniform dose (EUD) and tumor control probability (TCP) assuming an invariant dose distribution have been investigated [5].

2.2.1.2 Displacement Errors and Setup Uncertainties

Displacement error and setup uncertainty are distinct but related concepts. Displacement errors, δ, are analogous to systematic (or preparation) errors and refer to a single instance of a given error. They may be simple (such as translational) or complicated (such as internal movement, deformational, rotational, shear or a combination thereof).

Setup uncertainties, δSD, are equivalent to random (or execution) errors and are defined by a collection of (potentially unknown) displacement errors with an associated probability distribution. The simplest model assumes that the setup uncertainty is normally distributed and limited to only translational displacements errors [6,7]. In such cases, the setup uncertainty is more conveniently described by its standard deviation. For the purposes of this paper, a setup uncertainty of 1 cm implies normally distributed displacement errors with a standard deviation (SD) of 1 cm (i.e. δSD=1 cm). The incorporation of non-normally distributed displacement errors, such as the periodic breathing motion (where the time at exhalation is greater than at inhalation) [8], have also been studied.

2.2.2 Invariant and Variant Dose Distributions

By invariant dose distribution, we mean that the dosimetric effects of contour changes and tissue inhomogeneities as a function of the displacement error, δ, are negligible. If invariance is assumed, only one dose distribution is calculated (at δ=0 cm) and it undergoes a rigid body transformation to simulate various displacement errors. The invariant dose-volume histogram (DVH) for a given volume of interest (VOI) can undergo a “blurring” (i.e. convolution) function rather than a recalculation for different setup uncertainties. This greatly decreases the calculation time.

The major limitation of dose blurring and related methods is the implicit assumption that the dose distribution remains invariant, independent of contour changes and tissue inhomogeneities for different displacement errors. For some sites, such as the prostate, this may be a reasonable approximation since the target volume is centrally located and
the dose distribution is not significantly influenced by changes in the surface contour or by tissue inhomogeneity (with the exception of metallic hip prostheses).

Other sites, however, violate the assumption of an invariant dose distribution. Lung primaries are one of the most difficult sites to calculate an accurate dose distribution. The lower density of the lung and its associated electron disequilibrium influence the delivered dose compared to the homogeneous situation. Laterally scattered electrons travel farther in less dense media and consequently, increase the beam penumbra. The low isodose surfaces bows out while the high isodose surfaces bow in. As a result, larger penumbral margins are needed around the target volume to account for this effect (compared to the homogeneous situation) [9].

As Figure 2-1 shows, the assumption of invariance, although convenient, is unrealistic in lung tumors. A displacement error is deliberately introduced by shifting the beam 3.6 cm anteromedially from the original isocenter in the transaxial plane, perpendicular to the central beam axis. Figure 2-1B is a simple translational displacement of the original reference dose distribution (i.e. invariant) while Figure 2-1C recalculates the dose distribution for the new beam orientation (i.e. variant). The DVHs of the CTV between the invariant and variant situations are approximately the same for small displacement errors but will tend to diverge for larger displacement errors as variant effects become more important. Displacement errors will affect the DVH and the main cause of dosimetric differences between different displacement errors is attributable to the shift in the dose distribution (which is exactly what happens with an invariant shift). However, there is a smaller second order variant effect attributable to tissue inhomogeneity and contour changes as well. Different dose calculation algorithms correct for these variant effects differently. Hence, the results will, in part, depend on the dose calculation algorithm used.

Variant effects pose several technical difficulties. Sophisticated dose algorithms, such as Monte Carlo [10] or collapsed cone [11-13], are needed to accurately model the variant effects of tissue inhomogeneity but they are substantially slower. In some cases, tissue inhomogeneity can significantly influence the dose and up to a 20% deviation in absorbed dose in lung tumors irradiated with small fields between different inhomogeneity correction algorithms is possible [14,15]. In a paper by Xing et al [16], a method is described to separate and quantify the individual dosimetric effects of these errors and uncertainties.

Perhaps the simplest way to incorporate variant effects into the dose calculation is a “brute force” approach. The variant dose distribution is laboriously recalculated for each displacement error but this proves impractical for routine use. An accurate modelling of setup uncertainty effects requires a large number of dose calculations corre-
responding to many different displacement errors to achieve reasonable statistical accuracy [17,18].

2.3 PURPOSE

The primary purpose of this paper is to quantify the influence and significance of the variant effects of contour changes and tissue inhomogeneities with respect to setup uncertainties between target dose-volume histograms (DVHs) by comparing both the invariant and variant situations.

The secondary purpose is to describe and verify an interpolation method to address the problem of the prohibitive number of recalculation required for variant setup uncertainties.

2.4 MATERIALS AND METHODS

2.4.1 Clinical Patient

A lung cancer patient was selected as the test case having the steep contour changes and a significant amount of tissue inhomogeneity found inside the irradiated field. To exaggerate the variant effects, only a single right anterior oblique beam is planned. The patient is CT scanned (GE Medical Systems, Waukesha, WI, USA) with 5 mm thick slices in the supine position to include all visible lung. Volumes of interest, namely the GTV, CTV, PTV and external contour, are delineated on the treatment planning system. The GTV is defined as all visible tumor on the standard CT window settings and delineated by an experienced staff radiation oncologist. The CTV is defined as the GTV with a 0 cm margin. A 0.5 cm margin is added isotropically around the CTV to generate the PTV. Such a small margin would generally not be used clinically in the treatment of lung primaries but was deliberately chosen here to exaggerate the variant effects.

2.4.2 Treatment Planning

We used a three-dimensional treatment planning system (PINNACLE v. 6.0g, ADAC Laboratories, Milpitas, CA, USA) with a collapsed cone dose algorithm to accurately model the effects of heterogeneous media [11]. The dose grid resolution is $2 \times 2 \times 2 \text{ mm}^3$ and the calculation dose grid dimensions are $205 \times 159 \times 159$ voxels centered on the target.

