Pol J Med Phys Eng. 2008;14(3):135-149.
PL ISSN 1425-4689
doi: 10.2478/v10013-008-0012-2
website: http://www.pjmpe.waw.pl

Zdenek Moravek, Ludwig Bogner

# Entropy as a quality descriptor for the dose distribution — theory and practice for the patient target volume

Poliklinik fur Strahlentherapie, Universitat Regensburg,
Franz-Josef-Strauss Allee 11,
D-93042 Regensburg Germany
e-mail: zdenek.moravek@klinik.uni-regensburg.de

The most common and established way to evaluate the quality of a radiotherapy plan is to use the dose-volume histogram (DVH). The evaluation of the DVH, however, is a subjective procedure. This may not be crucial as long as the two plans are significantly different. In the case of several plans obtained with different planning or optimisation strategies the differences are often subtle and therefore a more objective comparison method is desirable. A commonly used approach is based on evaluation of the conformity index, however we show how it can fail for plans of similar quality.

Therefore we propose a new method based on the similarity of DVH to statistical distributions, which can be characterised uniquely by their entropy. The concept is defined separately for target volumes, where it is derived from the Fermi-like distribution, and for organs at risk, where the traditional approach is also considered in its derivation. The artificial illustratory and clinical examples show the properties of the entropy as the quality descriptor and compare it to the conformity index. The examples are focused to the patient target volumes, where the advantage of the concept is more evident.

**Key words:** plan evaluation, IMRT, DVH, conformity index.

## Introduction

The evaluation of a treatment plan is one of the crucial steps in radiotherapy as it determines the final plan from a set of possible ones. This is especially difficult in

optimisation procedures for intensity modulated radiotherapy (IMRT), where several plans with slight modifications of the constraints have to be compared. The choice is a process that depends on the individual treatment case and criteria defined by the clinician. Though there is no general agreement about these criteria due to the complexity of the problem [1], some objectives are considered superior to the others. The most commonly used criteria are homogeneous dose to the target volume and conformity to dose-volume constraints in the case of organs at risk [2–3].

The evaluation of the plan can be done in several ways with different effectiveness. The traditional procedure is based on comparison of isodose contours, which is a time-consuming and sometimes ambiguous task [1]. Others are based on computer algorithms that process dose and tissue information to return a simple descriptor. As the most popular the conformity index, conformation number, and dose-volume histogram (DVH) [4] are used. The conformity index examines volumes within a given isodose contour that belong to the target or not. There are various definitions, the following on [5–7]. Conformation number gives the volume fraction of the target that lies within a specified contour [8]. The DVH displays the volume fraction of a given region with the dose above a specified value [9].

Other descriptors are available, but they usually explore the properties of the three above — functional DVH (a variation of DVH that incorporates the non-uniform distribution of functional subunits into the dose-volume consideration [10]), effective DVH (using the biological effective dose to define DVH [11, 12]), etc.

All these descriptors suffer from the two basic limitations that we demonstrate on artificial examples shown in Figure 1. First, Figure 1a shows two target volume DVHs with one, A, being more suitable for cold regions and the other, B, for hot regions. A descriptor like conformity index will select one or the other according to what isodose level is chosen.

Second, Figure 1b shows two DVHs for an organ at risk with one, A, violating the first dose-volume constraint but then dropping fast to zero and the other, B, that meets all constraints but comes to zero much slowlier. Common descriptors disregards the violation of constraints completely.

For a given DVH, several methods of evaluation are in hand. The most common method is visual inspection [13]. However, we come up against the lack of objectivity. Alternatively, values may be inspected at several positions of the DVH [14]. Then, contradictory answers may be obtained at different points of the curve, as illustrated in

