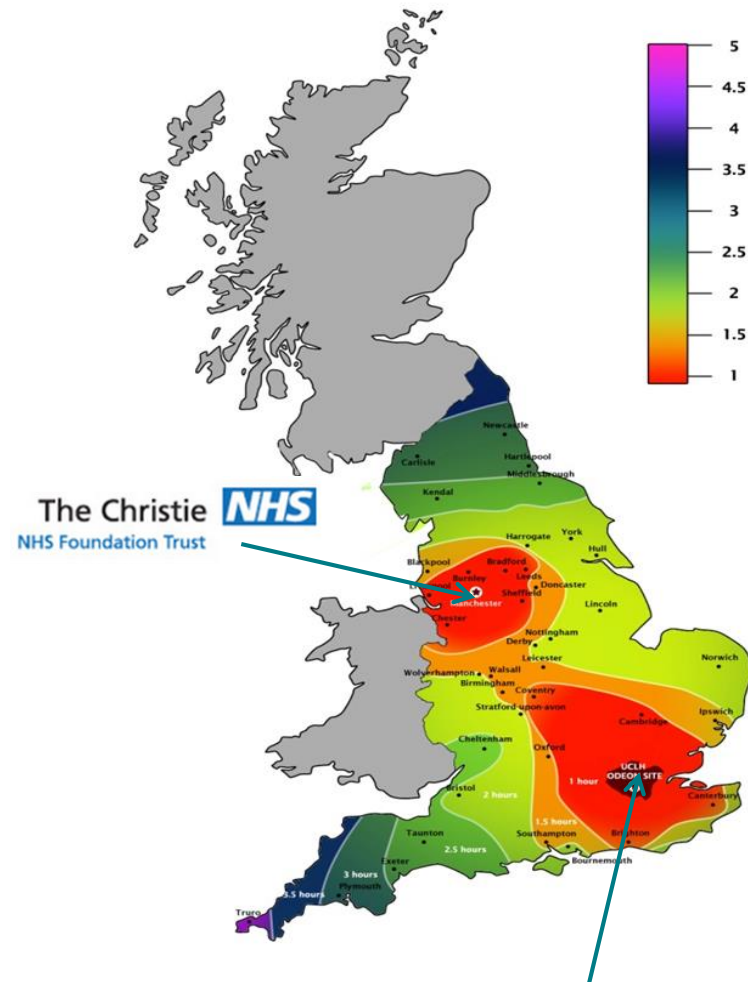


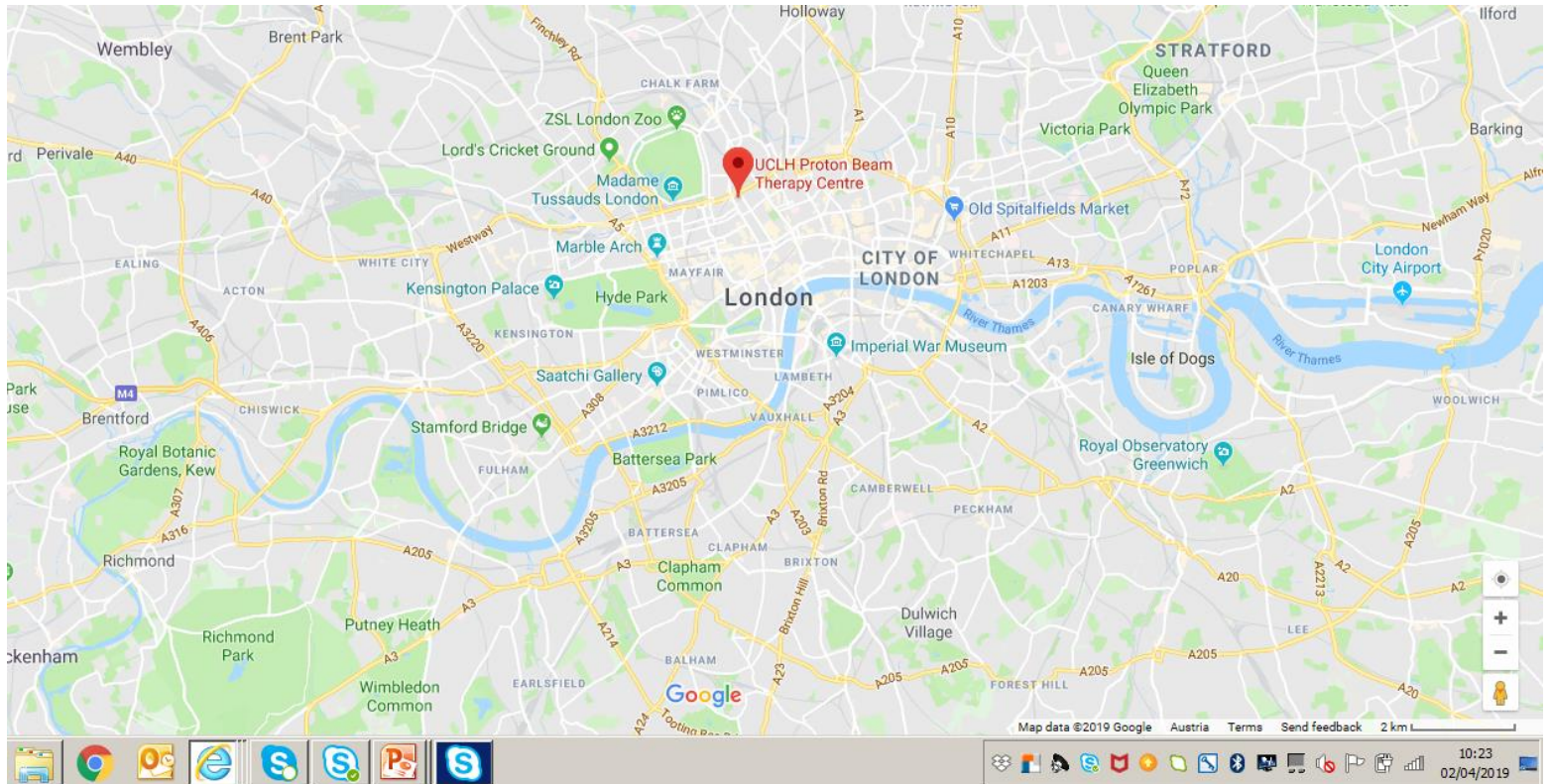
Treatment Planning NTCP Models

Sarah Gulliford
2nd April 2019

- In 2012 NHS England confirmed the two National PBT centres would be the Christie Hospital in Manchester and UCLH in London



UCLH Proton Beam Therapy Centre

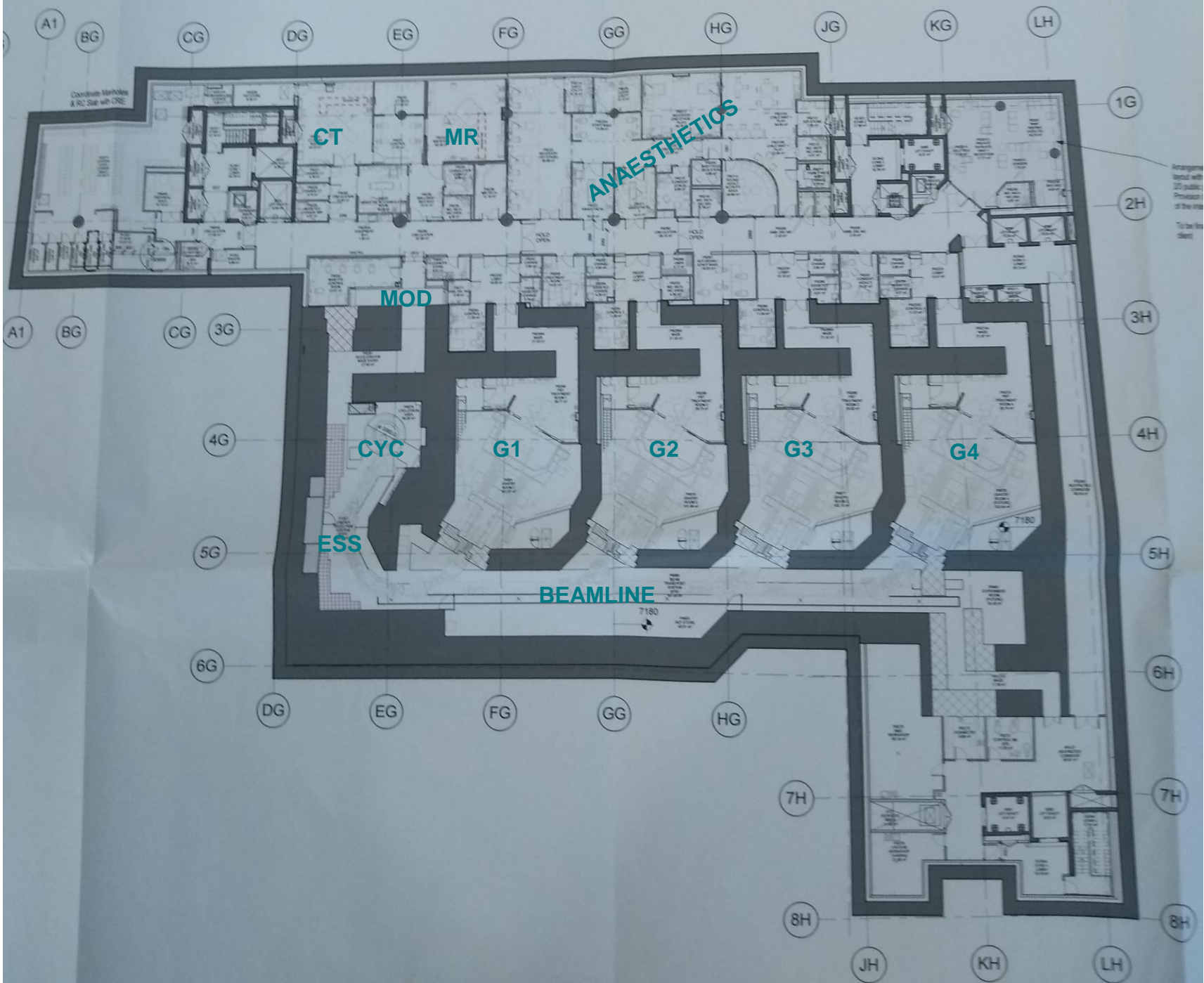




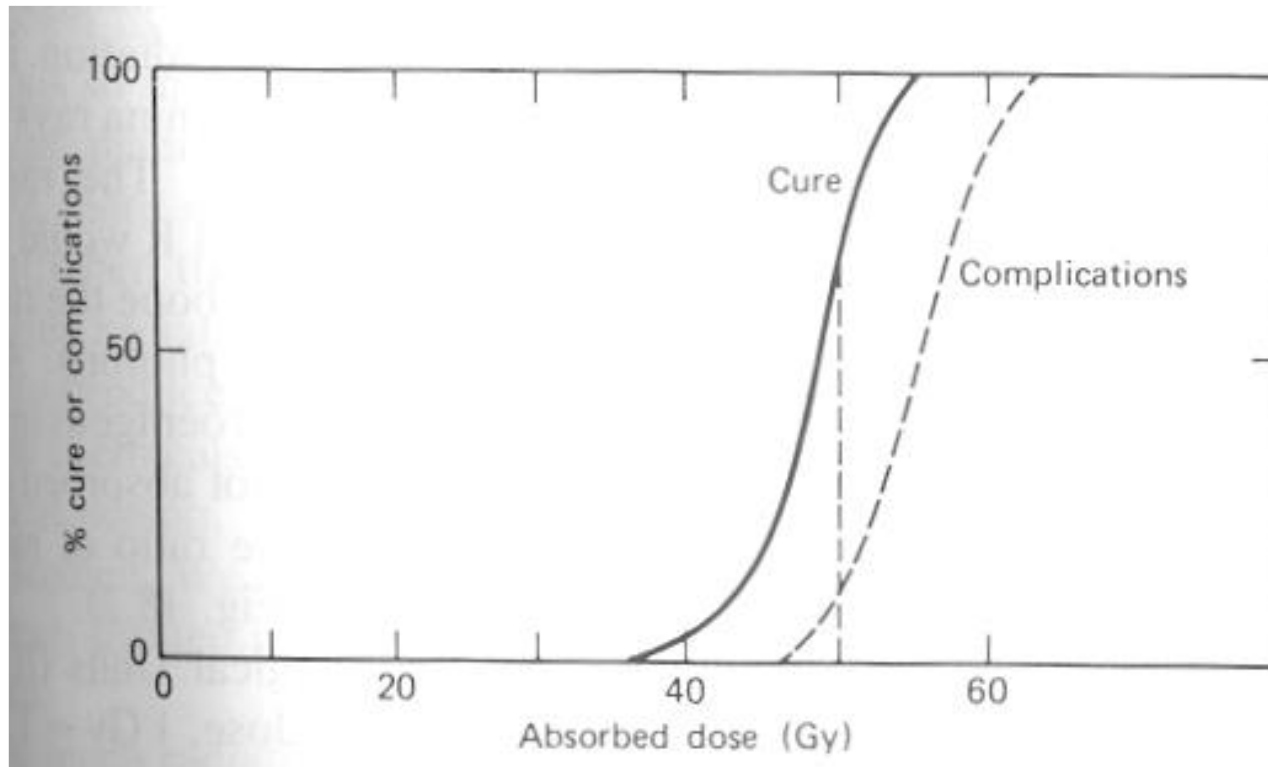
London

Section of Gantry 4 being lowered





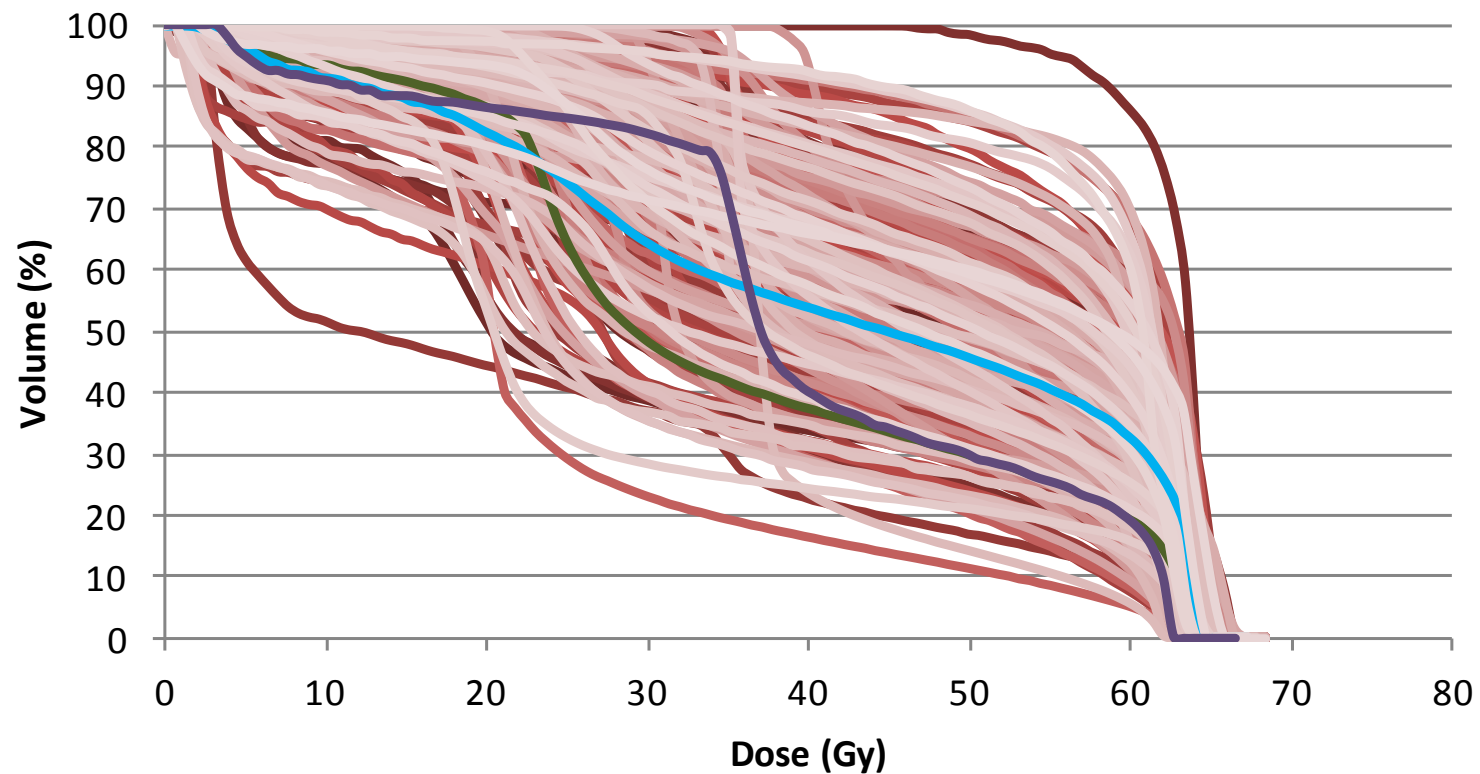
Where do we start?



