Parameters for and Use of NTCP Models in the Clinic

Ellen D. Yorke
Memorial Sloan-Kettering Cancer Center
New York City, NY
Supported in part by NCI Grant PO1-CA-59017

Modern radiation therapy does very well at shaping dose distributions
- For tumor control, tumor localization + sufficiently high dose (understanding tumor dose-response) are key

- For each normal tissue of interest, we know the planned dose distribution for the entire organ volume
  - IGRT will soon give information about delivered doses

What dosimetric features should be restricted to keep normal tissue complication risk low?
- Grading schemes: 1 is mild → 5 is lethal
- ‘Low’ is ≤ Grade 2 in any grading scheme
- What guidance is available?
- How does it apply to individual clinical practice?

Usual assumption
Probability of a particular complication has a sigmoidal increase with dose or a dose-related metric.

NTCP=Normal Tissue Complication Probability

TD50: Dose for 50% complication probability

Doses for 50% & 5% complication probability at 5 years
- Conventional fractionation: 1.8-2 Gy/Fx

- For many complications, the TD's increase if less of the organ is irradiated
  - Partial Irradiation of volume fraction v of organ
  - Tabulated TD50/5 and TD5/5 for v=1/3, 2/3 and 1.0

The 1991 Data Set

- Emami et al (9 authors – 7 MD's, 2 PhD's)
  - Part of NCI-funded CWG on 3D planning: IJROBP 21, 109-122, 1991

- Literature review up to 1991, 28 complications
  - Predates 3D-CRT era; even DVHs were new

- Tabulated estimates of TD50/5 and TD5/5
  - Doses for 50% & 5% complication probability at 5 years
  - Conventional fractionation: 1.8-2 Gy/Fx

- For many complications, the TD's increase if less of the organ is irradiated
  - Partial Irradiation of volume fraction v of organ
  - Tabulated TD50/5 and TD5/5 for v=1/3, 2/3 and 1.0

- Companion paper fit volume effects to power law and NTCP to Lyman model
  - Burman et al, IJROBP 21, 123-135
Volume effects according to 1991 report

Partial organ irradiation

- Zero dose
- Uniform Dose
- Volume fraction: 1 - v
- Volume fraction: v

For many complications, the iso-complication dose depends **inversely** on the irradiated volume fraction.

1991 to Now

- CT-simulation is routine
  - Plus MRI, PET, 4D-CT
- 3D-CRT is the norm, IMRT explodes
- 3D plan evaluation tools
- Complex dose distributions
- Huge amount of published information
  - Some listed in handout bibliography
  - Noisy data, various grading schemes, different calculation methods and plans – tough to sort out
- October 2007: QUANTEC
  - AAPM/ASTRO funded workshop on NTCP
  - Consensus guidance for clinical use of NTCP studies

**Definitions**

DVH in terms of absolute dose (Gy), % volume
- \( D_{\text{highest}} \): dose that encloses \( v\% \) organ volume
- \( D_{\text{max}} \): maximum dose in structure
- \( V_{\text{ref}} \): % volume receiving \( \geq \) dose \( D \)
  - Similar definitions for 'dose' in % of a reference value or absolute volume

- Linear Quadratic (LQ) model:
  - The smaller \( \alpha/\beta \), the greater is sensitivity to dose per fraction
- **Biological Effective Dose (BED)**
  \[
  \text{BED} = D \left(1 + \frac{d}{\alpha/\beta}\right)
  \]
  \( d = \) dose per fraction, \( D = \) total dose

**Volume effects are important!**

DVH plan evaluation, dose-volume optimization

- **Complications with a weak volume dependence**
  - Restrict \( D_{\text{max}} \) or \( D_{\text{ref}} \) to limit complication risk
  - Constrain maximum dose in optimization
- **Complications with strong volume dependence**
  - To limit complication risk, restrict mean dose or dose-volume features as guided by clinical studies
  - Dose-volume or mean dose constraints in optimization
- **Intermediate volume dependence**
  - \( V_d \) for variety of doses in plan evaluation, optimization
  - In the future - perhaps detailed spatial aspects of dose distribution (high dose clusters, circumferential irradiation) or dose-surface histograms would do better.
**NTCP Models**