The test patient is treated using 8 MV photons and the beam aperture is automatically shaped with multileaf collimators set to conform tightly around the PTV with a 0.6 cm penumbra margin using beam's-eye view to ensure the PTV receives at least 95% of the prescribed dose (Figure 2-1). No other beam modifying devices are used. The prescribed number of monitor units is kept constant for every displacement error at 516 monitor units per beam.
2.4.2.1 Displacement Errors

Different displacement errors are simulated by shifting the isocenter in the transaxial plane perpendicular to the central beam axis with the same beam parameters. The dose distribution is recalculated with the collapsed cone dose calculation algorithm for 15 displacement errors (δ=±3.6, ±2.4, ±1.5, ±1.2, ±0.9, ±0.6, ±0.3, and 0 cm). The differential DVH (dDVH) of the CTV is determined for each displacement error. This set of dDVHs is called dDVH calc and represents the "true" variant calculated DVHs.

The invariant dDVHs are determined by calculating the dose distribution at δ=0 cm and shifting it in exactly the same direction and magnitude as the isocenter shifts described above for dDVH calc. The CTV's dDVH for each displacement is calculated (for a total of 15 dDVHs). Obviously, dDVH calc at δ=0 and dDVH invar at δ=0 cm are identical. This set is called dDVH invar and represents the resulting DVHs when the dose distribution is assumed to be invariant to translational displacements, ignoring the effects of contour changes and tissue inhomogeneities.

2.4.2.2 Setup Uncertainties

The effect of setup uncertainties is derived analytically from the displacement errors. Setup uncertainties are assumed to be random and normally distributed. Therefore, the dDVH for displacement errors at least within ±2 SD of the setup uncertainty is necessary to calculate the setup uncertainty effect (to include approximately 95% of all the possible displacement errors). If the dDVH as a function of δ is known, then one can derive other clinically relevant parameters such as tumor control probability (TCP) and equivalent uniform dose (EUD) as a function of δ. The effect of a given setup uncertainty, δSD, on EUD is determined by generating a set of normally distributed δ with a mean of 0 cm and a SD equal to δSD. The EUDs for all these different displacement errors are then averaged together. By repeating the process for different SDs, the relationship between EUD as a function of δSD can then be derived. Because EUD, as a function of dDVH, is non-linear, individual dDVHs cannot be simply averaged together to calculate the overall treatment EUD. It is also worth noting that the mean EUD is a population averaged EUD and represents the average from either an infinite number of treatment fractions or an infinite number of finite-fractioned courses of treatment. The EUD for any finite fractioned course of treatment may deviate from the average EUD and from any EUD for any single δ. However, for fractionation schedules larger than 15 fractions, the expected difference between the mean sample EUD and the mean population EUD is negligible.

The statistical accuracy depends on the number of displacement errors used to represent a given setup uncertainty. Obviously the greater the number of displacement errors used, the more accurate the approximation becomes. The 15 dDVHs calculated for the discrete displacement errors above is inadequate due to the inaccuracies associated
with such a small number of displacement errors. Furthermore, these inaccuracies are exacerbated for small SD since the number of usable dDVH calc curves is less than 15. For example, if δSD=0.5 cm then approximately 95% of the normally distributed δ will be between -1 cm and +1 cm. However, we only have 7 calculated dDVHs (at δ=±0.9, ±0.6, ±0.3, and 0 cm) that are within the range of ±1 cm.

The accuracy can be improved by obtaining dDVHs at even smaller displacements, say in 0.1 cm increments, but at a cost of more dose distributions calculations for more displacement errors. The time required to calculate the variant dose distribution for 21 displacement errors is greater than 30 hours. Ideally, we would like to know the dDVH for the invariant and variant situations at any arbitrary δ and, at the same time, avoid the time consuming process of recalculation.

We propose a method to interpolate the dDVH at any arbitrary displacement error. Invariant dDVHs are generated by interpolating between dDVH invar and variant dDVHs, by interpolating between dDVH calc. Only setup uncertainties along a single axis are investigated to simplify calculations and to allow easier visualization of the data since the resulting interpolant is a simple surface (and not a hypersurface).

With interpolation, it is relatively easy to calculate the EUD as a function of δSD. Seventy-three dDVHs for both the invariant and variant dose distributions are interpolated from -3.6 cm to +3.6 cm in 0.1 cm increments. Once the corresponding EUDs are calculated, one can obtain the EUD directly as a function of δ. Since this curve is simple and continuous, it can also be interpolated. This second interpolation allows one to find the EUD for any arbitrary δ (between -3.6 cm and +3.6 cm) without having to determine the new dDVH each time.

It also provides a method of estimating the EUD for the dDVH calc curves. Because the dDVH calc consists of only 15 curves, only 15 EUDs (corresponding to their displacement errors) are known. By interpolating these 15 input data points, one can estimate the EUD for the “true” variant dose distribution (i.e. calculated) as a function of displacement error. Details of the interpolation method follow.

2.4.3 Interpolation

Interpolation is a useful tool to minimize work. As an analogy, we can use cartoon animation to illustrate the general principles. “Extremes” and “inbetweens” were used to increase workload efficiency. A skilled animator would draw only the “extremes” or keyframes in a scene. The celluloid would then be sent to less skilled animators who would animate (or interpolate) the frames in between the keyframes (hence, inbetweening).

In much the same manner, reference “keyframe” differential DVHs (dDVHs) can be calculated and “in between” dDVHs can be interpolated. The accuracy of the interpo-
SECTION 2.4 MATERIALS AND METHODS

Interpolation is dependent on several factors: the accuracy of the input reference data, the behaviour of the input reference data as a function of the input variables and the appropriateness of the interpolating method.

The accuracy of the calculated input reference data is dependent on the dose calculation algorithm. In all treatment planning systems, there is a trade-off between dose accuracy and computational speed. More complicated algorithms require more time to calculate more accurate dose distributions. If inaccurate data is used for the interpolation, then the resulting interpolant will be inaccurate as well.

