Published in final edited form as:

Proc SPIE Int Soc Opt Eng. 2010 January 1; 7627: 76270Z–. doi:10.1117/12.844189.

Therapy Operating Characteristic (TOC) Curves and their Application to the Evaluation of Segmentation Algorithms

Harrison H. Barrett, Donald W. Wilson, Matthew A. Kupinski, Kasarachi Aguwa, Lars Ewell^a, Robert Hunter, and Stefan Müller^b

College of Optical Sciences and Department of Radiology, University of Arizona, Tucson AZ

^a Dept. of Radiation Oncology, University of Arizona, Tucson AZ

^b University of Duisberg-Essen, Essen, Germany

Abstract

This paper presents a general framework for assessing imaging systems and image-analysis methods on the basis of therapeutic rather than diagnostic efficacy. By analogy to receiver operating characteristic (ROC) curves, it utilizes the Therapy Operating Characteristic or TOC curve, which is a plot of the probability of tumor control vs. the probability of normal-tissue complications as the overall level of a radiotherapy treatment beam is varied. The proposed figure of merit is the area under the TOC, denoted AUTOC. If the treatment planning algorithm is held constant, AUTOC is a metric for the imaging and image-analysis components, and in particular for segmentation algorithms that are used to delineate tumors and normal tissues. On the other hand, for a given set of segmented images, AUTOC can also be used as a metric for the treatment plan itself. A general mathematical theory of TOC and AUTOC is presented and then specialized to segmentation problems. Practical approaches to implementation of the theory in both simulation and clinical studies are presented. The method is illustrated with a brief study of segmentation methods for prostate cancer.

Keywords

Radiation therapy; segmentation; registration; operating characteristic

1. INTRODUCTION

It is now generally accepted, at least in the SPIE Medical Imaging community, that image quality must be assessed on the basis of the ability of a user to perform medically of scientifically relevant tasks with the image data.¹ For diagnostic tasks, many evaluation methods related to receiver operating characteristic (ROC) curves have been developed and used to assess imaging systems and reconstruction algorithms. Much less has been done with the estimation tasks that arise in the context of radiotherapy.

The overall goal of radiotherapy is to maximize the probability of destroying tumors while minimizing damage to surrounding normal tissues. Many sophisticated approaches to this goal, including three-dimensional conformal radiotherapy² (3DCRT), intensity-modulated radiotherapy^{3–5} (IMRT) and image-guided radiotherapy⁶ (IGRT), have been developed for this purpose, but they all depend on information extracted from images of the patient. Moreover, even after the treatment plan is finalized, there is still a tradeoff between the probability of tumor control and the probability of normal-tissue complications, both of which increase when the overall intensity or duration of the treatment beam is increased.

Much work has gone into estimating these probabilities,^{7–12} and standardized modules for their calculation are freely available.^{13, 14} The form of the output from such modules is depicted in Fig. 1. The graphs on the left are for a hypothetical therapy plan where the operating point (overall dose) can be chosen for high probability of tumor control, denoted $\text{Pr}(\text{TC})$, and minimal probability of normal-tissue complications, $\text{Pr}(\text{NTC})$. In the graphs on the right, however, $\text{Pr}(\text{NTC})$ is higher for any given choice of $\text{Pr}(\text{TC})$ than in the graphs on the left, so these curves correspond to a less effective treatment plan.

The similarity of the curves in Fig. 1 to plots of true-positive fraction (probability of detection) and false-positive fraction (probability of false alarm) vs. decision threshold will be evident to anyone conversant with signal detection.¹ It is then an easy step to replot Fig. 1 as $\text{Pr}(\text{TC})$ vs. $\text{Pr}(\text{NTC})$ with the overall dose (intensity or duration of the treatment beam) as a parameter. This plot, initially introduced by Moore and Mendelsohn¹⁵ and analyzed in ROC terms by Metz et al.,¹⁶ is referred to as the Therapy Operating Characteristic or TOC curve. The TOC curves corresponding to Fig. 1 are shown in Fig. 2.

This paper uses the area under the TOC curve (AUTOOC) as a figure of merit for radiation therapy planning and for the imaging components that affect the plan. An AUTOOC of 1.0 is ideal, and AUTOOC = 0.5 corresponds to a treatment scenario in which the probability of tumor control can be increased only at the expense of an identical increase in the probability of normal-tissue complications.