- We want to characterise the response of normal tissues so that we can design treatments that minimise the risk of toxicity and maximise the probability of “cure”

Inhomogeneity of dose distribution

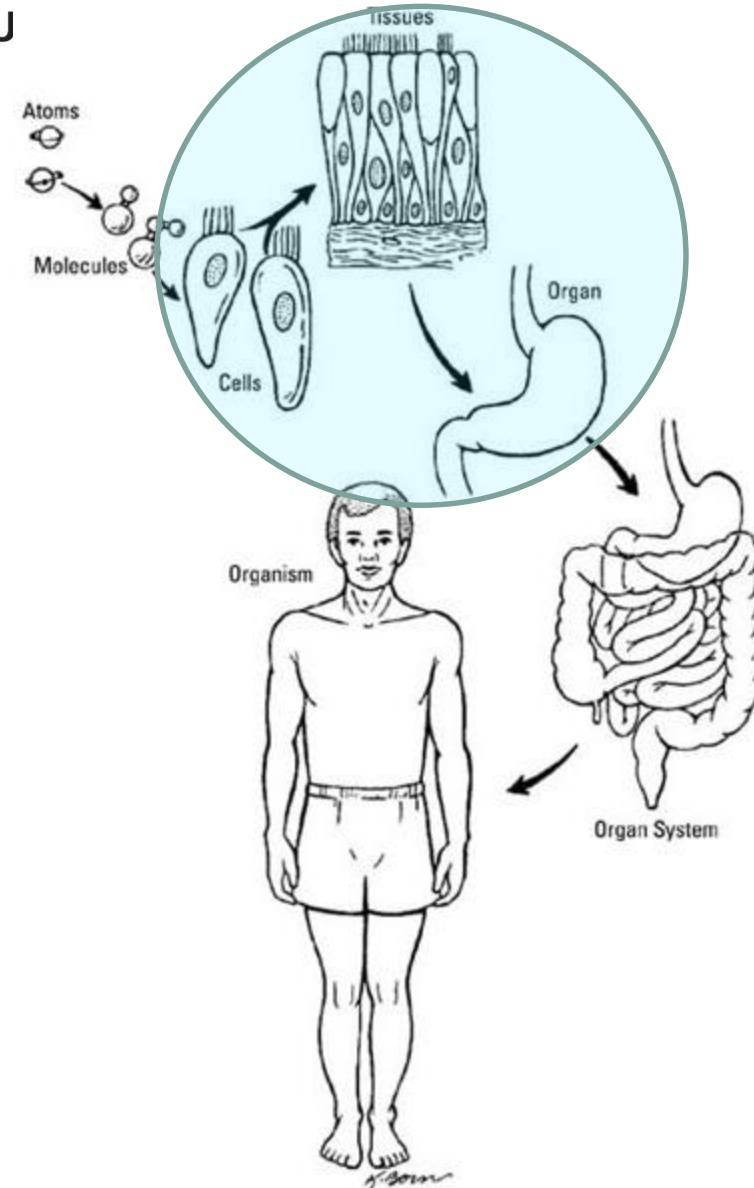
Dose Volume Histograms (Rectum)



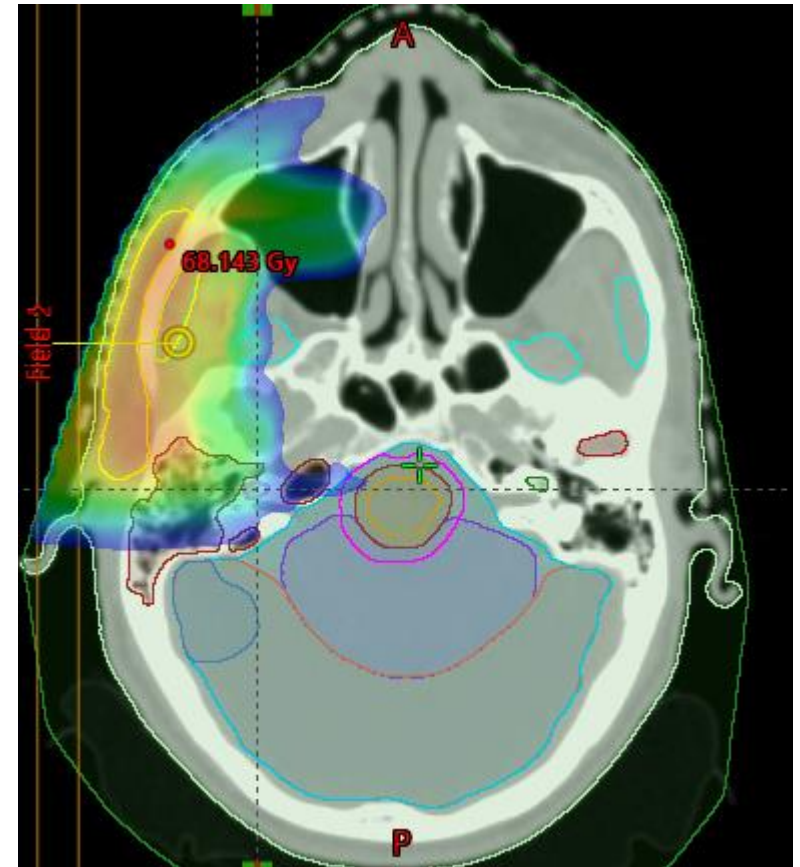
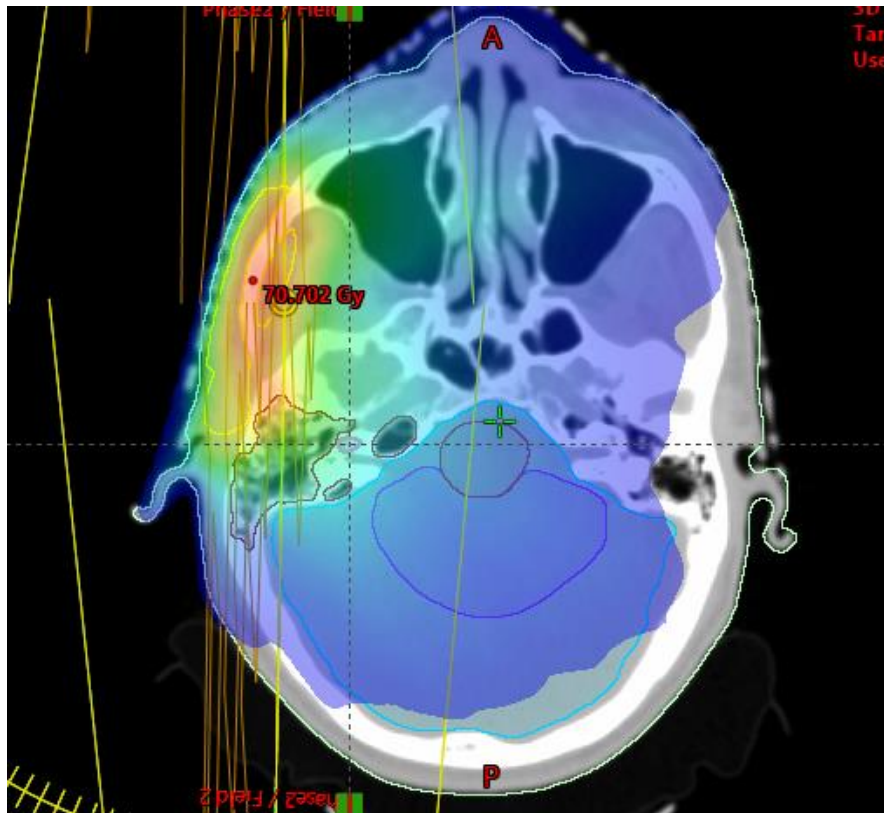
U

Normal tissue toxicity

- Cells
- Tissues
- Organs



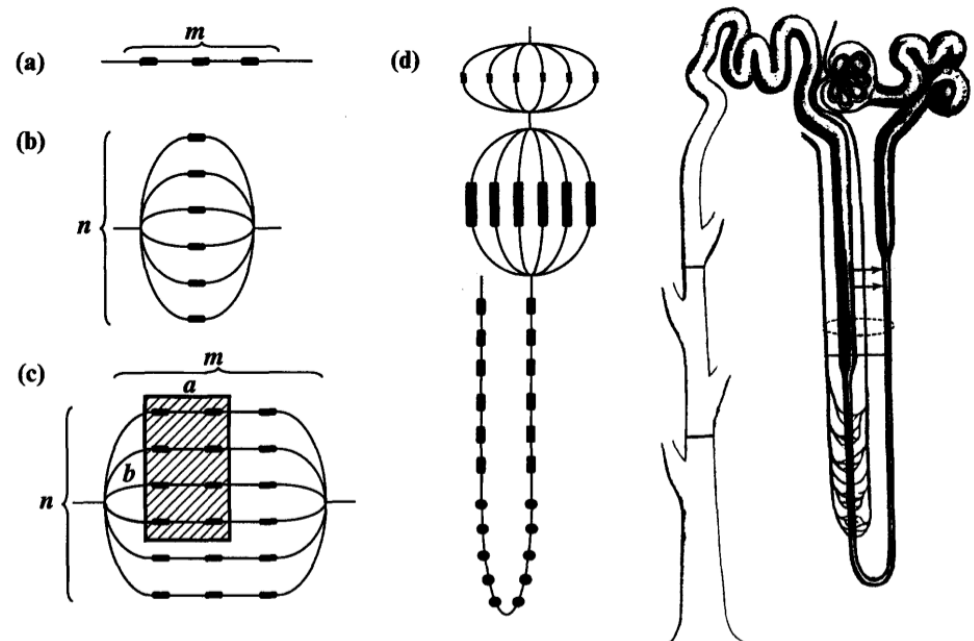
Plan Comparison



Organ architecture

- Each organ different in terms of FSU interaction
- Parallel type response like a rope: it can perform its function even if some strands break.

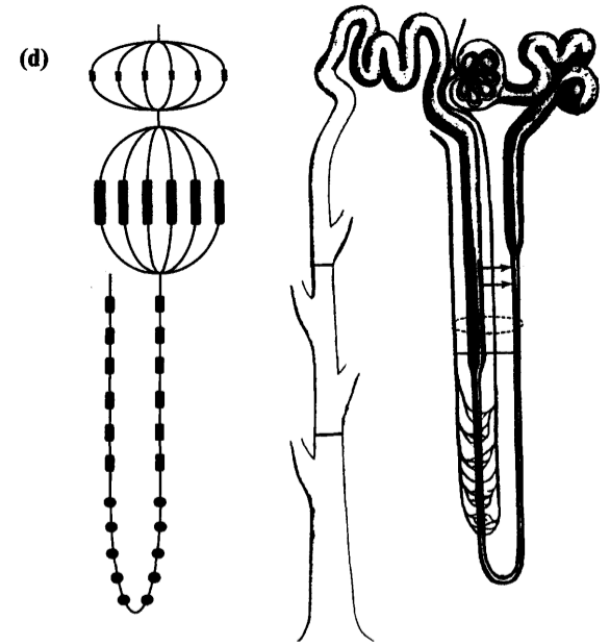
- Large volume effect
- Serial type response like a chain: function is lost if any one link breaks.
 - Small volume effect



Organ architecture

In some tissues FSU are anatomically defined structures

“An example of the parallel-serial model applied to a functional subunit of kidney, a nephron, is shown in d). The first parallel structure is the capillary system inside the glomerular capsule, followed by the capsule itself and the limbs and Henle's loop.



In other tissues there is no anatomical demarcation of the FSU

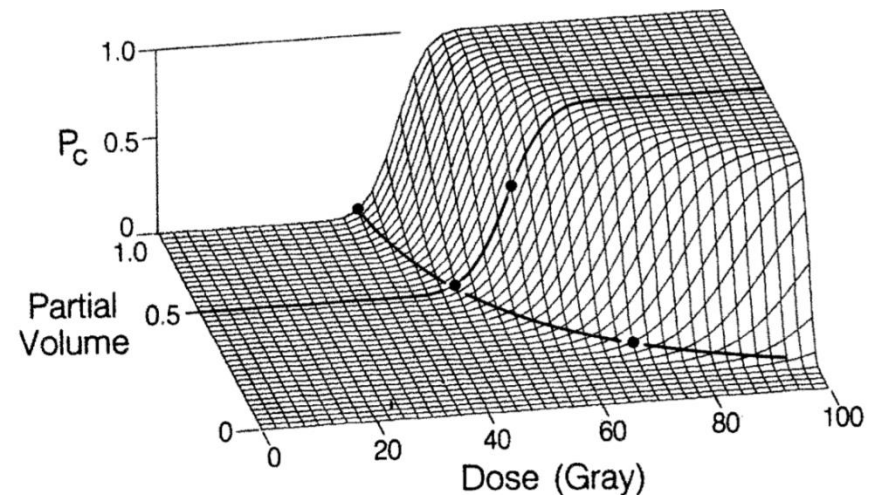
Skin
Mucosa
Spinal cord

Volume effect

Sigmoid dose-response for partial (uniform) organ irradiation:

The volume effect depends on:

- Functional reserve
- Migration of cells
- Stochastic tissue damage
- Inflammatory response?