- **Statistical models**
  - What dosimetric features correlate with complication?
    - medical variables (chemo, co-morbidities) may be included
  - Dose distributions may be corrected for dose-per-fraction (LQ model)
  - Correlation alone doesn’t tell what limits to use
    - Sigmoidal curves from logistic regression
    - Multivariable models include several factors in a single equation for NTCP
  - Results of statistical models are used
    - RTOG Nasopharynx protocol 0225, RTOG 9311 Lung protocol, $D_{\text{max}}$ for cord $\leq 45-50\text{ Gy}$

- **Empirical models**
  - Equations with parameters fit to outcomes data sets
  - No mechanistic foundation (except LQ if used)
  - Do study set characteristics apply to your practice?
    - Examples: Lyman model, gEUD

- **Semi-mechanistic models**
  - Tissue architecture, as well as cellular radiosensitivity, determine NTCP
  - Parameters chosen to fit clinical data
  - Do study set parameters apply to your practice?
    - Examples: Serial (critical element) model, Parallel (critical volume) model, Relative seriality model
  - All models rely heavily on statistical analysis
    - These papers are NOT easy reading!

**Lung: Acute Radiation Pneumonitis (RP)**

- Severe RP - steroids, oxygen or worse
- 20-25% incidence of RTOG Gr 3 accepted
- Usual RP onset $\leq 6$ months from tx start

  - Total organ ($v=1$) is the pair of lungs
  - Strong volume effect:
    - $T_{D5}(1)=17.5\text{ Gy}, T_{D5}(1/3)=45\text{ Gy}$
  - Low tolerance doses: $T_{D50}(1)=24.5\text{ Gy}, T_{D5}=17.5\text{ Gy}$
  - Most calculations not inhomogeneity corrected

- Since 1991 Review: Seppenwoolde et al, Sem Rad Onc 11, 247-258
  - Mean lung dose (MLD) is good predictive metric
  - $\alpha/\beta$ $\sim 2.4\text{ Gy}$ (fractionation sensitive)
  - $T_{D50}(1) \sim 28\text{ Gy}$ (inhomogeneity corrected)

- MLD$<20\text{ Gy}$ at conventional fractionation
  - Patients without special medical factors

- Low dose to large volumes surprisingly significant
  - $V_{20\%,13\%}$ correlate with RP
  - Medium doses also ($V_{20\%}, V_{40\%}$): not $D_{\text{max}}$
  - Importance of absolute volumes at low - medium doses?
  - References in handout

- Evidence for regional sensitivities
  - Lower lung dose much more significant than upper lung dose
Rectal complications

- Rectal bleeding, ulceration, stricture, fistula
  - Can set in early or late (up to ~ 4 years after treatment)
  - 10-20% moderate rectal bleeding clinically accepted
- 1991: High TD; small volume effects
- Since 1991-motivated by prostate IMRT
- Restricting % volume ≥ 75 Gy is very important
- But the volume effect is quite complex
  - Intermediate doses (V30-V72), mean dose; High dose clusters? Dose-surface histograms?
- α/β ~3-6 Gy
- Other considerations: co-morbidities, motion

Lyman Model

\[ NTCP = (2\pi)^{-1/2} \int_{-\infty}^{\infty} \exp(-t^2/2) \, dt \]

- 4 parameters
  - n - volume dependence; TD50(1) - dose sensitivity;
  - m - slope; reference volume for v=1 (often whole organ)
- TDC(v)=Dose for c% complications, partial irradiation of v
- Equivalent uniform dose to whole organ
  \[ D_{\text{eff}} = (\sum_v (D_v / D_{\text{ref}})^{1/a})^{1/a} \]
- Partial irradiation of effective volume to reference dose
  \[ v_{\text{eff}} = (\sum_v (D_v / D_{\text{ref}})^{1/a})^{1/a} \]

Lyman Model Parameters

- Recent Lyman parameters for many complications:
  - radiation induced liver disease (RILD), RP, xerostomia (n=1)
  - rectal bleeding, pericarditis, acute esophagitis
- Parameters affect model predictions for same dose distribution

Generalized Equivalent Uniform Dose (gEUD, EUD)

\[ gEUD = \left( \sum_v (D_v)^a \right)^{1/a} \]

- Normal tissue complication: a > 0; gEUD=Lyman D_{50}, n=1/a
- Different complications of a tissue can have different a’s
- For tumor control, a=0
- One formalism handles normal tissues and tumors

