RADIATION DOSE–VOLUME EFFECTS IN THE LUNG

LAWRENCE B. MARKS, M.D.,* SOREN M. BENTZEN, D.SC.,† JOSEPH O. DEASY, PH.D.,‡ FENG-MING (SPRING) KONG, M.D., PH.D.,§ JEFFREY D. BRADLEY, M.D., IVAN S. VOGELIUS, PH.D.,† ISSAM EL NAQA, PH.D., † JESSICA L. HUBBS, M.S.,* JOOS V. LEBESQUE, M.D., PH.D.,jj ROBERT D. TIMMERMAN, M.D.,† MARY K. MARTEL, PH.D.,# AND ANDREW JACKSON, PH.D.**

*Department of Radiation Oncology, University of North Carolina, Chapel Hill, NC; †Department of Human Oncology, University of Wisconsin School of Medicine, Madison, WI; ‡Department of Radiation Oncology, Alvin J. Siteman Cancer Center, Washington University School of Medicine, St. Louis, MO; §Department of Radiation Oncology, University of Michigan, Ann Arbor, MI; †Department of Radiation Oncology, The Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands; ‡Department of Radiation Oncology, University of Texas Southwestern, Dallas, TX; §Department of Radiation Physics, M. D. Anderson Cancer Center, Houston, TX; and **Department of Medical Physics, Memorial Sloan-Kettering Cancer Center, New York, NY

The three-dimensional dose, volume, and outcome data for lung are reviewed in detail. The rate of symptomatic pneumonitis is related to many dosimetric parameters, and there are no evident threshold “tolerance dose–volume” levels. There are strong volume and fractionation effects.

1. CLINICAL SIGNIFICANCE

Radiotherapy (RT) plays an important role in the treatment of several tumors in and around the thorax. Clinically significant symptomatic radiation pneumonitis (RP) occurs in approximately 5–50%, 5–10%, and 1–5% of patients irradiated for cancers of the lung, mediastinal lymphatics, and breast, respectively (1, 2), and is one of the most common clinical toxicities in these patients. The risk of RP limits the delivered dose for some and may thus hamper tumor control. A large fraction of patients experience subclinical RT-induced injury (e.g., reductions in formal pulmonary function tests and/or radiologic changes) that may be chronic and reduce the patient’s reserve to deal with future cardiopulmonary stresses.

2. ENDPOINTS

Several endpoints can be used to define RT-induced lung injury (Table 1). In the context of quantitative analysis of normal tissue effects in the clinic (QUANTEC), consideration is limited to the endpoint of symptoms—arguably the most clinically meaningful endpoint for patients. Approximately 80% of RP is clinically manifest within 10 months of RT.

The scoring of symptomatic RP presents several challenges: (1) Dyspnea is nonspecific and can also be caused by, for example, anemia, cardiac arrhythmia, infection, and tumor. In a prospective clinical study, 28% of patients suspected of having RP also had ongoing medical conditions confounding the diagnosis (3). (2) Toxicity grading systems often consider the medical interventions (e.g., steroid use). Therefore, physicians who are more apt to prescribe steroids may note a higher reported rate of pneumonitis. Steroid use is Grade 3 in the Radiation Therapy Oncology Group (RTOG) scoring system but Grade 2 in several other systems. (3) Treatment-induced tumor shrinkage may improve overall lung function (especially for central lesions compressing regional airways/vessels), thus perhaps masking the effects of RT on the normal lung. (4) The relevant grade of symptoms is controversial. Grade 1 RP is common and is often not clinically significant. More severe RP is more clinically relevant, but its lower incidence limits the statistical power of analysis based on severe events.

3. CHALLENGES DEFINING VOLUMES

The lung is usually considered as a single, paired organ (total lung tissue) rather than as separate ipsi- and contralateral lungs. Because lung volumes vary with breathing, there is...
ambiguity in defining its dose–volume histogram (DVH)-

based parameters. In the articles herein reviewed, dosimetric

information was mostly based on CT images obtained during

free breathing. The dosimetric parameters would change had

these scans been obtained at specific phases of the respiratory

cycle. Segmentation of a thoracic scan can be challenging.

