Big Data Meets Medical Physics Dosimetry
Advanced Computer Integrated Surgery
Seminar for Project IX:
Foundations and Future Challenges

Fumbeya Marungo
Presenter

John Rhee
Team Member

Hilary Paisley
Team Member

Todd McNutt, PhD
Mentor

Scott Robertson, PhD
Mentor

March 13, 2014

1 Introduction

At the core of radiation therapy planning is calculating normal tissue complication probability (NTCP). NTCP it the likelihood that a given radiation treatment plan will result in a particular toxicity (complication).

Project IX's goal is to use big data analytical techniques to improve NTCP estimation in xerostomia (dry mouth) due parotid gland irradiation. Two areas identified in the Project IX's plan for knowledge discovery are understanding the problem domain, and performing data reduction and transformation.

To better understand the problem domain, the Section 2 presents the four papers — Lyman (1985); Emami et al. (1991); Burman et al. (1991); Kutcher & Burman (1989) — that underlie the LKB model; this model is the currently the conventional approach to calculating NTCP.

Section 3 presents a recent approach to data reduction for NTCP calculation using PCA. This paper is presented both as an example of newer approaches, as well as to introduce a technique we intend to use in the project.

The concluding section, Section 4, summarizes the LKB model, and the PCA approach presented.

2 Foundations — Lyman, Kutcher, and Berman (LKB)

2.1 Introduction

In the early mid-1980's to mid-1990's the standard method for assessing NTCP the work from Rubin & Casarett (1972). This method assumed that radiation treatment provided a single uniform dose to an entire volume. For a given toxicity, tolerance doses were provided that represented 5% and 50% five year toxicity probabilities (TD_5 and TD_{50}) based on previous clinical experience.

At that time, advances in radiation therapy and three dimensional treatment planning was rapidly growing. As clinicians were gaining greater flexibility in targeting disease while managing exposure to tissue, the whole volume assumption did not reflect the new clinical reality. There was a growing need for an NTCP approach that modeled partial volume irradiation.

To address this need, Lyman (1985) put forward a model for relating partial irradiation NTCP to whole volume NTCP data. Additionally Emami et al. (1991) provided clinical data for partial volume NTCP in a variety of toxicities. Burman et al. (1991) used the data to parameterize the model for the different toxicities. Kutcher & Burman (1989) then presented a method for interpreting a patient's dose volume histogram using the model.

Combined, these four works constitute the current conventional approach to partial volume NTCP assessment; The approach is now known as the Lyman-Kutcher-Berman (LKB) model. The papers cover a large number of regions and toxicities, this section limits specific descriptions to Project IX's focus — xerostomia due to irradiation of the parotid gland.

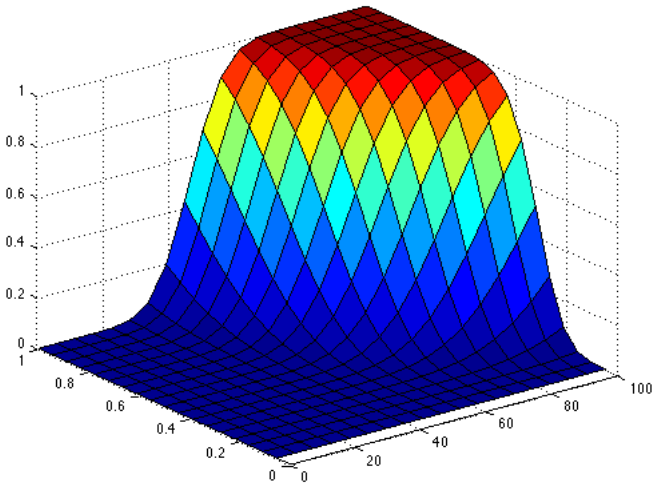


Figure 1: Partial Volume NTCP surface for a parotid using Lyman (1985) model with parameters from Burman et al. (1991). x = Dose in Gy, y = Volume proportion, and z = NTCP

