Int. J. Radiation Oncology Biol. Phys. Vol. 21, pp. 137–146
Printed in the U.S.A. 

0360-3016/91 $3.00 + .00


● *Original Contribution*

# HISTOGRAM REDUCTION METHOD FOR CALCULATING COMPLICATION PROBABILITIES FOR THREE-DIMENSIONAL TREATMENT PLANNING EVALUATIONS

G. J. KUTCHER, PH.D.,[1] C. BURMAN, PH.D.,[1] L. BREWSTER, M.S.,[1]
M. GOITEIN, PH.D.[2] AND R. MOHAN, PH.D.[1]

[1]Memorial Sloan-Kettering Cancer Center, New York, NY 10021; and [2]Massachusetts General Hospital, Department of Radiation Medicine, Boston, MA 02114 and Harvard Medical School

**New tools are needed to help in evaluating 3-D treatment plans because of the large volume of data. One technique which may prove useful is the application of complication probability calculations. A method of calculating complication probabilities for inhomogeneously irradiated normal tissues is presented in this paper. The method uses clinical estimates of tolerance doses for a few discreet conditions of uniform partial organ irradiation, an empirical fit of a continuous function to these data, and a technique (the effective volume method) for transforming nonuniform dose-volume histograms into equivalent uniform histograms. The behavior of the effective volume histogram reduction method for various boundary conditions is reviewed. The use of complication probabilities in evaluating treatment plans is presented, using examples from an NCI 3-D treatment planning contract.**

**3-D treatment planning, Complication probability calculations, Treatment planning evaluation, Scoring.**

## INTRODUCTION

Treatment plans have traditionally been evaluated by reviewing dose distributions on one or a few cross sections of the patient. Usually the field arrangements are straightforward, and decisions on a given treatment plan or a comparison between plans, while not unambiguous, are guided by tradition and experience. The relative ease in evaluating 2-D treatment plans cannot be maintained in the 3-D case for two reasons. On the one hand, the amount of data that should be reviewed for a full three-dimensional evaluation is considerably greater. On the other hand, with the new dose planning systems, it is possible to obtain unfamiliar beam arrangements; two examples are noncoplanar fields and plans using dynamic scanning techniques. The dose distributions in these cases may be quite different from those for traditional treatment plans, and therefore an extrapolation of clinical experience from traditional situations may be difficult.

Reducing the amount of data is one solution to the problem. Dose-volume histograms (5, 9, 18) (DVHs) have proved useful in this respect. However, DVHs do not answer the second problem discussed above. In addition, it is sometimes difficult to choose between dose-volume histograms which cross one another. Another approach to the evaluation problem is to use normal tissue complication probability calculations (1, 2, 11–13, 15–17, 19, 20) and tumor control probabilities (1, 2, 4, 8). Such calculations substantially condense the dose distribution data, and in principle are capable of aiding in the extrapolation of clinical knowledge from simple to more complex cases.

The purpose of this paper is to present a methodology for calculating normal tissue complication probabilities under conditions of inhomogeneous irradiation; another aim is to present the results of these calculations for plans at several sites, and to demonstrate how normal tissue complication probabilities may be used for treatment planning evaluation, and for the analysis of treatment planning problems. The accuracy of the complication probability calculations and their behavior under different planning conditions are also discussed.

## METHODS AND MATERIALS

Our program for calculating normal tissue complication probabilities for nonuniform normal organ irradiation is comprised of three parts. In the first part, the clinically known or accepted tolerance doses for some specific conditions including both uniform whole and partial organ irradiation are obtained. A compilation of these tolerances and a discussion is contained in a paper by Emami *et al.*

Supported in part by NCI Contracts N01 CM-47316, N01 CM-47695, N01 CM-47696, and N01 CM-47697.

Reprint requests to: G. J. Kutcher, Memorial Sloan-Kettering Cancer Center, Medical Physics, 1275 York Ave., New York, NY 10021.

(7). Their aim was to obtain, preferably from the literature, or from their best clinical estimates, tolerance doses for 5% and (generally) 50% complication probabilities for whole, and (generally) 2/3 and 1/3 uniform partial organ irradiation. These data comprise a basic set on which interpolations and extrapolations are performed to obtain probabilities for doses and volumes not in the initial data set.

In order to apply the very limited data to the broad range of situations encountered in clinical practice, the data are fit to a four parameter model suggested by Lyman (12). The normal tissue complication probability, NTCP, for a uniform dose, D, to a volume, V, of the organ is given by

$$\mathrm{NTCP} = 1/\sqrt{2\pi} \int_{-\infty}^{t} \exp(-t^2/2)\, dt \qquad (1)$$

where

$$v = V/V_{ref} \qquad (2)$$

$$t = (D - TD_{50}(v))/(m \cdot TD_{50}(v)) \qquad (3)$$

and $TD_{50}(v)$ is the tolerance dose for a 50% complication probability for uniform irradiation to the partial volume v. It is related to the tolerance dose for whole organ irradiation (v = 1) through

$$TD(1) = TD(v) \cdot v^n \qquad (4)$$

$V_{ref}$ is taken to be the volume of the organ where possible, otherwise it is some arbitrary reference volume. The parameters, n and m, and $TD_{50}$ are obtained by fitting the tolerance data to Eq. 1. Details of the procedure and results of fitting the tolerance data of Emami (7) are discussed by Burman (3). In general, the slope parameter m is approximately in the range of 0.1 to 0.2, and the volume factor n varies from 0.1 to 0.7. Figures 1a and 1b show examples of complication probability distributions vs. dose for an organ with a small volume effect, the small intestine (n = 0.15), and one with a large volume effect, the parotid gland (n = 0.7). Using Eqs. 1–4 it is possible to obtain, for any tissue, the complication probability for uniform irradiation to any partial organ volume once n, m, $V_{ref}$, and $TD_{50}$ are known.