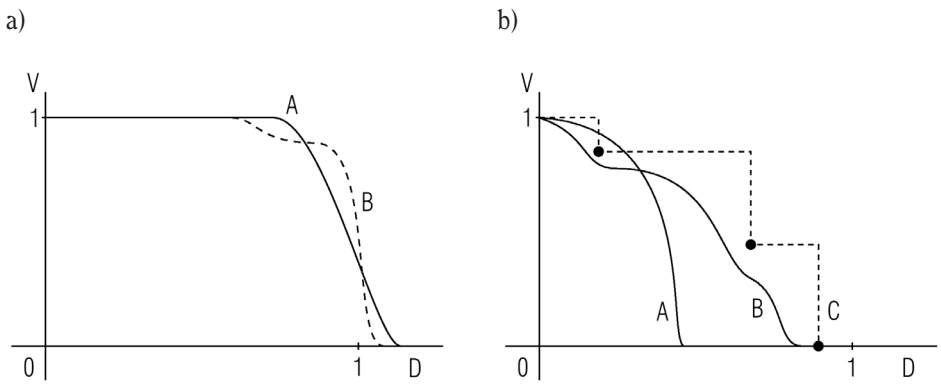


**Figure 1.** Dose ambiguity demonstrated on artificial DVHs. a) Typical target volume DVH — A and B curves show different quality in different regions; the ideal curve would be a step function at D=1. b) Typical DVH of an organ at risk with three constraints. Curve A violates first constraint but accommodates well with the others, curve B gives a satisfactory result, curve C is the worst acceptable (critical) DVH

Figure 1 and the result of the evaluation depends on the choice of the evaluation point. We call this aspect the DVH ambiguity.

The missing or superfluous area in the DVH [1], area under the graph [11], or least square difference to the ideal DVH can also be used as a measure of quality. In all cases the problems illustrated in Figure 1 persist — two DVHs may have such measures equal, but due to different distributions of cold/hot spots their quality differs. We call this aspect the area balance.

The source of DVH ambiguity and area balance is a consequence of the linearity of the evaluation methods used. There are other methods that evaluate the dose distribution in a nonlinear way such as the TCP-based models. These models derive inactivation probability of a volume element (from quadratic cell-response approximation) and combine the results for the whole target [15]. Therefore they bring in both, biology properties and nonlinearity, into the description. On the other hand, the models require precise parametrisation and more profound mathematical background, which makes them harder to be used.

In this work we propose a simpler nonlinear method for DVH evaluation that exploits the similarity of DVHs and the statistical distributions in physics. More precisely, it evaluates the DVH by one number derived from the shape of the curve with a well

defined algorithm. The inclusion of all DVH points in the calculation of the descriptor resolves the previously discussed problems.

## Methods

The method should consider the vector of DVH values and transform it into a unique scalar property. In physics or theory of information such a transformation can be done on the basis of knowledge of a statistical distribution $p(x)$ of quantity $x$ and the Boltzmann formula:

$$S = -\int_0^\infty p(x) \ln\Big(p(x)\Big)\, dx \quad . \tag{1}$$

This concept has already been used in radiotherapy as a part of objective function in the plan optimisation [16]. The authors used entropy as an additional penalty for dose heterogeneity in the target volume searching for a state with maximum entropy. In this paper, the concept of entropy is adopted in a different way — it is a search for the minimum entropy configuration. Additionally, two extensions are implemented to make the concept more useful for the comparisons.

First, in physics the deviations from the completely ordered configuration of a system are penalised symmetrically. It means that the entropy is combined of equal contribution due to increased population of one state and decreased of another. This is a consequence of the fact that the distribution is normalised. In case of DVH however, the distribution is not normalised and does not need to be normalised. In the first case, the norm of DVH itself reports of its quality (as in area based descriptors), in the other it may become important to treat hot and cold spots differently [7, 17]. For this reason Equation 1 is extended by a weighting function that accounts for such asymmetry

$$S = -\int_0^\infty w(x)\, p(x) \ln\Big(p(x)\Big)\, dx \quad . \tag{2}$$

where $p(x)$ is now the effective statistical distribution. The second aspect of using entropy for comparison is that the effective statistical distribution is considered to be the DVH itself,

$$p(x) \approx V(D) \tag{3}$$

As the nature of the DVH and its evaluation differs for the target volumes and organs at risk, both are considered in independent subsection. In short — the difference is manifested by the form of the weighting function.

## Target volume

As our understanding of biological response to irradiation is limited, we usually adopt the concept of homogeneous dose distribution in the target volume [2]. In this concept the ideal DVH is a step-like function, which resembles the Fermi-Dirac distribution at zero temperature, a state with minimum entropy. Any perturbation to the step-like function is understood in terms of increasing temperature. The measure of such perturbation is uniquely characterised by the entropy of the system. The weighting function in case of the target volume defines the importance of a perturbation to the prescribed value of the DVH in its different regions. This way hot and cold spots can be ranked differently. The form of the weighting function reflects a subjective strategy selected for the particular set of DVHs to be compared. Typically

- $w(D) = 1$. Then cold and hot spots are evaluated equally.
- Step function at $V_{ref} = 1$. Then either cold or hot spots are not penalised at all.
- Symmetric Gaussian function $w(D) = 1 - \exp\left(-(V - V_{ref})^2\right)$. Then the big deviations are penalised more than small ones, still symmetrically.
- Sigmoidal function at $V_{ref} = 1$, $w(D) = \dfrac{2}{1 + \exp\left(-(V - V_{ref})^2\right)} - 1$. Then both are penalised asymmetrically according to prescribed criteria (expressed as parameters of the sigmoidal function).

