Parameters for the Lyman Kutcher Burman (LKB) model of Normal Tissue Complication Probability (NTCP) for specific rectal complications observed in clinical practise.

Authors

Sarah L Gulliford PhD[1*], Mike Partridge PhD[1], Matthew R Sydes CStat[2], Steve Webb DSc[1], Phil M Evans DPhil[1], David P Dearnaley FRCR[3]

*** Corresponding Author**

Sarah L Gulliford PhD

Telephone +442086613320

Fax +442086433812

Email sarahg@icr.ac.uk

1) Joint Department of Physics, Institute of Cancer Research and Royal Marsden NHS Foundation Trust, Sutton and London, UK

2) Cancer Group, MRC Clinical trials unit, London, UK

3) Department of Academic Urology, Institute of Cancer Research and Royal Marsden Hospitals, Sutton and London, UK

Shortened Title: LKB Parameters for specific rectal toxicities.

Conflict of Interests

Matthew Sydes is employed by the trial sponsor (Medical Research Council UK).

Abstract

Background and Purpose:

The Normal Tissue Complication Probability (NTCP) for rectum is usually defined for late rectal bleeding. This study calculates NTCP parameter values for additional rectal toxicity endpoints observed in clinical practise.

Materials & Methods:

388 patients from the multicentre MRC-RT01 prostate conformal radiotherapy trial were used to derive independent Lyman Kutcher Burman model (LKB) parameters for 5 late rectal toxicity endpoints: rectal bleeding, proctitis, stool frequency, loose stools and rectal urgency. The parameters were derived using maximum likelihood estimation. Bootstrap and leave-one-out methods were employed to test the generalisability of the results for use in a general population.

Results:

A consistent pattern of increasing value of TD50(1) for Grade 2 toxicity only compared to Grade 1 & 2 toxicity was observed for all endpoints. Parameter values varied between endpoints (particularly for the volume parameter n). TD50(1), m and n were 68.5 Gy (95% CI)(66.8-70.8), 0.15 (0.13-0.17) and 0.13 (0.10-0.17) respectively for G2 rectal bleeding. Bootstrap and leave-one-out results showed that the rectal bleeding and proctitis parameter fits were extremely robust.

Conclusions:

The variation between the values derived for different endpoints may indicate different patho-physiological responses of the rectum to radiation. Therefore different parameter sets would be required to predict specific rectal toxicity endpoints.

Keywords: Prostate radiotherapy, NTCP, LKB, Rectal Complications

Introduction

The Lyman Kutcher Burman (LKB) model is probably the most well known method for predicting Normal Tissue Complication Probability (NTCP) for a radiotherapy treatment plan. The model was developed by Lyman[1] for heavy charged particle beams where partial volumes of homogenous dose could be achieved and adapted for conventional radiotherapy through the histogram reduction work of Kutcher[2] and parameter values of Emami[3] and Burman[4]. There are three parameters in the LKB model. TD50(1) represents the dose for a homogenous dose distribution to an organ at which 50% of patients are likely to experience a defined toxicity within 5 years. m is related to the standard deviation of TD50(1) and describes the steepness of the dose-response curve and n indicates the volume effect of the organ being assessed. The recent publication of the Quantec Report[5] has brought together much of the literature and experience of normal tissue toxicity. A dedicated article on radiation induced rectal injury[6] includes a summary of published parameter values for the LKB model. Amongst the publications included are data from MD Anderson[7] where the LKB model was used to predict rectal bleeding from a cohort of 128 patients of which 29 reported rectal bleeding ≥ RTOG Grade 2. Values of TD50=53.6 Gy, m=0.156 and n=3.91 were obtained for a calculation based on effective dose (n is the reciprocal of the calculation of effective volume for small values only). Sohn et al[8] obtained the LKB parameters using equivalent uniform dose (EUD) of TD50(1)=78.4 Gy m=0.108 and a=11.9 again for Grade 2 or greater rectal bleeding (51/319 patients). Rancati et al[9] found values of TD50(1)=81.9 Gy, m=0.19 n=0.23 for a modified Grade2/3 for rectal bleeding observed in 38/547 patients. The recent publication of a mixture Lyman model using data from RT0G 94-06[10,11] reported parameter values of TD50(1)= 76.1Gy m=0.146 and n=0.077 in a cohort where 148/1010 reported ≥RTOG Grade 2. A meta-analysis of the these papers [7-9,11] in the Quantec review[6], derived parameter values of TD50(1)=76.9 Gy, m=0.13 95% CI (0.10-0.17) and n=0.09 95% CI (0.04-0.14) for Grade2 rectal bleeding which are similar to the original Emami values for rectum[3] (TD50(1)=80Gy, m=0.15, n=0.12).