2.2 The model — Lyman (1985)

2.2.1 Description

Lyman (1985) assumes the following:

1. The toxicity (complication) probability $NTCP$ derives from a normally Gaussian Normally distributed value, t, where:
$$t = \frac{D - TD_{50}}{\sigma(V)}, \text{ where } , \sigma(V) = TD_{50}(V)$$

2. The tolerance dose for a partial volume is a power function of the whole volume tolerance dose, that is:
$$TD_{50}(V) = \frac{TD_{50}(1)}{V^n} \tag{1}$$

m, n, and TD_{50} parameters, D is the dose, and V is the portion of the volume that is irradiated. Therefore, after parameterization we can have a partial volume model where $NTCP : D \times V \rightarrow (0, 1)$. Figure 1 is a fully parameterized NTCP surface. m governs the surfaces steepness; n governs the trade-off between volume and dose; and TD_{50} governs the dose and volume levels where NTCP begins to increase.

2.2.2 Summary

By assuming that a partial volume's NTCP is a power function of a whole volume dose, and that toxicity is normally distributed with a mean of the whole dose TD_{50}, Lyman (1985) presents a model applying the previous whole dose uniform volume to calculate partial volume NTCP.

While volume can vary, the model only applies to a uniform dose. In addition it requires parameterization.

2.3 Parameterization — Emami et al. (1991); Burman et al. (1991)

2.3.1 TD_{50} — Emami et al. (1991)

With almost 2,700 citations, Emami et al. (1991) the standard reference source for partial irradiation tolerance doses. The method was to report TD_5 and TD_{50} values for $\frac{1}{3}$, $\frac{2}{3}$, and whole organ volumes based on a combination of sources. Depending on availability results we derived from hard and soft data, as well as experience-based clinician estimates.

In the case of xerostomia, Emami et al. (1991) found precise estimates difficult to determine due to wide range of findings in the literature. The authors concluded that the traditional TD_5 level of $5,000$ cGy and TD_{50} level of and $7,000$ cGy were too high based on clinical experience and a review of the literature. The study settled on an the following: irradiating less then 50% of the organ volume created a near zero NTCP; tolerance dose values for two-third and whole volumes were identical at $TD_5 = 3,200$, $TD_{50} = 4,600$, and $TD_{100} = 5,000$.

2.3.2 m and n — Burman et al. (1991)

Burman et al. (1991) used the Emami et al. (1991) data to fit to Lyman (1985) for a variety of toxicities. The curve fitting was not a purely computational process; the work described performing data fitting "by eye"

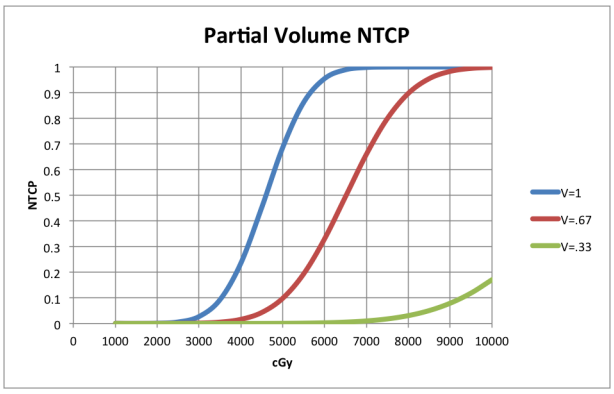


Figure 2: Partial Volume NTCP based on Burman et al. (1991)

For xerostomia, the resulting parameter values were $n = 0.70$, $m = 0.18$, and $TD_{50} = 4600$. Figure 2 presents the resulting cumulative NTCP probability curves for whole, two-thirds, and one-third partial volume irradiation. Note that the whole volume curve closely tracks the TD_5 and TD_{100} values provided in Section 2.3.1