Interpolation works best for mathematically well-behaved input reference data. By this, we mean the known input values describe a (hyper-)surface that is smooth, simple and continuous. This assumption is valid as long as the voxels (and its associated dose grid) and the magnitude of the isocenter displacement are kept reasonably small. The calculated dDVH will have less noise and less partial volume effect with smaller voxels and a finer dose grid, resulting in a smoother curve.

Many different methods of interpolation exist. The appropriateness of a given method depends on the required precision of the approximation as well as the relative speed of the interpolation. For this study, a cubic spline interpolating function is used. The primary assumption of well-behaved input data is that they describe a smooth (hyper-)surface. Splines are advantageous in that both the first and second derivatives of the interpolant are continuous and smooth.

### 2.4.3.1 Multilevel B-spline Functions

The method described by Lee et al.[19] uses multilevel B-spline functions to interpolate scattered data. The paper focuses primarily on three dimensional (3D) data points but the technique is easily extensible to any set of multi-dimensional data. Scattered data points can be used as input and the degree of precision in the approximation can be adjusted by changing the number of lattice iterations and/or the size of the control lattice.

Conceptually, every data point defines a point on a surface or a hypersurface (for points with a dimensionality greater than 3). Each point is assumed to have exactly one dependent variable (in the case of DVHs, the relative volume) and at least one independent variable (such as the relative dose and/or displacement error). The B-spline takes these scattered input points and creates a control lattice. The control lattice has the same number of dimensions as the number of independent variables but can be considered dimensionless. This lattice can then be used to find the dependent value of any arbitrary set of independent values (within the boundaries of the lattice).
Figure 2-2. B-spline interpolation and the effects of the number of iterations (iter) and lattice size (latsz) on the interpolant using the 5 input data points \((x, y, z): (0,0,2), (1,1,1), (1,-1,1), (-1,1,1), \) and \((-1,-1,1)\). Increasing the number of iterations reduces the residual error between the interpolant and the input data. Increasing the initial lattice size decreases the "stiffness" of the interpolant surface.

Figure 2-2 illustrates how the locality and precision of the interpolation is dependent on the lattice size and number of iterations. The lattice size controls the range of influence of the input data and is related to the deformability or "stiffness" of the interpolant surface. For small lattice sizes (Figure 2-2A), the interpolant can be considered as a stiff surface. For large lattice sizes (Figure 2-2F), it can be considered elastic with highly local effects from the input data.

With each iteration, the lattice size is doubled in all dimensions and the interpolation reduces the residual error between the input data and the interpolant. The residual error of the interpolant will tend towards zero at those points that correspond to the input reference data. Because we assume the input data is smooth, simple and continuous,
Section 2.4

Small initial lattice sizes and numerous iterations are preferred. By minimizing the area of the (hyper-)surface of the interpolant fitted through the scattered input data, we ensure the interpolant is as smooth and simple as possible. A full description of multilevel B-splines is beyond the scope of this paper and the reader is referred to the original paper [19] for further details.

2.4.3.2 Implementation

The B-spline interpolating function is implemented within MATLAB v.6.0.0.88 Release 12 (The MathWorks Inc., Natick, MA, USA) on a 350 MHz Pentium III personal computer with 512 MB of RAM.

All the interpolations are iterated 8 times between an initial starting lattice size of [5 x 1] and a final lattice size of [1280 x 256]. Because a Gaussian setup uncertainty is assumed, the dDVHs curves with the highest probability will be closest to δ=0 cm. The reference input dDVHs used for the interpolation are non-equally spaced, clustered around δ=0 cm, to minimize their number.

The input reference data used for the interpolation consists of 7 dDVH curves (at δ=±3.6, ±1.5, ±0.6 and 0 cm from dDVH calc). Afterwards, 15 dDVH curves with identical displacements as dDVH calc are interpolated for comparison. This set is called dDVH interp and represents an interpolated version of dDVH calc. A perfect interpolation would have identical dDVH interp and dDVH calc at every displacement error.

2.4.4 Evaluation and Analysis

The dDVH calc, as calculated using the collapsed cone dose algorithm, is assumed to be the “true” dose. It is compared to the invariant dDVHs in order to quantify the magnitude of the variant effects. It is also compared to the interpolated dDVHs to determine the accuracy and precision of the interpolation. To measure the differences, the root mean square (RMS) of the differences in relative volume (i.e. dDVH interp-dDVH calc) quantifies the overall accuracy of interpolation and the maximum absolute difference quantifies the upper limit of the interpolation’s accuracy.

Another metric used is the equivalent uniform dose (EUD) which estimates the biologic effect due to dose inhomogeneity [20]. The same concept can be used to quantify the effect of differences in the dDVH curves for different displacement errors and setup uncertainties. We adopt a reference dose per fraction of 2 Gy, a surviving fraction at 2 Gy of 0.5, and an α-value of 0.35 Gy⁻¹. A homogeneous clonogenic tumour density is assumed in all cases. The dose is normalized to ensure the EUD is approximately 2 Gy at δ=0 cm (actually 2.03 Gy).

The EUD as a function of displacement error is calculated by interpolating 73 dDVHs from -3.6 to +3.6 cm in 0.1 cm increments and determining the EUD for each dis-
placement error. The EUD as a function of setup uncertainty SD, as discussed in Section 2.4.2.2, is calculated by generating 105 different displacement errors that are normally distributed with a standard deviation equal to the setup uncertainty SD. dDVHs for these displacement errors are interpolated and the corresponding EUD is then calculated and averaged. Because the displacement error ranges from -3.6 to +3.6 cm, only the setup uncertainties from 0 to 1.8 cm (i.e. 2 SD=3.6 cm) in 0.1 cm increments are calculated.

2.5 RESULTS

Figure 2-3. A plot of the planning target volume's differential dose-volume histogram (dDVH) curves as a function of displacement error (δ) assuming an invariant dose distribution (A, dDVH invar) as well as the difference (B, dDVH invar-dDVHcalc) between the invariant and calculated dDVH curves.
Table 2-1. The maximum absolute differences (DIFF) and root mean square (RMS) deviations between the invariant (invar) and interpolated (interp) dose-volume histograms compared to the calculated dose-volume histograms for different displacement errors (DISP). Two types of dose-volume histograms are shown: the differential dose-volume histograms (dDVHs), and the integral dose-volume histograms (iDVHs). All values are given in percent of relative dose (% DOSE) or relative volume (% VOL), depending on the histogram. The differences converge to zero at $\delta=0$ cm for all histograms. Under the interp columns, zeros are found at the displacement errors corresponding to the reference input dDVHs used for the interpolation (at $\delta=\pm3.6, \pm1.5, \pm0.6$, and 0 cm).