Effect of various weighting strategies is demonstrated in Figure 2. It shows the weighting functions and how they modify the analysed DVH.

The strategy in all cases is based on purely physical criteria, however a biological model can be used to define the weighting function as well.

Properties of this measure are the same as of the physical entropy — it acquires minimum for the ideal step-like distribution. Any perturbation manifests itself in increase of $S$ with amplitude given by the disorder brought about by this perturbation.

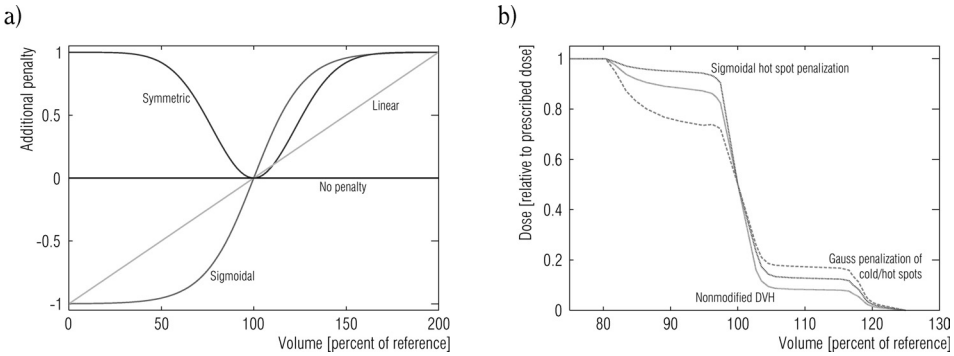


**Figure 2.** Effect of weighting: a) Several types of weighting functions and additional penalty they produce. b) Two different weighting strategies modify the original DVH — Gaussian penalisation of hot and cold spots pronounces deficiency of the DVH and sigmoidal function that tolerates cold spots (improving the DVH in the respective region) and penalises hot spots (making the DVH worse)

Two different configurations of a system may have the same value of $S$. They are then assumed to be equivalently ordered (in our words, they have the same quality). This quantity is therefore able to judge the situation in Figure 1a according to the measure of order (thus quality) of the curves.

## Organ at risk

For organs at risk, the radiological concept is to keep the dose as low as possible in the organs, at least to comply to a set of predefined dose-volume constraints. Therefore there is no definite ideal curve, but rather a critical curve that separates regions within which any curve is either acceptable or not. The shape of the critical curve is a sequence of step functions placed at the dose-volume constraints' positions (Figure 3), and the number of constraints can vary in the compared plans.

The fact that there is no ideal curve and that number of constraints is not constant prevents us from using the entropy as for target volumes. We therefore extend the area based descriptors to account for constraint violation. The extension must be so that the analysis of examples in Figure 3 is the following: curves A, B, which are very different,

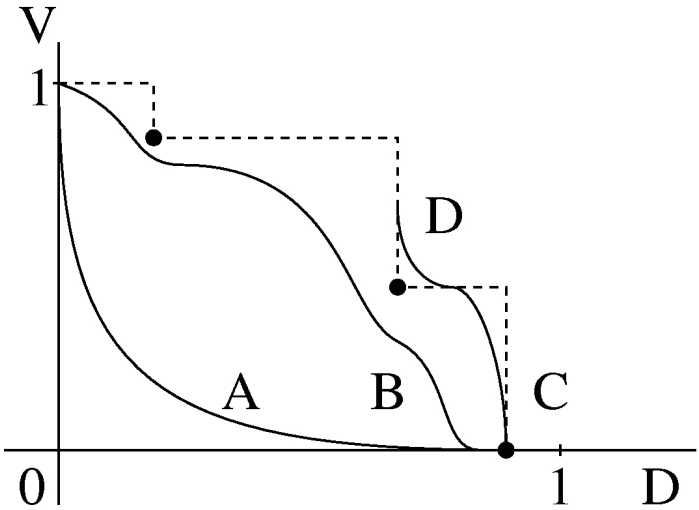


**Figure 3.** Demonstration of problems with defining entropy for organs at risk. Curve A describes a plan, where virtually no dose is delivered to the organ. Curve B shows a dose in the organ but well within the constraints. Curve C shows the worst acceptable (critical) result and curve D is a plan that slightly violates the constrains

have difference in descriptor values smaller than curves C, D. These are quite similar, but D violates a constraint.