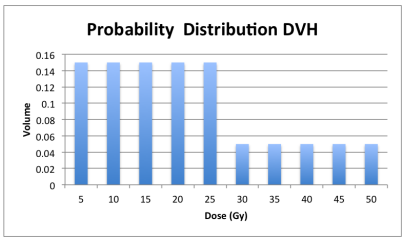
2.4 Modelling Nonuniform Doses — Kutcher & Burman (1989)

2.4.1 The Dose Volume Histogram

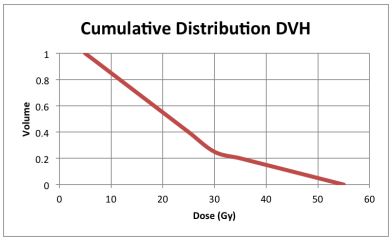
Kutcher & Burman (1989) uses the concept of a dose volume histogram (DVH) to represent nonuniform doses across the volume. The DVH represents the volume proportion that a given dose level. In practice cumulative distribution DVHs are typically used instead probability distribution DVHs. Both are simply different visualizations of the same data.

Figure 3a presents a probability distribution DVH in which 15% of the volume receives a 5 Gy dose, another 15% receives a 10 Gy, and so on until 30 Gy; then 5% receives a 30 Gy dose, and so on until 50 Gy. The corresponding cumulative distribution DVH in Figure 3b as "y portion of the volume has received at least x dose."

The DVH provides a two dimensional representation of partial dosage; spatial location is removed. In order to apply Lyman (1985), the DVH must be reduced to single scalar pair of dose and volume values.

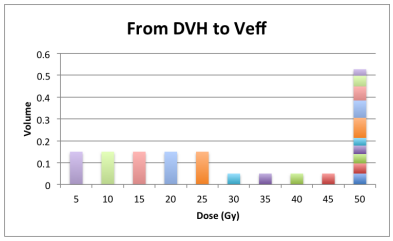


(a)

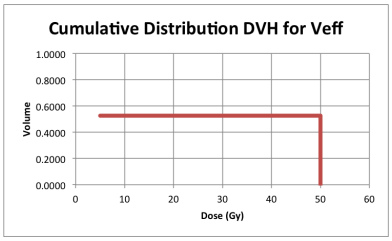


(b)

Figure 3: A probability distribution DVH and the corresponding cumulative distribution DVH.



(a)



(b)

Figure 4: Corresponding histograms for V_{eff}, D_{max}.

2.4.2 Reducing the DVH, Maximum Dose, Effective Volume

To reduce the DVH to a single dose volume pair, Kutcher & Burman (1989) assumes that the power relationship in Equation 1 holds across both dosage and volume variation in the organ. Simply put, a small volume of a large dose is equivalent to a large volume of a small dose. This relationship is represented by:

$$\Delta V_{eff} = \Delta V_i (\frac{D_i}{D_{max}})^{\frac{1}{n}} \tag{2}$$

Equation 2 states that the portion of the volume V_i receiving dose D_i is equivalent to a smaller volume V_{eff} — the effective volume — receiving the maximum treatment dose D_{max}. Therefore the effective dose is:

$$V_{eff} = \sum_i V_i (\frac{D_i}{D_{max}})^{\frac{1}{n}} \tag{3}$$

Using Equation 3, it is possible to transform an arbitrary DVH into a uniform dose, D_{max}, over a smaller effective volume, V_{eff}. 4 visualizes the operation. All of the non-maximum bin volumes are reduced effective volume and placed on the maximum dose (lower doses on top). The result is $V_{eff} = 0.53, D_{max} = 50$. The corresponding cumulative DVH is constant at the effective volume.

2.4.3 Summary

Combined, the four papers present cover the essential elements of LKB. Lyman (1985) uses a power function to relate partial volume to whole volume TD_{50} values. Emami et al. (1991); Burman et al. (1991) provide parameter values. Finally, Kutcher & Burman (1989) provides a method of applying the model to non-uniform doses via DVH.