<table>
<thead>
<tr>
<th>DISP (CM)</th>
<th>dDVH (% VOL) DIFF</th>
<th>RMS</th>
<th>iDVH (% VOL) DIFF</th>
<th>RMS</th>
<th>iDVH (% DOSE) DIFF</th>
<th>RMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3.6</td>
<td>2.01</td>
<td>0.00</td>
<td>0.27</td>
<td>0.00</td>
<td>15.63</td>
<td>5.98</td>
</tr>
<tr>
<td>-2.4</td>
<td>0.85</td>
<td>0.60</td>
<td>0.22</td>
<td>0.10</td>
<td>15.49</td>
<td>4.33</td>
</tr>
<tr>
<td>-1.5</td>
<td>0.60</td>
<td>0.00</td>
<td>0.16</td>
<td>0.00</td>
<td>10.65</td>
<td>3.78</td>
</tr>
<tr>
<td>-1.2</td>
<td>0.60</td>
<td>0.11</td>
<td>0.13</td>
<td>0.02</td>
<td>8.12</td>
<td>1.50</td>
</tr>
<tr>
<td>-0.9</td>
<td>0.56</td>
<td>0.07</td>
<td>0.10</td>
<td>0.02</td>
<td>6.45</td>
<td>1.28</td>
</tr>
<tr>
<td>-0.6</td>
<td>0.35</td>
<td>0.00</td>
<td>0.07</td>
<td>0.00</td>
<td>4.29</td>
<td>0.00</td>
</tr>
<tr>
<td>-0.3</td>
<td>0.19</td>
<td>0.04</td>
<td>0.03</td>
<td>0.01</td>
<td>2.33</td>
<td>0.65</td>
</tr>
<tr>
<td>0.0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.3</td>
<td>0.11</td>
<td>0.06</td>
<td>0.02</td>
<td>0.01</td>
<td>2.26</td>
<td>0.45</td>
</tr>
<tr>
<td>0.6</td>
<td>0.44</td>
<td>0.00</td>
<td>0.11</td>
<td>0.00</td>
<td>11.15</td>
<td>0.00</td>
</tr>
<tr>
<td>0.9</td>
<td>0.30</td>
<td>0.10</td>
<td>0.08</td>
<td>0.02</td>
<td>9.04</td>
<td>1.23</td>
</tr>
<tr>
<td>1.2</td>
<td>0.36</td>
<td>0.12</td>
<td>0.11</td>
<td>0.02</td>
<td>12.34</td>
<td>0.64</td>
</tr>
<tr>
<td>1.5</td>
<td>0.44</td>
<td>0.00</td>
<td>0.13</td>
<td>0.00</td>
<td>14.71</td>
<td>0.00</td>
</tr>
<tr>
<td>2.4</td>
<td>0.73</td>
<td>0.52</td>
<td>0.19</td>
<td>0.07</td>
<td>20.13</td>
<td>3.80</td>
</tr>
<tr>
<td>3.6</td>
<td>1.46</td>
<td>0.00</td>
<td>0.20</td>
<td>0.00</td>
<td>16.72</td>
<td>6.44</td>
</tr>
</tbody>
</table>
Figure 2-4. A plot of the planning target volume’s differential dose-volume histograms (dDVHs) as a function of displacement error (δ) for: (A) the calculated dDVHs (dDVH calc), (B) the interpolated dDVHs (dDVH interp) and (C) the difference (dDVH interp−dDVH calc). The dDVH interp is plotted for δ from -3.6 to +3.6 cm in 0.1 cm increments. The other curves, dDVH calc and dDVH interp−dDVH calc, are plotted for δ=±3.6, ±2.4, ±1.5, ±1.2, ±0.9, ±0.6, ±0.3, and 0 cm. Note the waterfall plot axes shown in Figure 2-4C are reversed. Because the plot is floating, the corner closest to the viewer at dose=0 % corresponds to δ=-3.6 cm.
The RMS deviation quantifies these differences. Table 2-1 lists the RMS difference between dDVH invar and dDVH calc (i.e. dDVH invar-dDVH calc) and dDVH interp and dDVH calc (i.e. dDVH interp-dDVH calc) as a function of displacement error. The interpolated variant dDVHs (compared to the invariant dDVHs) better approximates the calculated variant dDVHs. The deviations increase with increasing displacements.

2.5.1 Invariant Dose Distributions

Figure 2-3A shows the invariant dDVHs of the CTV (dDVH invar). The most prominent difference between the invariant and the calculated dDVHs is seen in the low dose bins at δ=-3.6 cm. Figure 2-3B plots the differences (i.e. dDVH invar-dDVH calc) and highlights the differences particularly in the low dose bins at the displacement errors’ extremes. The differences converge at δ=0 cm since the curve is identical for both dDVH calc and dDVH invar. Further away from δ=0 cm, the differences become more prominent.

2.5.2 Variant Dose Distributions

The maximum absolute difference in relative volume between the interpolated and calculated curves for the differential and integral variant DVHs over the range of displacements investigated is 0.6% and 4.3% (both at δ=-2.4 cm). The maximum absolute difference and RMS deviation in relative dose between the interpolated and calculated curves for the integral DVHs (iDVHs) over the range of displacements investigated is 16.9% and 2.5% at δ=-0.3 cm and -1.2 cm, respectively (see Section 2.10). The relative dose is read perpendicular from volume axis, parallel to the dose axis of the iDVH.