If the dose distribution is not a simple, idealized example in which one section of an organ receives a uniform and known dose and the next receives no dose at all, then Eqs. 1–4 do not apply. To use them it would be desirable to convert the nonuniform histogram into an equivalent uniform one. Two methods for transforming a nonuniform histogram have been suggested, the first by Lyman (12) (the interpolation method) and the other by Kutcher (11) (the effective volume method). In the former method, the histogram is transformed into a uniform one with a volume equal to the whole organ volume, and a peak dose less than

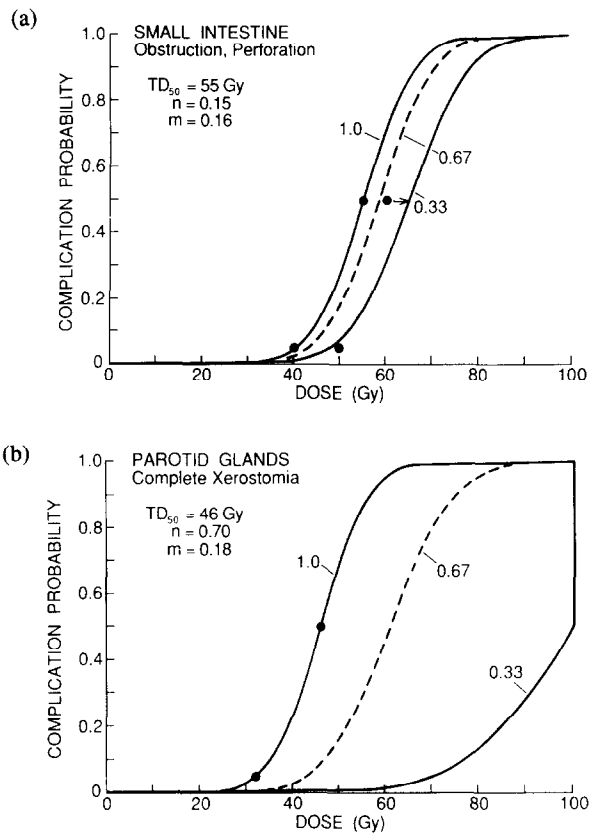


Fig. 1. (a) Complication probability vs. dose for obstruction and perforation of the small bowel for uniform whole, one-half and one-quarter irradiation. (b) Complication probability vs. dose for complete xerostomia of the parotid gland for uniform whole, two-thirds, and one-third irradiation.

or equal to the maximum dose to the organ. In the latter technique, the transformed uniform histogram has a volume less than or equal to the whole organ volume, and a peak dose equal to the maximum dose delivered to the organ.

Reference to Figure 2 will clarify the two methods. The interpolation method represents a nonuniform dose-volume histogram by a series of steps (Fig. 2a), and then proceeds using the observation that the complication probability for the last two steps of a dose-volume histogram has a value somewhere between the probability for the next to last step $P(D_i,V_i)$ and the last step $P(D_m,V_i)$. Once this probability $P(D',V_i)$ is known, the last two steps can be replaced by a single step shown in the figure. The position of the interpolated step (that is the position of D′) may be determined through

$$P(D',V_i) = (V_i - V_m)/V_i \cdot P(D_i,V_i) + V_m/V_i \cdot P(D_m,V_i) \qquad (5)$$

by solving for D′. Equation 5 states that the probability as a function of dose can be linearly interpolated in volume. This procedure is continued step by step until the entire

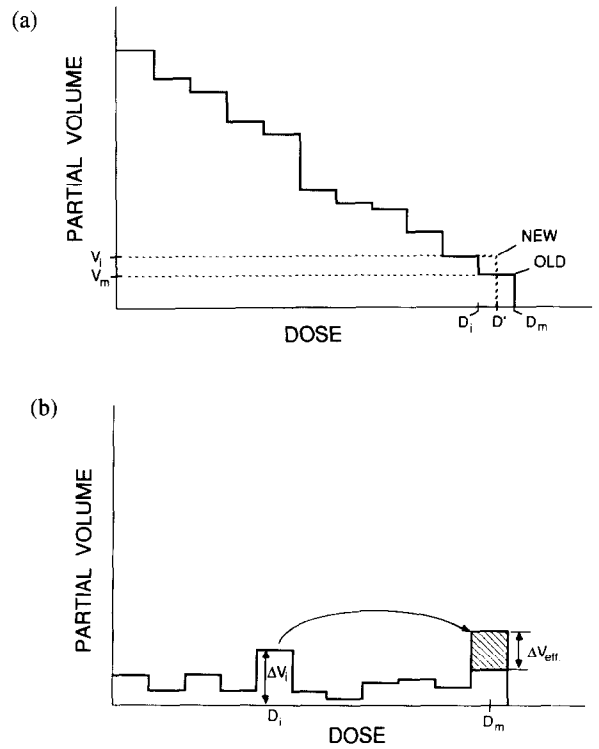


Fig. 2. (a) Dose-volume histogram reduction using linear interpolation. (b) Dose-volume histogram reduction using the effective volume method.

histogram has been reduced to a histogram with only one bin. In a similar manner, the interpolation can be done over dose, or over both dose and volume.