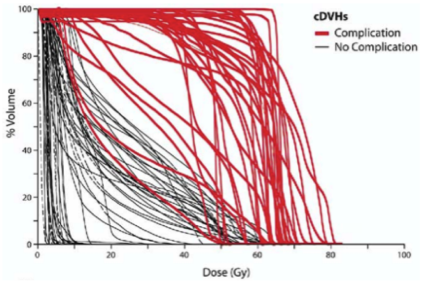
2.5 Discussion

By accounting for nonuniform partial volume irradiation, the LKB model is a marked improvement in NTCP prediction over the preceding tolerance dose approach. However, the model is lacking in robustness relative to the complex nature of treatment planning.

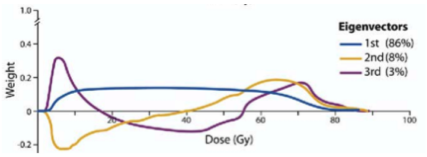
LKB assumes that a high dose treatment in any small area of the organ volume has the same effect on NTCP as a lower dose spread over a larger volume. The use of DVH removes dose location as a feature. Another concern is the standard parameterization for LKM. Both Emami et al. (1991) and Burman et al. (1991) note the limited amount of hard data available. This lack of data implies a level of uncertainty in the resulting parameters.

On a more general level, actual NTCP is a multi-factor problem. Features that impact outcome are not limited to treatment decisions; prescribed medications, family history, and social factors may all inform on toxicity risk.

Many of LKB's limitations were acknowledged by the authors. However, creating richer models require accumulated clinical experience, substantial data, and processing power. All three were generally unavailable more than two decades ago as three dimensional treatment planning was beginning.



(a) Set of DVHs



(b) DVH PCA eigenvectors

Figure 5: Parotid data, copied from Dawson et al. (2005)

3 A Recent NTCP Analysis Example — Dawson et al. (2005)

3.1 Introduction

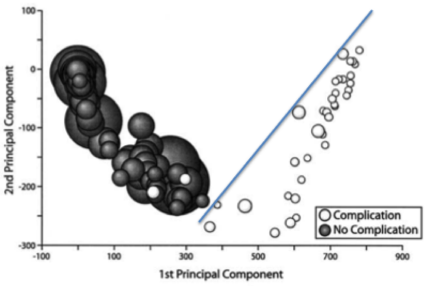
As we discuss in Section 2.5, LKB does not account for dose location. This is an area that the Project IX team wishes to explore. One approach is creating to create 125 subsections in each parotid glad by evenly dividing the regions of interest into fifths along each dimension and creating DVHs for each subregion.

Using Kutcher & Burman (1989)'s effective volume presents the problem of generating a value in each subregion for the parameter n in Equation 1. In addition, there is the more general question of whether better alternatives exist to reducing the DVH to a single effective volume, and the maximum dose.

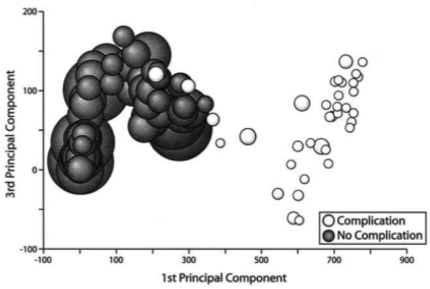
3.2 Description

Dawson et al. (2005) reports on performing principal component analysis (PCA) on parotid and liver DVHs. The parotid dataset contains 39 patients (see Figure 5). The first principal component (PC) captures the middle of the dosage range; the second represents increases in the high dosage region and accompanied by decreases in the low dosage region. The third component represents an increase in both the extremes of the dosage range, with a decrease in the middle range.

In the parotid case, PC1 and PC2 combined capture 94% of the variance. Dawson et al. (2005) reports that the relationship between PC1 and xerostomia incidence is statistically significant. In addition, Figure



(a) First vs. second principal components provides a linear seperation of all but two complication cases (line added).