Although rectal bleeding is the most frequently reported rectal toxicity, it is not necessarily the most prevalent and a number of other endpoints are known to concern patients. It would therefore be useful to be able to predict a range of toxicities that are commonly observed including rectal bleeding and quality of life related issues such as continence and bowel habits. Less data are available for these endpoints. However, Peeters et al[12] fitted the LKB model for rectal bleeding,

stool frequency and fecal incontinence (anal canal only) with significant differences in the estimated parameter values compared to rectal bleeding. The incidence of complication was 5%, 6% and 7% respectively. The values for rectal bleeding were TD50(1)=81 Gy, m=0.4, n =0.13, stool frequency TD50(1)=84 m=0.24 and n =0.39 and faecal incontinence TD50(1)=105 Gy, m=0.46 and n=7.48

The ability to fit the parameters of the model hinges on the availability of detailed dosimetric information and corresponding clinical follow up data with accurate reporting of toxicity. This study presents results of maximum likelihood estimation to derive parameter values for the LKB model fitted with data from five specific rectal toxicity endpoints. In addition, the robustness of the parameters when generalised to a wider population was also investigated.

Methods and Materials

Rectal NTCPs were calculated for 388 patients treated with prostate radiotherapy as part of the MRC RT01 trial (ISRCTN 47772397)[13,14]. The trial randomised the radiotherapy prescription to 64Gy or 74Gy. All patients received 64Gy using either a 3 or 4 field beam arrangement. Those patients randomised to 74Gy receive an extra 10Gy boost to the prostate only delivered using either a 4 or 6 field beam arrangement. Fields were conformed using either low melting point alloy blocks or MLCs. The field arrangement and conformallity method were chosen by the participating centres. The treatment planning CT was performed with an empty rectum. The rectum was outlined from the anus taken at the level of the ischial tuberosities or 1cm below the PTV whichever was more inferior to the recto-sigmoid junction and considered as a solid organ (including rectal filling). All rectal contours were reviewed by a single observer. The LKB model was fitted to five different toxicity endpoints; these were rectal bleeding (RMH toxicity score), proctitis (RTOG), stool frequency (LENT/SOMA) which were clinician reported and loose stools (UCLA-PCI) and rectal urgency (UCLA-PCI) which were patient reported. As previously reported [15] the grading schemes were unified to a common grading scheme where the grading was designed to consider the impact on patients. In each case the fit was made separately for Grade 0 vs. Grade 1&2 (G1&2) i.e. none vs. mild/moderate/severe and Grade 0&1 vs. Grade 2 (G2) i.e. none/mild vs. moderate/severe. In each case the maximum grade recorded over the entire length of follow-up was used. Patients who

experienced a defined endpoint prior to treatment were excluded from the parameter fitting for that endpoint only. Details of the number of patients included for each endpoint and the grade of toxicity reported are presented in table 1.