An appropriate extension is to evaluate each region related to a constraint independently and to decompose the result into two parts — one related to the difference to the critical curve, $S_A$ and the other related to the constraint, $S_C$. The first part is standard area based descriptor, most easily the area

$$S_A = \int_0^\infty p(D)\, dD \ . \tag{4}$$

For $S_C$ the simplest choice is to use 0 when a constraint is not violated and the area below the critical curve otherwise. The total entropy related to the organ at risk is then the sum of all contributions,

$$S = S_A + \sum_{i \in constraints} S_C^i \ . \tag{5}$$

The procedure is in case of organ at risks similar to those described in [1, 11], however the use of the critical curve gives us more flexibility to penalise multiple violation of the constraints.

## Plan evaluation

A plan in general consists of several target volumes and organs at risk. Therefore it is characterised by a set of entropy values. The selection of a plan can be made according to their sum (additivity of entropy), weighted sum or just according to a subset of them. The selection must always come with a defined strategy for treatment of the hot/cold spots.

## Results and Discussions

### Effects of the weighting function

To investigate the effect of the weighting function we created three dose distributions — one with a big cold spot (CS), another with a small cold and hot spot (MED) and the last with a big hot spot (HS). The respective DVHs are shown in Figure 4. The weighting function was chosen to be linear in the transition area (relative dose ≈ 1.0),

$$\begin{aligned} w(x) &= c_1 \quad , \quad x \in [0, 0.8]; \\ w(x) &= a \cdot x + b \quad , \quad x \in [0.8, 1.2]; \\ w(x) &= c_2 \quad , \quad x \in [1.2, \infty]. \end{aligned} \tag{6}$$

**Table 1.** Entropy values for different weighting strategies

| Strategy | NO_OVERDOSE | NO_UNDERDOSE | No weighting |
|---|---|---|---|
| Cold spot | 0.005303 | 0.008769 | 0.014072 |
| Medium | 0.013565 | 0.008865 | 0.022396 |
| Hot spot | 0.021199 | 0.009019 | 0.030222 |

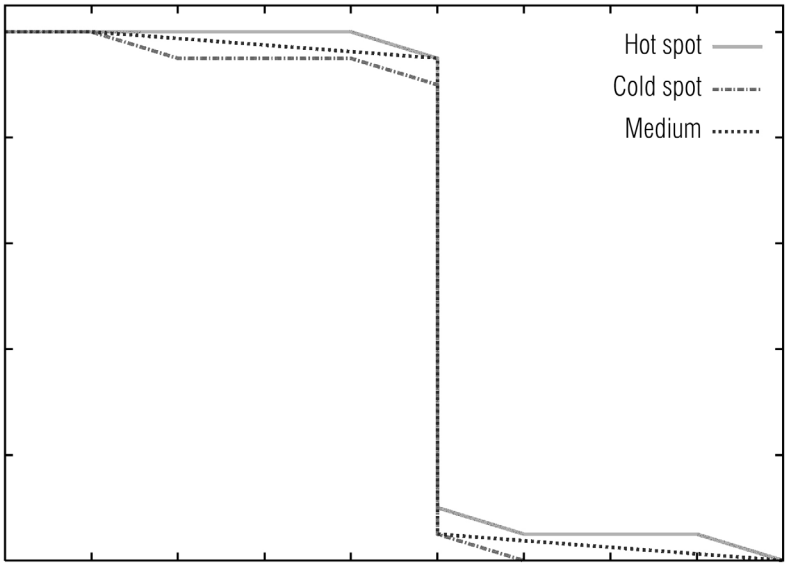


**Figure 4.** Model DVHs for the demonstration of effects of the weighting function on evaluation of the hot and cold spots in the DVH

In first approach, NO_OVERDOSE, the accent is put to penalise more the hot spot. This can be achieved with the following coefficients, $c_1 = 0, a = 2.5, b = -2, c_2 = 1$. For the contrary approach, NO_UNDERDOSE, the coefficients are $c_1 = 1, a = -2.5, b = 3$, $c_2 = 0$. Entropy calculations are summarised in Table 1, additionally values for no weighting are included as the reference.