In Sec. 2 we lay out the full theory of TOC with as much generality as we can muster. The result is that $\text{Pr}(\text{TC})$ and $\text{Pr}(\text{NTC})$ are each described by three nested integrals, where each integral is over a huge or even infinite number of variables and the integrands are unknown. In Sec. 3 we make some simplifying assumptions applicable to the goal of using AUTOOC to evaluate segmentation algorithms, and in Secs. 4 and 5, respectively, we show how AUTOOC can be estimated from phantom images or real clinical images. An illustrative application to prostate radiotherapy is presented in Sec. 6.

2. GENERAL THEORY

The imaging system is modeled as a continuous-to-discrete (CD) mapping¹ where the object is a function of continuous variables and the image is a set of discrete numbers such as voxel values. We assume that the tumor and all normal organs of interest have well-defined boundaries which can be represented as mathematical surfaces in a 3D space. For patient j , these surfaces are specified by a (potentially infinite) set of parameters denoted as $\mathbf{B}_j$.

The surface descriptions contained in $\mathbf{B}_j$ are, of course, unknown, and it is the goal of a segmentation algorithm to estimate them in some way. The resulting estimate of the surfaces for patient j is denoted $\hat{\mathbf{B}}_j$. Unlike the continuous descriptions in $\mathbf{B}_j$, the estimates in $\hat{\mathbf{B}}_j$ specify a set of voxels or surface tessellations that define the surfaces in a discrete sense.

The data used for estimating the boundaries consist of one or more images. Typically, the normal-organ boundaries and the tumor boundaries will be estimated from a CT or MR image, but there is an increasing interest in including biological information estimated from PET images into the treatment planning process. For generality, we denote the available image data set for patient j as $\mathbf{G}_j$, but we assume that this set must always include a CT study because x-ray attenuation data are needed for treatment planning.

From the estimated boundaries and the original CT data, a dosimetrist or physicist will derive a treatment plan which specifies the beam profiles, energies and exposures. Implementation of this plan will result in a dose distribution for patient j denoted $D_j(\mathbf{r})$, where $\mathbf{r}$ is a three-dimensional position vector in the patient's body. We assume that the dose distribution scales by the same factor at each position $\mathbf{r}$ as the exposure time or beam current is varied. If we assume further, for fractionated treatments, that the exposure is varied by the same proportion in each session, we can write

$$D_j(\mathbf{r}) = \alpha D_{0j}(\mathbf{r}), \quad (1)$$

where $D_{0j}(\mathbf{r})$ is the dose distribution at some reference exposure and α is the ratio of an actual exposure to the reference exposure. The horizontal axis in Fig. 1 can now be interpreted as $\alpha - 1$. For notational economy, we denote $\alpha D_{0j}(\mathbf{r})$ as $\mathbf{D}_j$, with the dependence on α implicit.

In passing, we note that scaling the exposure time is not precisely the same as varying beam current because of the dose-rate-dependent balance of tumor kill, repair and regrowth processes, but these effects are neglected in this initial paper.

The stochastic quantities that control the probabilities of tumor control and normal-tissue complications are noise in the image data; randomness in the estimated boundaries derived from a given image set; uncertainty in the actual delivered dose distribution, and of course the stochastic nature of the tissue response itself. We can describe all of these effects abstractly in terms of conditional probability density functions (PDFs).

The multivariate PDF for the image noise for patient j , denoted $\text{pr}(\mathbf{G}_j/j)$, describes the effect of the noise in the raw projection data as propagated through the reconstruction algorithm; for much more on this PDF and copious references on the subject, see Chapter 15 of Barrett and Myers.¹

The PDF for the estimated boundaries given the image data, denoted $\text{pr}(\hat{\mathbf{B}}_j/\mathbf{G}_j)$, would be a huge-dimensional delta function for an automatic segmentation algorithm because repeated trials with the same images would always deliver the same boundaries, but human analysts would not be so repeatable. We therefore retain $\text{pr}(\hat{\mathbf{B}}_j/\mathbf{G}_j)$ as an unspecified PDF for now.