The maximum absolute difference in relative volume between the invariant and calculated curves for the dDVHs and iDVHs over the range of displacements investigated is 2.0% and 20.1% at δ=-3.6 and +2.4 cm, respectively (Figure 2-4). The maximum absolute difference and RMS deviation in relative dose between the invariant and calculated curves for the iDVHs over the range of displacements investigated is 65.2% and 5.8% at δ=+0.6 cm and +3.6 cm, respectively. The differences in the integral DVHs are larger since the differences are cumulative (Table 2-1).
Figure 2-5. A plot of the planning target volume’s equivalent uniform dose (EUD) for the invariant (dotted line), variant (solid line) and calculated ('x') differential dose-volume histograms (dDVHs) as a function of: (A) displacement error ($\delta$) and (B) setup uncertainty standard deviation ($\delta_{SD}$). Only 15 calculated EUD points are plotted in Figure 2-5A since only 15 variant dDVHs were calculated. A reference dose per fraction of 2 Gy, a survival fraction at 2 Gy of 0.5, an $\alpha$ of 0.35 Gy$^{-1}$, and a homogeneous clonogen density is assumed.

2.5.3 Equivalent Uniform Dose

The EUD as a function of displacement error and setup uncertainty are calculated to estimate their biologic effect. The maximum EUD for Figure 2-5 is 2.03 Gy. Seventy-
three dDVH invar and dDVH interp curves are interpolated from -3.6 cm to +3.6 cm in 0.1 cm increments and their respective EUDs are calculated and plotted in Figure 2-5A. In contrast, only 15 dDVH calc curves are available so only 15 EUD points are plotted. The maximum absolute difference in EUD for invar and interp compared to calc is 0.13 and 0.05 Gy both at $\delta=+2.4$ cm, respectively. The RMS of their differences is 0.06 and 0.02 Gy.

Figure 2-5B plots the relationship between EUD invar, EUD interp and EUD calc as a function of setup uncertainty. The maximum absolute difference in EUD for invar and interp compared to calc is 0.02 and 0.01 Gy (or 1.0 and 0.5% of the maximum EUD) at SD=1.6 and 1.8 cm respectively. The RMS of these differences is 0.01 and 0.006 Gy.

Besides the EUD, we also compared the mean, maximum and minimum dose and found that the corresponding errors for these dose parameters are very small.

2.6 DISCUSSION

2.6.1 Validity

2.6.1.1 Appropriateness of the Clinical Example

The treatment plan of the clinical example is purposely made suboptimal. There are several reasons why the clinical lung patient presented is a “worst case” example. First, the treatment uses only a single beam. In multiple beam arrangements, the other beams will tend to compensate for the dosimetric differences from contour changes and tissue inhomogeneities, lessening their effects.

Secondly, the margins defined around the target are very tight. The margin width between the CTV and PTV is 0.5 cm, which is smaller than what is typically used at our institution. Thus, any displacement error greater than 0.5 cm will have a significant detrimental effect on the DVH of the CTV.

Thirdly, this example represents a patient with a relatively small target volume. The maximum dimensions of the PTV, as seen in beam's eye view, are 7.3 cm by 7.6 cm. Smaller target areas will tend to exaggerate the dosimetric differences in the DVH for a given displacement error.

Fourthly, the beam direction that was selected maximizes the effects of contour changes and tissue inhomogeneities. Suppose the beam is oriented along the right lateral direction. If the beam is displaced posteriorly, then the increased attenuation of the beam from the contour changes is partially offset by the decreased attenuation from the greater amount of lung tissue in the field. If, however, the beam approaches the target from the right anterior oblique direction, then the contour changes and tissue inhomogeneities do not compensate each other and their effects are additive. Fifthly, the
applied displacement errors are calculated up to 3.6 cm, which is larger than what is typically seen in real clinical patients.

Lastly, the EUD as a function of setup uncertainty is calculated as a weighted average of dDVHs and their corresponding EUDs. This will tend to underestimate the "true" EUD as determined by the cumulative voxel dose inside the CTV as a function of the setup uncertainty. Strictly speaking, since the relationship between the EUD and setup uncertainty is not linear, a simple average should not be used. EUDs, as calculated from the dDVHs, ignore spatial information of individual voxels. Using our 1D example, without loss of generality for higher dimensions, it is apparent that relatively large displacement errors will result in the relative underdosage of the CTV. However, all else being equal, the positive and negative displacement errors will tend to underdose different volumes of the CTV such that the overall cumulative dose will be higher than otherwise expected. Consequently, the EUD calculated from the cumulative dose as a function of the setup uncertainty will tend to be greater than the EUD calculated as a weighted average of dDVHs and their corresponding EUDs. As we are using a "worst-case" example, this is not a significant difficulty. This, however, highlights some of the potential problems encountered if one is unaware of the limitations of dDVHs.

\[ S = \pm 2.4 \text{ cm} \quad \delta = \pm 0.3 \text{ cm} \]

Figure 2-6. A plot of the planning target volume's differential dose-volume histogram (dDVH) curves comparing the calculated (solid line) and the interpolated (dotted line) curves at a displacement error (\( \delta \)) of +2.4 cm (A) and +0.3 cm (B).

### 2.6.1.2 Accuracy

Figure 2-4A-B show the interpolated dDVH (dDVH interp) surface using only 7 input reference dDVHs from ±3.6, ±1.5, ±0.6 and 0 cm. With the interpolation, it is possible to obtain dDVHs at any arbitrary displacement error and Figure 2-4B shows dDVHs in 0.1 cm increments. To verify the overall accuracy of the interpolant, the collapsed cone algorithm calculated the "true" dDVH (dDVH calc) at 15 displacement errors (\( \delta = \pm 3.6, \pm 2.4, \pm 1.5, \pm 1.2, \pm 0.9, \pm 0.6, \pm 0.3, \) and 0 cm). The differences between the corresponding displacement errors (i.e. dDVH interp–dDVHcalc) are plotted on Figure 2-4C. The
difference between the curves converges to zero at the displacements that correspond to the input reference data at \( \delta = \pm 3.6, \pm 1.5, \pm 0.6 \) and 0 cm. This is a function of the number of iterations and is a property of the interpolation method itself. Visual inspection reveals the largest differences between the interpolated and calculated curves are found at \( \delta = \pm 2.4 \) cm.

