Complication probability for radiation pneumonitis (RP) after stereotactic body radiotherapy (SBRT)

Jimm Grimm\textsuperscript{a,}, David Palm\textsuperscript{b}, Suresh Senan\textsuperscript{c}, Jinyu Xue\textsuperscript{d}

\textsuperscript{a}Holy Redeemer Hospital, Meadowbrook, PA
\textsuperscript{b}London Regional Cancer Program, London, ON, Canada
\textsuperscript{c}VU University Medical Center, Amsterdam, The Netherlands
\textsuperscript{d}Cooper University Hospital, Camden, NJ

*Correspondence to: Jimm Grimm, Ph.D., DABR, Holy Redeemer Hospital, Bott Cancer Center, 1648 Huntingdon Pike, Meadowbrook, Pa 19046, USA.

Conflict of Interests Notification:

None of the authors has received any funding for this research. Dr. Senan’s department has received research funding from Varian Medical Systems, and he has received speaker’s honoraria from Varian, but not directly related to the analysis in this manuscript. Dr. Palma is the recipient of a Clinician-Scientist grant from the Ontario Institute for Cancer Research, funded through the Government of Ontario. Dr. Grimm developed and holds intellectual property rights to the DVH Evaluator software tool which is an FDA-cleared product in commercial use, and which has been used for this analysis.

NOTICE: this is the authors' version of a work that was accepted for publication in the Journal of Radiosurgery & SBRT, www.OldCityPublishing.com/JRSBRT/JRSBRT.html. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in PUBLICATION, [VOL#, ISSUE#, (DATE)] DOI#
Abstract:
Purpose/Objective:
To determine clinically relevant SBRT/stereotactic ablative radiotherapy (SABR) dose tolerance limits for RP based on statistical analysis of outcomes data.

Materials and Methods:
Eighteen consecutive patients who were treated using volumetric modulated arc therapy (RapidArc) for lung tumors exceeding 80cc were assessed. Clinical outcomes have been published elsewhere, and here we present a normal tissue complication probability (NTCP) analysis. The dose volume histogram (DVH) reduction techniques of total lung V20Gy, V15Gy, V10Gy, V5Gy and mean lung dose (MLD) were each analyzed, as well as ipsilateral lung V5Gy and contralateral lung V5Gy, using the DVH Evaluator software tool. The framework of the Lyman Model was used except that each DVH reduction method was analyzed independently instead of using the power-law relationship for volume dependence. Model parameters were fitted using Maximum Likelihood.

Results:
RP was reported in 5 patients (CTC Grade 2 in 3, and Grade 3 in 2). Total lung V5Gy and contralateral lung V5Gy were the best predictors of RP (p < 0.0001 for both). For V5Gy, the 10% risk level for Grade 2-3 RP was 27.9% for total lung and 21.8% for contralateral lung. For V20Gy, the 25% risk level is 10.5% of total lung.

Conclusions:
Analysis of RP endpoints has identified total lung V5Gy and contralateral lung V5Gy as the best predictors of RP following SBRT when delivered with RapidArc. These findings are based on limited clinical data, and longer follow-up in larger patient cohorts is required in order to determine more accurate dose tolerance limits.

Key words: Dose tolerance limits; Adverse events; Stereotactic body radiotherapy; Pneumonitis; DVH Evaluation

[See the full version of this article at www.OldCityPublishing.com/JRSBRT/JRSBRT.html]
Introduction

Reliable model-based normal tissue complication probability (NTCP) estimates for V20Gy or V5Gy are unavailable for stereotactic body radiation therapy (SBRT). A recent publication has given detailed dose-volume outcomes data for SBRT treatments of large stage I-II lung tumors in eighteen consecutive patients [1], but did not determine complication probabilities associated with the various dosimetric parameters. Since SBRT is increasingly being applied for the treatment of large and central tumors, it is important to determine the most reliable dose tolerance limits to guide clinical practice. The goal of this study was to estimate these dose tolerance limits based on analysis of clinical outcomes data.

Materials and Methods

Patient characteristics and details of the treatments have already been described in [1, 2] and are only briefly overviewed here. Eighteen consecutive patients who were treated using volumetric modulated arc therapy with RapidArc™ (Varian Medical Systems, Palo Alto, Ca) for lung tumors exceeding 80cc were assessed. The mean planning target volume (PTV) volume was 137cc, with range 87-286cc. Risk-adapted fractionation schemes of either five fractions of 11 Gy or eight fractions of 7.5 Gy were used, depending on T stage and estimated risk of normal tissue toxicity [3, 2]. All plans were normalized such that the nominal fraction dose corresponded to the 80% isodose line. Dose was calculated using the Anisotropic Analytical Algorithm (AAA), accounting for tissue inhomogeneity, with a standard grid resolution of 2.5 mm. The PTV was subtracted from the assessed lung contours. Median follow-up time was 12.8 months. The evaluation of toxicity was scored using Common Toxicity Criteria for Adverse Events (CTCAE) 4.0. Radiation pneumonitis (RP) was reported in 5 patients (Grade 2 in 3, and Grade 3 in 2).