Simplicity → use in plan comparison and optimization

- One tunable parameter for tumors and NTCP
- With two more parameters, get a sigmoidal function

*Wu Q et al., LRDBP 53 and Med Phys 32
**Tissue architecture models**

- See Withers et al., IJROBP 14, 791-799
- H1: Organs are made of functional subunits (FSUs)
  - FSU radiosensitivity
  - FSU organization (tissue architecture)

- Assume: FSUs respond independently
  - LQ model takes care of temporal/fractionation effects

- None of these models are mechanistic at the cellular or anatomic level. Are they better than statistical or empirical models?

---

**Simple differential DVH**

<table>
<thead>
<tr>
<th>a</th>
<th>gEUD (Gy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-20</td>
<td>34.8</td>
</tr>
<tr>
<td>0.2</td>
<td>49.3</td>
</tr>
<tr>
<td>1</td>
<td>50</td>
</tr>
<tr>
<td>20</td>
<td>60.9</td>
</tr>
<tr>
<td>100</td>
<td>67.9</td>
</tr>
</tbody>
</table>

---

**Serial (critical element) model**
- FSUs arranged “in series”
- NTCP: probability at least one FSU killed
- Weak volume dependence (like small n Lyman model)
- Infrequently used: $D_{\text{max}}$ preferred

**Parallel (critical volume) model**
- FSUs work in parallel:
- Complication only if $\geq$ critical fraction $f^*$ are damaged
- For general DVH, $(D_i, v_i)$
  - $f_{\text{dam}} = \sum v_i p(D_i)$
  - $p(D)$ is the probability that a dose $D$ damages an FSU

**Strong volume dependence** — fit to clinical data for RP, RILD, xerostomia, radiation nephritis.
- Seldom used: Users prefer the other models described here.
- Probably neither is “true”

**Qualitative difference: Lyman vs parallel model**

- Lyman: $D_{\text{eff}} = D v^p$, $n=1$
  - As $D \uparrow$, $D_{\text{eff}} \uparrow$, even if $v$ small
  - If $D_{\text{max}} \gg TD50(1)$, NTCP $= 100$

- Parallel: $f_{\text{dam}} = v p(D)$
  - As $D \uparrow$, $p(D) \rightarrow 1.0$ and $f_{\text{dam}} \rightarrow v$
  - NTCP can be small
    - depends on $f^*$ distribution
    - If $v < f^*$, predicted NTCP is zero, regardless of dose
- Probably neither is “true”
Relative seriality model

- Assume complication has serial and parallel aspects
  - Each parallel part made of serial components
- 4-parameter model
  - Published parameters for RILD, RP, late cardiac mortality, xerostomia, rectal bleeding, esophageal stricture

- NTCP_{WO}(D) = whole organ response
  \[ P(D) = NTCP_{WO}(D) = 2 - \exp\left(\gamma (1 - D/D_{50})\right) \]
  \( \gamma \) determines slope; \( D_{50} \) the radiosensitivity
  - LQ model for dose per fraction effects \( (\alpha/\beta) \)
- For general DVH, \( (D, v) \)
  \[ NTCP = \left(1 - \Pi_v \left(1 - P(D)\right)^s \right)^{1/s} \]
- \( s \) is the seriality parameter
  - \( s=1 \): serial structure; weak volume dependence
  - \( s <<1 \): strong volume dependence

Different models or different parameter sets, applied to the same dose distribution generally predict different absolute and even relative complication rates
- They can rank competing plans differently
- Model parameters may depend on non-dosimetric variables
  - chemotherapies, co-morbidities, age, gender, societal backgrounds (which may influence co-morbidities)
- Nonetheless, judicious use of NTCP model information can make treatment plans more consistent/less planner dependent.
  - We’ve probably all been using such information in daily practice.