The PDF for the actual dose for patient j given the estimated boundaries and the image data is denoted $\text{pr}(\mathbf{D}_j|\hat{\mathbf{B}}_j, \mathbf{G}_j, j)$. The estimated boundaries and the variations in x-ray attenuation coefficient in the CT data determine the treatment plan and hence the intended dose distribution, but this PDF allows the actual dose to deviate in an unknown, random way from the intended one and to depend also on other characteristics of the patient such as fine details not captured in the image data.

Finally, the probability (not PDF) of tumor control, which depends on the actual dose distribution, the actual boundaries and possibly other biological factors such as oxygenation and perfusion, can be denoted $\text{Pr}(TC|\mathbf{D}_j, \mathbf{B}_j, j)$. The notation here is a trifle redundant because conditioning on the patient is implicitly specifying the true organ boundaries, but we retain the dependence on $\mathbf{B}_j$ to emphasize its importance.

With these definitions, we can apply elementary rules of conditional probability to write the probability of tumor control for patient j , in full generality, as

$$\text{Pr}(TC|j) = \int d\mathbf{D}_j \int d\hat{\mathbf{B}}_j \int d\mathbf{G}_j \text{Pr}(TC|\mathbf{D}_j, \mathbf{B}_j, j) \text{pr}(\mathbf{D}_j|\hat{\mathbf{B}}_j, \mathbf{G}_j, j) \text{pr}(\hat{\mathbf{B}}_j|\mathbf{G}_j) \text{pr}(\mathbf{G}_j|j). \quad (2)$$

Similarly, the general expression for the probability of normal-tissue complications is

$$\text{Pr}(NTC|j) = \int d\mathbf{D}_j \int d\hat{\mathbf{B}}_j \int d\mathbf{G}_j \text{Pr}(NTC|\mathbf{D}_j, \mathbf{B}_j, j) \text{pr}(\mathbf{D}_j|\hat{\mathbf{B}}_j, \mathbf{G}_j, j) \text{pr}(\hat{\mathbf{B}}_j|\mathbf{G}_j) \text{pr}(\mathbf{G}_j|j). \quad (3)$$

Note that the probabilities given by (2) and (3) depend on both the actual and the estimated boundaries; the estimated boundaries determine the treatment plan, but given a plan and hence a dose distribution, it is the actual boundaries that determine $\text{Pr}(TC|j)$ and $\text{Pr}(NTC|j)$.

3. APPLICATION TO SEGMENTATION

The general equations (2) and (3) can be used to describe and evaluate almost any aspect of the therapy-planning process, including evaluation of planning algorithms and the dose computations used therein; evaluation of imaging systems and reconstruction algorithms; and study of the accuracy and reproducibility of humans performing manual segmentation. This paper, however, focuses on evaluation of segmentation algorithms, so we can introduce some additional assumptions about the other links in the therapeutic chain.

Most importantly, we can neglect the potential inaccuracies in dose calculations and assume that the actual dose distribution, given the plan, is precisely what the therapy-planning program says it is. There are two main justifications for this assumption. First, this critical aspect of radiotherapy has been widely studied, and the dose-computation algorithms have been continually refined and validated by Monte Carlo studies.¹⁷ One can surely have confidence in them if an accurate image model is used. Second, the dose delivered by a beam of high-energy photons is rather non-local; a photon that undergoes a photoelectric or Compton interaction at a specific point in the tissue delivers its dose over a range of a centimeter or more, and there are many such interactions, so noise or limited spatial resolution in the image may have little significant effect on the dose distribution.

The importance of this assumption is that we can eliminate one of the three integrals in (2) or (3). If there are no significant inaccuracies in dose calculation, we can write

$$\text{pr}(\mathbf{D}_j | \hat{\mathbf{B}}_j, \mathbf{G}_j, j) = \delta [\mathbf{D}_j - \mathbf{D}(\hat{\mathbf{B}}_j, \mathbf{G}_j)], \quad (4)$$

where $\mathbf{D}(\hat{\mathbf{B}}_j, \mathbf{G}_j)$ is the dose distribution (interpreted as a function of continuous spatial variables) returned by the planning algorithm. The delta function can then be used to perform the integral over $\mathbf{D}_j$, and (2) becomes

$$\text{Pr}(TC|j) = \int d\hat{\mathbf{B}}_j \int d\mathbf{G}_j \text{Pr} [TC | \mathbf{D}(\hat{\mathbf{B}}_j, \mathbf{G}_j), \mathbf{B}_j, j] \text{pr}(\hat{\mathbf{B}}_j | \mathbf{G}_j) \text{pr}(\mathbf{G}_j | j). \quad (5)$$