Statistical Modeling

The original Lyman model [4] uses a Gaussian distribution to estimate NTCP in terms of the normalized slope $m$ and the $TD_{50}(V)$ 50% tolerance dose (TD) for a given partial volume ($V$):

$$NTCP = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2} dx$$

(1)

where $t = (Dv - TD_{50}(V))/(m \times TD_{50}(V))$, $m$ is the normalized slope, and $Dv$ is the dose to the given volume $V$. Instead of assuming a power-law relationship to model the volume dependence as Lyman did, in this work we analyze the volume separately for each specified dose level to the critical structure.
To determine statistical dose tolerance limits from the pneumonitis data [1] we used the maximum likelihood parameter fitting technique [5] because of its effectiveness in extracting the most information from limited datasets such as this. Maximum likelihood principles were derived by R. A. Fisher [6, 7] and have been proven in many instances to be theoretically optimal. The maximum likelihood parameter fitting technique was applied to NTCP modeling by Jackson et al. [5] about 20 years ago:

$$L(\gamma_1, \gamma_2, \cdots) = \prod_{m \text{ complications}} NTCP_m(\gamma_1, \gamma_2, \cdots) \times \prod_{n \text{ no complications}} (1 - NTCP_n(\gamma_1, \gamma_2, \cdots))$$

For the model in Eq. (1), only two parameters need to be solved, $\gamma_1 = m$ and $\gamma_2 = TD_{50}(V)$. Confidence intervals were determined using the profile likelihood method [8, 9], with the value of 0.495 from the $\chi^2_1$ distribution as the threshold to obtain a contour on the three-dimensional log-likelihood function corresponding to the 68% confidence region. The effects of the confidence region on the dose response curve were determined by considering all combinations of $m$ and $TD_{50}(V)$ within this contour, and plotting the full extents as the dashed green confidence intervals in Fig. 1.

Frequently, NTCP is considered in terms of dose, with probability plotted on the y-axis as a function of dose on the x-axis, and the resulting treatment planning constraints are often called “dose tolerance limits.” This perspective implies that the volume ($V$) in Eq. (1) has been set to a fixed value. However, lung dose tolerance is typically specified conversely as $Vd$, where $Vd$ is the volume exceeding a specified dose $d$. As an example, for conventional fractionation the recent Quantitative Analysis of Normal Tissue Effects in the Clinic (QUANTEC) report [10] has recommended that the V20Gy should be kept below 30-35% to limit the risk of Grade 2-3 RP to 20% or less. The same formulation in Eqs. (1) and (2) can be used to analyze this by exchanging dose and volume in the equations. This analysis was performed using the DVH Evaluator software tool (DiversiLabs, Huntingdon Valley, Pa) which was developed by the first author.

**Results**

The strongest predictors of RP were total lung V5Gy and contralateral lung V5Gy; both of these had $p$-values less than 0.0001 [1]. Ipsilateral V5Gy also significantly correlated with clinical symptoms ($p=0.004$). Although the total lung V20Gy did not reach statistical significance for prediction of RP ($p=0.1$), it has been used in ROSEL and RTOG lung protocols for SBRT [11, 12, 13] and is in common clinical use. For these reasons, in this work we applied Eqs. (1) and (2) to the V5Gy for total, SBRT RP Tolerance Limits.
contralateral, and ipsilateral lung, as well as to the V20Gy for total lung, as shown in Fig. 1. The same rationale was used to analyze all of the dose-volume data from Fig. 3 of [1], and for each metric Table 1 shows the dosimetric values corresponding to 10%, 25%, or 50% predicted risk of RP.

Fig. 1. Dose-Response modeling of total lung V20Gy and V5Gy for total, ipsilateral, and contralateral lung.
Table 1. Maximum likelihood fitted results for 10%, 25%, and 50% risk levels for all dose-volume criteria from Fig. 3 of [1].