Reference values show that HS is almost twice worse than CS and considerably worse than MED. However, NO_UNDERDOSE approach that should not penalise HS brings indeed the three values much closer. NO_OVERDOSE approach, on the other hand, increases the differences in the entropies.

## Clinical examples

The examples shown here are intended to demonstrate the use of our approach when compared to standard methods. For the sake of demonstration, we considered hot and cold spots in the same way ($w(D) = 1.0$) and we allowed only a reduced number of ROIs and constraints. We show two examples, a simple one, where the answer can be obtained with the standard methods and is confirmed by the new one, and the second case, where

the unambiguous answer can only come from the new approach. For the standard method we chose the conformity index as defined by Paddick,

$$\mathrm{CI} = \frac{\mathrm{TV}_P^2}{\mathrm{TV} \cdot \mathrm{PIV}} \tag{7}$$

where CI is the conformity index, stands for patient target volume within the specified isodose curve, TV is the total patient volume and PIV is for the total volume covered by the isodose curve.

Comparison of different strategies in optimisation of a head-neck case. In this section we deal with a photon IMRT of a head-neck case, which includes the target volume and a spinal cord as an organ at risk. We also add a margin of unspecified tissue (UT) around the target as another organ at risk (OAR) to avoid the hot spots close to the target.

We compare three plans obtained by different methods for the same patient and constraint setup:

- Monte Carlo calculation of a set of segments obtained from the commercial system (denoted as TMS);
- Monte Carlo calculation of the same set of segments with weights optimised to obtain required dose distribution, segment kernel optimisation — denoted as SKO, [18];
- Monte Carlo calculation of fluence maps defined by inverse kernel optimisation, denoted as IKO, [19].

For Monte Carlo calculations we used code developed by M. Fippel [20], called XVMC. The segments were obtained from TMS© (TMS, Nucletron, former Helax) planning system.

**Table 2.** Comparison of entropy and the conformity index (CI) values for given isodose (here at 95% of the prescribed dose)

| Method | CI(95) | Entropy |
|---|---|---|
| TMS | 0.55 | 0.046 |
| SKO | 0.57 | 0.038 |
| IKO | 0.66 | 0.035 |

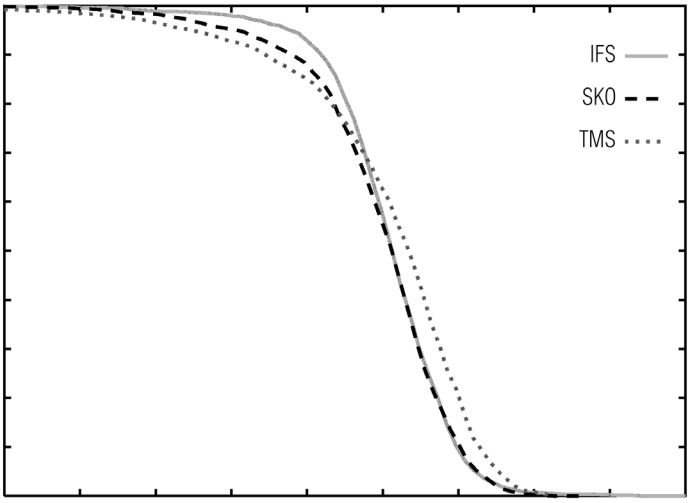


**Figure 5.** Three different optimisation strategies — TMS is a set of segments obtained from the commercial TMS©, SKO is a reoptimization of TMS weights for the segments, IKO is inverse kernel optimisation of the fluence maps. All three methods are based on Monte Carlo dose calculations: Target DVHs

The results can be easily evaluated by simple visualisation, as there is no ambiguity in the DVHs, Figure 2. We use this example to illustrate the reliability of the entropy approach.

Table 2 shows entropy values as well as the values of the conformity index (CI) for patient target volume (PTV) and isodose contour at 95% of the target average. We use the definition given by [5]. The observed values are in agreement with those of entropy calculation.

Beam setup study in proton therapy. Beam setup is a process in which an optimal orientation of a given number of beams is found with respect to the patient geometry. We started the optimisation process with two beams and kept the information from the initial state of the beam angle optimisation, one of the intermediate iteration state and one of the final state. With a pencil beam proton algorithm we calculated the relevant dose distributions and compared their DVHs of the target volume.