Another reasonable approximation is to neglect the biological subtleties of tissue damage by ionizing radiation. As noted above, $\text{Pr}(TC | \mathbf{D}_j, \mathbf{B}_j, j)$ can depend on patient characteristics not captured by the image data; that was the reason for the final conditioning on j . In fact, we have little way of knowing what these characteristics are in clinical practice, and we certainly could not incorporate them realistically in a simulation study. Thus we opt to delete that last variable and write $\text{Pr}(TC | \mathbf{D}_j, \mathbf{B}_j, j) = \text{Pr}(TC | \mathbf{D}_j, \mathbf{B}_j)$.

Now we note that the product of the last two conditional probabilities in (5) can be rewritten as a joint probability:

$$\text{pr}(\hat{\mathbf{B}}_j | \mathbf{G}_j) \text{pr}(\mathbf{G}_j | j) = \text{pr}(\hat{\mathbf{B}}_j, \mathbf{G}_j | j). \quad (6)$$

Thus (5) becomes

$$\text{Pr}(TC|j) = \left\langle \text{Pr} [TC | \mathbf{D}(\hat{\mathbf{B}}_j, \mathbf{G}_j), \mathbf{B}_j] \right\rangle_{\hat{\mathbf{B}}_j, \mathbf{G}_j}, \quad (7)$$

where the notation indicates that the quantity in angle brackets is to be averaged over the joint distribution of the subscripted quantities, $\hat{\mathbf{B}}_j$ and $\mathbf{G}_j$.

4. IMPLEMENTATION WITH SIMULATED IMAGES

The expression in (7) suggests a Monte Carlo algorithm that can be implemented with simulated data. Suppose we begin with an anatomical phantom to represent a particular patient j and we generate K random image samples, differing only by the realization of the Poisson noise in the raw projection data. The k^{th} such image sample is denoted $\mathbf{G}_j^{(k)}$ and the corresponding boundary estimate is denoted $\hat{\mathbf{B}}_j^{(k)}$. The pair $(\hat{\mathbf{B}}_j^{(k)}, \mathbf{G}_j^{(k)})$ is a sample from the joint distribution $\text{pr}(\hat{\mathbf{B}}_j, \mathbf{G}_j | j)$; only one index k is needed to specify the sample because $\hat{\mathbf{B}}_j^{(k)}$ is generated directly (though not necessarily deterministically) from $\mathbf{G}_j^{(k)}$.

Because we started with an anatomical phantom, the true boundaries $\mathbf{B}_j$ are known, and we can use a standard therapy planning algorithm and standard TCP and NTCP modules to evaluate the quantity in angle brackets in (7). The Monte Carlo estimate of $\Pr(TC|j)$ is then

$$\Pr(\hat{TC}|j) = \frac{1}{K} \sum_{k=1}^K \Pr \left[TC | \mathbf{D}(\hat{\mathbf{B}}_j^{(k)}, \mathbf{G}_j^{(k)}), \mathbf{B}_j \right], \quad (8)$$

where the probability on the righthand side is interpreted as the output of the TCP module.

Similarly,

$$\Pr(\hat{NTC}|j) = \frac{1}{K} \sum_{k=1}^K \Pr \left[NTC | \mathbf{D}(\hat{\mathbf{B}}_j^{(k)}, \mathbf{G}_j^{(k)}), \mathbf{B}_j \right]. \quad (9)$$

The notation doesn't show it, but the Monte Carlo estimates of (8) and (9) are still functions of α , so the computation gives one point on the TOC curve. Repeating for several values of α as in an ROC study generates the TOC curve for phantom j ; the study can be repeated for several phantoms to get an average TOC curve and AUTO C.