The effective volume method proceeds from the conjecture that each volume element of the differential dose-volume histogram independently obeys the same dose-volume relationship as the whole organ: that relationship, for organs which are represented by Lyman's four parameter, is given by Eq. 4. If this is the case, then each step of the histogram shown in Figure 2b of height $\Delta V_i$ and dose $D_i$ is adjusted to one with smaller volume $\Delta V_{eff}$ and dose $D_m$ through

$$(\Delta V_{eff})_i = \Delta V_i \, (D_i/D_m)^{1/n} \qquad (6)$$

The compression of each section of the histogram is repeated until a single bin is obtained with dose $D_m$, and volume

$$V_{eff} = \Sigma \, (D_i/D_m)^{1/n} \, \Delta V_i \qquad (7)$$

This single bin in the differential dose-volume histogram corresponds to a uniform cumulative dose-volume histogram with dose equal to $D_m$. The choice of $D_m$ as the dose to which each element $\Delta V_i$ is transformed (Eq. 6) assures that $V_{eff}$ (Eq. 7) is always less than or equal to the volume of the whole organ.

These calculations are augmented in two ways. First, the best estimate of $V_{ref}$ is that derived from CT data. However, the CT scans may not contain the entire volume of the organ. In that case we advocate using an expected or standard volume when the calculated volume is less than some fraction of the standard volume. Secondly, we introduce the notion of a critical volume, namely that volume which can receive an arbitrarily high dose without a complication for the same end point. If the volume of tissue which receives any dose at all (equal to the second bin of the DVH) is less than that critical volume, then the complication probability is set to zero.

## RESULTS AND DISCUSSION

*Relative behavior of histogram reduction*

The calculation of complication probabilities for nonuniform irradiation cannot be compared to clinical data, since few such data are available. Since histogram reduction is an extrapolation of probabilities from the uniform to the nonuniform radiation condition, the technique should be shown to satisfy a number of boundary conditions. At least this will assure that the calculation behaves properly under known conditions.

1. The energy bin size used for the histogram reduction should be chosen so that the calculations are insensitive to bin size changes.
2. If the histogram reduction is applied to a uniform histogram then it should yield the expected complication probability for uniform irradiation.
3. When small low dose areas (cold spots) are contained in a histogram, the calculated NTCP should be less than the NTCP for the histogram without the cold spot. When hot spots are introduced the calculated value should increase. In addition, as the size of hot or cold spot approaches zero, the calculated NTCP should approach that for the uniform histogram.
4. The values of n vary between zero and unity (3). Therefore:
   (a) Since small values of the parameter n for normal tissues or organs means a high dependence of NTCP with volume (or equivalently a high dependence of complication probability with dose), as n approaches zero, the histogram reduction technique should provide that the calculated probability is determined by the largest dose in the histogram.
   (b) As n approaches unity, it follows from Eq. 4 that if the dose and volume of a histogram are modified so that their product is unaltered, then the calculated complication probability should remain unaltered. Therefore, when n is equal to unity, the complication probability should be correlated with the integral dose.
5. When histogram A, for an organ, is enclosed by histogram B for the same organ, the calculated NTCP for organ A must be less than or equal to the calculated NTCP for organ B.

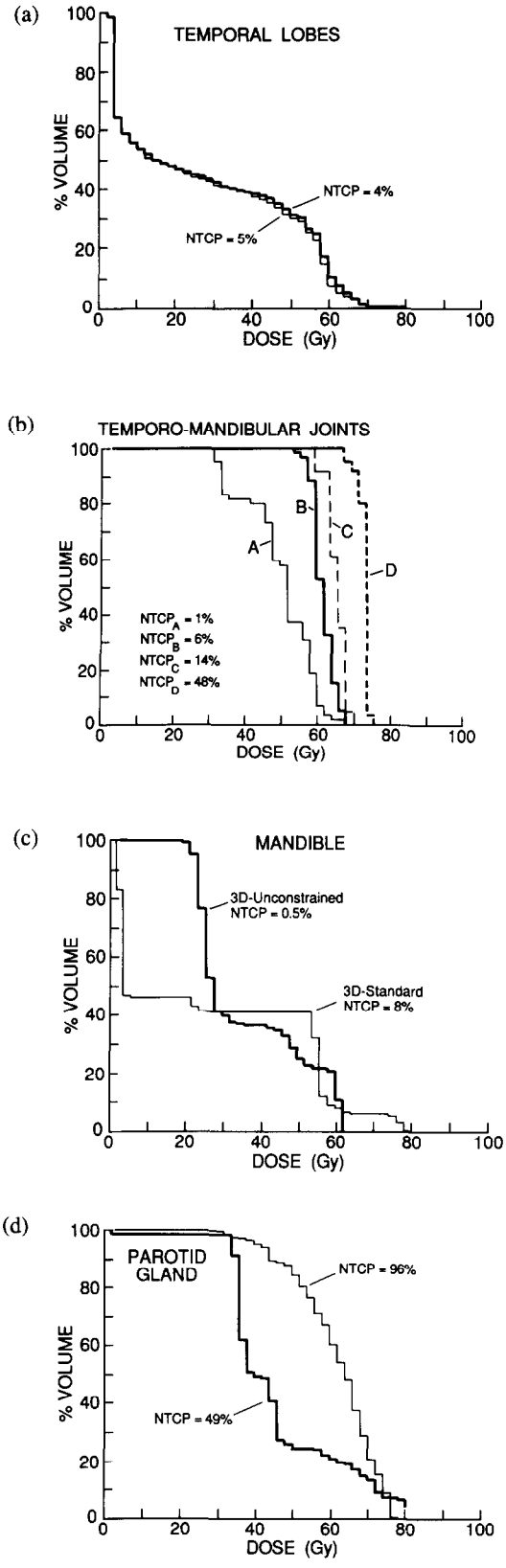


Fig. 3. (a) Dose-volume histograms for the temporal lobes for two treatment planning techniques. (b) Dose-volume histograms for left temporo-mandibular joint for 2 different treatment planning techniques. (c) Dose-volume histograms for the mandible for two treatment planning techniques. (d) Dose-volume histograms for the parotid gland for two treatment planning techniques.