The similarity of the DVHs in particular cases makes the visual evaluation uncertain (Figure 6). However the entropy calculation revealed, that the plan quality improves in the process though not being a direct optimisation objective.

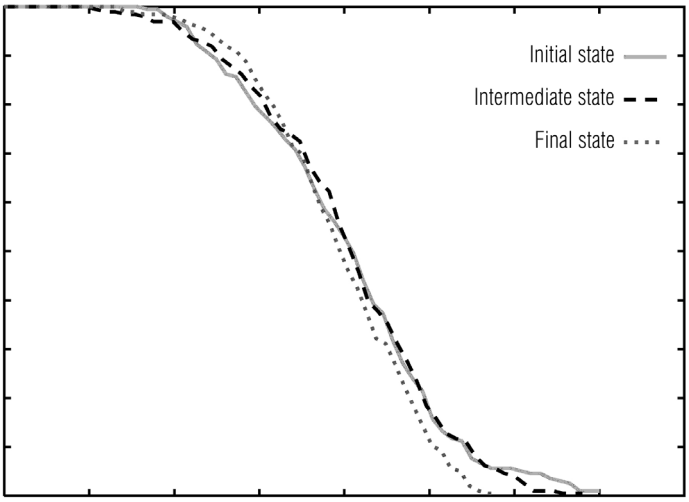


**Figure 6.** Proton beam setup optimisation in intensity modulated proton therapy: Target DVHs for several optimisation steps

Table 3 gives values of the entropy and the conformity index for PTV and two isodose contours (90% and 95% of the target average). It can be seen how the conformity index may be sensitive to its external parameter, the isodose level, and for the two levels it gives ambiguous answers. In first case, the first plan is shown to be the best one, in the other it is the third plan. This in turn reflects the fact that DVH curves themselves have different quality in different dose regions.

**Table 3.** Entropy and conformity index of PTV for selected isodoses (90% and 95%)

| Method | Entropy | CI(90) | CI(95) |
|---|---|---|---|
| S(Initial) | 0.0521 | 0.91 | 0.76 |
| S(Mediate) | 0.0485 | 0.87 | 0.77 |
| S(Final) | 0.0387 | 0.89 | 0.82 |

The entropy, on the other hand, is parametrised by its weighting function, Equation 2, and when applied to different plans, the relative quality is rather insensitive to the parameters as shown in the discussion of the weighting function effects. In fact, as long as the type of the weighting function remains unchanged (either to penalise more the hot spots or the cold spots) so remains the result of the mutual relative comparisons, Table 3.

## Conclusions

We described a new method useful for deeper exploration of information hidden in the DVH. The method was designed for cases where usual inspection may fail (i.e. when the DVH of two apparently different dose distributions are relatively similar) and is not meant to replace the standard ones. Thus, it comes in hand for comparison of various optimisation processes of the same problem. Here the step-to-step difference cannot be observed in any way but on bases of the presented description. Also a comparison of several plans can be more easily presented as a table of descriptor values than as a set of overlapping curves.

We derived a scalar property which is uniquely connected to the quality of a DVH in terms of physical optimisation. This measure is based on the widely used criteria — homogeneous dose to the target and minimum dose to the organ at risk using the analogy of DVHs and statistical distributions in physics. The algorithm comprises different weights given to under/over dose regions and includes the dose-volume constraints.

When compared to other scalar descriptors, the entropy concept does not disregard different importance of different DVH parts as the integration of the DVH curve uses non-linear prescription. The effects of weighting different regions of the DVH were shown. The descriptor is able to implement easily predefined strategies for treating hot and cold spots specifically.

The method, like the others, does not solve the following drawback: it can be used to give an answer to the question ‘which DVH is better’, but not to the question ‘by how much one DVH is better than the other’. It is a consequence of the nonlinearity in calculation of its value. It could be resolved if we knew the behaviour of the entropy as a function of defined number of voxels with a defined deviation of dose from the prescribed value in both directions.

Another obstacle is related to the fact that the entropy depends on the size of the system. There are three consequences for entropy-based descriptors:

– two different cases cannot be compared, unless some way of normalisation is applied;
– two different sets of prescribed values (reference dose for target volume and dose-volume constraints for organs at risk) cannot be compared at all;
– unlike the statistical distributions that were used for inspiration, the DVHs are not normalised. This property is on the other hand positive when various analysis of one case are compared.