Although the differences appear large in this figure, they are relatively small when examining the dDVH curves individually. Figure 2-6 compares the interpolated and the calculated dDVHs for \( \delta = +2.4 \) cm and +0.3 cm to demonstrate visually how the accuracy of the interpolant is affected by the distance between the input reference dDVHs. The interpolated curve at \( \delta = +2.4 \) cm is primarily a function of the input dDVHs at \( \delta = +3.6 \) and +1.5 cm while the interpolated curve at \( \delta = 0.3 \) cm mainly uses the reference dDVHs at \( \delta = 0 \) and +0.6 cm.

We expect the interpolant to be less accurate between input reference dDVHs that are farther apart. For example, the interpolated dDVH curve at \( \delta = \pm 2.4 \) cm primarily relies on the input dDVH curves at \( \delta = \pm 3.6 \) cm and \( \pm 1.5 \) cm. The assumption of smoothness between these input dDVHs is only partially correct and therefore adds to the inaccuracy of the interpolant. In this case, the assumption of smoothness between input dDVH curves is valid when the displacements are less than the PTV margin (i.e., <0.5 cm). Any displacement less than 0.5 cm would result in a dDVH very similar to the original dDVH. That does not imply the input reference dDVHs must be in increments of 0.5 cm or less.

Although the optimal number and spacing of the reference input dDVHs are not specifically investigated, the staggered, non-equidistant displacements used for the interpolation will tend to minimize the number of input reference dDVHs required for a reasonably accurate interpolant. The probability density for the setup uncertainties is centered around the mean (\( \delta = 0 \) cm) with a given standard deviation. When calculating the average EUD for a given setup uncertainty, the most accurate interpolated dDVHs will be found where the input reference dDVHS are closest together (i.e. \( \delta = 0 \) cm).

Only a 1D displacement error (and not a true 3D displacement error) is studied and one could argue the results will tend to underestimate the real dosimetric effects. However, given all the “worst case” conditions described above, it seems unlikely that a full 3D displacement error interpolation would materially affect the results and conclusions, particularly if more realistic setup uncertainties are kept in mind (i.e. SD<1 cm). The variant effect of contour changes and tissue inhomogeneities is most significant in the displacement direction selected. Even with a full 3D displacement implementation, moving the isocenter along the superficial-deep direction is primarily dependent on the inverse square law and is not highly dependent on tissue inhomogeneity and contour
changes. Similarly, moving the isocenter along the superior-inferior direction is not as dependent on variant effects since the long axis of the thorax is cylindrical in shape.

To verify a 2D displacement error with a similar spacing used in this 1D displacement error would require approximately 225 dDVH calc curves (15 dDVH calc curves squared). The recalculations of 225 dose distributions is impractical due to the prohibitively long computation time required. In principle, however, the interpolation for 2-(or more) D displacement errors should be as accurate as the 1D displacement errors as long as the input reference data are similarly smooth.

To interpolate displacement errors in 3D, 3D reference dDVHs are required. For example, one could use isocenter shifts at $\delta=(\pm1, \pm1, \pm1)$ and $(0,0,0)$. These 9 reference dDVHs would take approximately 13 hours to calculate (using a similar dose grid as this study). The calculation time required is long but determining the effect of variant 3D setup uncertainties accurately, by any other means, will likely take substantially longer to calculate.

The largest RMS difference in relative dose between the interp and calc dDVH is 2.5% with a maximum absolute difference of 16.9%. The large absolute difference is from the relatively flat portion that corresponds with the steep gradient region of the dose distribution of the DVH curve. These numbers compare favourably to the recommended relative dose tolerance of 2-4% for the low dose gradient and up to 15% for steep dose gradient regions [21]. We conclude that the interpolated dDVHs accurately approximate the “true” variant calculated dDVHs with respect to displacement error, given the number and spacing of the input reference dDVHs used. In comparison, the largest RMS difference and maximum absolute difference in relative dose between the invar and calc dDVH is 5.5% and 65.2% respectively which is outside the recommended relative dose tolerance. The application and validity of such recommendations are elaborated in the Appendix.

Clearly, the accuracy of the interpolation can be improved by using more input reference dDVHs spaced closer together but in the interest of efficiency, the lowest number of reference dDVHs that yields an acceptably accurate interpolation is preferred.

### 2.6.2 Efficiency

To assess how computationally efficient interpolation is compared to a straight recalculation, one can examine and compare the calculation time between the two methods. In dDVH calc, 15 dose distributions are required, consuming approximately 22 hours of dedicated calculation time. In contrast, dDVH interp requires 7 input reference curves that took almost 10 hours to calculate. These 7 dDVHs are then used for the interpolation and approximately 6.2 minutes is required to generate a control lattice. It takes less than 1 minute to interpolate the corresponding 15 dDVH curves. On the basis of these
time figures, interpolating is about 2 times more computationally efficient than recalculation.

However, once the control lattice is created, then one can find the dDVH for any arbitrary displacement within the lattice boundary. Approximately 82 minutes are required to interpolate 106 dDVH points (or about 1 200 interpolation points per minute). This translates to 0.08 minutes per dDVH curve for the interpolation (ignoring the time required to calculate the input reference dDVHs) compared to 88 minutes per dDVH curve for recalculation using the collapsed cone dose algorithm. Thus, interpolation is more than 1 000 times faster than recalculation in this example. It can be made even faster if larger dose bins are used and if the MATLAB code is more rigorously optimized.

The gain in relative computational efficiency using the interpolation compared to recalculation is related to: the dose calculation algorithm, the number of input reference dDVHs and the number of interpolated points required. Very sophisticated, time-consuming dose calculation algorithms applying Monte Carlo simulations would benefit the most from interpolation. However, when a choice needs to be made between accuracy and speed, interpolation may provide a workable compromise.