5. IMPLEMENTATION WITH CLINICAL IMAGES

If we wish to use real clinical images rather than simulations to generate TOC curves, we need a gold standard for the true boundaries. One approach is to have two or more experienced radiation oncologists painstakingly segment the images and agree on the resulting boundaries for the tumor and the critical organs. By analogy to the use of expert panels in standard ROC studies, the resulting $\mathbf{B}_j$ can then be adopted as the gold standard for evaluation of the performance of less experienced (or less painstaking) oncologists or of automatic contouring algorithms. It is also possible to begin with normal images, segment them manually in order to delineate the critical normal organs, and then add realistic simulated tumors with known boundaries.

Another problem in applying the theory introduced in this paper to real data is that we do not have the option of collecting multiple images on the same patient, so we cannot implement the factor $\Pr(\mathbf{G}_j|j)$ in (2), (3) or (5). We can proceed with the general probabilistic formulation of Sec. 3, therefore, only if we assume that, at the dose levels normally used in clinical CT, the image noise is not the limiting factor in segmentation accuracy. We can check this assumption by adding more noise to the raw CT data before reconstruction and determining at what noise level AUTO C is significantly affected, or we can revert to simulation.

If image noise is indeed negligible (for segmentation purposes) at clinical dose levels, then by analogy to (4) we can write

$$\Pr(\mathbf{G}_j|j) = \delta(\mathbf{G}_j - \overline{\mathbf{G}}_j), \quad (10)$$

where $\bar{\mathbf{G}}_j$ is the mean or noise-free image of patient j ; in practice, we simply estimate $\bar{\mathbf{G}}_j$ with the actual observed image for that patient.

With (10), (5) becomes

$$\Pr(TC|j) = \int d\hat{\mathbf{B}}_j \Pr \left[TC | \mathbf{D}(\hat{\mathbf{B}}_j, \bar{\mathbf{G}}_j), \mathbf{B}_j, j \right] \Pr(\hat{\mathbf{B}}_j | \bar{\mathbf{G}}_j). \quad (11)$$

How we apply this equation depends on whether the segmentation is manual or automatic. If we have N people performing manual segmentation, and the n^{th} of these produces estimated boundaries $\hat{\mathbf{B}}_{jn}$, then an estimated probability of tumor control for patient j , with manual segmentation and the assumptions already incorporated into (11), is given by

$$\Pr(\hat{TC}|j) = \frac{1}{N} \sum_{n=1}^N \Pr \left[TC | \mathbf{D}(\hat{\mathbf{B}}_{jn}, \bar{\mathbf{G}}_j), \mathbf{B}_j, j \right]. \quad (12)$$

A similar expression holds for the probability of normal tissue complications, so a patient-specific TOC curve is readily generated as described in Sec. 4. Further averaging over patients will yield an average TOC and AUTO.

Evaluation of automatic segmentation algorithms is even easier. These algorithms are never uncertain but usually wrong – relative to the assumed gold standard. In estimation-theoretic terms, they have zero variance but nonzero bias. If algorithm m produces boundaries $\hat{\mathbf{B}}_{jm}$ for patient j , $\Pr(TC)$ for that algorithm and patient is given by

$$\Pr(\hat{TC}|j, m) = \Pr \left[TC | \mathbf{D}(\hat{\mathbf{B}}_{jm}, \bar{\mathbf{G}}_j), \mathbf{B}_j, j \right]. \quad (13)$$

Again, this expression depends implicitly on α , and a similar expression holds for

$\Pr(\hat{NTC}|j, m)$, so a TOC curve specific to algorithm m and patient j is readily constructed. Averaging over patients then provides a task-based metric for the particular algorithm.

6. EXAMPLE: APPLICATION TO PROSTATE CANCER

Rectal bleeding is a significant normal-tissue complication in radiotherapy of prostate cancer.¹⁸ For this reason, we chose a clinical case of prostate cancer to illustrate the use of TOC curves with clinical images for a single patient. The prostate, seminal vesicles, bladder and rectum had previously been manually segmented, and we considered these boundaries to be the gold standard.