The illustrative application of these measures was demonstrated for two clinical cases. The values obtained from them were in perfect agreement with intuitive evaluation (if feasible).

## Acknowledgements

This work was supported by Bayerisches Staatsministerium fur Umwelt, Gesundheit und Verbraucherschutz.

## References

[1] Viggars D A, et al. The objective evaluation of alternative treatment plans III: The quantitative analysis of dose volume histograms.
Int J Rad Oncol Biol Phys. 1992; 23: 419–427.

[2] Goitein M, Niemierko A. Intensity modulated therapy and inhomogeneous dose to the tumour: a note of caution. Int J Rad Oncol Biol Phys. 1996; 36: 519–522.

[3] Ragazzi G, Cattaneo GM, Fiorino C et al. Use of dose-volume histograms and biophysical models to compare 2D and 3D irradiation techniques for non-small cell lung cancer. Brit J Radiology. 1999; 72: 279–288.

[4] Niemierko A, Goitein M. Dose-volume distributions: a new approach to dose-volume histograms in three-dimensional treatment planning. Med Phys. 1994; 21(1): 3–11.

[5] Paddick I. A simple scoring ratio to index the conformity of radiosurgical treatment plans. J Neurosurg. 2000; 93(S3): 219–222.

[6] Knoos T, Kristensen I, Nilsson P. Volumetric and dosimetric evaluation of radiation treatment plans: Radiation conformity index.
Int J Rad Oncol Biol Phys. 1998; 42: 1169–1176.

[7] ICRU Prescribing, Recording and Reporing Photon Beam Therapy (Supplement to ICRU Report 50). ICRU Report 62, International Commission on Radiation Units and Measurements. 1999.
[8] van't Riet A, Mak A, Moerland M, Elders L, van der Zee W. A conformation number to quatify the degree of conformality in brachytherapy and external beam irradiation: Application to the prostate. Int J Rad Oncol Biol Phys. 1997; 37: 731–736.
[9] Drzymala RE. Dose-volume histograms. Int J Rad Oncol Biol Phys. 1991; 21: 71–78.
[10] Lu Y, Spelbring DR, Chen G. Functional dose-volume histograms for functionally heterogeneous normal organs. Phys Med Biol. 1997; 42: 345–356.
[11] Jani AB, Hand CM, Pelizzari CA, Roeske JC, Krauz L, Vijayakumar S. Biological-effective versus conventional dose volume histograms correlated with late genitourinary and gastrointestinal toxicity after external beam radiotherapy for prostate cancer: a matched pair analysis. BMC Cancer. 2003 May 13; 3: 16.
[12] Alber M, Nusslin F. Tools for the analysis of dose optimisation: I. Effect-volume histogram. Phys Med Biol. 2002; 47: 2451–2458.
[13] Fu W, Wang LH, Zhou ZM, Dai JR, Hu YM, Zhao LJ. Comparison of conformal and intensity-modulated techniques for simultaneous integrated boost radiotherapy of upper esophageal carcinoma. World J Gastroenterol. 2004; 10: 1098–1102.
[14] Gibbons JP. Technical note: The effect of the 4-mm collimator output factor on gamma knife dose distributions. J Applied Clinical Medical Physics. 2003; 4: 386–389.
[15] Webb S, Nahum AE. A model for calculating tumour control probability in radiotherapy including the effect of inhomogeneous dose distribution and clonoginic cell density. Phys Med Biol. 1993; 38: 653–666.
[16] Wu X, Zhu Y. A maximum-entropy method for the planning of conformal radiotherapy. Med Phys. 2001; 28 (11): 2241–2246.
[17] ICRU Prescribing, Recording and Reporing Photon Beam Therapy. ICRU Report 50, International Commission on Radiation Units and Measurements. 1993.
[18] Rickhey M, Moravek Z, Bogner L. Inverse treatment planning and integration of segmentation procedures. Z Med Phys. 2008; 18(3): 163–169.
[19] Bogner L, Hartmann M, et al. Application of an inverse kernel concept to Monte Carlo based IMRT. Med Phys. 2006;33: 4749–4757.
[20] Fippel M Fast Monte Carlo dose calculation for photon beams based on the VMC electron algorithm. Med Phys. 1999;26: 1466–1475.