(b) Calculating effective volume, copied from Kutcher et al. (1991)

Figure 6: PCA plots with bubble size indicating saliva flow rate, copied from Dawson et al. (2005)

6 visually indicates that including PC2 may provide additional information gain.

3.3 Discussion

Dawson et al. (2005) demonstrates that potential may lie in creating more robust NTCP prediction models using sophisticated data analysis techniques. The two false negatives that are identifiable in Figure 6a have DVH curves that are indistinguishable from normal curves. However, there are many DVHs with complications that have low maximum doses and complication free DVHs with maximum doses on the high end (see Figure 5a). Both cases my lead to less accurate NTCP predictions using LKB.

4 Conclusion

As part of Project IX's plan requirements to understand the problem domain, this paper present the foundational papers on the LKB model. This model is the current standard for three dimensional NTCP assessment.

LKB proceeded from a previous tolerance dose based model that assumed a single fixed dosage applied to an entire organ volume. To address the need to assess NTCP of a partial volume irradiation, Lyman (1985) developed a power law based relationship between whole and partial volumes. Burman et al. (1991) used data from Emami et al. (1991) to parameterize the model. Kutcher & Burman (1989) presents a method for adapting nonuniform dosage distribution to the model using the DVH and the assumption that

effective volumes of the maximum dose is equivalent to larger actual volumes of smaller doses. Kutcher & Burman (1989) then uses this assumption to transform a DVH into a single scaler pair of effective volume and maximum dose.

The LKM model's robustness is limited due to its underlying assumptions. The DVH approach does not include dose location as a factor in calculating NTCP. As part of Project IX involves including spatial dose data in computing NTCP by calculating individual DVH's for 125 individual subregions as well as the entire volume. Based on Dawson et al. (2005) we propose using PCA to create two PCs per DVH.

References

Bentzen, S. M., Constine, L. S., Deasy, J. O., Eisbrunch, A., Jackson, A., Marks, L. B., Haken, R. K. T., & Yorke, E. D. (2010). Quantitative analysis of normal tissue effects in the clinic (QUANTEC): An introduction to the scientific issues. *International Journal of Radiation Oncology Biology Physics*, *76*(3), S3–S9.

Burman, C., Kutcher, G., Emami, B., & Goitein, M. (1991). Fitting of normal tissue tolerance data to an analytic function. *International Journal of Radiation Oncology* Biology* Physics*, *21*(1), 123–135.

Dawson, L. A., Biersack, M., Lockwood, G., Eisbruch, A., Lawrence, T. S., & Ten Haken, R. K. (2005). Use of principal component analysis to evaluate the partial organ tolerance of normal tissues to radiation. *International Journal of Radiation Oncology* Biology* Physics*, *62*(3), 829–837.

Emami, B., Lyman, J., Brown, A., Cola, L., Goitein, M., Munzenrider, J., Shank, B., Solin, L., & Wesson, M. (1991). Tolerance of normal tissue to therapeutic irradiation. *International Journal of Radiation Oncology* Biology* Physics*, *21*(1), 109–122.

Kutcher, G., Burman, C., Brewster, L., Goitein, M., & Mohan, R. (1991). Histogram reduction method for calculating complication probabilities for three-dimensional treatment planning evaluations. *International Journal of Radiation Oncology* Biology* Physics*, *21*(1), 137–146.

Kutcher, G. J., & Burman, C. (1989). Calculation of complication probability factors for non-uniform normal tissue irradiation: The effective volume method gerald. *International Journal of Radiation Oncology* Biology* Physics*, *16*(6), 1623–1630.

Lyman, J. T. (1985). Complication probability as assessed from dose-volume histograms. *Radiation Research*, *104*(2s), S13–S19.

Rubin, P., & Casarett, G. (1972). Direction for clinical radiation pathology. the tolerance dose. Tech. rep., Univ. of Rochester, NY.