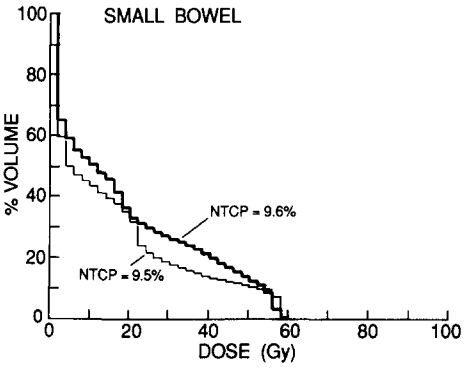


Fig. 4. Dose-volume histograms for the small bowel for two treatment planning techniques for treatment of the rectum.

It is shown in detail by Kutcher and Burman (11) that the effective volume method satisfies all these conditions. It is also demonstrated that, with the exception of certain extreme conditions, organs which obey a power law relationship, the effective volume calculations agree, with linear interpolation when the latter is interpolated over both dose and volume. This seems to imply that NTCP calculations are not overly sensitive to the technique of histogram reduction, which lends some confidence to the procedure. In addition, it was also demonstrated that a bin size of 1 Gy to 2 Gy is adequate for most histogram reduction calculations. All calculations that follow in this report were performed using the effective volume method with a dose bin size of 2 Gy.

In Figure 3 the relative behavior of the histogram reduction technique is demonstrated for several plans, all of which were developed for treatment of the nasopharynx. In all cases 50 Gy is prescribed to microscopic disease, and 70 Gy to the primary. In Figure 3a two dose-volume histograms which almost superimpose on one another are shown for the temporal lobes. As expected, the NTCPs for the two are almost identical. In Figure 3b four dose-volume histograms for the left temporo-mandibular joint for different treatment planning techniques are presented. As can be seen, these dose-volume histograms enclose one another. As the curves move to the left, the calculated values of the NTCPs monotonically decrease as they should. Figure 3c shows two dose-volume histograms for the mandible, one for a standard parallel opposed treatment and one using multiple fields. In this case, the dose-volume histograms cross one another in three places. Since the volume effect (n) is small for this organ, the high dose region plays a dominant role, and the histogram with the largest volume between 65 Gy and 80 Gy yields the largest calculated NTCP. In contradistinction, Figure 3d demonstrates that for a tissue with a large volume effect (n = 0.7), such as the parotid gland, the high dose region (75 Gy and 80 Gy) is not the overriding factor. In this case the histogram with the larger volume at moderate doses has the higher compli-

Table 1. NTCP (%) for treatment planning for carcinoma of the nasopharynx: patient 1

| | A | | B | | C | | D | |
|---|---|---|---|---|---|---|---|---|
| Organ | Standard | Unconstrained | Standard | Unconstrained | Standard | Unconstrained | Standard | Unconstrained |
| Parotid | 91 | 86 | 99 | 49 | 99 | 92 | 98 | 96 |
| Mandible | 399 | 15 | 28 | 22 | 25 | 13 | 10 | 15 |
| TM joint rt. | 379 | 339 | 45 | 0.1 | 48 | 389 | 369 | 0.1 |
| Mid ear rt. | 12 | 12 | 5 | 3 | 32 | 5 | 2 | 0.1 |
| Lens. lt. | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| Optic chiasm | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| Brain stem | 27 | 30 | 0.5 | 3 | 41 | 15 | 3 | 0.5 |
| Spinal cord | 5 | 5 | 3 | 7.5 | 28 | 17 | 7 | 4 |
| Prob. for no complication | 4.6 | 2.6 | 0 | 35 | <1 | 4.3 | 1.3 | 3.3 |
| Preferred plan* | preferred | preferred | | preferred | | preferred | | preferred |
| Ranking† | | 4 | | 1 | | 2 | | 3 |

* Clinically preferred plan in pairwise comparison of 3-D standard and 3-D unconstrained plans.
† Clinical ranking of the four 3-D unconstrained plans, where 1 is best.

cation probability. If the value of the parameter n for the parotid, that is the volume dependence, was similar to that of the mandible, then the difference in volume at moderate doses (below 70 Gy) would not be so important. In another manner, this is demonstrated in Figure 4 for two dose-volume histograms for the small bowel for treatment of the rectum. Histogram B is slightly above histogram A between 56 Gy and 60 Gy and considerably below for most of the histogram below 56 Gy, yet the complication rates for A and B are the same. This follows from the observation that the small bowel has small volume dependence ($n = 0.11$). It is clear from Figures 3c, 3d, and Figure 4 that the volume dependence, n, plays an important role in the calculation of NTCP.

*Accuracy*

Although the behavior of the effective volume method has been shown to be consistent, this does not in the least imply that the calculations are accurate. In fact, the accuracy of the NTCPs depends upon: 1) the quality of the clinical tolerance data; 2) how well Eqs. 1–4 represent the underlying biological processes; 3) how well the histogram reduction technique models the underlying biological processes; and 4) by other factors in the calculation, for example, the use of critical and standard volumes. The first two issues are discussed in the reports by Emami (7) and Burman (3). The critical volumes used in the calculations were obtained as best estimates by the clinical staff on the NCI contract. The critical volume serves the purpose of introducing a volume threshold to the probability calculations; a sharp threshold is not necessarily obtained in Eqs. 1–4. However, in general the complication probability calculations are insensitive to the choice of the critical volume for realistic treatment planning problems. The