The treatment plan actually used clinically with this segmentation provided the dose distribution for the tumor and normal organs. In the notation of (13), this procedure provided $\mathbf{D}(\hat{\mathbf{B}}_{j1}, \bar{\mathbf{G}}_j)$ with the best manual segmentation being denoted as algorithm $m = 1$. Because this algorithm was taken as the gold standard, we set $\mathbf{B}_j = \hat{\mathbf{B}}_{j1}$ in (13) and used the TCP and NTCP modules of Gay and Niemierko¹³ to compute the probability of tumor control and the

probability of rectum complications from the cumulative dose-volume histograms. The resulting dose-response curves are shown on the left in Fig. 3.

Then we artificially expanded the boundaries of the prostate by 5 mm in all directions and made a new treatment plan with wider beams as if these expanded boundaries were correct. Equivalently, we increased the PTV (planned treatment volume) by 5 mm in all directions.

This new segmentation was denoted $m = 2$ and a new dose distribution, $\mathbf{D}(\hat{\mathbf{B}}_{j2}, \hat{\mathbf{G}}_j)$, was derived. With the true boundaries still taken as $\mathbf{B}_j = \hat{\mathbf{B}}_{j1}$, we again used the modules of Gay and Niemierko to compute $\Pr[TC/\mathbf{D}(\hat{\mathbf{B}}_{j2}, \hat{\mathbf{G}}_j), \hat{\mathbf{B}}_{j1}, j]$, and similarly for NTCP. The resulting dose-response curves for this case are shown on the right in Fig. 3.

The two sets of dose-response curves in Fig. 3 are remarkably similar, and one might think that the intentional segmentation error made no difference at all. When the data from Fig. 3 are replotted as TOC curves in Fig. 4, however, the difference becomes evident. The gold-standard manual estimation yields $AUTOC = 0.894$, and the erroneous segmentation yields $AUTOC = 0.857$. There is no need to assign error bars to these figures of merit or to inquire whether the difference is statistically significant because there is nothing statistical about them. The TCP and NTCP modules compute the probabilities from the input dose-volume histograms and the input biological parameters for the organs in question, and there is no variance in the results. The output probabilities are no doubt biased, in the sense that they may not agree well with clinical outcomes, but the bias is the same for both segmentation algorithms. In this example, the manual segmentation was the gold standard, and the intentionally erroneous segmentation gave higher probability of rectal complications at all tumor-control probabilities, as expected. We made no effort to use biologically accurate input parameters to the modules, and we speculate that the rank ordering of segmentation algorithms will be invariant to the choice of these parameters in most situations.

7. SUMMARY AND CONCLUSIONS

We have analyzed therapy operating characteristic curves and proposed the use of the area under the curve as a figure of merit for comparison of segmentation algorithms. The sensitivity of this method of comparison was demonstrated for a single clinical case of prostate cancer.

TOC curves have many other applications as well. They can be used to compare different imaging instruments, reconstruction algorithms or image-registration methods that may go into the treatment-planning process. They can also be used to compare and optimize the planning algorithms themselves, especially in simulation studies where there is no uncertainty about the true organ boundaries. Finally, they can be used to evaluate methods to overcome practical clinical problems such as patient positioning or motion during treatment.

In all of these cases, it seems likely that the comparison of instruments, algorithms and treatment methodologies will be insensitive to the biological accuracy of the TCP and NTCP calculations. If, however, that biological accuracy can be clearly established, TOC curves for individual patients can, in principle, be used to determine the operating point (overall

dose) that is best for that patient. In this case, multiple TOC curves for different normal organs and degrees of complication can be used to choose the right balance for the patient.

Acknowledgments

We have benefited significantly from discussions with Julian Rosenman and Stephen Pizer at the University of North Carolina, and we thank Charles Metz for calling our attention to references ¹⁵ and ¹⁶. This research was supported by the National Institutes of Health under grants R37 EB000803 and P41 EB002035.