The LKB model was taken from the original publications

$$NTCP = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-t^2/2} dt \quad (1)$$

where

$$u = \frac{D - TD50(V)}{m \times TD50(V)} \quad (2)$$

$$TD50(V) = TD50(1) / V^n \quad (3)$$

TD50(1) is the tolerance dose for a homogenous dose distribution to an organ at which 50% of patients are likely to experience a defined toxicity within 5 years, TD50(V) is the tolerance dose for a partial volume V. The parameter m multiplied by TD50(V) approximates the standard deviation of volume V and n indicates the volume effect of the organ being assessed. n = 0 indicates a completely serial structure where the maximum dose dominates outcome and n=1 indicates a parallel structure where the mean dose is related to outcome. D is the maximum dose of the DVH to ensure V<1[16]. Histogram reduction was performed to calculate the effective volume V according to the method described in Kutcher et al[2]

$$V = \Sigma_i (Di/D)^{1/n} \Delta V_i \quad (4)$$

where Di is the dose defined for each bin in a differential dose volume histogram and D is the maximum dose to the organ. ΔV_i is the volume in a specific dose bin i. The dvh data were available

in bins which were 1% of the prescription dose. The bin size therefore varied depending on the arm of the trial which the patient was randomised to.

Maximum likelihood estimation (MLE)[7,9,12,17] was employed to find the best fit values of the parameters TD50(1), m and n of the NTCP model for the known binary outcomes y(i) of the available data by maximising the natural log of the likelihood (LLH) that the fitted model describes the data correctly. Fits were made separately for the 5 specific rectal toxicities and the 2 Grades of complication, mild (G1&2) and moderate/severe (G2 only).

$$LLH(TD50(1), m, n) = \sum_{y(i)=1} \ln(NTCP(TD50(1), m, n)) + \sum_{y(i)=0} \ln(1 - NTCP(TD50(1), m, n))) \quad (5)$$

Confidence intervals for the optimal fit parameters were obtained using profile likelihood estimation[18,19].

It is acknowledged that the parameter fits are specific to the data used for fitting therefore to generalise the fit to a wider population a bootstrap method was employed[20].1000 different cohorts of 388 patients were generated from the patient data using sampling with replacement. The LKB model was refitted for the 5 endpoints for each of the 1000 sampled populations using maximum likelihood estimation. In addition the effect of removing individual patients was investigated using a leave-one-out approach where maximum likelihood estimation was repeated with each individual case omitted iteratively.

Finally since NTCP is often used for ranking, the non parametric Mann Whitney U test was calculated using SPSS (SPSS Inc Chicago Illinois vs 15) to test for a statistically significant difference in the NTCP values of the group of patients who reported a specific endpoint compared to those patients who did not.

Results

The maximum likelihood estimation of the LKB parameters obtained for the entire patient cohort for each endpoint and toxicity level considered are shown in table 2 along with associated confidence intervals. The parameter values for m fitted to G1&2 toxicity indicate large variability in the patient data. However the values of m are much smaller for the LKB fits to G2. The TD50(1) values for G2 are all higher than the corresponding fits for G1&2 but in most cases still lower than the original Emami value of 80Gy. The values of n are reasonably consistent between the two fits for each endpoint but vary between endpoints with loose stools and rectal urgency both having a much less serial response than rectal bleeding and proctitis, where the results are a little higher than the Quantec value of 0.09. The results for G1&2 stool frequency have a large value for m with a wide confidence interval which suggests a poor fit to the model. It was not possible to derive a maximum likelihood estimation for Grade 2 stool frequency as the value of log likelihood was still tending to a maximum at very high values of TD50(1). This is demonstrated in figure 1 which compares how the log likelihood varied over a range of TD50(1) and n values for G2 rectal bleeding and stool frequency for the optimal value of m. It is easy to observe the region of best fit for the rectal bleeding data however for stool frequency, TD50(1) tends to a maximum around 300Gy but a range of nearly 50Gy shows comparatively little variation. It is also interesting to note the shape of the rectal bleeding distribution; a swathe of results which indicate results close to the best log-likelihood value highlight interrelationship in the model between the TD50(1) and n parameters.