Table 2. NTCP (%) for treatment planning for carcinoma of the nasopharynx: patient 2

| | A | | B | | C | | D | |
|---|---|---|---|---|---|---|---|---|
| Organ | Standard | Unconstrained | Standard | Unconstrained | Standard | Unconstrained | Standard | Unconstrained |
| Parotid | 24 | 40 | 83 | 72 | 78 | 58 | 61 | 47 |
| Mandible | 0.5 | 1 | 8 | 0.5 | 3 | 2 | 5 | 1 |
| TM joint rt. | 5 | 5 | 48 | 3 | 42 | 40 | 37 | 11 |
| Mid ear rt. | 17 | 11 | 15 | 13 | 38 | 23 | 11 | 11 |
| Lens. lt. | 82 | 73 | 0 | 0 | 100 | 0.1 | 0.1 | 2 |
| Optic chiasm | 60 | 60 | 4 | 0.1 | 0.1 | 0.5 | 1 | 5 |
| Brain stem | 18 | 12 | 2 | 1 | 27 | 2 | 0.5 | 3 |
| Spinal cord | 3 | 4 | 4 | 4 | 7 | 4 | 2 | 2 |
| Prob. for nocomplication | <1 | <1 | 6.5 | 22 | <1 | 18 | 20 | 29 |
| Preferred plan* | preferred | preferred | | preferred | | preferred | | preferred |
| Ranking† | | 3 | | 1 | | 2 | | 1 |

* Clinically preferred plan in pairwise comparison of 3-D standard and 3-D unconstrained.
† Clinical ranking of the four 3-D unconstrained plans, where 1 is best.

standard volume, on the contrary, may be important. This occurs because, for incompletely scanned patients, the calculated volume is replaced by the standard volume. The criteria used for the calculations in this report was to replace the calculated volume by the standard volume when the former is less than 70% of the latter. The standard volumes were taken either from ICRP (10), or where data was not available, from complete CT scans of other patients. The potential uncertainty in the organ volumes for incompletely scanned patients may affect the accuracy of the NTCP calculations for those tissues with a high volume dependence. However, even if an error is introduced by the use of the wrong standard volume, the relative ranking for a sequence of histograms (for different plans on the same patient) would be unaffected.

One method of assessing the overall accuracy of the complication probability calculations is to review NTCPs for a range of treatment plans, particularly standard or traditional treatment plans where there is some clinical appreciation of the complication rates. Tables 1 and 2 list complication probabilities for two patients for a number of major organs for treatment plans for the nasopharynx for both 3-D standard and 3-D unconstrained beam configurations. As part of a collaborative effort to evaluate 3-D treatment planning, each of four institutions developed one 3-D standard plan (i.e., a plan using the institution's usual beam configuration) and one 3-D unconstrained plan (i.e., a plan using a beam configuration considered most suitable, including noncoplanar beam arrangements) for each of two patients. The prescribed dose, as previously discussed, was 50 Gy for microscopic disease, and 70 Gy for the primary and enlarged nodes.

The accuracy of the probability calculations may be gauged by a review of the 3-D standard beam plans. In general, the values of the NTCPs were considered reasonable for plan evaluation by the clinicians who reviewed the plans, but modifications were definitely considered necessary. For example, the estimated likelihood of mandible necrosis for the first patient is most probably too high, and the brain stem and cord complication probabilities are too high in all cases.

The latter can be explained from these considerations: although $TD_5$ is given as 50 Gy, the choice of the parameter n ($n = 0.07$) is so low that when even a very small volume of the spinal cord receives a high dose, the complication probability is dominated by that dose. Since the target volume for the nasopharynx and enlarged nodes rests precariously close to the cord (in the first patient the target volume extends posterior and lateral to the brain stem and spinal cord), small sectors of the cord, which, however, do not encompass its full cross section, are likely to receive a high dose. An analysis of this type forces a reconsideration of the input data for the spinal cord. In retrospect, the value of n derived from the data of Emami *et al.* (7) may, on one hand, reflect a conservative bias (that is, a decision on the part of clinicians to prevent much reduction in the "tolerance dose" as smaller sections of the cord are involved), but this may also be unrealistic and lead to unnecessarily high calculated complication probabilities. On the other hand, since dose-volume histograms throw away spatial information, it may make them inappropriate for calculating complication probabilities in a situation, such as the one under consideration, where the dominant consideration may be that only a small fraction of the cord cross section is in the high dose region.

Table 3. NTCP (%) for small bowel obstruction and perforation in radiation treatments of rectal cancer

| | Patient 1 | | Patient 2 | |
|---|---|---|---|---|
| Institution | Standard | Unconstrained | Standard | Unconstrained |
| A | 14 | 16 | 10 | 10 |
| B | 10 | 11 | 7 | 11 |
| C | 20 | 16 | 20 | 12 |
| D | 10 | 11 | 8 | 7 |

Complication rates for the small bowel for treatment of the rectum to 56 Gy are given in Table 3. The probabilities appear too high, perhaps by a factor of two. This follows given that the volume of the small bowel was obtained from multiple CT scans, on each of which the area drawn was not the actual bowel wall, but rather the total area of potential bowel positions. This leads to a larger estimated volume for the indicated small bowel, and thus to higher calculated probabilities. It will be necessary to fine tune the input data for each of the organs in order to obtain a self-consistent set of calculated probabilities.

Just as the tolerance doses for uniform irradiation are uncertain, so too are the parameters used for the empirical fitting function. Of particular interest is the uncertainty in the volume effect (parameter n), since this parameter is derived from tolerance data for partial organ irradiation which is probably less dependable than the whole organ tolerance doses. It is shown in Burman (3) that an uncertainty of $\pm 2$ Gy in the tolerance dose for partial organ irradiation leads to an uncertainty in n of approximately $\pm$ 0.08. What is the effect of this uncertainty on calculated NTCPs for nonuniform organ irradiation, and how does this affect the ranking of complication probabilities between different plans?