Although 7 input dDVHs were used for the interpolation, in more realistic clinical situations, displacement errors greater than 3.6 cm are unlikely to occur. The practical range of setup uncertainties is patient and site specific but the vast majority of cases would be limited to much less than 1.8 cm. In a review by Hurkmans et al [22], the stated achievable setup uncertainty (1 SD) for lung tumors is less than 3.5 mm. If a similar spacing is used, then only 3 input dDVHs (e.g. $\delta=\pm 0.7$, 0 cm) would be required to cover 2 SD (2×3.5 mm). However, this 3.5 mm SD is the best achievable setup uncertainty so it likely underestimates the typical setup uncertainty found in actual practice. A more prudent approach would be to use more than ±2 SD.

**2.6.3 Clinical Relevance**

Once the dosimetric differences between the invariant and variant situations have been quantified, the question of clinical relevance naturally arises. Clearly there are differences (as demonstrated by the RMS deviation) but what effect, if any, might it have on outcome? To answer this, the EUD can be used as a rough estimate of differing biologic effects from an inhomogeneous dose distribution. The EUD has been criticized for being unrealistic and an overly simplistic approach [23,24]. Nonetheless, the EUD remains a useful measure when comparing inhomogeneous dose distributions and is more robust to parameter uncertainty than TCP.

In our analysis, we determine the EUD for different displacement errors and the mean EUD for setup uncertainties. The divergence of these curves represents the true vari-
ant effect on EUD. Patient specific factors such as the actual contour shape and exact amount of lung within and without the irradiated fields will vary for different anatomies and geometries and it is unlikely any variant effects can be generalized for all patients.

However, some general observations can be made. Firstly, variant effects are smaller with smaller displacement errors and setup uncertainties. Secondly, the variant effect increases the EUD as the volume of irradiated lung increases. The decreased attenuation and increased lateral scatter (relatively) increases the effective absorbed dose and this results in a higher EUD. This is the reason why displacement errors directed towards more irradiated lung (i.e. $\delta > 0$ cm) have variant EUDs greater than invariant EUDs and vice versa.

Thirdly, asymmetric variant effects will tend to cancel out when the mean EUD is calculated as a function of setup uncertainty. The variant effects are asymmetric since there is little lung in the irradiated field for the negative displacement errors and more lung in the beam for positive displacement errors. As explained above, this asymmetry will cause the invariant EUDs to be greater than variant EUDs for negative displacement errors and vice versa for positive displacement errors. The variant effects as a function of displacement error are quite clearly shown in Figure 2-5A.

However, somewhat counter-intuitively, when the mean EUD as a function of setup uncertainty is determined, this asymmetry will tend to lessen the difference between the invariant and variant mean EUD. Since the setup uncertainty is normally distributed, there will be approximately the same number of positive and negative displacement errors. The invariant EUD's "overdosage" for negative displacement errors and "underdosage" for positive displacement errors tend to compensate each other and will tend to minimize the difference in EUD as a function of setup uncertainty between the variant case.

The importance of this effect is demonstrated in the mean EUD for the calculated dDVHs. If Figure 2-5A is studied carefully, one can see calculated EUD points overlap the interpolated EUD curve at the displacement errors that correspond to the reference input dDVHs (at $\delta = \pm 3.6$, $\pm 1.5$, $\pm 0.6$, and 0 cm) which is expected. But for the other corresponding displacement errors, the calculated EUD appears slightly greater than the interpolated EUD. Even though this difference appears relatively small with respect to the displacement error, in Figure 2-5B, the difference between the calculated and interpolated curves has approximately the same magnitude as the difference between the interpolated and invariant curves. Unlike the invariant curve in Figure 2-5A, the displacement errors for the calculated curve do not tend to compensate each other (compared to the interpolated curve). The implication is that symmetric variant effects
have a stronger influence on the EUD as a function of setup uncertainty than asymmetric variant effects.

In general, the variant effects are relatively small, particularly with respect to setup uncertainty where the maximum absolute difference is ≤1% of the maximum EUD, suggesting that the incorporation of contour changes and tissue inhomogeneities (i.e. variant effects) can be ignored. Since the variant effects on EUD are negligible for this “worst case” scenario, then, by implication, they can also be ignored for all “better case” scenarios. Although these results need to be interpreted with caution, it appears that the effects of a variant dose distribution on EUD may be ignored in the majority of clinical cases.

The results also suggest invariant dose blurring convolution-type functions could be used, even for extreme situations such as this clinical example, without significantly impacting the accuracy of the calculated EUD. Engelsman et al [9] also investigated the effect of random setup uncertainties on the EUD for lung tumors and found that increasing the random setup uncertainty from 0 to 1.5 cm reduced the EUD 7.3% (70 to 64.9 Gy). In our data, we found a EUD reduction of 12.8% for the variant (2.03 to 1.77 Gy) and 13.3% (2.03 to 1.76 Gy) for the invariant situation when increasing the setup uncertainty from 0 to 1.5 cm. Not surprisingly, our “worst case” example has larger relative reductions in the EUD. Substantial work in this area has been published by several authors [4, 9, 25-28].

This analysis concerns only the target volume and ignores the surrounding normal tissue. Although the effect on adjacent organs at risk was not the aim of this study, larger setup uncertainties will clearly spread the dose outside the target, potentially leading to more normal tissue complications.

### 2.6.4 Application

Given that the EUD and DVH can be calculated assuming either an invariant or a variant dose distribution, when can the variant effects on the target volume be ignored and when are they important? The significance of the variant effects depends strongly on what is being evaluated, namely displacement errors versus setup uncertainties as well as DVHs and EUDs. Setup uncertainties are more robust to variant effects compared to displacement errors.

With respect to DVHs, variant effects can be significant even for displacement errors as small as 0.6 cm. If DVHs are being used to evaluate the merit of a particular treatment plan, then, in principle, variant effects should be included for all displacement errors. However, if the target volume’s EUD is used to evaluate the treatment plan, then the variant effects could be excluded, particularly for small setup uncertainties. In general,
although the dosimetric differences may not be significant, variant calculations are more accurate than the invariant approach.