There is uncertainty regarding how much of the bronchus

should be defined as “lung,” and the lung edges may vary

with the window/level setting. Thus, volume-based param-

ters will vary between investigators. The accuracy of any au-

tosegmenting tools should be carefully assessed, especially to

ensure that portions of atelectatic lung or tumor at soft-tissue

interfaces are not inadvertently omitted from the lung.

During RT planning, the total lung volume is usually de-

fined to exclude the gross tumor volume (GTV). Excluding

the planning target volume (PTV) rather than the GTV

from the lung volume may reduce the apparent lung exposure

(because normal lung within the PTV but outside the GTV

will be excluded) and may increase interinstitutional vari-

ations (because PTV margins may vary).

During treatment there may be changes in GTV, with cor-

responding changes in normal tissue anatomy. Thus, plans

defined on the basis of pre-RT imaging may not accurately

reflect the degree of normal lung exposure. Although this

effect has not been widely considered, presumably tumor

shrinkage (with movement of normal lung into space previ-

ously occupied GTV) will increase normal lung exposure re-

lative to pre-RT plans. Similarly, changes in pleural effusions

and re-aeration of lung regions can cause anatomic and func-

tional changes. Indeed, the ability to predict changes in lung

function according to pre-RT dosimetric data is reduced in

patients with tumor-associated airway obstruction (i.e., those

most likely to experience re-aeration during therapy) (4).

4. REVIEW OF DOSE–VOLUME DATA

The literature on dose–volume parameters and pneumoni-

tis is extensive: for this review we identified >70 published

articles. The results are inconsistent, both for the best predic-

tive metrics and significant comorbid factors.

Table 1. Example endpoints for radiotherapy-induced lung

injury (and approximate incidence)

<table>
<thead>
<tr>
<th>Regional</th>
<th>Global</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical</td>
<td></td>
</tr>
<tr>
<td>Bronchial stricture (&lt;3%*)</td>
<td>Shortness of breath (5–50%)</td>
</tr>
<tr>
<td>Subclinical</td>
<td></td>
</tr>
<tr>
<td>Radiologic abnormalities (e.g., computed tomography, perfusion/ventilation scans) (20–80%)</td>
<td>Pulmonary function tests, 6-min walk test, blood gases, exercise capacity†</td>
</tr>
</tbody>
</table>

Example endpoints used to study radiotherapy-induced lung injury can be broadly segregated as shown.

* Uncommon with conventional fractionation and doses. More common with brachytherapy, high total doses, and/or hypofractionation.

† Many patients experience declines in functional assessments, but the magnitude of the decline is variable.

Fig. 1. Meta-analysis of reported n values (volume parameter) for the Lyman-Kutcher-Burman (LKB) model using an inverse-var-

iance (IV) weighting method. Recovery of variance estimates from the 95% confidence interval (CI) and use of approximately

±2*sigma instead of 1.96*sigma gave rise to small deviations in the derived 95% CI as compared with the literature reported values.

Data estimated from references 47–49. Fixed = fixed effect model. The n value reflects the manner in which dose–volume parameters

lead to complications. A lower value of n suggests that the tissue is sensitive to hot spots (e.g., an organ structured in “series”),

whereas a higher value of n (closer to 1.0), suggests that the risk is more related to the volume of an organ irritated (e.g., “parallel”

structure).

Lyman-Kutcher-Burman DVH reduction scheme and mean lung dose

The most widely used normal tissue complication proba-

bility model for RP is the Lyman-Kutcher-Burman (LKB) model. This model has three parameters: a position param-

ter, TD50, a steepness parameter, m, and the volume expo-

sure parameter, n (where n = 1 the model reverts to mean lung dose [MLD]). Although TD50 is strongly dependent on the grade

of RP being considered, n is often regarded as a tissue char-

acteristic. Figure 1 shows a meta-analysis of reported n values; it does not include the study by Rancati et al. (5),

which used only the ipsilateral lung. The best estimate for n is 1.03 with standard deviation 0.17 (95% confidence inter-

val [CI], 0.67, 1.39), the test for heterogeneity of the datasets is not significant, and I² is zero. The TD50 values cannot be

pooled in a meaningful way, because the various reports analy-

ized considered varying grades of RP.