Presented in Table 4 are NTCPs for 3-D unconstrained plans for the nasopharynx for four institutions. The calculations are given for the standard set of values of n [Burman (3)] and for n increased or decreased by 0.08. The values of the NTCPs in the table reveal that there is a substantial change in the NTCPs for organs with a small volume effect, and almost no change in the NTCPs for organs with a large volume effect. However, it is clear from the table that the ordering of the NTCPs between plans, for a given organ, does not change with n in almost every case. An exception is the brain stem, where there is a

Table 4. NTCP (%) vs. n for 3-D unconstrained treatment plans for the nasopharynx

| Organ | Institution | n | NTCP | n | NTCP | n | NTCP |
|---|---|---|---|---|---|---|---|
| Spinal cord | A | 0.05 | 4 | 0.13 | 1 | 0.001 | 21 |
| | B | | 8 | | 4 | | 21 |
| | C | | 17 | | 5 | | 66 |
| | D | | 4 | | 2 | | 21 |
| Brain stem | A | 0.05 | 30 | 0.13 | 11 | 0.001 | 62 |
| | B | | 3 | | 0.7 | | 11 |
| | C | | 15 | | 3 | | 69 |
| | D | | 0.4 | | 0.1 | | 2 |
| Retina right | A | 0.20 | 0.0 | 0.28 | 0.0 | 0.12 | 0.00 |
| | B | | 0.0 | | 0.0 | | 0.00 |
| | C | | 0.0 | | 0.0 | | 0.00 |
| | D | | 0.1 | | 0.1 | | 0.2 |
| Optic chiasm | A | 0.25 | 2 | 0.33 | 2 | 0.17 | 2 |
| | B | | 0.0 | | 0.0 | | 0.0 |
| | C | | 0.0 | | 0.0 | | 0.0 |
| | D | | 0.1 | | 0.1 | | 0.2 |
| Lens right | A | 0.30 | 0.0 | 0.38 | 0.0 | 0.22 | 0.00 |
| | B | | 0.0 | | 0.0 | | 0.00 |
| | C | | 0.0 | | 0.0 | | 0.00 |
| | D | | 10 | | 8 | | 14 |
| TM joint left | A | 0.07 | 4 | 0.15 | 3 | 0.001 | 19 |
| | B | | 5 | | 1 | | 28 |
| | C | | 46 | | 44 | | 60 |
| | D | | 10 | | 6 | | 38 |
| TM joint right | A | 0.07 | 3 | 0.15 | 3 | 0.001 | 5 |
| | B | | 0.1 | | 0.0 | | 2 |
| | C | | 8 | | 5 | | 28 |
| | D | | 0.1 | | 0.0 | | 1 |
| Parotid glands | A | 0.70 | 8 | 0.78 | 86 | 0.62 | 87 |
| | B | | 49 | | 47 | | 52 |
| | C | | 92 | | 92 | | 92 |
| | D | | 96 | | 96 | | 96 |
| Mandible | A | 0.07 | 15 | 0.15 | 3 | 0.001 | 70 |
| | B | | 22 | | 4 | | 98 |
| | C | | 13 | | 4 | | 60 |
| | D | | 15 | 4 | 69 | | |
| Mid ear left | A | 0.25 | 8 | 0.33 | 8 | 0.17 | 10 |
| | B | | 6 | | 4 | | 10 |
| | C | | 59 | | 59 | | 59 |
| | D | | 3 | | 2 | | 4 |

difference in the NTCPs between plans A and C for n = 0.05 and n = 0.13, but not for n = 0.001. Otherwise, the relative ordering of plans for different organs is insensitive to the accuracy of the volume effect (parameter n) or equivalently the accuracy of the partial organ tolerances (TD(v)) used to derive n. The probability calculations may be inaccurate, but that inaccuracy seems consistent in that plans are still ranked in the same order. Since the values of n used in Table 4 cover a range from the smallest (n = 0.07) to the largest (n = 0.7) volume effect, the conclusions drawn for this nasopharynx case should also apply to other sites.

*Paired organs*

Paired organs present a special problem in using NTCPs for evaluating treatment plans. In general, when dealing with a pair of organs, the complication probability for one of the organs expressing a complication is different than for

both organs. Clearly, it is much worse if both organs of the pair, rather than one, have a high likelihood of complication. It is therefore important to know these likelihoods. For the purposes of the NCI contract we calculated NTCPs for each organ independently in all but one case. For example, the calculated NTCP for the right ear did not depend in the slightest on the radiation injury (or any other injury) to the left ear. The parotid glands, however, were considered by the clinicians to respond to radiation as a single organ; the NTCP was therefore obtained by calculating the effective volume for the combined DVH for both parotid glands. It is likely that some paired organs (for example the kidneys) act neither wholly independently, nor as a single entity, but rather respond to radiation in some intermediate manner. For the purpose of evaluating treatment plans, at least as a first approximation, it is probably adequate to treat paired organs (with their associated end points) as either one unit or as entirely independent units. In the latter case, it is helpful to calculate the related quantities: the probability that both organs realize the complication, and the probability that only one organ realizes the complication. For a further discussion see Munzenrider *et al.* (14).