Table 3 details the results for 1000 bootstrap cohorts to test the idea that a different selection of cases might influence the parameters. The results show that the mean values of TD50(1) and m are reasonably close to the exact fit to the patient cohort for all 5 endpoints. For rectal bleeding and proctitis this is also true for the volume parameter n. However, with the exception of G1&2 rectal urgency, the values for the other endpoints show much larger values of n and relatively large standard deviations. It was possible to derive parameters for G 2 stool frequency using the bootstrap method, indicating that the model fit is very sensitive to the selection of cases. The large standard deviations in the result represent a large variation in the optimal parameter fits. This implies that the incidence of stool frequency (as it is described and reported in this dataset) is likely to have a weak dependence on dosimetry. For the other 4 endpoints, further analysis was performed to investigate the effect on the LKB parameters when each individual case was left out of the MLE. The parameter values obtained are summarised in figure 2. The fitted parameter values for rectal bleeding and proctitis remained consistent, however for the other endpoints, the

parameter values varied as different cases were removed. The exact NTCP parameter values, derived using the entire cohort of patients, were able to discriminate between the cohort of patients with and without toxicity with a statistical significance at the level $p<0.01$ for rectal bleeding, proctitis, loose stools and rectal urgency (for both G 1&2 and G 2 only).

Discussion

The maximum likelihood estimation of TD50(1) for rectal bleeding for both G1&2 and G2 only are significantly less than the Quantec value of 76.9Gy. The RT01 trial was conducted in an era when conformal radiotherapy was first being implemented and the only rectal constraint in the trial protocol was that the maximum dose to the rectum should not exceed the prescribed dose. As such the volume of rectum at a range of dose levels may be larger compared to more contemporary studies which included constraints. Variations in the grading and reporting of late toxicities may also account for the difference between the results. Most of the papers in the Quantec analysis reported on either RTOG rectal bleeding or global score which are perhaps slightly more severe endpoints than the Grade 2 toxicity fitted here. The advantage of using milder toxicity is that it is more relevant to the patient population. It can be argued, should the model be for the few worst cases or more general experience?

The fitted values for rectal bleeding of m and n are in better agreement with the Quantec values of 0.13 and 0.09 respectively. All 3 parameters have narrow confidence intervals and have been shown to be resilient to the effect of variations in fitting data, as shown using the bootstrap and leave one out methods. This confirms a good fit to the data and the serial nature of the response of the rectum to radiation which results in rectal bleeding and proctitis.

The value of TD50(1) steps up as the severity of complication increases for each endpoint. Parameter values for TD50(1) and m for loose stools and rectal urgency are not dissimilar to those obtained for rectal bleeding and proctitis for the exact fit of the patient data. However the fitted value for n is consistently larger for the other endpoints and is more marked for rectal urgency. The

bootstrap results are similarly consistent for n although the standard deviations are generally larger for these endpoints. An increase in n equates to a fit that takes in more of the DVH than just the highest doses perhaps representing composite values from different patient groups with different patho-physiological responses to radiation. For example reduced absorption of the rectal mucosa or neurovascular damage impairing musculature [Fiorino R&O Pelvis review ref] Figure 1 is a reminder that variation in n is also related to variation in TD50(1) we have previously postulated that if large areas of rectum receive intermediate doses this may inhibit repair to surrounding high dose regions [RT01 constraints paper]

The only endpoint other than rectal bleeding that can be compared to other published literature is stool frequency, however large uncertainties are demonstrated in the results reported here, this may be in part due to the subjective nature in reporting stool frequency and confounding causes other than radiotherapy. For comparison the parameters obtained by Peeters et al were for >6 times per day compared to Grade 2 which was >5 times per day. The parameters obtained were TD50(1)=84 Gy m=0.24 and n=0.39. These results are significantly different to the bootstrap results obtained for G2 stool frequency however in both cases the value of n indicates a less serial response.

The results presented here emphasise the benefit of bootstrap and leave-one-out analysis where the effect on the wider population can be tested. Since in general the number of positive events is small for each end point (especially at G2) it is interesting to observe through the leave-one-out analysis how much influence a single case may have. The reasons for this are many and varied. Issues related to the accuracy in reporting toxicity have the potential to skew data as do confounding causes of late effects which are plausible for most of the endpoints fitted here. Although the 388 patients included in the study was less than half of those enrolled in the trial there was no obvious bias in the available data, it is expected that if more patients from the cohort had been available the parameter fits would have had smaller confidence intervals.