Nb volume effect \neq FSU in planning terms

- Quantitative Analysis of Normal Tissue Effects in the Clinic (QUANTEC).
- IJROBP 76 (3) Supplement 2010

Introductory Papers

History/Overview/Scientific Issues

Application of QUANTEC metrics/models into clinical practice

Organ-Specific Papers

- Brain
- Optic Nerve/Chiasm
- Brain Stem
- Spinal Cord
- Ear
- Parotid
- Larynx/Pharynx
- Lung
- Heart
- Esophagus
- Liver
- Stomach/Small Bowel
- Kidney
- Bladder
- Rectum
- Penile Bulb

Vision Papers

True Dose
Imaging
Biomarkers
Data Sharing
Lessons of QUANTEC

Each with 10 sections

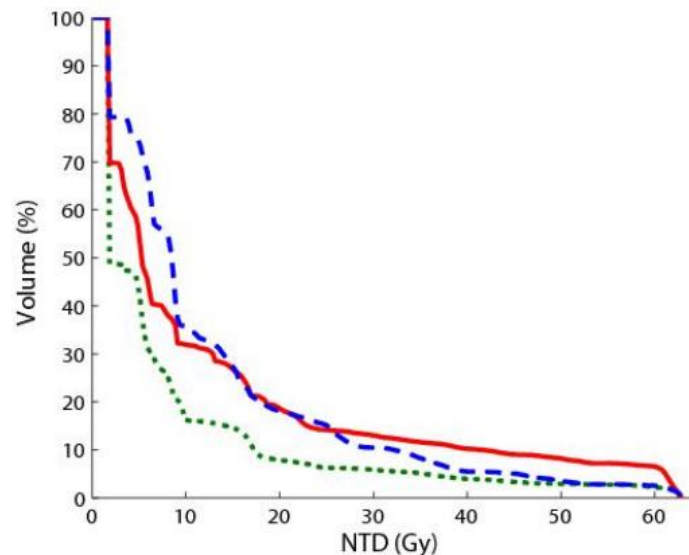
- Clinical Significance**- Describes the clinical situations where the organ is irradiated, and the incidence/significance of organ injury.
- Endpoints**- Describes the different endpoints often considered when assessing injury, the impact of endpoint-selection on the reported injury rates, the challenges/utilities of different endpoints, and the time course of organ injury.
- Challenges Defining Volumes**- Describes how the organ is typically defined (or segmented) on treatment planning images. Includes a discussion of uncertainties/challenges in organ definition (e.g. changes in organ volume/shape during therapy), and the associated impact on DVH's and dose/volume/outcome analyses.
- Review of Dose/Volume Data**- A comprehensive summary of reported 3D dose/volume data for clinically-relevant outcomes.
- Factors Affecting Risk**- Other clinical factors affecting the risk of injury are noted (e.g. age, combined modality therapy, dose fractionation).
- Mathematical/Biological Models**- Models that have been used to relate 3D dose/volume data to clinical outcomes are summarized, along with associated model parameters, limitations and uncertainties.
- Special Situations**- Most of the data discussed relates to conventional fractionation. This section describes situations where the presented data/models may not apply (e.g. hypo-fractionation).
- Recommended Dose/Volume Limits**- The available information is condensed into meaningful dose/volume limits, with associated risk rates, to apply clinically.
- Future Toxicity Studies**- Describes areas in need of future study.
- Toxicity Scoring**- Recommendations on how to score organ injury.

Common dose parameters

- Lots of data on normal tissue response have been collected over the last 20 years.
- Different dose parameters correlate with toxicity
 - Maximum dose
 - Mean dose
 - V_x (the volume receiving at least x Gy)
- The relevant dose depends on the organ architecture, the size of the functional reserve.
- No reserve: maximum dose?
- Large reserve: V_x or mean dose?

Dose volume histogram (DVH)

- (Cumulative) dose volume histograms (DVHs) are often used to evaluate treatment plans.
- Useful for comparing alternative plans
- If two DVHs overlap it is not clear which is better.
- DVH needs to be reduced to a single parameter.



Regression-based NTCP models

- $$\text{NTCP} = \frac{1}{1 + e^{-(\beta_0 + x_1\beta_1 + \dots)}}$$
- Model with a continuous output between 0 & 1.
- Coefficients fitted with multivariate logistic regression

Radiotherapy planning

Parotid gland

Spinal cord

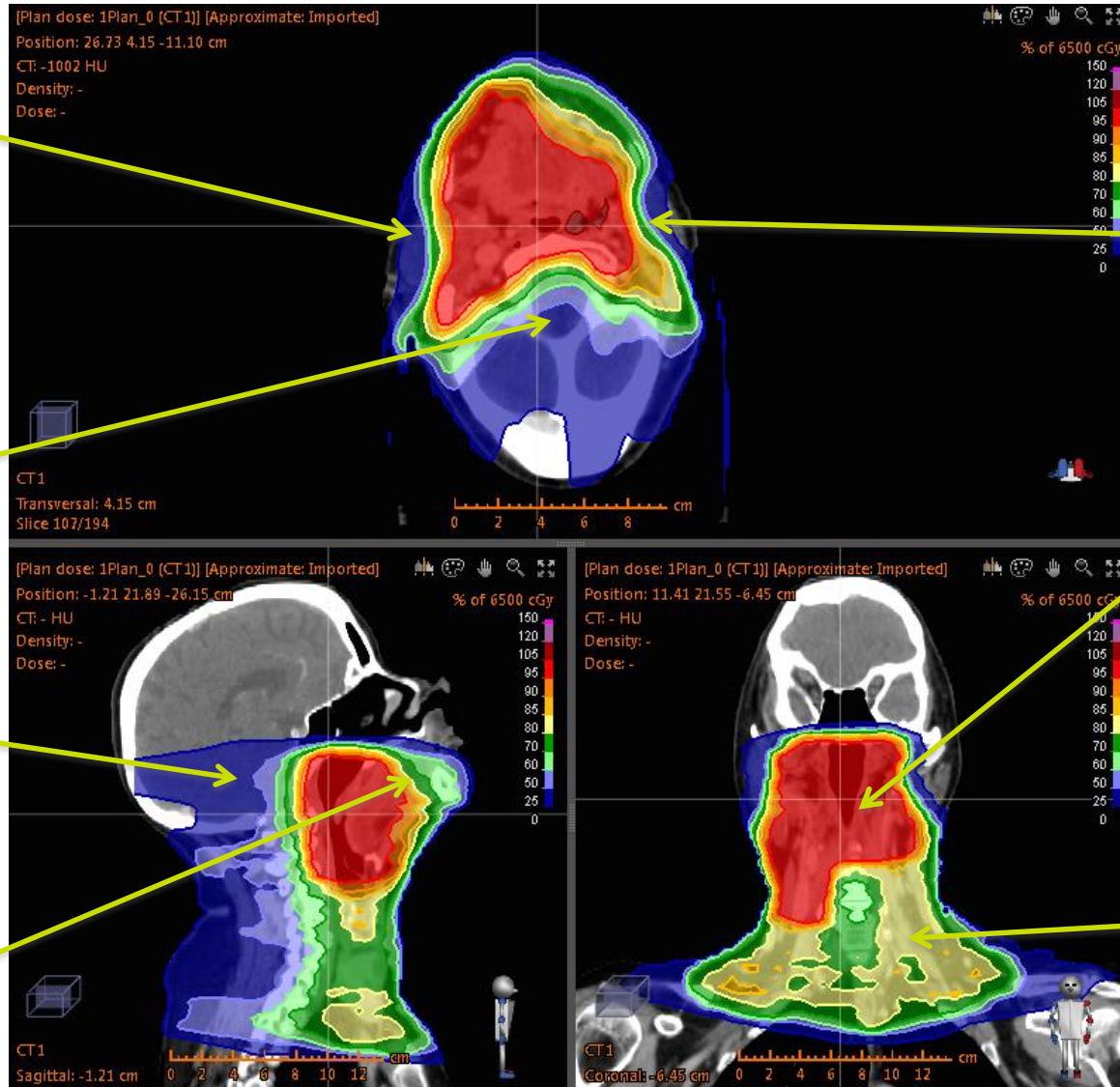
Brain stem

Oral cavity

Parotid gland

Primary planning target volume

Nodal planning target volume



Clinical problem

- Oral mucositis
- Dysphagia
- Severe impact on quality of life
- Acute and often transient
- Consequential late effects
- Limits dose-escalation and accelerated fractionation



PARSPORT	71	Oropharynx, hypopharynx	Bilateral; Conventional, IMRT	No
COSTAR	78	Parotid gland	Unilateral; Conventional, IMRT	No
Dose Escalation	30	Hypopharynx, larynx	Bilateral; IMRT	Yes
Midline	117	Oropharynx	Bilateral; IMRT	Yes
Nasopharynx	36	Nasopharynx	Bilateral; IMRT	Yes
Unknown Primary	19	Unknown primary	Bilateral; IMRT	Yes

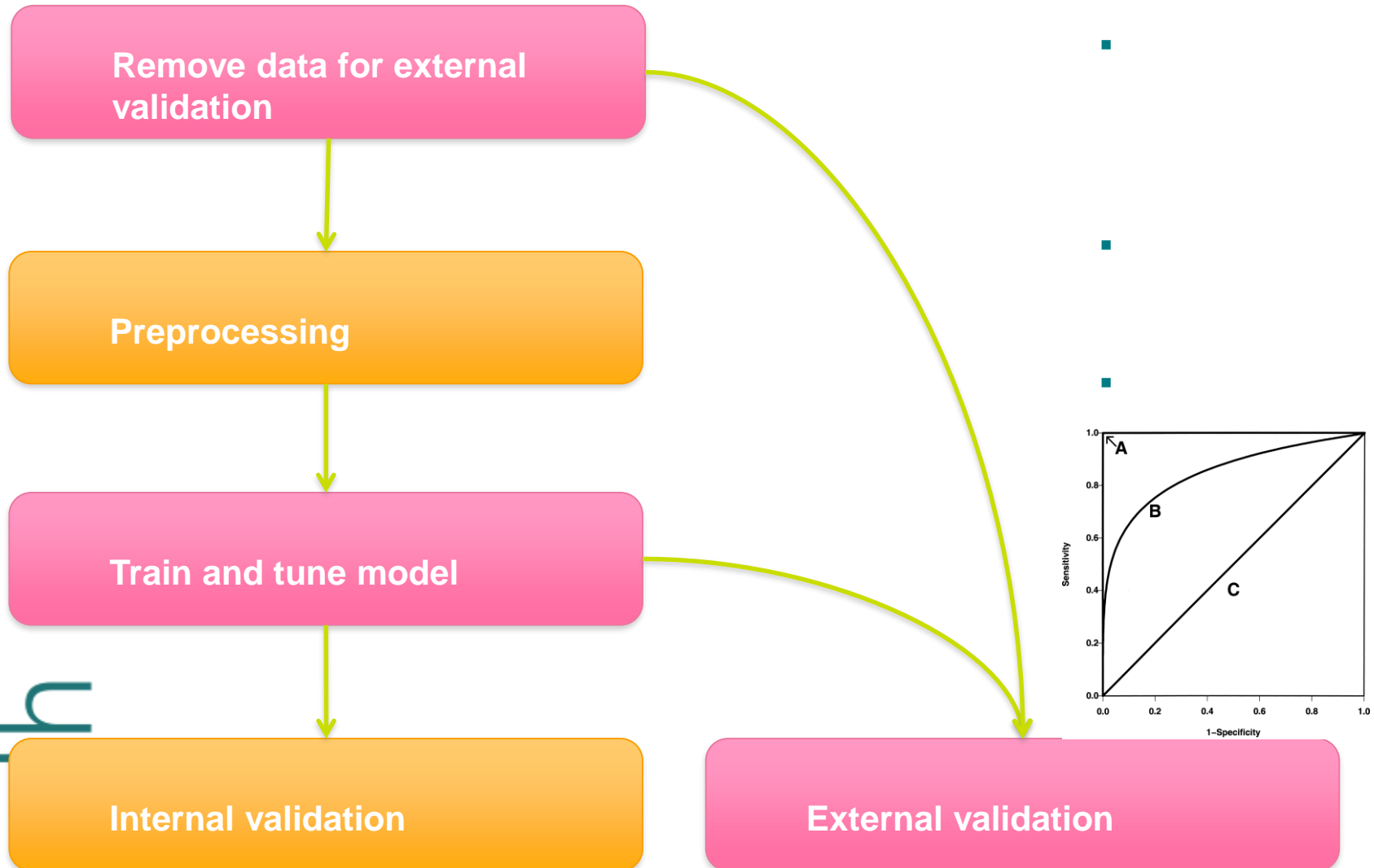
Toxicity scoring

Clinical oral mucositis	Erythema of the mucosa	Patchy ulcerations	Confluent ulcerations	Tissue necrosis; significant spontaneous bleeding
Dysphagia	Symptomatic, able to eat regular diet	Symptomatic and altered eating/swallowing; IV fluids indicated < 24 hours	IV fluids, tube feedings, or TPN indicated \geq 24 hours	Life threatening consequences

- Prospectively measured at baseline, weekly during and 1, 2, 3, 4 and 8 weeks post-radiotherapy
- Patients with missing data excluded
- **Peak grade < 3 vs \geq 3**

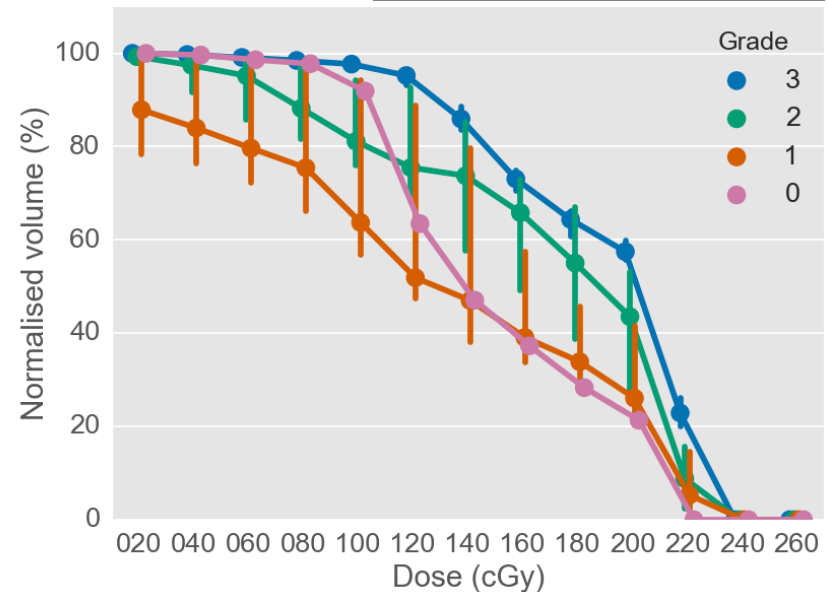
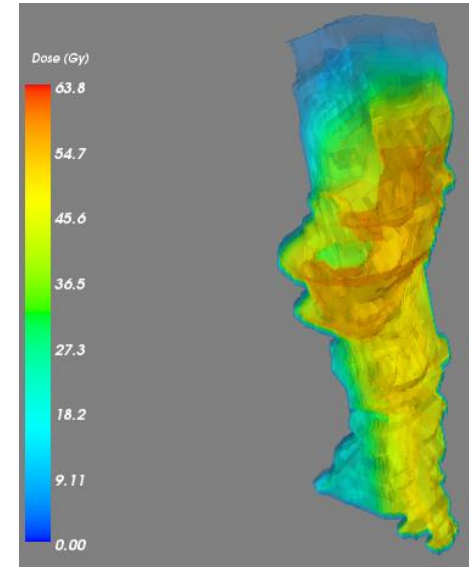
Clinical data