The MLD model is widely considered owing to its simplic-

ity and effectiveness. It was the metric used by the large multi-institutional analysis of Kwa et al. (6) and often per-

forms as well as more complex models. Figure 2A shows a lo-

gistic regression fitted to RP vs. MLD data from all published studies of a significant size that had extractable complication

rates binned by mean dose. Some of the variation around the fitted curve is possibly explained by differences in patient se-

lection, as well as differences in the grade of RP reported in the various studies. Nevertheless, there is a relatively small

68% confidence interval (stippled lines). A similar fit using the probit model (equivalent to fitting the Lyman model with n fixed at 1) gives an essentially identical response function in the region of the data. The gradual increase in dose re-

sponse suggests that there is no absolute “safe” MLD below which there is no pneumonitis. The clinically acceptable risk

of RP—and therefore the associated planning constraint on MLD—will depend on the risk/benefit ratio in the individual case.

A number of non–DVH-based factors may affect the risk of RP (see “Factors Affecting Risk”). Finally, it is likely
that the MLD–RP relationship may have lower predictive power for “nonstandard” dose distributions not included in these analyses, for example after stereotactic body radiotherapy (SBRT), Intensity-Modulated Radiation Therapy (IMRT), or proton therapy.

Dose–volume threshold analyses

Various Vx values (percentage lung volume receiving ≥x Gy) are associated with RP risk (Fig. 2B). The observation that a variety of dose levels are predictive suggests that there is no sharp dose threshold below which there is no risk. Within individual datasets there are usually strong correlations between the different dosimetric parameters (e.g., V5 and V20), and thus this may partly obscure any “optimal” threshold. Furthermore, the correlations between dosimetric parameters are technique dependent, and readers should carefully assess the similarity of their treatment technique to the historical reports before using any of these limits as clinical constraints.

Radiotherapy-induced dyspnea appears more commonly in patients with lower- vs. upper-lobe tumors and may be better correlated with RT doses to the lower vs. upper lung (7–11). An analysis that combined institutional data with RTOG 93-11 (n = 324) concluded that RP is much better predicted (at least for that dataset) according to MLD and positional dependence of the high-dose region as opposed to MLD alone (12).

5. FACTORS AFFECTING RISK

Several patient- and treatment-related factors have been inconsistently reported to correlate with the risk of developing RP. Vogelius and Bentzen (unpublished data) applied standard meta-analysis methodology to eight factors with meaningful data. In summary, there was no significant evidence for an association between RP and GTV laterality (left vs. right lung), comorbidity, or gender. Younger patients, typically defined as <60 or <70 years of age, had a lower risk of RP than older patients. Surgery had a just-significant p value, but the test for heterogeneity was significant (p = 0.03), suggesting that the variation among studies cannot be explained by chance alone. Thus, at present, the reduced rate of RP in patients undergoing surgery remains controversial. Interestingly, current smokers had a significantly reduced risk of developing RP.

Chemotherapy

Many systemic agents have known pulmonary toxicities (13) and may exacerbate RT-induced injury. The varying drugs, doses, and schedules (e.g., sequential or concurrent) make any synthesis of data from multiple studies generally not feasible. On the basis of general experience, adding chemotherapy might be expected to increase the risk of RP.
Nevertheless, the agents most commonly used with RT for lung cancer, such as cisplatin, carboplatin, paclitaxel, and etoposide, have not been consistently shown to increase the risk of pneumonitis (7, 11, 14–16). More modern agents have been associated with high rates of pulmonary toxicity when used concurrently with thoracic RT (e.g., docetaxel and gemcitabine) (1, 17, 18).