<table>
<thead>
<tr>
<th>Criteria</th>
<th>10% G2-3 Risk</th>
<th>25% G2-3 Risk</th>
<th>50% G2-3 Risk</th>
<th>m</th>
</tr>
</thead>
<tbody>
<tr>
<td>V20Gy Total Lung</td>
<td>5.0%</td>
<td>10.5%</td>
<td>16.5%</td>
<td>0.5428</td>
</tr>
<tr>
<td>V15Gy Total Lung</td>
<td>8.0%</td>
<td>13.5%</td>
<td>19.6%</td>
<td>0.4631</td>
</tr>
<tr>
<td>V10Gy Total Lung</td>
<td>15.2%</td>
<td>20.2%</td>
<td>25.7%</td>
<td>0.3171</td>
</tr>
<tr>
<td>V5Gy Total Lung</td>
<td>27.9%</td>
<td>32.5%</td>
<td>37.6%</td>
<td>0.2000</td>
</tr>
<tr>
<td>V5Gy Ipsilateral Lung</td>
<td>41.1%</td>
<td>47.9%</td>
<td>55.3%</td>
<td>0.2000</td>
</tr>
<tr>
<td>V5Gy Contralateral Lung</td>
<td>21.8%</td>
<td>25.4%</td>
<td>29.3%</td>
<td>0.2000</td>
</tr>
<tr>
<td>MLD Total Lung</td>
<td>5.8Gy</td>
<td>6.8Gy</td>
<td>7.9Gy</td>
<td>0.2000</td>
</tr>
</tbody>
</table>

The boldfaced values of Table 1 have been extracted and rounded to use as clinical reference points, as shown in Table 2.

Table 2. Selected clinical reference values from Table 1.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>25% G2-3 Risk (Preferred)</th>
<th>50% G2-3 Risk (Deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>V20Gy Total Lung</td>
<td>10%</td>
<td>15%</td>
</tr>
<tr>
<td>V5Gy Total Lung</td>
<td>33%</td>
<td>40%</td>
</tr>
<tr>
<td>V5Gy Contralateral Lung</td>
<td>25%</td>
<td>30%</td>
</tr>
</tbody>
</table>

If unconstrained, the slope for the NTCP curve of the mean lung dose (MLD) and the three V5Gy criteria for this dataset would approach infinity. The steepest published normalized slope that we have seen for any SBRT NTCP model for any lung criteria is 0.2 for MLD [14], so for this preliminary analysis we set all four of these to 0.2, as seen in the last column of Table 1 (note that this value is still somewhat steeper than the optimized value for V10Gy). We look forward to results from datasets involving more patients, but it is still clinically worthwhile to consider these preliminary results as soon as they emerge. The $TD_{50}(V)$ values would not be affected much by this slope; it would affect the lower risk results the most.

The 68% confidence intervals in Fig. 1 are reasonably tight in the region where we have data, but quickly diverging elsewhere, most clearly visualized in the uppermost subplot of Fig. 1. In the cases
where we had to fix \( m = 0.2 \), the confidence intervals were generated both with and without this constraint and the combined worst case was plotted so the full impact of the uncertainty of the slope can be seen. The confidence intervals show that the statistical analysis is the most reliable where most of the clinical data exists, which is the most clinically relevant range.

**Discussion**

Ideally the dose tolerance limits for most Grade 3 adverse events following radiotherapy are specified in terms of TD 5/5 and TD 50/5 dose tolerance limits; the 5% and 50% risk levels at 5 years. Since most of the adverse events observed in this study are Grade 2, we have instead used the 25% and 50% risk levels. The main findings in this study are that the total lung V5Gy and contralateral lung V5Gy best correlated with symptomatic radiation pneumonitis, and estimates for the former as well as for the V20Gy, are summarized in Table 2. These dose tolerance limits are within a clinically useful range, but caution is still advised due to the limited dataset and limited length of followup, and these findings should be validated in a separate dataset.

There are insufficient cases in this dataset to accurately assess the effect of the two alternative fractionation schemes on the complication rate, but that is an important topic of future research. If \( \alpha / \beta = 3 \), the linear quadratic (LQ) model predicts that 20Gy in 8 fractions is equivalent to 17.1Gy in 5 fractions, so for the critical structure doses discussed in this work the difference is only about 3Gy in 5 fractions, and there is at least that much uncertainty in this data, as may be seen by examining the confidence intervals in Fig. 1. Therefore we have kept the two fractionation schemes combined together as was done in the prior publication [1], in order to provide an important early estimate of the desired doses and to spur future research to study more cases to gather sufficient data for the necessary statistical power to answer these questions in more detail.

**Comparison to other publications**

[See the full version of this article at www.OldCityPublishing.com/JRSBRT/JRSBRT.html]

**Conclusion**

[See the full version of this article at www.OldCityPublishing.com/JRSBRT/JRSBRT.html]
References


[See the full version of this article at www.OldCityPublishing.com/JRSBRT/JRSBRT.html]