- Age
- Sex
- Primary disease site
- Definitive radiotherapy vs postoperative radiotherapy
- Concomitant treatments
 - Induction chemotherapy
 - Concurrent chemotherapy regime
- No smoking, alcohol or genetic data



Dysphagia modelling

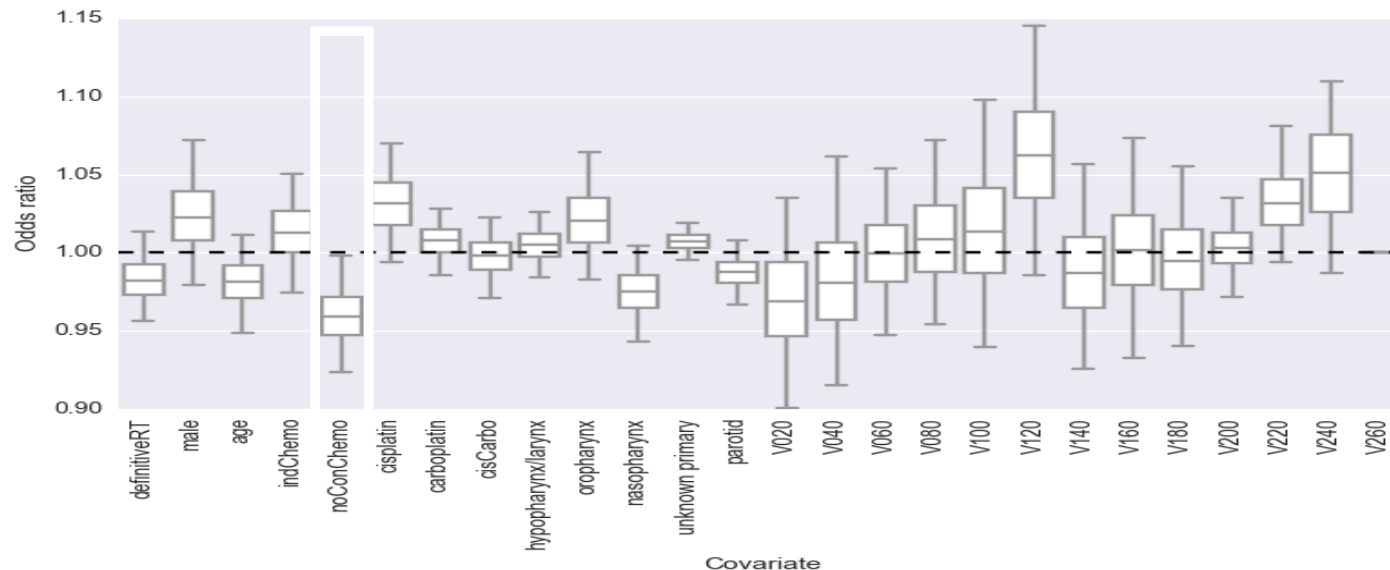
- Feeding tube dependence
 - Pharyngeal mucosa
 - Dose-volume
-
- Training data –
 - Independent external validation with University of Washington cohort – **24 patients**



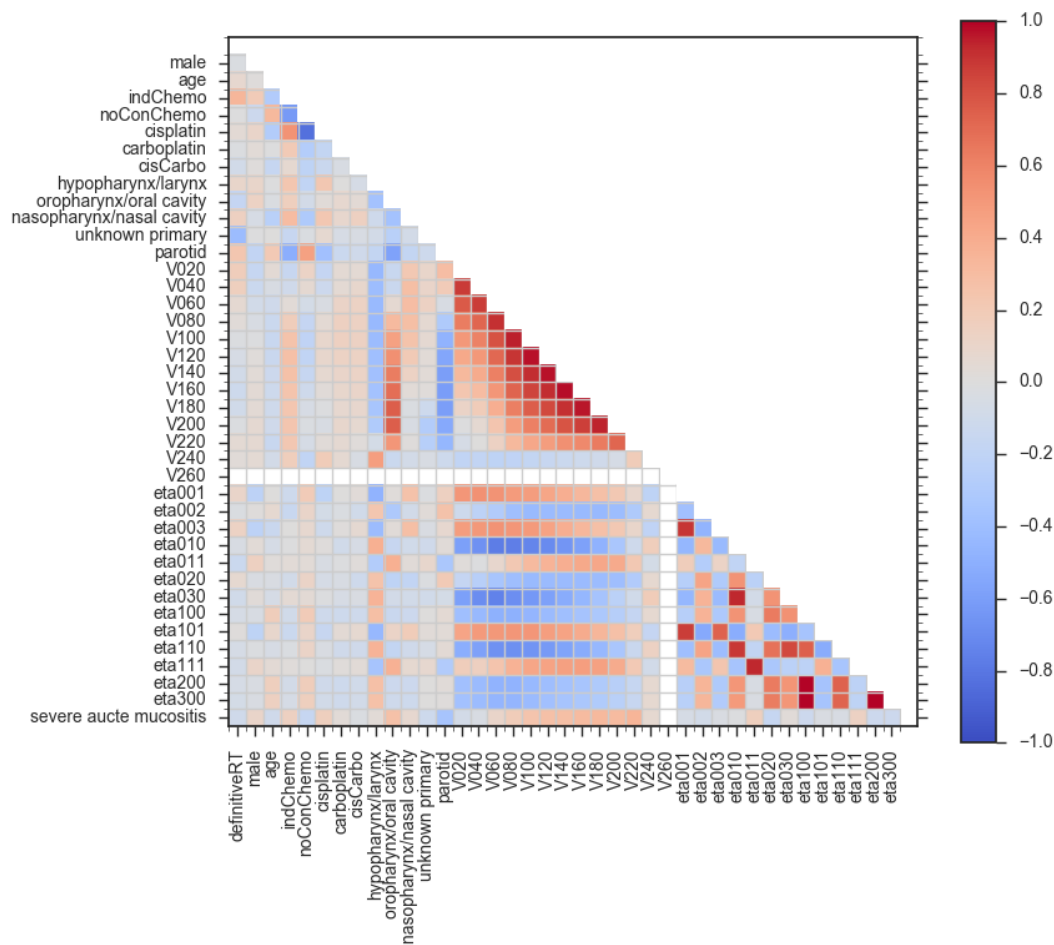
Dysphagia

Model chosen	Logistic regression	Logistic regression
Internal validation AUC (mean (s.d.))	0.77 (0.08)*	0.76 (0.08)
External validation AUC	0.82	

*selected



Correlated Variables



Conclusions

- Trained and validated predictive model of swallowing dysfunction
- High performance model for severe dysphagia
- - No concurrent chemotherapy has highest absolute regression coefficient
- - V70 has highest regression coefficient out of dose-volume features

Lyman Model 1985

$$NTCP = \frac{1}{\sqrt{2\Pi}} \int_{-\infty}^t e^{-t^2/2} dt$$

where

$$t = \frac{D - TD_{50}(v)}{m * TD_{50}(v)}$$

$$TD_{50}(v) = TD_{50}(1) / V^n$$

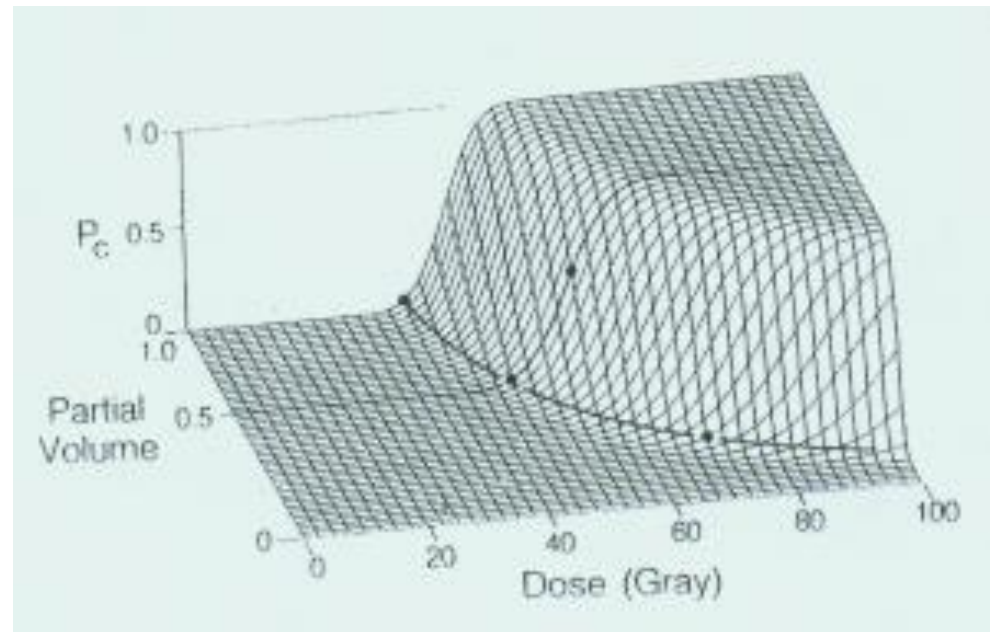
D maximum dose to the structure

TD₅₀(1) tolerance dose for 50% incidence of complication

m slope of TD₅₀(1)

n indicates serial/parallel nature of the structure

v effective volume if the structure was irradiated uniformly to D

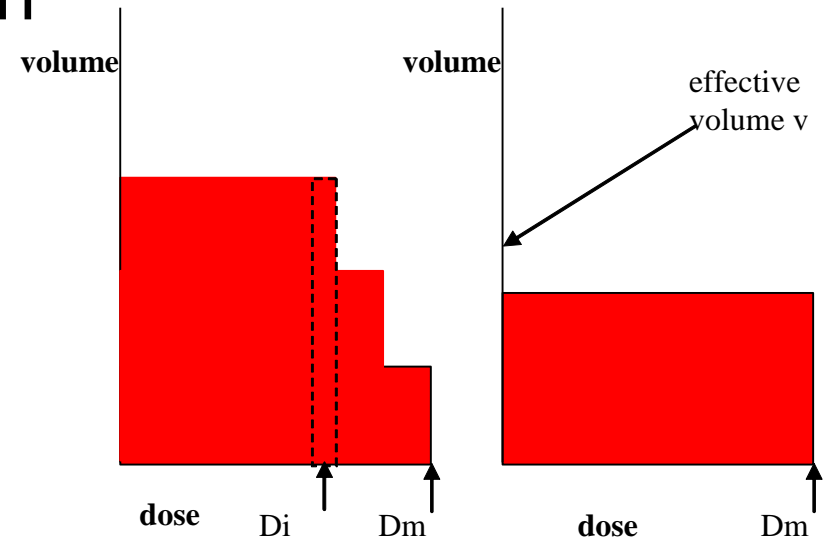


Lyman, J.T. Complication probability as assessed from dose-volume histograms.

Radiat. Res. Suppl 8, S13-S19 (1985).

■ Dose Volume Histogram Reduction

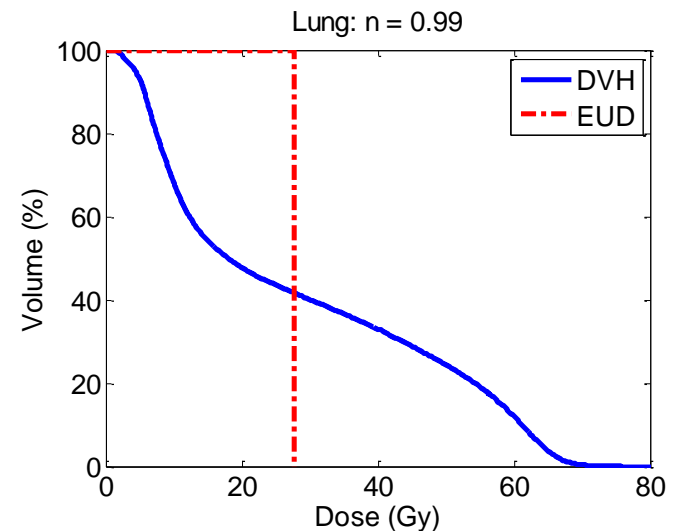
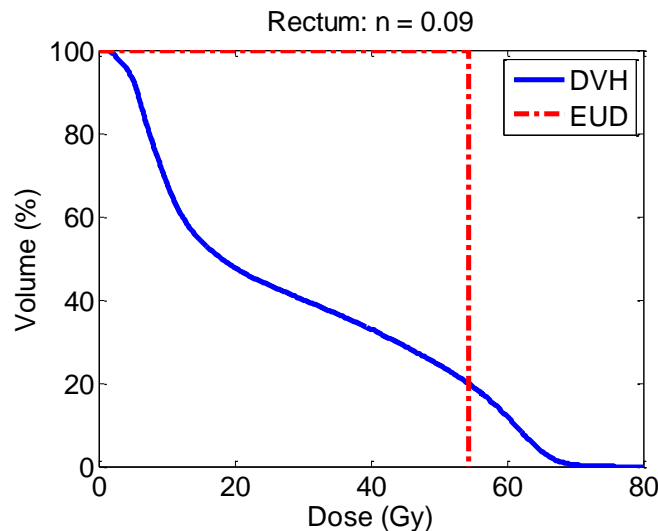
$$v = \sum (D_i / D_m)^{1/n} \Delta v_i$$



Kutcher, G.J. et al, Histogram reduction method for calculating complication probabilities for three-dimensional treatment planning evaluations. *Int. J. Radiat. Oncol. Biol. Phys.* 21, 137-146 (1991).