*Evaluation of treatment plans*

The use of complication probabilities as an aid in evaluating and analyzing treatment plans will be illustrated for two sites, the nasopharynx and the rectum. Tables 1 and 2 list the complication probabilities for a number of organs for the nasopharynx for 3-D standard and 3-D unconstrained plans developed by each of the four institutions. Graphical dose distributions and dose-volume histograms were evaluated by a physician and a physicist. First, a blind comparison of 3-D standard and 3-D unconstrained plans was carried out pairwise for each institution's plans (the institution was also unknown) and, based upon clinical experience, the better of the plans was chosen. This is noted under the appropriate plan in the tables. In addition, the 3-D unconstrained plans were compared blindly, one with another, and the clinical rankings are indicated by numbers from 1 to 4 (1 is preferred). In addition, the probability for the absence of any complication, that is the unweighted product of (1 − NTCP) for each organ, is also given in Tables 1 and 2.

Although there is a good correlation in this example between the clinical rankings of the plans and the likelihood of no complication, this calculated quantity is inadequate for two reasons. First, tumor control rates have not been included, although they are clearly an important consideration. However, since in this case all plans had the same goal for target volume coverage, they are likely to have been comparable in this regard. Secondly, the gravity of each complication probability is not accounted for in the calculations of the likelihood of no complication. For example, although the parotid gland has a high complication rate, its importance is much less than the lower, but nonzero, complication rates for the cord or brain stem. The product of (1 − NTCP) for each organ does not account for this weighting, and because the NTCP for the parotid gland is larger, it tends to dominate the product.

Table 5. NTCP (%) as a function of energy for small bowel obstruction and perforation in radiation treatments of rectal cancer

| Institution | Technique | Energy | Patient 1 | Patient 2 |
|---|---|---|---|---|
| A | 2 Field | 18 MV | 20 | 20 |
| B | 3 Field | 15 MV | 14 | 10 |
| C | 4 Field | 10 MV | 10 | 7 |
| D | 4 Field | 25 MV | 10 | 8 |
| D | 4 Field | 10 MV | 12 | 8 |

The importance of the need for tumor control probabilities is shown in the following example of planning radiation therapy for rectal cancers. In Table 3, complication probabilities are given for small bowel injury for 3-D standard and 3-D unconstrained plans. For this site, the prescription was 46 Gy for microscopic disease—for the large pelvic fields—and 56 Gy for the rectum. It can be seen that the scores for the 3-D standard and 3-D unconstrained plans are uncorrelated with the complication probability for small bowel necrosis. This is because there is only one critical structure, the small bowel, which receives a significant dose and it surrounds the target volume. In designing the best plan, the planners probably chose to optimize the target coverage (maximize tumor control) rather than minimize the small bowel dose. Except for Institution C, where plans went from two fields (3-D standard plan) to multi-field (3-D unconstrained plan), the small bowel complications do not change appreciably between plans. Whether the correct choice between target coverage and bowel dose has been made for these plans cannot be assessed by these calculations alone.

The value of NTCP calculations is also demonstrated in Table 5 for treatment planning for the rectum as a function of technique for 3-D standard beam arrangements. The NTCPs for small bowel obstruction and perforation are given as a function of the large pelvic field arrangement and energy, for each of the two patients. It should be noted that as the number of fields increases, the corresponding small bowel complication probabilities decrease for both patients. In addition, the energy seems to play no role between 10 MV and 25 MV in reducing small bowel complications. This conclusion must be made with the caveat that the boost technique for each institution was different, and the primary target volume is surrounded by small bowel. Therefore, small differences in the high dose treatment volume for the primary tumor bed could affect the complication probability. However, Institution D carried out the same technique at both 10 MV and 25 MV, and there was little change in the NTCPs. The results imply that, at least for the multi-field techniques used, energy plays little or no role between 10 MV and 25 MV in reducing small bowel complications in this site.

It should be evident from the discussion in this report that, even if the probability calculations were known to greater accuracy, such calculations cannot replace inspection of dose distributions and the use of dose-volume histograms. Furthermore, the overall assessment of a plan is hampered by uncertainties as to how each probability should be weighted; however, the importance of complication probability calculations for reviewing individual treatment plans, and analyzing the results of many plans, is also evident. The value of NTCP in condensing data is perhaps best understood by considering the review of multiple plans, such as the four 3-D unconstrained plans presented in Tables 1 and 2. An analysis without NTCPs would entail reviewing eight sets (one for each organ) of dose-volume histograms in each of which there would be four overlying dose-volume histograms. In addition, dose distributions for each of the four plans would need to be reviewed, with a possibility of 4 × 20 such distributions being necessary. Under these conditions, it is very difficult to compare and remember the order and importance of the dose distributions and to assess the possible tissue toxicities for each of the organs. However, Tables 1 and 2 are, in effect, a condensed representation of that analysis. It may be valuable, after reviewing the dose distributions and dose-volume histograms and confirming the clinical reasonableness of the calculations of NTCPs, to use a tabular representation of them. The availability of this "plan synopsis" permits the physician to concentrate on considering the gravity of the possible tissue toxicities and the tradeoffs involved in the decision of which plan to choose. Alternatively, it is possible to review the NTCPs initially in order to see if there are any problem areas. Then the relevant dose distributions and dose-volume histograms could be analyzed to discover the causes of any high NTCPs.

## CONCLUSIONS

A methodology for the calculation of normal tissue complication probabilities has been outlined. It consists of collecting clinical data on tolerance doses for uniform whole and partial organ irradiation, fitting the data to a four parameter function, and transforming a nonuniform histogram into an equivalent uniform histogram. The consistency and reproducibility of the methodology, which depends mainly upon the histogram transformation technique, was demonstrated for the effective volume method. However, the accuracy of the calculations depends both on the quality of the input data and the appropriateness of the model. Because of the large uncertainties in the values of $TD_{50}$ and especially of the volume effect (the parameter n), and because the models are as yet largely untested and of uncertain value, it is advisable not to accept calculated NTCPs as representations, and much less as accurate ones, of the underlying complication probabilities. Rather, the calculations should be used in a relative sense for comparing alternate treatment plans. In this approach, some, but not all, of the errors in the calculations may be factored out. At the same time, it should be possible to fine tune the complication probability calculations in order to obtain more accurate results. This may be accomplished by reviewing calculated NTCPs for various organs for standard treatment plans, and adjusting the input parameters or the model, if necessary.