The Lyman model has traditionally been coupled with a histogram reduction method to account for the heterogeneous dose distribution received by normal tissues. It was developed in an environment where partial volumes of homogenous dose were more prevalent and applied in an era

where the ability to spare normal tissues was limited. The advances in delivery techniques allow us to create sculpted dose distributions using inverse optimisation. The dose-volume histogram reduction methods are generally insufficient to fully characterise these dose-distributions since they take all the available dosimetric information and condense it to a single value which may not be representative of the response of the rectum to the dose distribution. A correction for dose per fraction (on a bin by bin basis) was not included the histogram reduction model since the recent publication by Tucker et al[21] demonstrated that for fractionation near to 2Gy there was no significant difference in the parameters derived for late rectal toxicity using the LKB model. In addition to the model uncertainties, variation between the planned dose-distribution to the rectum and the treated dose-distribution introduce uncertainty in to the dosimetric data. However in a reasonably large cohort including patients from a large number of centres there is unlikely to be a systematic error and random errors are likely to reduce the statistical power of results rather than skew them.

It is reassuring that, tested on a set of varied conformal dose distributions, the LKB model is still applicable for proctitis and rectal bleeding. However for other endpoints, the poorer fits may be due to variability in reporting of toxicity or perhaps an insufficient characterisation of the dose distribution to the rectum.

Conclusion

It has been shown that the dose-volume response of the rectum is different for different endpoints and that quality of life related issues such as stool frequency, rectal urgency and loose stools may not be fully predicted by the classic n value of the LKB model for rectum. In addition, the degeneracy of the model can lead to parameter fits being influenced by single cases. This should be fully explored when deriving parameter sets for clinical use.

Figure Legends

Figure 1. Maximum Likelihood Estimation values plotted as a function of LKB parameters TD50(1) and n with fixed value of m fitted to a) rectal bleeding and b) stool frequency. Both plots use the best fit value of m which were 0.15 and 0.6 respectively. The effect of m is illustrated with the significant difference in the range of TD50(1) values displayed on the x axis. The relatively large value of m observed for stool frequency indicates that clinical data fits poorly to the LKB model resulting in similar MLE estimates for a wide range of TD50(1) and n parameters. The deceptively better MLE value for stool frequency results from the small number of Grade 2 cases for stool frequency 29 (vs 54 for rectal bleeding.)

Figure 2. Distribution of leave-one-out results obtained using Maximum Likelihood Estimation to fit parameters TD50(1) (a), m(b) and n (c) to the LKB model for specific Grade 2 rectal toxicity endpoints.

Acknowledgments

The trial was supported and coordinated by the UK Medical Research Council, Institute of Cancer Research and the Cancer Research UK (CR-UK) Section of Radiotherapy Grant number C46/A2131. We acknowledge NHS funding to the NIHR Biomedical Research Centre.

[1] Lyman JT. Complication probability as assessed from dose-volume histograms. Radiat.Res.Suppl 1985;8:S13-S19.

[2] Kutcher GJ, Burman C, Brewster L, Goitein M Mohan R. Histogram reduction method for calculating complication probabilities for three-dimensional treatment planning evaluations. Int.J.Radiat.Oncol.Biol.Phys. 1991;21:137-146.

[3] Emami B, Lyman J, Brown A, et al. Tolerance of normal tissue to therapeutic irradiation. Int.J.Radiat.Oncol.Biol.Phys. 1991;21:109-122.

[4] Burman C, Kutcher GJ, Emami B Goitein M. Fitting of normal tissue tolerance data to an analytic function. Int.J.Radiat.Oncol.Biol.Phys. 1991;21:123-135.