Equivalent uniform dose

$$EUD = \sum_i \left(D_i^{\frac{1}{n}} \frac{V_i}{V_{tot}} \right)^n$$



Parameter ' n ' is organ (endpoint) specific

- Low (≈ 0): small volume effect, $EUD \approx \text{max dose}$
- High ($= 1$): large volume effect, $EUD = \text{mean dose}$

Maximum Likelihood Estimation

$$\begin{aligned} \text{LLH}(\text{TD50}(1), m, n) = & \sum_{y(i)=1} \ln(\text{NTCP}(\text{TD50}(1), m, n)) \\ & + \sum_{y(i)=0} \ln(1 - \text{NTCP}(\text{TD50}(1), m, n)) \end{aligned}$$

- **y(i) known outcome**

Bootstrapping

- Sample available data (with replacement) to create different populations
- Test generalisability of fit.
- Results shown for 1000 Bootstrapped cohorts

Fitting NTCP models :an example

- RT01 Prostate Radiotherapy Trial
Patients randomised to 64 Gy vs 74 Gy
- Accrual from Jan 1998 to Dec 2001
- Patients randomised from 23 centres
throughout the UK, 1 in NZ, 1 in Victoria

Fitting NTCP models :an example

- Rectal Bleeding -RMH
- Proctitis-RTOG
- Stool Frequency- Lent Som
- Loose Stools -UCLA QoL
- Rectal Urgency -UCLA QoL

0=none 1=mild 2=moderate/severe toxicity

1000 Bootstrapped cohorts fitted using
maximum likelihood estimation.

Bootstrap Approach

Endpoint (toxicity grades)	TD50(1) (Gy)	SD	<i>m</i>	SD	<i>n</i>	SD
Rectal bleeding (G1&2)	59.2	3.5	0.29	0.12	0.17	0.12
Proctitis (G1&2)	57.3	4.1	0.33	0.15	0.2	0.17
Stool frequency (G1&2)	62.6	6.3	0.6	0.21	0.36	0.27
Loose stools (G1&2)	59.9	4.6	0.34	0.15	0.38	0.3
Rectal urgency (G1&2)	54.1	5.4	0.52	0.22	0.4	0.31
Rectal bleeding (G2)	68.9	4.2	0.16	0.05	0.18	0.14
Proctitis (G2)	68.3	4.6	0.22	0.07	0.17	0.08
Stool frequency (G2)	155.4	27.4	0.46	0.07	0.29	0.31
Loose stools (G2)	74.0	7.2	0.25	0.09	0.45	0.36
Rectal urgency (G2)	70.1	8.0	0.37	0.12	0.45	0.26

Bootstrap Approach

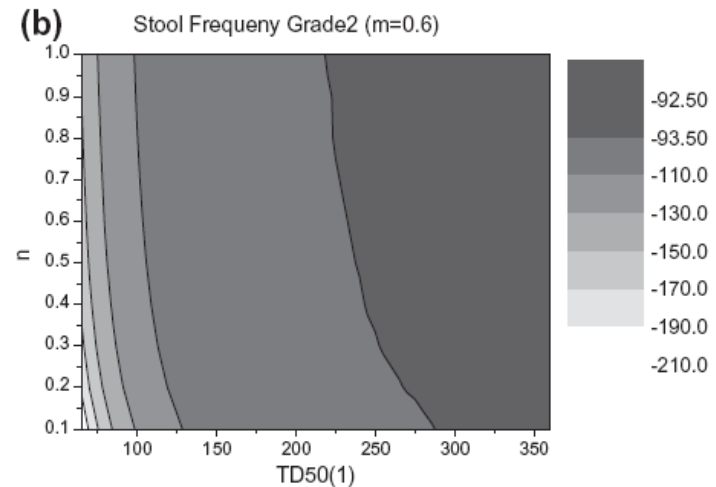
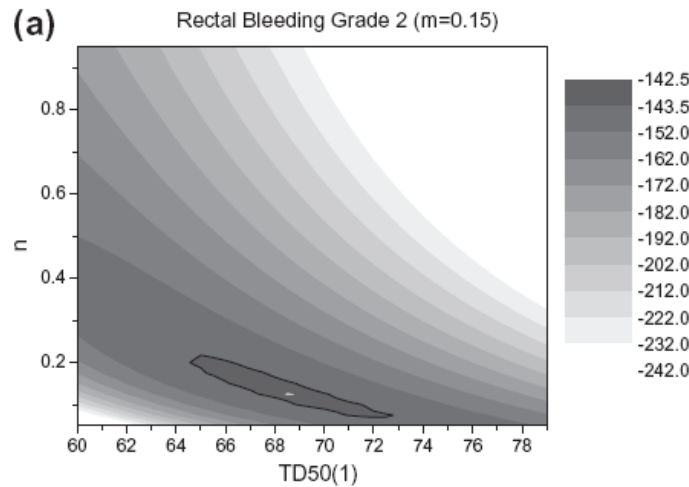
Endpoint (toxicity grades)	TD50(1) (Gy)	SD	<i>m</i>	SD	<i>n</i>	SD
Rectal bleeding (G1&2)	59.2	3.5	0.29	0.12	0.17	0.12
Proctitis (G1&2)	57.3	4.1	0.33	0.15	0.2	0.17
Stool frequency (G1&2)	62.6	6.3	0.6	0.21	0.36	0.27
Loose stools (G1&2)	59.9	4.6	0.34	0.15	0.38	0.3
Rectal urgency (G1&2)	54.1	5.4	0.52	0.22	0.4	0.31
Rectal bleeding (G2)	68.9	4.2	0.16	0.05	0.18	0.14
Proctitis (G2)	68.3	4.6	0.22	0.07	0.17	0.08
Stool frequency (G2)	155.4	27.4	0.46	0.07	0.29	0.31
Loose stools (G2)	74.0	7.2	0.25	0.09	0.45	0.36
Rectal urgency (G2)	70.1	8.0	0.37	0.12	0.45	0.26

Quantec Values Grade 2 Toxicity

TD50 76.9Gy $n=0.09$ $m=0.13$

Maximum Likelihood Estimation

Log-likelihood values



Taking empirical models further

- Accounting for confounding factors
 - Confounding factors include any non-dosimetric factors (X) influencing the outcome, as these are not modelled.
 - Health status, chemo/surgery, radiosensitivity etc.
 - Sharper slope, greater certainty

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

$$t = \frac{\varphi - \varphi_{50}}{m \cdot \varphi_{50}} + \beta_2 X_2 + \dots + \beta_n X_n$$

The Critical Volume

In the Critical Volume model the summary measure is a damaged volume, calculated using a local dose-effect function $E(D_i)$.

$$\text{Damaged volume} = \sum_i E(D_i) \frac{V_i}{V_{tot}}$$

$$E(D_i) = \frac{1}{1 + (D_{50}/D_i)^k}$$

D_i = total dose to bin i
 V_i = absolute volume in bin i
 V_{tot} = total organ volume

D_{50} = dose causing 50% local effect
 k = parameter for slope of curve

$$E(D_i) = \begin{cases} \frac{D_i/D_{50} - 1}{1 + (D_i/D_{50} - 1)^2} + 1/2 & \text{when } D_i < 2D_{50} \\ 1 & \text{when } D_i \geq 2D_{50} \end{cases}$$

Function 1:
2 parameters

Function 2:
1 parameter

$$E(D_i) = \begin{cases} 0 & \text{when } D_i < D_{50} \\ 1 & \text{when } D_i \geq D_{50} \end{cases}$$

Function 3:
1 parameter

The Relative Seriality model

The Relative Seriality model first calculates a probability of local damage for each dose bin in the DVH:

$$P(D_i) = 2^{-\exp\left(e\gamma\left(1-\frac{D_i}{D_{50}}\right)\right)}$$

D_i = total dose to bin i
 D_{50} = Dose causing 50% probability
 γ = slope of the curve

Then NTCP is estimated using the relative seriality parameter 's'.

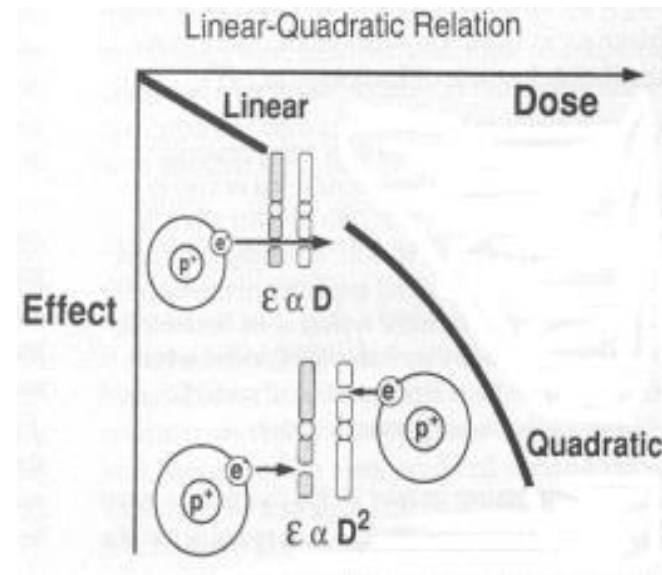
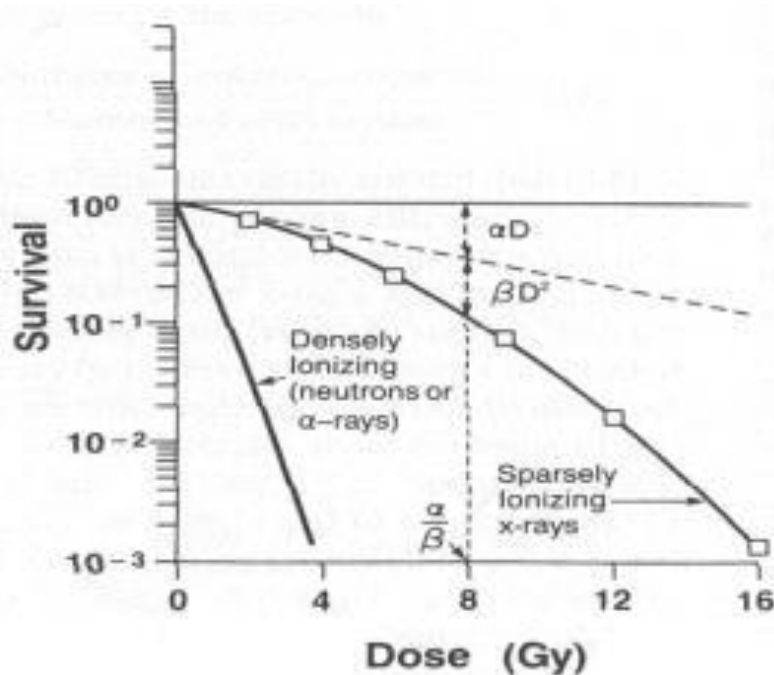
- Serial organs: $s = 1$, high doses to small volume important
- Parallel organs: $s \approx 0$, large irradiated volume important

$$\text{NTCP} = \left(1 - \prod_i (1 - P(D_i)^s)^{\frac{V_i}{V_{tot}}}\right)^{1/s}$$

V_i = absolute volume in bin i
 V_{tot} = total organ volume
 s = relative seriality parameter

Fractionation

Linear Quadratic (LQ) Model



$$S \sim \exp(-\alpha D - \beta D^2)$$

Acute responding tissues $\alpha/\beta \sim 10$ late responding tissues $\alpha/\beta \sim 3$ (classically)

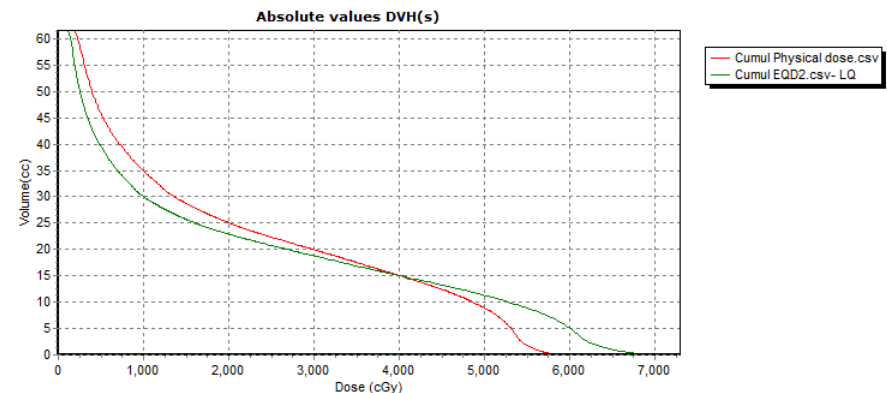
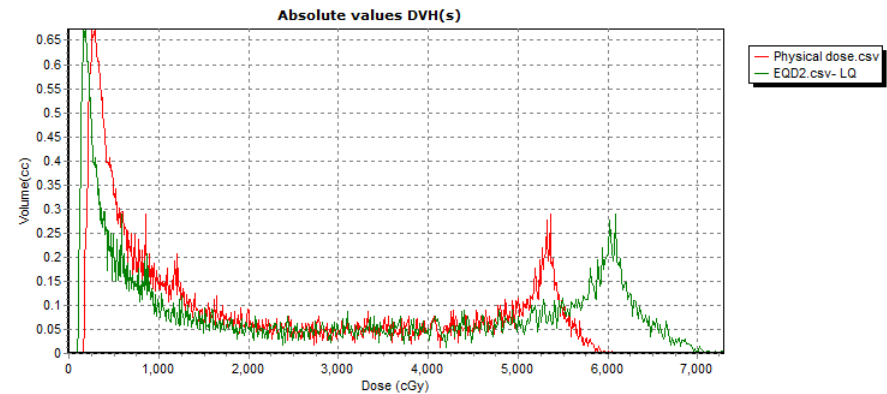
Fractionation correction!

Each bin in the DVH receives a different dose per fraction.

- Convert each dose-bin to 2Gy fraction equivalence (reference conditions).

$$BED = D \left[1 + \frac{d}{(\alpha / \beta)} \right]$$

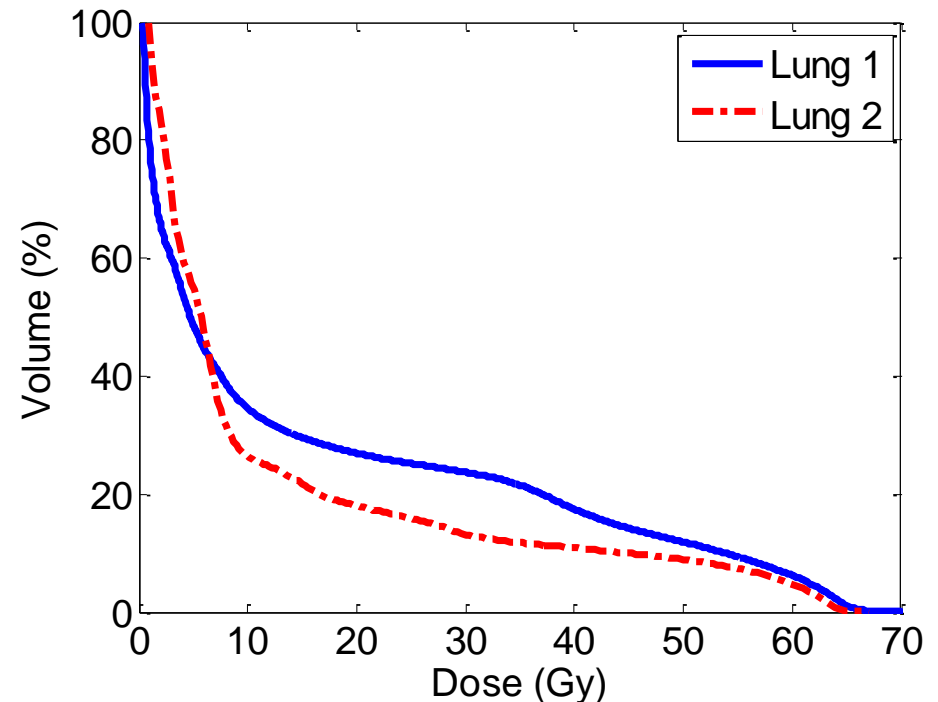
$$D_i^{2Gy} = D_i \left[\frac{d + (\alpha / \beta)}{2 + (\alpha / \beta)} \right]$$



Example

DVHs for 2 lung cancer patients

- Total lung – GTV
- Fractionation correction ($\alpha/\beta = 3$ Gy)
- NTCP estimated using 3 different models
- Parameter values from Seppenwoolde et al. 2003



	Patient 1	Patient 2
LKB	9.2%	5.4%
Critical Volume	9.9%	6.9%
Relative Seriality	8.5%	4.4%

When can I use NTCP?