References

1. Barrett, HH.; Myers, KJ. Foundations of Image Science. Wiley; New York: 2004.
2. Zelevsky MJ, Leibel SJ, Kutcher GJ, Fuks Z. Three-Dimensional Conformal Radiotherapy and Dose Escalation: Where Do We Stand? Seminars in Radiation Oncology. 1998; 8(2):107–114. [PubMed: 9516591]
3. Eisbruch A. Intensity-Modulated Radiation Therapy: A Clinical Perspective. Seminars in Radiation Oncology. 2002; 12(3):197–282. [PubMed: 12118384]
4. Hatano K, Araki H, Sakai M, et al. Current status of intensity-modulated radiation therapy (IMRT). Int J Clin Oncol. 2007; 12:408415.
5. Ahnesjo A, Hardemark B, Isacsson U, Montelius A. The IMRT information processmastering the degrees of freedom in external beam therapy. Phys Med Biol. 2006; 51:R381R402. [PubMed: 16790914]
6. Jaffray DA. Image-Guided Radiation Therapy: From Concept to Practice. Seminars in Radiation Oncology. 2007; 17(4):243–305. [PubMed: 17903701]
7. Chow JCL, Markel D, Jiang R. Technical Note: Dose-volume histogram analysis in radiotherapy using the Gaussian error function. Med Phys. 2008; 35(4):1398–1402. [PubMed: 18491534]
8. Romeijn HE, Dempsey JF. Intensity modulated radiation therapy treatment plan optimization. TOP. 2008; 16:215243.
9. Wu Q, Mohan R, Niemierko A, Schmidt-Ullrich R. Optimization of intensity-modulated radiotherapy plans based on the equivalent uniform dose. Int J Radiation Oncology Biol Phys. 2002; 52:224–35.
10. Stavrev P, Schinkel C, Stavreva N, Fallone BG. How well are clinical gross tumor volume DVHs approximated by an analytical function? Radiol Oncol. 2009; 43(2):132–135.
11. Deasy, JO.; El Naqa, I. Radiation Oncology Advances. Vol. Chap 11. Springer; 2008. Image-Based Modeling of Normal Tissue Complication Probability for Radiation Therapy.
12. Karger, CP. New Technologies in Radiation Oncology. Vol. Chap 18. Springer; Berlin: 2006. Biological Models in Treatment Planning.
13. Gay HA, Niemierko A. A free program for calculating EUD-based NTCP and TCP in external beam radiotherapy. Physica Medica. 2007; 23:115–125. [PubMed: 17825595]
14. Warkentin BJ, Field C, Stavrev P, Schinkel C, Stavreva N, Fallone BG. A TCP-NTCP estimation module using DVHs and known radiobiological models and parameter sets. Journal of Applied Clinical Medical Physics. 2004; 5(1):50–63. [PubMed: 15753933]
15. Moore DH II, Mendelsohn ML. Optimal treatment levels in cancer therapy. Cancer. 1972; 30:97–106. [PubMed: 5040749]
16. Metz CE, Tokars RP, Kronman HB, Griem ML. Maximum likelihood estimation of dose-response parameters for therapeutic operating characteristic (TOC) analysis of carcinoma of the nasopharynx. Int J Radiat Oncol Biol Phys. 1982; 8:1185–1192. [PubMed: 7118618]
17. Chetty IJ, Curran B, Cygler JE, et al. Report of the AAPM Task Group No. 105: Issues associated with clinical implementation of Monte Carlo-based photon and electron external beam treatment planning. Med Phys. 2007; 34(12):4818–4853. [PubMed: 18196810]
18. Sohn M, Yan D, Liang J, et al. Incidence of late rectal bleeding in high-dose conformal radiotherapy of prostate cancer using equivalent uniform dose-based and dose-volume-based

normal tissue complication probability models. *Int J Radiation Oncology Biol Phys.* 2007; 67(4): 10661073.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1.

Dose-response curves for two hypothetical treatment plans. Left: Probabilities of tumor control and normal-tissue complications for a relatively effective plan. Right: Same for a less effective plan. In both, the abscissa is the deviation of the overall dose from some standard or nominal value (arbitrary units).

Figure 2. Therapy Operating Characteristic curves for the two treatment plans depicted in Fig. 1. The upper curve in this figure corresponds to the plan on the left in Fig. 1.

Figure 3.

Dose-response curves showing TCP for prostate tumor and NTCP for rectum vs. dose. Left: Data obtained with best manual segmentation and optimized plan. Right: Data obtained with new plan based on expanded prostate boundaries, but dose calculated for correct prostate boundaries.

Figure 4.

TOC curves derived from the dose-response curves of Fig. 3. The top curve, corresponding to best manual segmentation, has $AUTOC = 0.894$, and the lower curve, where the treatment plan was based on an erroneous segmentation, has $AUTOC = 0.857$.