[5] Marks LB, Ten Haken RK Martel MK. Guest editor's introduction to QUANTEC: a users guide. Int.J.Radiat.Oncol.Biol.Phys. 2010;76:S1-S2.

[6] Michalski JM, Gay H, Jackson A, Tucker SL Deasy JO. Radiation dose-volume effects in radiation-induced rectal injury. Int.J.Radiat.Oncol.Biol.Phys. 2010;76:S123-S129.

[7] Cheung R, Tucker SL, Ye JS, et al. Characterization of rectal normal tissue complication probability after high-dose external beam radiotherapy for prostate cancer. Int.J.Radiat.Oncol.Biol.Phys. 2004;58:1513-1519.

[8] Sohn M, Yan D, Liang J, Meldolesi E, Vargas C Alber M. 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.Radiat.Oncol.Biol.Phys. 2007;67:1066-1073.

[9] Rancati T, Fiorino C, Gagliardi G, et al. Fitting late rectal bleeding data using different NTCP models: results from an Italian multi-centric study (AIROPROS0101). Radiother.Oncol. 2004;73:21-32.

[10] Tucker SL, Dong L, Bosch WR, et al. Late rectal toxicity on RTOG 94-06: analysis using a mixture Lyman model. Int.J.Radiat.Oncol.Biol.Phys. 2010;78:1253-1260.

[11] Tucker SL, Dong L, Bosch WR, et al. Fit of a generalized Lyman Normal-Tissue Complication Probability (NTCP) model to grade >= 2 late rectal toxicity data from patients treated on protocol RTOG 94-06. International Journal of Radiation Oncology Biology Physics 2007;69:S8-S9.

[12] Peeters ST, Hoogeman MS, Heemsbergen WD, Hart AA, Koper PC Lebesque JV. Rectal bleeding, fecal incontinence, and high stool frequency after conformal radiotherapy for prostate cancer: Normal tissue complication probability modeling. Int.J.Radiat.Oncol.Biol.Phys. 2006.

[13] Sydes MR, Stephens RJ, Moore AR, et al. Implementing the UK Medical Research Council (MRC) RT01 trial (ISRCTN 47772397): methods and practicalities of a randomised controlled trial of conformal radiotherapy in men with localised prostate cancer. Radiother.Oncol. 2004;72:199-211.

[14] Dearnaley DP, Sydes MR, Graham JD, et al. Escalated-dose versus standard-dose conformal radiotherapy in prostate cancer: first results from the MRC RT01 randomised controlled trial. Lancet Oncol. 2007;8:475-487.

[15] Gulliford SL, Foo K, Morgan RC, et al. Dose-Volume Constraints to Reduce Rectal Side Effects From Prostate Radiotherapy: Evidence from MRC RT01 Trial ISRCTN 47772397. Int.J.Radiat.Oncol.Biol.Phys. 2009.

[16] Luxton G, Keall PJ King CR. A new formula for normal tissue complication probability (NTCP) as a function of equivalent uniform dose (EUD). Phys Med Biol 2008;53:23-36.

[17] Jackson A, Ten Haken RK, Robertson JM, Kessler ML, Kutcher GJ Lawrence TS. Analysis of clinical complication data for radiation hepatitis using a parallel architecture model. Int.J.Radiat.Oncol.Biol.Phys. 1995;31:883-891.

[18] Stryhn H Christensen J. Confidence intervals by the profile likelihood method, with applications in veterinary epidemiology. Proceedings of ISVEE X 2003.

[19] Venzon DJ. A Method for Computing Profile-Likelihood-Based Confidence Intervals. Appl.Statist. 1988;37 87-94.

[20] Deasy JO, Chao KS Markman J. Uncertainties in model-based outcome predictions for treatment planning. Int.J.Radiat.Oncol.Biol.Phys. 2001;51:1389-1399.

[21] Tucker SL, Thames HD, Michalski JM, et al. Estimation of alpha/beta for Late Rectal Toxicity Based on RTOG 94-06. Int J Radiat Oncol Biol Phys 2011;81:600-605.