- ICRU report 83: level 3 dose reporting
- Ranking treatment plans
- Evaluation of treatment data from clinical studies; derive parameter values locally
 - Better to use locally derived parameter values than published values
- Develop new potential techniques
- Dose prescription
- Radiobiological plan optimisation

Which model should I use?

- Preferably several
- The LKB model is the most commonly used model and has parameter values published for many organs/endpoints (e.g. QUANTEC).
- All the above models are empirical.
 - Statistical fits to clinical data
 - Biological interpretation of DVH reduction method should not be relied on blindly.
- The more parameters the more clinical data needed for parameter fitting. Up to 3 parameters often appropriate.
- Datasets often in the order of 100-300 patients. The more 'events', the more information to model on.

Important considerations

- The model is only as good as the data used for parameter fitting.
 - Well-designed study (unbiased, representative)?
 - Sample size, number of events
 - Uncertainties in dose distributions (incl. outlining, organ motion, organ definition)

Important considerations

- The model is only reliable for plans reasonably similar to the plans included in the study (also applies to patient specific factors).
- When a model is used without knowledge of these factors, or for a different technique/patient group, the same limitations apply as to empirical 'tolerance doses'.
- NTCP is continuous but the patient outcome is binary

Model-based selection

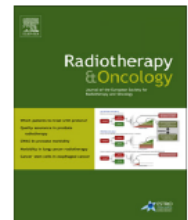
Radiotherapy and Oncology 107 (2013) 267–273



Contents lists available at SciVerse ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com



Proton radiotherapy

Selection of patients for radiotherapy with protons aiming at reduction of side effects: The model-based approach

Johannes A. Langendijk^{a,*}, Philippe Lambin^b, Dirk De Ruyscher^c, Joachim Widder^a, Mike Bos^d, Marcel Verheij^e

^aDepartment of Radiation Oncology, University Medical Center Groningen, University of Groningen, The Netherlands; ^bDepartment of Radiation Oncology (MAASTRO Clinic) & Research Institute GROW, University Hospital Maastricht, The Netherlands; ^cDepartment of Radiation Oncology, University Hospitals Leuven/KU Leuven, Belgium; ^dHealth Council of the Netherlands; ^eDepartment of Radiotherapy, The Netherlands Cancer Institute-Antoni van Leeuwenhoek Hospital, The Netherlands

1.02 in

Patient Selection

- Step 1 NTCP Models
 - Data from optimal photon treatment
- Step 2 In silico planning studies
 - Photon vs Proton
- Step 3 Estimation of clinical benefit

Step 3 Estimation of clinical benefit

270

Selection of patients for protons

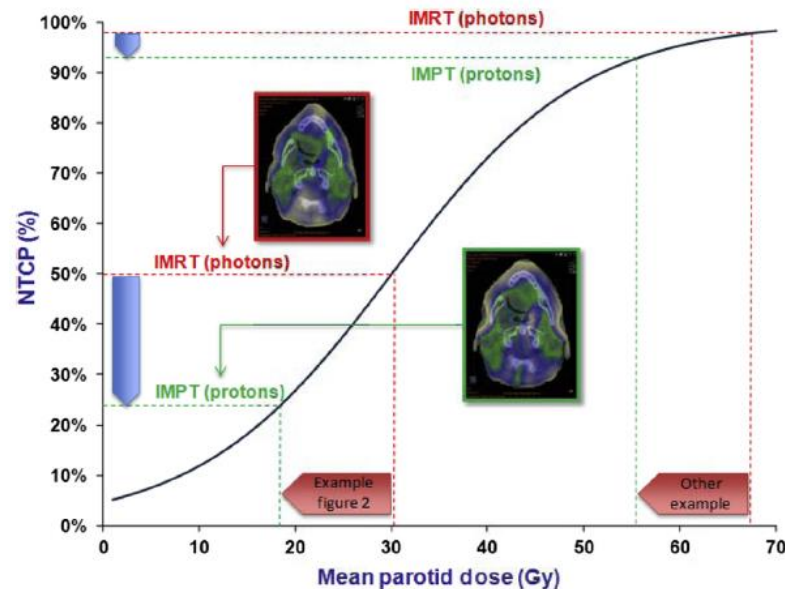
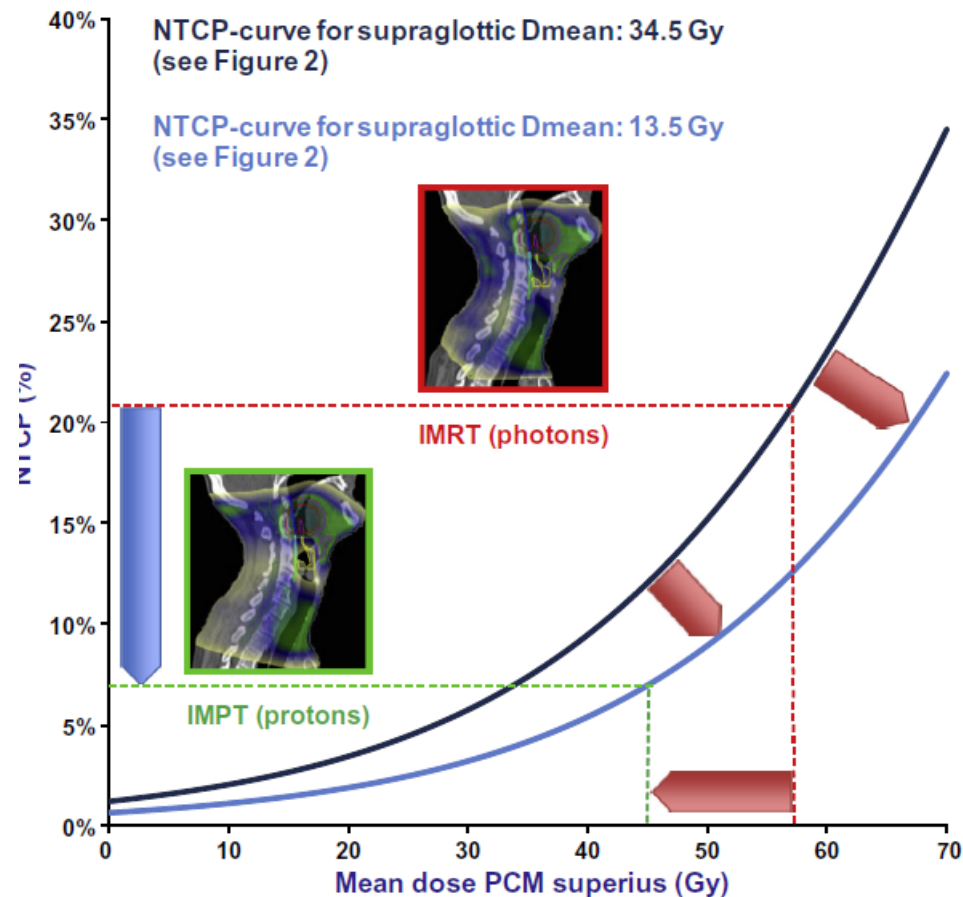


Fig. 3. Translation of the results of the individual ISPC-study depicted in Fig. 2 with regard to xerostomia. The reduction of the mean parotid dose from 30.1 Gy to 18.4 Gy (red arrow: example Fig. 2) corresponds with an estimated NTCP-value reduction for severe xerostomia from 50% to 24% according to the NTCP-model published by Semenenko. However, exactly similar absolute dose reductions (red arrow: other example) result in a minimal estimated NTCP-value reduction when the initial dose is much higher, due to the shape of the NTCP-curve.

Estimation of clinical benefit



NTCP data derived from proton therapy

Radiotherapy and Oncology 130 (2019) 164–171



Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com



Development and validation of NTCP models for acute side-effects
resulting from proton beam therapy of brain tumours

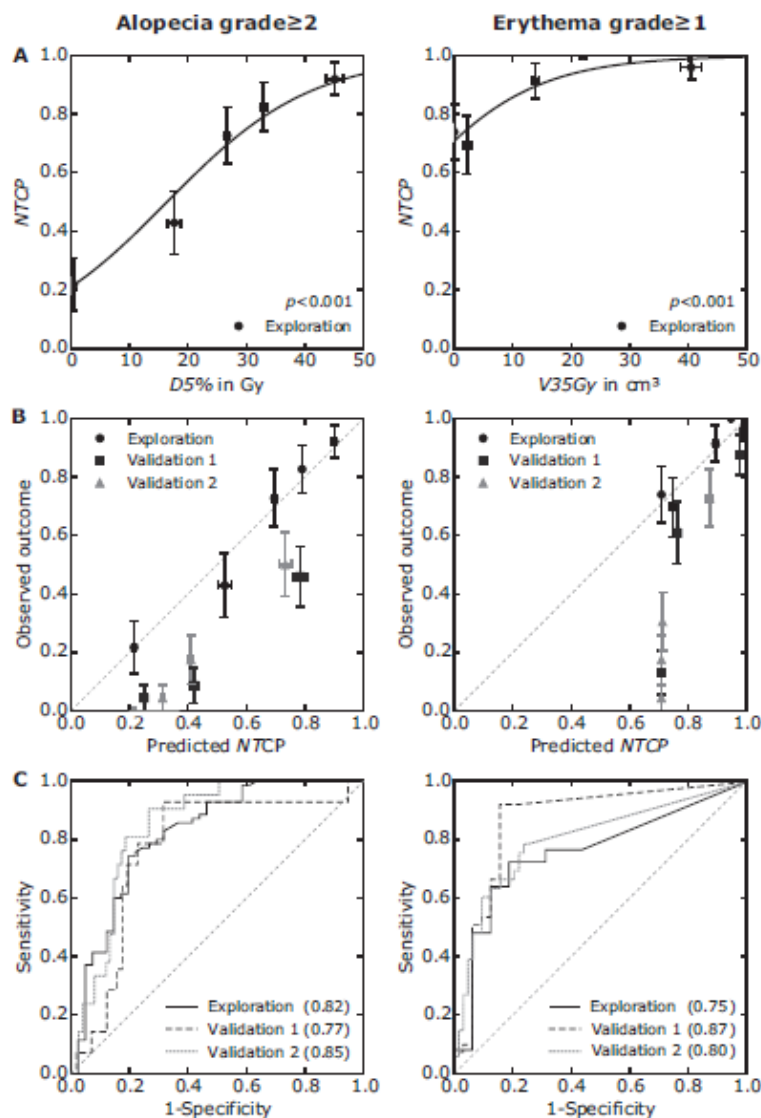


Almut Dutz^{a,b,c,*}, Armin Lühr^{a,b,c}, Linda Agolli^{a,d}, Esther G.C. Troost^{a,b,c,d,e}, Mechthild Krause^{a,b,c,d,e},
Michael Baumann^{a,b,c,d,e,f}, Xavier Vermeren^g, Dirk Geismar^{g,h,i}, Emily F. Schapira^j, Meghan Bussière^j,
Jillian E. Daly^j, Marc R. Bussière^j, Beate Timmermann^{g,h,i,k,1}, Helen A. Shih^{j,1}, Steffen Löck^{a,c,d,e,1}

Model	AUC	(95% CI)	β_0	(95% CI)	β_1	(95% CI)	p-Value	β_2	(95% CI)	p-Value
Erythema grade ≥ 1										
Model parameter					V35Gy Skin (cm ³)					
Exploration	0.75	(0.54–0.90)	1.00	(0.32–1.69)	0.09	(0.02–0.15)	0.008			
Validation 1	0.87	(0.77–0.95)								
Validation 2	0.80	(0.71–0.89)								
Erythema grade ≥ 2										
Model parameter					V35Gy Skin (cm ³)					
Exploration	0.77	(0.64–0.89)	–1.54	(–2.20–0.88)	0.06	(0.03–0.08)	<0.001			
Validation 1	*									
Validation 2	0.84	(0.77–0.91)								
Alopecia grade ≥ 1										
Model parameter					D2% Skin (Gy)					
Exploration	0.88	(0.73–0.99)	–0.94	(–2.14–0.27)	0.10	(0.05–0.15)	<0.001			
Validation 1	0.82	(0.70–0.92)								
Validation 2	0.84	(0.75–0.92)								
Alopecia grade ≥ 2										
Model parameter					D5% Skin (Gy)					
Exploration	0.82	(0.69–0.95)	–1.33	(–2.91–0.47)	0.08	(0.05–0.11)	<0.001			
Validation 1	0.77	(0.63–0.89)								
Validation 2	0.85	(0.76–0.92)								
Fatigue grade ≥ 1										
Model parameter					D2% Brain-CTV (Gy)			Gender		
Exploration	0.68	(0.50–0.84)	–0.90	(–2.30–0.51)	0.03	(0.00–0.06)	0.067	1.28	(0.33–2.23)	0.009
Validation 1	0.45	(0.31–0.61)								
Validation 2	0.52	(0.40–0.64)								

Abbreviations: AUC, area under the receiver operating characteristic curve; VxGy, absolute volume receiving x Gy; Dx% dose in x% of the volume of the organ at risk; CTV, clinical target volume.

* External validation not applicable due to zero incidence.