This apparent inconsistency highlights the importance in selecting an evaluator for a treatment plan and in distinguishing between displacement errors and setup uncertainties. Our results are consistent with other authors [4,9] who conclude that systematic errors are more significant than random errors. Ignoring the variant effects (as is true for invariant dose distributions) exaggerate the dosimetric differences for different displacement errors. However, this is not true for setup uncertainties since they partially compensate for these variant effects. For some parameters, such as EUD, the variant effects are so small they can be ignored. This is consistent with the observation that EUDs are less sensitive to dosimetric changes compared to DVHs.

2.6.5 Future Directions

The B-spline interpolation is extensible to multi-dimensional data. In principle, other geometric variables, such as rotation, could be interpolated. Other authors have used 6D transformation errors (3 translational and 3 rotational for x, y and z-axes) [5]. The next logical step would be to incorporate and test the interpolation approach with different variables. A more flexible interpolation is possible by using the dose grid directly. Instead of (volume, dose, displacement) dDVHs points as input (as per this paper), dose grid data points consisting of (dose, x, y, z, displacement) could be used. The required dDVH or isodose surface could then be calculated directly from the interpolated dose grid. Another advantage to this approach is that very complicated variables and effects can be modelled such as organ movement, organ deformation and CTV margin width.

The major disadvantage with the multilevel B-splines is that the control lattice size grows exponentially with each additional variable. Each iteration doubles every lattice dimension. Therefore, the number of lattice points increases by 2n with each iteration (where n is the number of independent variables). Several hundred megabytes of computer memory can be easily consumed during the generation of the control lattice so the number of variables and input data points should be kept as small as possible. It also follows that the selected independent variables be limited to those that are most significant. Practical computational limits make it difficult to test some of these models, particularly dose grid interpolation, due to the extremely large number of lattice points required.

2.7 Conclusions

Since a “worst case” example is selected, we conclude that, in the majority of clinical cases, the variant effects of contour changes, tissue inhomogeneities and setup uncer-
tainties on EUD are negligible. Interpolation is a valid, efficient method to approximate dDVHs.

**2.8 ACKNOWLEDGEMENTS**

The authors gratefully acknowledge the kind help provided by Joos Lebesque and Lambert Zijp.

**2.9 REFERENCES**


Appendix A

2.10 CLINICAL RECOMMENDATIONS ON DOSE TOLERANCE

One of the problems with the interpolation approach is qualifying the accuracy of the interpolant compared to the “true” value. Stated another way, when do the differences between the interpolated and the calculated dose-volume histogram (DVH) curves become significant? Or, more generally, how does one determine if the differences between 2 DVH curves are significant or not?

One method is to extrapolate the recommendations from other related situations to DVHs. For example, the tolerance allowed between the calculated and the measured dose could be used as a guideline to determine when dosimetric differences become clinically significant.

The recommended dose tolerance between calculated and measured relative dose calculations vary between 2-4%, depending on the complexity of the treatment setup and delivery as well as which part of the beam is being measured [21]. Measurements performed within the build-up region of the beam penumbra have steep dose gradients so small deviations in position can dramatically affect the measured dose. These recommendations are based on beam profile and percentage depth dose data.

One can, however, make some reasonable assumptions regarding these recommendations and apply them to DVHs. Beam profiles and percentage depth dose curves graphically describe the dose at number of given points. We can interpret the 2-4% dose tolerance recommendation as the dose at every point within the volume of interest (VOI) is within 2-4% of the measured dose. Thus, as a first order approximation, the calculated dose at every voxel within the VOI should also be within 2-4% of the measured dose, at least within the low dose gradient volume within the VOI.

For integral DVHs (iDVHs), the common interpretation is that it defines the relationship between some volume and some dose (i.e. 75% of the volume receives at least 50% of the prescribed dose). However, iDVHs can also be interpreted as all the voxels of the VOI ordered by dose. To illustrate this point of view, imagine a VOI with exactly 100 voxels with some dose distribution. Conceptually, every percent of the relative volume can be thought of as a “volume bin” (analogous to dose bins but along the other axis) and directly corresponds to a voxel. By reading the dose at each volume bin, one can find the dose for every voxel. In effect, each volume bin in the integral DVH (iDVH) behaves like a voxel. Strictly speaking, the maximum absolute difference in relative dose for between 2 iDVH curves should also lie between 2-4% if they are not clinically different.
One should also recognize that such a criterion, if applied to the volume bins of different iDVHs, is less stringent than if applied to each voxel due to the loss of spatial information in iDVHs. Because the voxel volume bins are ordered by dose, one cannot, in principle, determine the dose for an individual voxel. This is related to the degeneracy associated with iDVHs. The same iDVH can result from many different dose distributions and, therefore, iDVHs have a higher degree of freedom. As a consequence, it is possible for a given voxel to have a difference in relative dose greater than 4% but still have an iDVH with a difference less than 4%.

Another consideration is what the 2-4% recommended dose tolerance actually represents in practical terms. If the relative dose for all points within the VOI can be increased by 4% and the resulting iDVH is not considered clinically different, then one can use the RMS of the difference (i.e. RMS<2-4%) instead. The advantage with the RMS deviation is that a few volume bins can have a dose difference greater than 4% yet still result in an overall RMS less than 4%. This is particularly useful for regions with a steep dose gradient such as the beam penumbra. In fact, the recommendations recognize that for steep dose gradients, a tolerance of 2-4% is not always possible and allow up to a 15% difference (which corresponds to a distance of 3 mm).

However, if the differences in relative dose between the calculated and measured dose distributions are random and normally distributed, then the recommended tolerance of 2-4% should be viewed as an extreme value. In other words, most of the differences in relative dose should lie between ±2-4%. Under this assumption, the majority (e.g. ±2 SD=95%) of the absolute differences should be less than 2-4%. Therefore, one could interpret the standard deviation (SD) of the difference as 1-2%. This is equivalent to saying the RMS of the difference in relative dose between 2 iDVHs should be less than 1-2%.

Obviously this criterion is more stringent than maintaining an overall RMS of less than 4%. Which criterion is correct depends on which interpretation of the recommendation is preferred.