**Radiation dose, time, and fractionation**

Radiation pneumonitis has a relatively high fractionation sensitivity (19, 20); the best current estimate (±1 standard error of the estimate) of the \( \alpha/\beta \) ratio of the linear-quadratic model is 4.0 ± 0.9 Gy (21). For comparison, the upper bound of the 95% CI for \( \alpha/\beta \) for pulmonary fibrosis is 3.5 Gy. There is also a significant time factor for pneumonitis, with an overall best estimate of the dose recovered per day, \( D_{pr} \), of 0.54 ± 0.21 Gy/day. Several investigators have suggested methods to adjust the DVH to reflect the impact of fraction size (22, 23).

### 6. MATHEMATIC/BIOLOGIC MODELS

The association between RP risk and MLD (logistic fit to the data in Fig. 2(a) can be expressed as

\[
p = \frac{\exp(b_0 + b_1 \cdot MLD)}{1 + \exp(b_0 + b_1 \cdot MLD)}.
\]

Best-fit parameters (95% CI) are \( b_0 = -3.87 \) (−3.33, −4.49) and \( b_1 = 0.126 \) (0.100, 0.153) Gy⁻¹. These estimates yield a predicted TD₅₀ = 30.8 (28.7, 33.9) Gy and \( \gamma_{50} = 0.97 \) (0.83, 1.12) (this parameter represents the increase in response [measured as percentage] per 1% increase in dose, at the 50% dose–response level). A fit using the probit response function (equivalent to a fit of the Lyman model with \( n = 1 \)) yields TD₅₀ = 31.4 Gy (95% CI, 29.0, 34.7 Gy) and \( m = 0.45 \) (0.39, 0.51). The resultant response function is essentially identical to that of the logistic fit in the region occupied by the data. The curvature is slightly smaller, resulting in slightly larger TD₅₀ value. Both fits assumed heterogeneity corrected dose distributions (an approximate correction of 11% was applied to doses from studies using homogeneous calculations).

### 7. SPECIAL SITUATIONS

The data reviewed here are largely derived from patients who received partial-lung irradiation using conformal three-dimensionally planned external-beam RT with conventional fractionation (e.g., 1.8–2.0 Gy per fraction). Several special situations are discussed here.

**Whole-lung irradiation**

Near-uniform irradiation of both lungs occurs during total-body irradiation as conditioning for stem cell transplants, hemibody RT for diffuse metastases, and whole-lung irradiation for prophylaxis or treatment of pulmonary metastases from various malignancies. The risk of RP depends on total dose and fraction size (Fig. 3). The development of RP in these settings is an ominous sign, proving fatal in up to 80% of patients (24). The pathogenesis of RP, in particular after total-body irradiation, is relatively complex and depends on multiple patient- and treatment-related factors (25). There are consistent data supporting a protective effect of low dose rate and low dose per fraction. For a recent comprehensive review, see Sampath et al. (26).

**Hypofractionation**

Stereotactic body radiotherapy generally involves 1–5 large fractions (e.g., 14–30 Gy) given over 5–20 days (27, 28). The high-dose volumes are small, and dose gradients are typically uniformly steep, minimizing dose to surrounding critical structures. However, because numerous beams are used, there are large areas of lung receiving low to medium doses (28). Thus, the dose–volume characteristics of SBRT are quite different from those of conventional lung RT and deserve special consideration. Radiation pneumonitis is relatively uncommon after SBRT, usually <10% (28–30) but as high as 25% (31). Bronchial injury/stenosis, an unusual complication with conventional doses (32), has been associated with SBRT of perihilar/central tumors (28).

**Intensity-modulated radiotherapy for lung cancer**

The M. D. Anderson Cancer Center reported a lower rate of symptomatic Grade ≥3 pneumonitis in 68 patients treated with intensity-modulated radiotherapy (IMRT) compared with a historical control group of 222 receiving conventional three-dimensional conformal RT (33). The Memorial Sloan-Kettering Cancer Center recently noted an acceptable 11% rate of Grade ≥3 pneumonitis in 55 patients treated with IMRT (34). Postoperative IMRT for mesothelioma has been associated with a high rate of lethal pneumonitis (8–46%) (35–37), and extreme care should be used to limit lung irradiation in these cases (see next section).