Using NTCP models to quantify variation in RBE

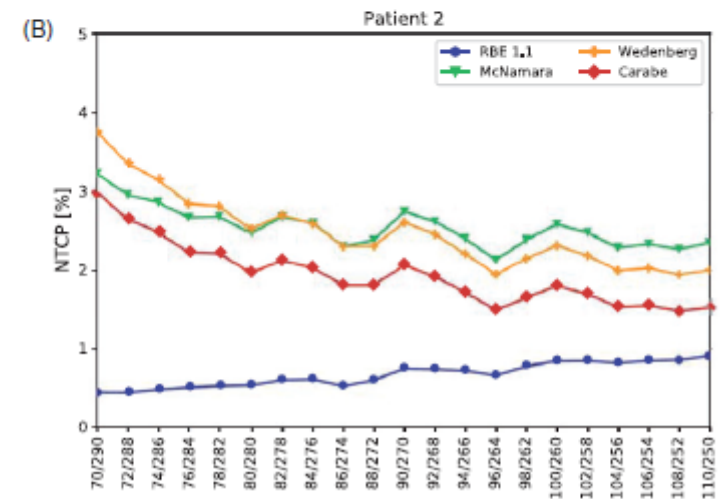
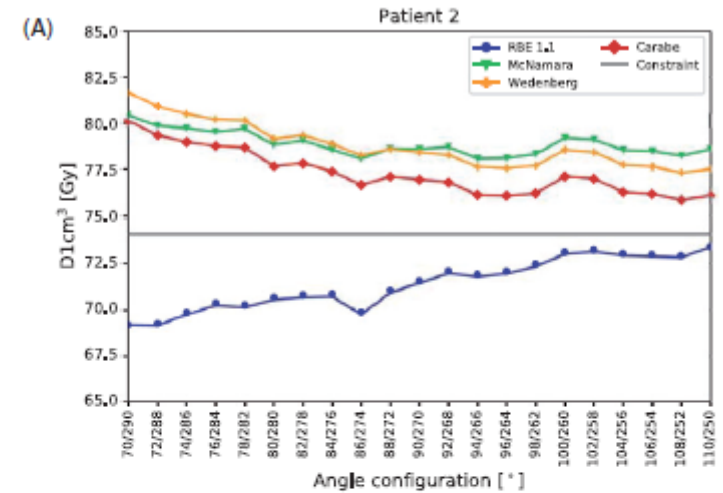
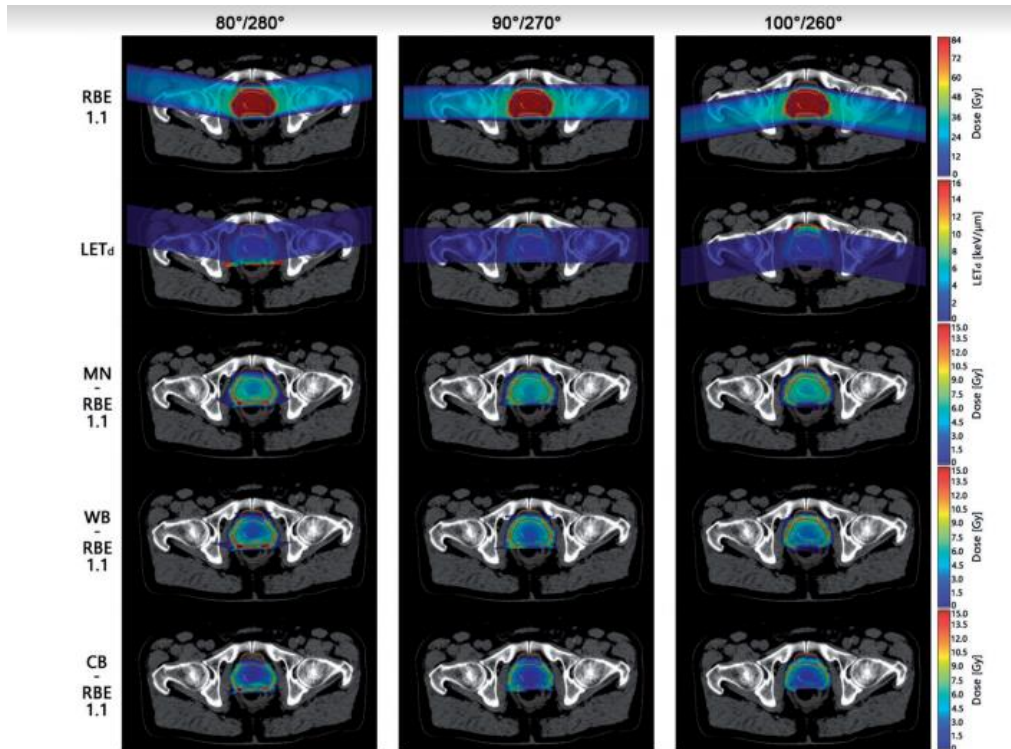


Biological dose and complication probabilities for the rectum and bladder based on linear energy transfer distributions in spot scanning proton therapy of prostate cancer

Jesper Pedersen, Jørgen B. B. Petersen, Camilla H. Stokkevåg, Kristian S. Ytre-Hauge, Stella Flampouri, Zuofeng Li, Nancy Mendenhall & Ludvig P. Muren

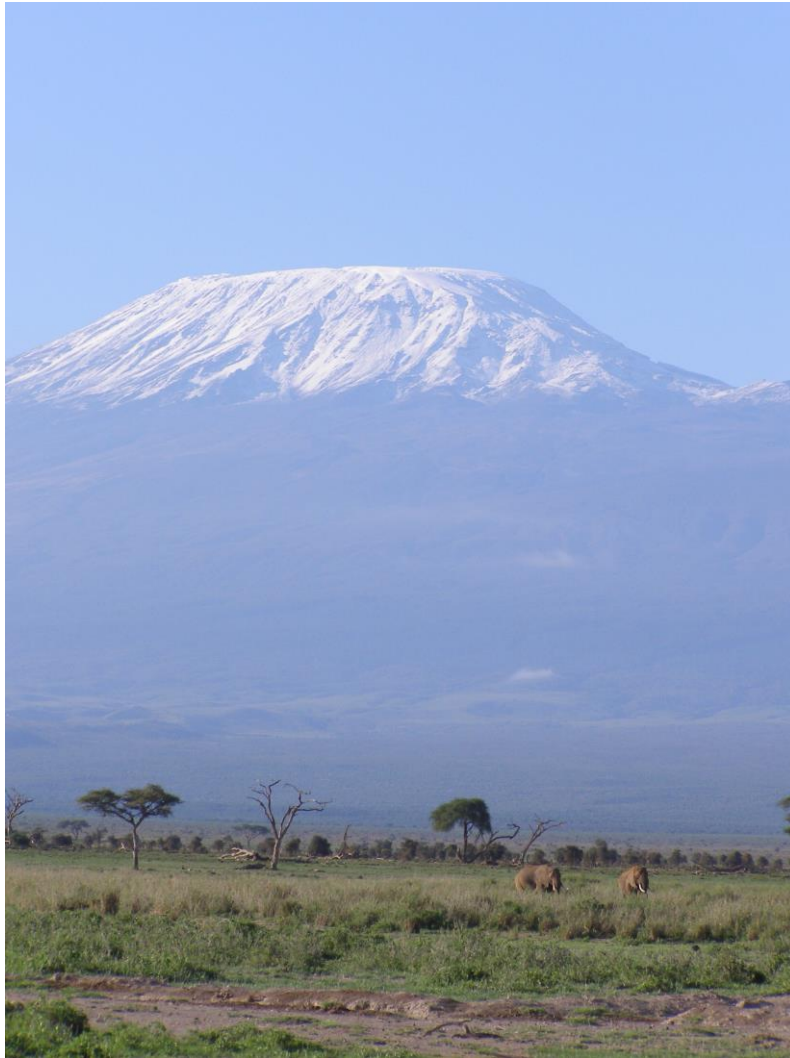
To cite this article: Jesper Pedersen, Jørgen B. B. Petersen, Camilla H. Stokkevåg, Kristian S. Ytre-Hauge, Stella Flampouri, Zuofeng Li, Nancy Mendenhall & Ludvig P. Muren (2017) Biological dose and complication probabilities for the rectum and bladder based on linear energy transfer distributions in spot scanning proton therapy of prostate cancer, Acta Oncologica, 56:11, 1413-1419, DOI: [10.1080/0284186X.2017.1373198](https://doi.org/10.1080/0284186X.2017.1373198)

To link to this article: <https://doi.org/10.1080/0284186X.2017.1373198>



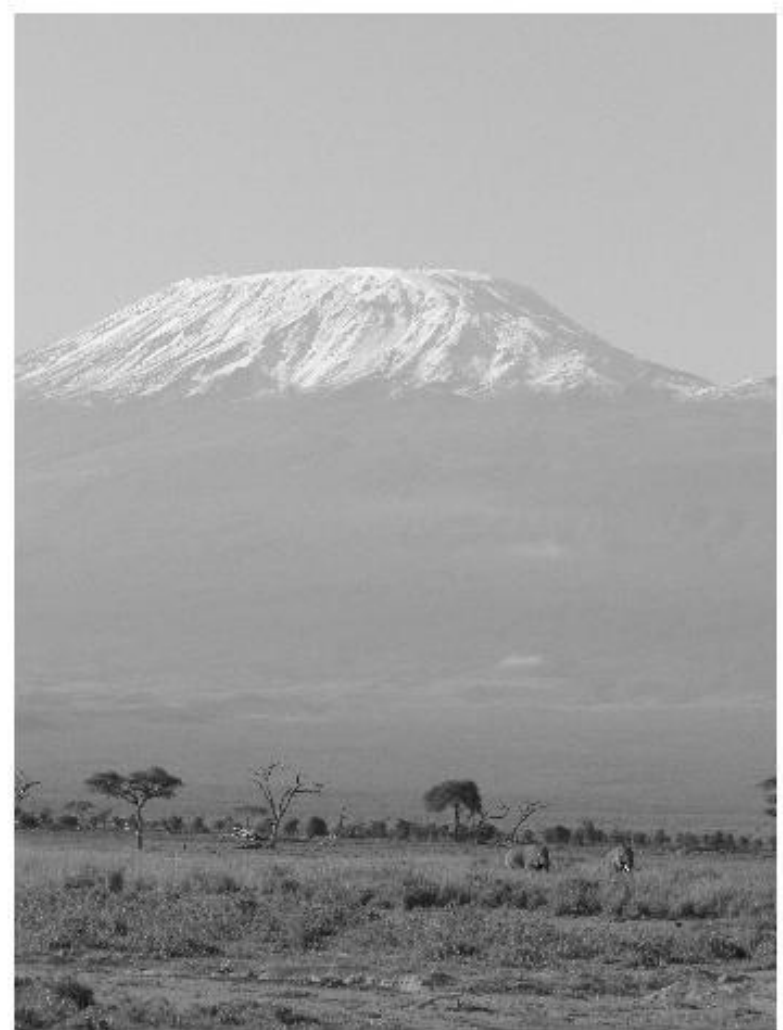


The curse of dimensions



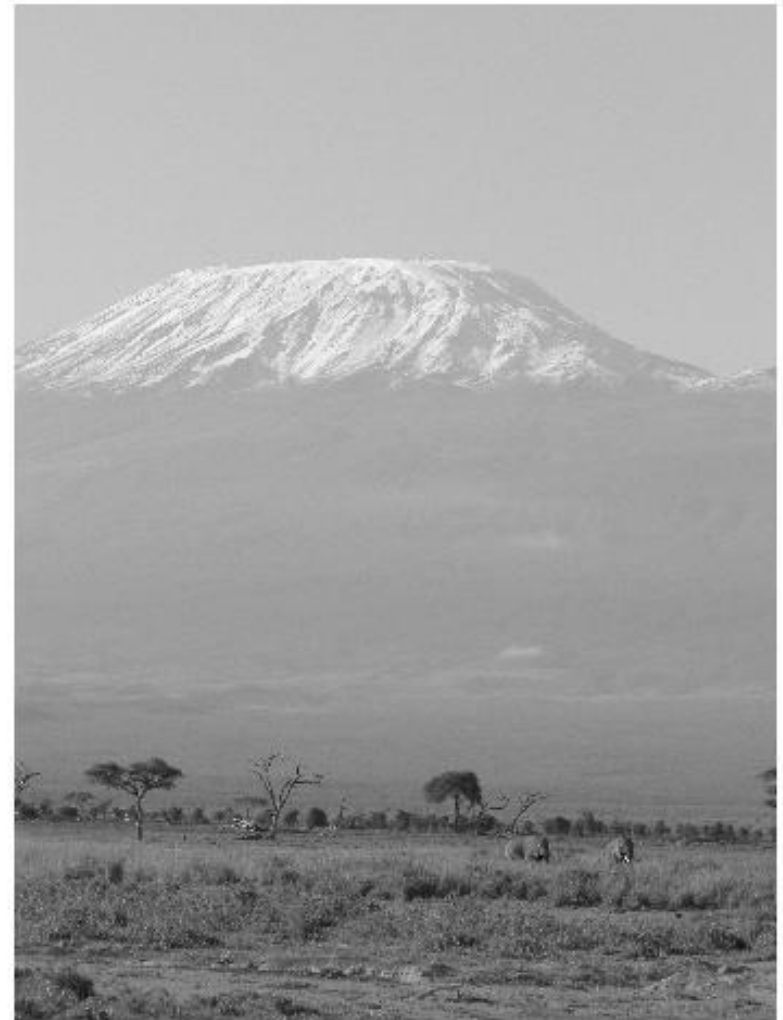
Data stored as a jpeg
3 dimensional array
2816x2112x3

The curse of dimensions

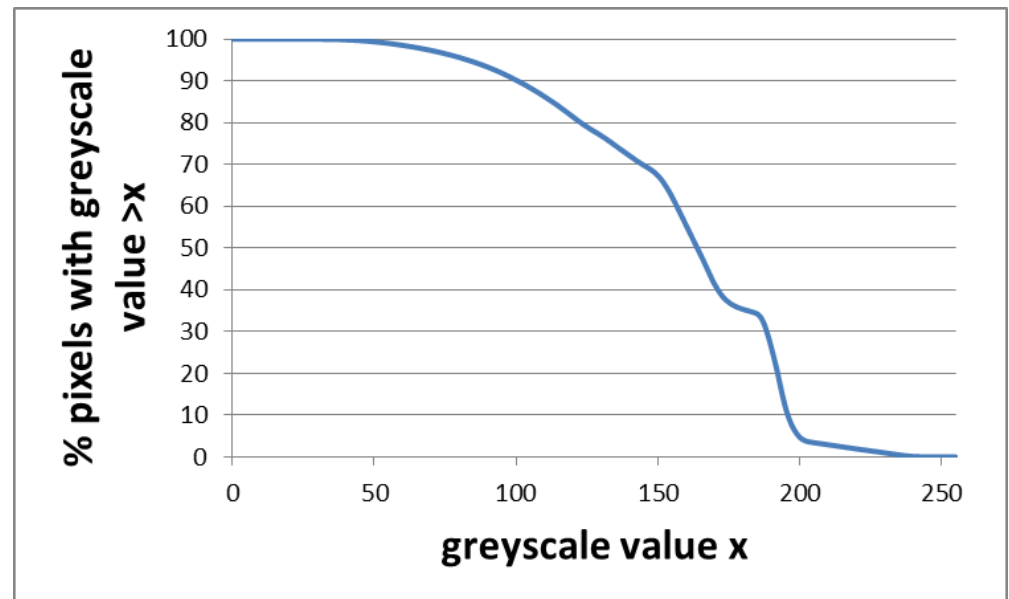
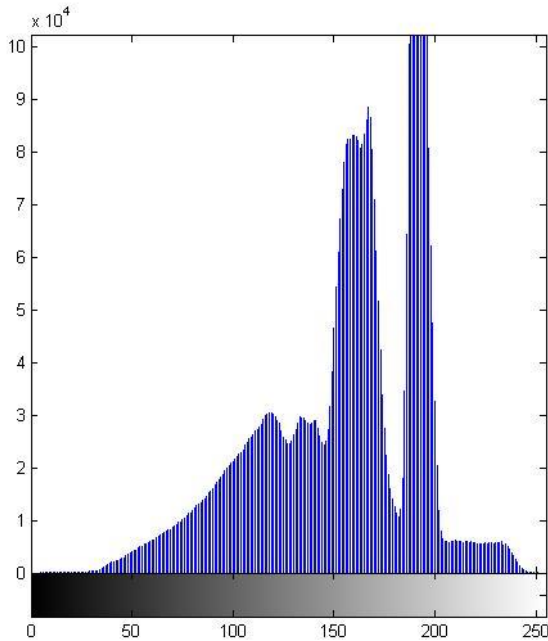