Complication probability calculations may prove useful for scoring and evaluating treatment plans. On the one hand, a set of NTCPs can be used, along with dose distributions and dose-volume histograms, to evaluate treatment plans. On the other hand, probability calculations can be the basis for numerical scoring functions. In addition to normal tissue complication probabilities, tumor control probabilities are also needed, as well as a representation of the relative weights to be assigned to both the complications and the tumor control probabilities. The proper weighting factors may prove difficult to obtain. Tumor control and normal tissue complication probabilities with appropriate weighting may also prove useful as the basis for calculating an objective function for computerized optimization. This approach should prove valuable for designing compensators in noncoplanar beam arrangements where traditional wedge techniques may not apply, and for automated plan optimization where time-consuming 3-D dose computations are needed.

## REFERENCES


1. Brahme, A.; Agren, A. K. Optimal dose distribution for eradication of heterogeneous tumors. Acta Oncol. 26:1–9; 1987.
2. Brahme, A. Dosimetric precision requirements in radiation therapy. Proceedings of Eighth International Conference on Use of Computers in Radiation Therapy; 1984:269–274.
3. Burman, C.; Kutcher, G. J.; Emami, B.; Goitein, M. Fitting of normal tissue tolerance data to an analytic function. Int. J. Radiat. Oncol. Biol. Phys. 21:123–135; 1991.
4. Bush, M.; Rosenow, U. Dose volume relationships in computer applications. In: Sternick, E., ed. Radiation oncology. Hanover, NH: University Press of New England: 1976: 279–285.
5. Chen, G. J. Y.; Austin-Seymour, M.; Castro, J. C.; Colher, J. M.; Lyman, J. T.; Pithick, S.; Saunders, W. M.; Fink, S. R. Dose volume histograms in treatment planning evaluation of carcinoma of the pancreas. Proceedings of Eighth International Conference on Use of Computers in Radiation Therapy. IEEE ISB NO-8181-0559-6; 1984:264–268.
6. Dritschilo, A.; Chaffer, J. T.; Bloumon, W. D.; Marck, A. The complication probability factor: A method of selection of treatment plans. Br. J. Radiol. 51:370–374; 1978.
7. Emami, B.; Lyman, J.; Brown, A.; Coia, L.; Goitein, M.; Munzenrider, J. E.; Shank, B.; Solin, L. J.; Wesson, M. Tolerance of normal tissue to therapeutic irradiation. Int. J. Radiat. Oncol. Biol. Phys. 21:109–122; 1991.

8. Goitein, M. The probability of controlling an inhomogeneously irradiated tumor. In: Zink, S., ed. Evaluation of treatment planning for particle beam radiotherapy. Bethesda: National Cancer Institute; 1987:1–13.
9. Goitein, M.; Miller, T. Planning proton therapy of the eye. Med. Phys. 10:275–283; 1983.
10. International Commission on Radiological Protection No. 23. Report of the Task Group on Reference Man. Oxford: Pergamon Press; 1975.
11. Kutcher, G. J.; Burman, C. Calculation of complication probability factors for nonuniform normal tissue irradiation: The effective volume method. Int. J. Radiat. Oncol. Biol. Phys.16:1623–1630; 1989.
12. Lyman, J. T. Complication probability as assessed from dose volume histograms. Rad. Res. 104:S-13–S-19; 1985.
13. Lyman, J. T.; Wolbarst, A. B. Optimization of radiation therapy III: A method of assessing complication probabilities from dose volume histograms. Int. J. Radiat. Oncol. Biol. Phys. 13:103–109; 1987.
14. Munzenrider, J. E.; Brown, A. P.; Chu, J. C. H.; Coia, L. R.; Doppke, K. P.; Emami, B.; Kutcher, G. J.; Mohan, R.; Purdy, J. A.; Shank, B.; Simpson, J. R.; Solin, L. J.; Urie, M. M. Numerical scoring of treatment plans. Int. J. Radiat. Oncol. Biol. Phys. 21:147–163; 1991.
15. Mizoe, J.; Von Essen, C. F. Complication probability analysis of late rectal reactions of patients treated with negative ions. SIN (Swiss Institute for Nuclear Research) Medical Newsletter No. 6; 1984.
16. Schultheiss, T. E.; Orton, C. G.; Peck, R. A. Models in radiotherapy: Volume effects. Med. Phys. 10:410–415; 1983.
17. Schultheiss, T. E.; Orton, C. G. Models in radiotherapy: Definition of decision criteria. Med. Phys. 12:183–187; 1985.
18. Shipley, W. U.; Tepper, J. E.; Prout, G. R.; Verhey, L. J.; Mendiondo, O. A.; Goitein, M.; Koehler, A. M.; Suit, H. D. Proton radiation as boost therapy for localized prostatic carcinoma. JAMA 241:1912–1915; 1979.
19. Wolbarst, A. B.; Chin, L. M.; Swensson, G. J. Optimization of radiation therapy: Integral-response of a model biological system. Int. J. Radiat. Oncol. Biol. Phys. 8:1761–1769; 1982.
20. Wolbarst, A. B. Optimization of radiation therapy III: The critical voxel model. Int. J. Radiat. Oncol. Biol. Phys. 10:741–745; 1984.