### 8. RECOMMENDED DOSE/VOLUME LIMITS

Recommending dose/volume limits is challenging because there are no clear and consistent “thresholds” for candidate metrics (i.e., the response function is often gradual), and
the “acceptable” risk level varies with the clinic scenario. Radiotherapy fields for lung cancer may be appropriately large for target coverage; physicians and patients often need to accept the significant pulmonary risks. Furthermore, there are marked interpatient variations in pre-RT lung function that may impact symptom development, and tumor-related dysfunction may improve after RT.

Despite these caveats, it is prudent to limit V20 to ≤30–35% and MLD to ≤20–23 Gy (with conventional fractionation) if one wants to limit the risk of RP to ≤20% in definitively treated patients with non–small-cell lung cancer. Similar guidelines for other parameters can be extracted from the figures. Limiting the dose to the central airways to ≤80 Gy may reduce the risk of bronchial stricture (30). In patients treated after pneumonectomy for mesothelioma, it is prudent to limit the V5 to <60%, the V20 to <4–10%, and the MLD to <8 Gy (see Miles et al. [37] for detailed review).

9. FUTURE TOXICITY STUDIES

Progress regarding the predictors of RT-induced lung injury requires further understanding of the following.

Endpoint interaction

The study of RT-induced lung injury is confounded by the use of ambiguous endpoints. Many scoring systems combine radiologic, functional, and symptomatic criteria to define a “global score.” Because each endpoint may have different dose–volume dependence, this approach may be counterproductive. Therefore, we recommend that further study of lung injury explicitly consider symptomatic, functional, and radiographic endpoints separately.

Impact of clinical factors

The impact of clinical factors (e.g., pre-RT functional status, tobacco use) and systemic agents (e.g., chemotherapy) on the risk of RP needs further study.

Organ interactions

Some pre-clinical data suggest that there may be interactions between the lung and heart in the development of RT-associated dyspnea. In rats, the respiratory rate after thoracic RT was related to the volume of lung and heart irradiated (38–40).

Impact of an in situ lung cancer on the risk of radiation-induced lung injury

The data for whole-lung radiation is derived essentially from patients without primary lung cancers (e.g., elective lung RT for sarcoma), vs. fractionated partial lung radiation, often derived from patients with gross lung cancers. The confounding effect of tumor in the lung makes the study of RT-induced lung injury extremely challenging. Indeed, in several studies, the ability to predict for RT-induced lung injury is improved in patients without large central or occluding tumors. Thus, it might be relevant to develop separate predictive models in patients with intact intraparenchymal lung tumors vs. those without such a lesion (i.e., postresection RT for lung cancer, or RT for other thoracic tumors).

Radiation response modifiers

Amifostine is a thio-organic prodrug believed to scavenge harmful free radicals mediating RT-induced injury. Several randomized studies in patients receiving RT for lung cancer note a reduction in RP in the amifostine arm (41–43), although the largest study (from RTOG) was negative (44). However, this study has been criticized because the drug was administered once daily (4 days/week) whereas the RT was delivered twice daily (5 days/week), and thus 60% of the RT fractions were delivered without the protector. Such mixed results, combined with the acute toxicities of amifostine (nausea/vomiting, hypotension, infection, and rash), have dissuaded many from using it in routine practice. One small randomized study demonstrated a protective effect of pentoxifylline, but pentoxifylline is not currently used in routine clinical practice (45).

Biomarkers

Additional work is needed to assess the predictive ability offered by biomarkers (see Bentzen et al. in this issue), such as transforming growth factor β (measured before and/or during RT) (46).

10. TOXICITY SCORING

A Late Effects of Normal Tissue–Subjective, Objective, Management, and Analytic (LENT-SOMA)-type scoring system is recommended because it explicitly considers symptomatic, functional, and radiographic endpoints individually. A global score can be generated, but the granular data can be maintained.

REFERENCES