The curse of dimensions

Data stored as 2D
matrix 2816 by 2112

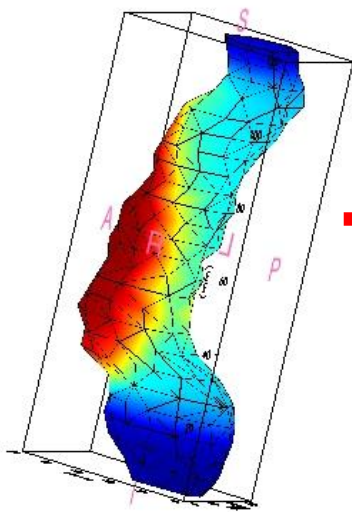


The curse of dimensions

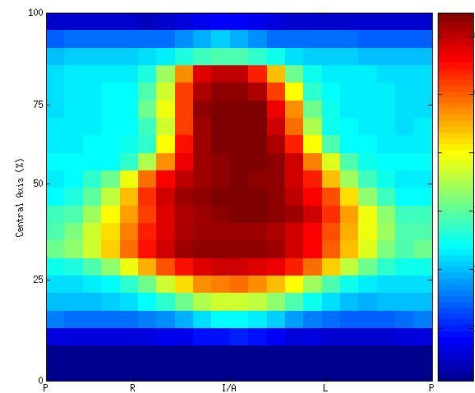


Rectum unfolding and feature extraction

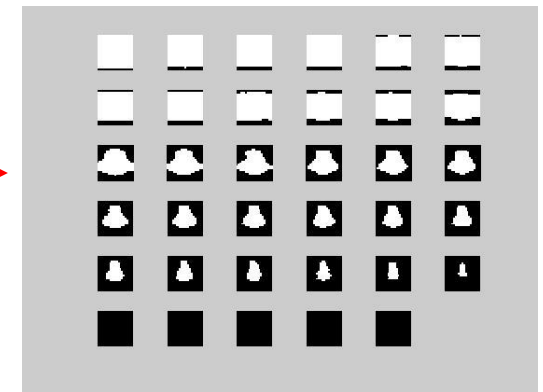
- Extract a **limited** set of **interpretable** features
- Unfold rectum and project dose on **2D map**
- Consider **binary maps** for feature extraction



3D dose distribution

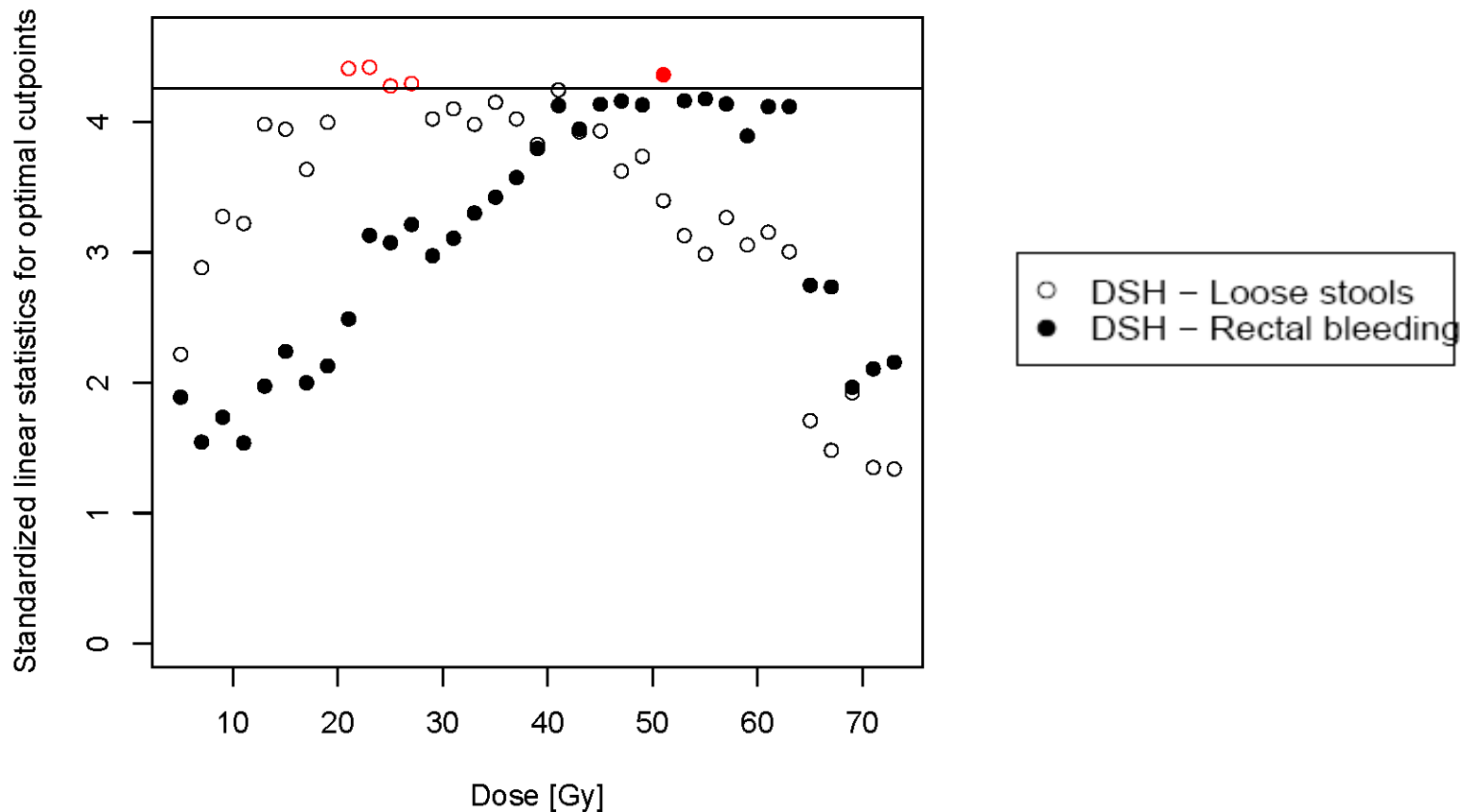


Primary DSM



Binary maps corresponding to 35 dose levels between 5 Gy and 73 Gy

Results

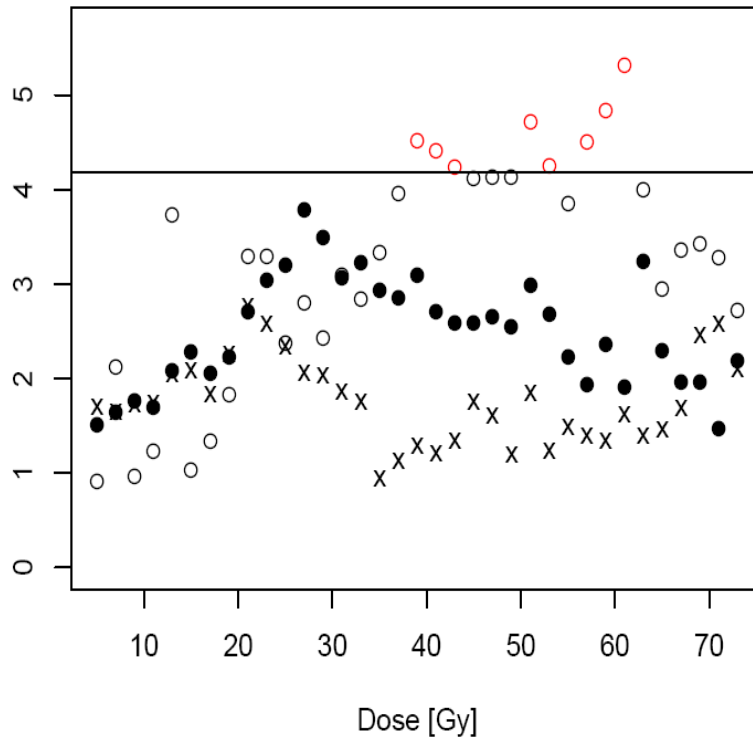


Maximally selected standardized T for the 35 bins of the DSH

Results *cont.*

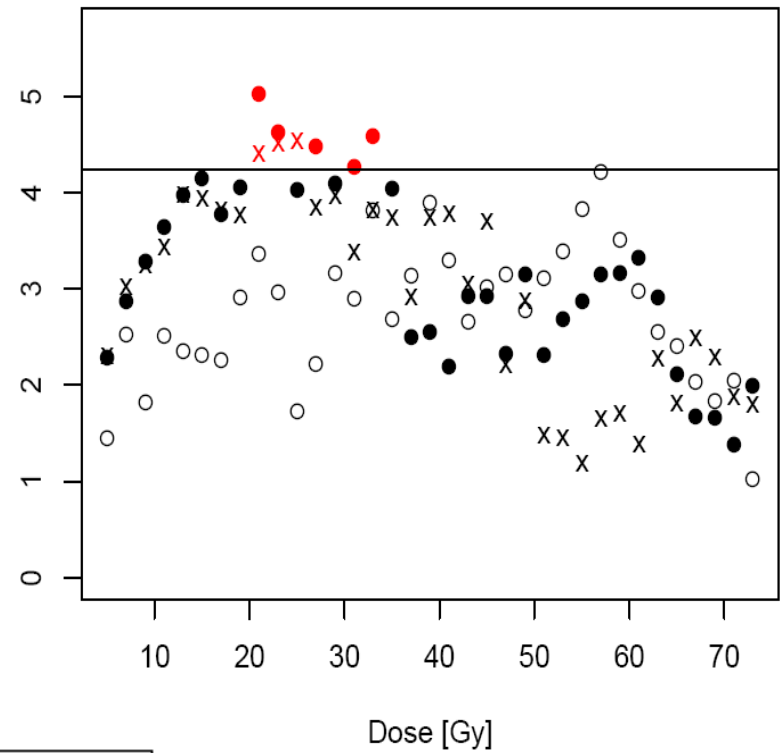
Standardized linear statistics for optimal cutpoints

Rectal Bleeding



Loose stools

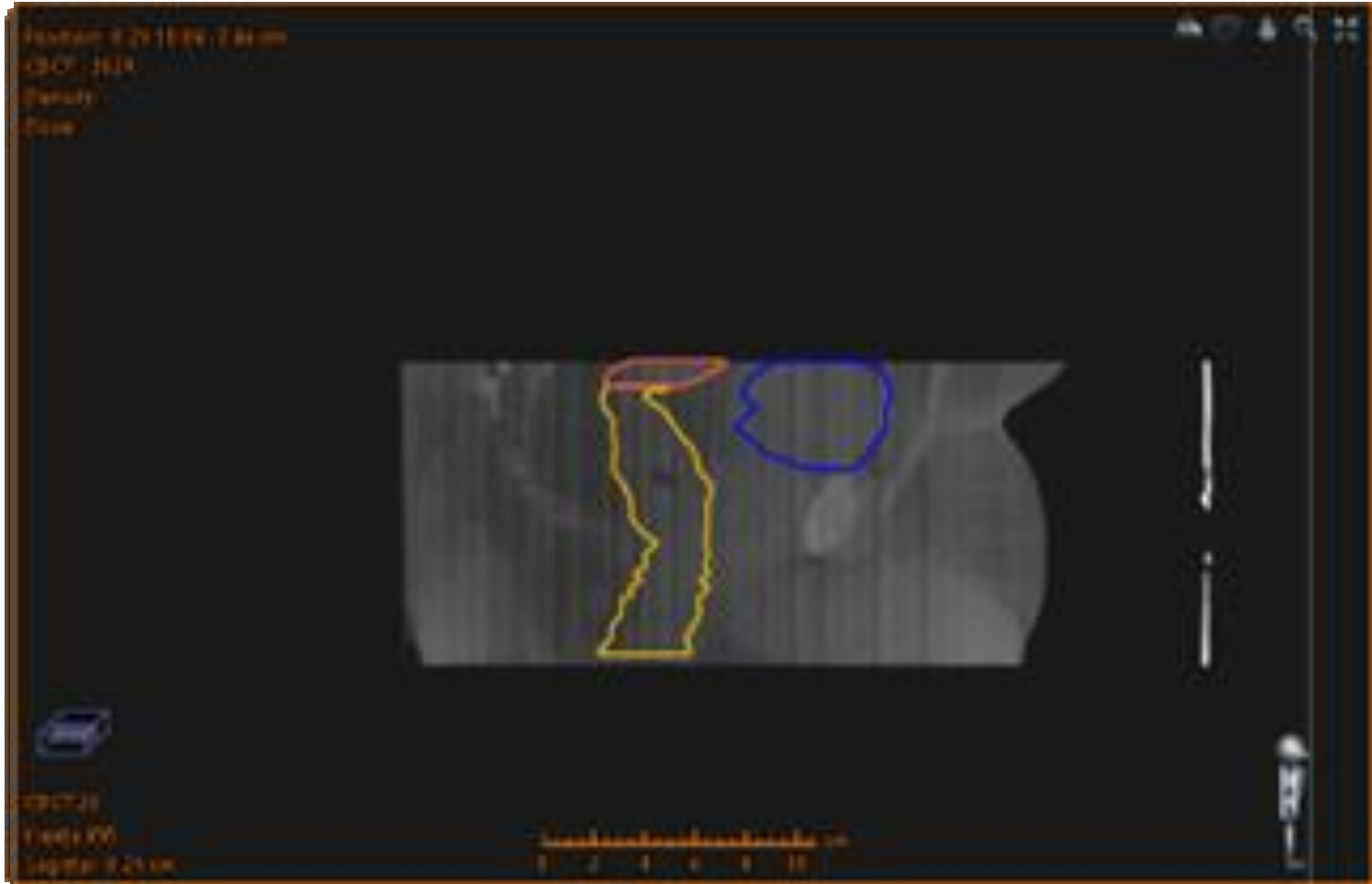
Standardized linear statistics for optimal cutpoints



- Lateral extent
- Longitudinal extent
- x Eccentricity

Cone Beam Series

CHHIP 20# Prostate



Final Comments

Need to be cautious when applying constraints or predictive model derived from one technique to patients treated with a different technique.

Know the limitations of the model and accept that data doesn't always fit the model

Consider the definition of toxicity used to derive the model and what is relevant to the clinic.

All constraints and models should be independently validated.

Acknowledgements

Prof David Dearnaley
Dr Julia Murray
Dr Florian Buettner
Dr Jamie Dean
Prof Chris Nutting
Mr Andy Poynter