MSKCC Prostate IMRT

- Rectal wall contoured ~5 mm sup to ~5 mm inf of PTV
- PTV = prostate+6 mm posteriorly; prostate+1 cm otherwise
- Dose, dose-volume constraints: quadratic score function
- Strict rectal wall evaluation constraints based on analysis of in-house data (Swanson et al, IMCPEP 47; Jackson et al, IMCPEP 49)
  - \( V_{T30} \leq 30\% ; V_{T50} \leq 53\% \), No hot-spots in rectal wall
- Very low rate of > Grade 2 rectal bleeding
  - <2% at 8 yr; prescription=81 Gy (Zalewski et al, J Urol 2006)
- PSA-relapse-free survival by risk group: 89%, 78%, 67%
- Maximum prescription dose 86.4 Gy (Zelefsky, ASTRO 2006)

MSKCC Lung cancer planning

- Lyman model, 1991 parameters
  - NTCP \leq 25% (\( d_{eff} = 20-21 \) Gy) and/or
- Parallel mode: parameters from Ten Haken
  - \( f_{dam} \leq 0.28 \)
  - and/or at MD’s discretion; Rx lowered so at least one is met
- Intrinsic model differences give planners grief!
- Observed \( \geq \) RTOG Grade 3 RP is \( \leq 15\% \)
Radiobiological Indices in Optimization

- Disclaimer: Very little personal experience
- To go to “Level 2”, I anticipate a learning curve
- Literature: Evidence for gEUD advantages
  - Different functions to put gEUD into score function
  - Equivalent or better normal tissue protection
  - Target EUD higher, distributions less homogeneous than conventional
    - Is homogeneity beneficial without detailed information about clonogen location within target?
    - Dummy normal structure to prevent overdosing target volume
- Other NTCP models could be similarly included
- “In the works” with TPS vendors

Spare

Evolution

2D

3D-CRT

IMRT
Statistical Models: Radiation Myelitis

• Protection against myelitis mandatory
  – Delivered dose may be >planned due to setup error, scatter
• 1991: T D5(1)~47-50 Gy, T D50(1)~70 Gy

Very weak volume dependence for myelitis

• Updated information Schultheiss et al, IJROBP 31
  – Weak volume dependence confirmed
  – ω/β~2 Gy (sensitive to dose per fraction)
  – T D5~57-60 Gy but we routinely keep D max ≤ 45-50 Gy
  – Slow damage repair component (> 8 hrs)

• Wiggle room for retreatment?
  – Lit search, very small # pts (~60) Nieder et al, IJROBP 66
  – ω/β=2 Gy for C, T spine, 4 Gy for L-spine
  – No myelopathies reported for total BED<120 Gy²
    • time between courses >6 months; individual BED’s < 98 Gy; small risk for total BED<135.5 Gy²

Most normal organs suffer several different radiation complications; there are several different scoring systems

Acute Radiation Pneumonitis
  NCI Grades
  0 None
  1 Radiographic changes (RC), asymptomatic or symptoms not requiring steroids
  2 RC and steroids or diuretics
  3 RC and symptoms requiring oxygen
  4 RC and assisted ventilation (AV)
  5 Death

RTOG Grades
  0 None
  1 Mild dry cough, dyspnea at exertion
  2 Persistent cough, dyspnea at rest
  3 Severe cough, steroids and/or intermittent oxygen
  4 Severe respiratory insufficiency, continuous oxygen or AV
  5 Death

Late lung toxicity (RTOG/EORT)
  0 None
  1 Asymptomatic or mild (dry cough), slight radiographic appearances
  2 Moderate symptomatic fibrosis or pneumonitis (severe cough, low grade fever, patchy radiographic appearances
  3 Severe symptomatic fibrosis or pneumonitis (severe cough); dense radiographic changes
  4 Severe respiratory insufficiency; continuous O₂; assisted ventilation
  5 Death

Relative seriality model

Difference between relative seriality and power law predictions

Slope of whole organ response increases with γ
Volume dependence increases as s decreases

Parallel (Critical Volume) Model

• FSUs work “in parallel”
• Probability of damaging a single FSU=p(D)
  At least 2 parameters for p(D)
• Complication if ≥ critical fraction, f *, of FSUs are damaged
  – NTCP=probability that ≥ f* are damaged
  f*=25%
  XXXX XXX or any other combo XXXX
  XXXX vs XXXX 4 or more damaged XXXX
  XXXX XXXX XXXX

• Fraction damaged= f dam = \sum_i p(D_i)
• Strong volume dependence predicted
  – Applied to RP, RILD, xerostomia, nephritis
  – Need at least 4 parameters for NTCP calculation
  – Users prefer Lyman, relative seriality models, gEUD or Vₙ for plan evaluation and optimization