Summarizing our knowledge of normal tissue tolerances: the progress and future directions of QUANTEC*.

ANDREW JACKSON
Memorial Sloan-Kettering Cancer Center

*QUAntitative Normal TissuE models in the Clinic

Purpose of QUANTEC

- Both AAPM and ASTRO recognized ($$):
- Need for a systematic overhaul of our understanding of normal tissue tolerances
- For use in clinical treatment planning and optimization

History of QUANTEC

- 2006 AAPM Science Council
  - Ellen Yorke and Rock Mackie
  - Steering Committee: Deasy, Bentzen, Yorke, Ten-Haken, Jackson, Marks, Eisbruch, Constein
- 2007 1st QUANTEC meeting in Madison Wisconsin
  - Initial review of tolerances of involving physicists, bio-statisticians and physicians.
- 2008 Special Issue of Red Journal
  - in preparation, to be published in the fall.

Organ specific QUANTEC articles

- Significance of injury
- Clinically relevant endpoints
  - Time course
  - Ambiguities
- Anatomic definitions
  - Variations in contouring practice
- Review of literature on dose-volume dependence of endpoints
  - Level of evidence

Organ specific QUANTEC articles

- Patient related risk factors
- Models of the data
  - Limitations
- Caveats concerning models
- Dose-volume and model dependent limits for clinical use
- Future studies –
  - Additional knowledge required to improve toxicity prediction
  - Endpoint scoring and data capture in future studies

Organs Included

- Brain, Brainstem, Cord, Optic Structures, Ear
- Salivary Glands, Larynx & Pharynx
- Lung, Heart, Esophagus
- Liver, Kidney, Small Bowel,
- Rectum, Bladder, Penile Bulb

Overview of Results for Specific Organs

- Great variety in quality and quantity of data
- Consistency issues:
  - Endpoints studied
  - Physical issues (organ motion, contouring issues, dosimetry)
  - Models/D-V constraints reported
- Some organs are more equal than others
  - Lung: ~70 studies
  - Ear: ~4 studies of dose response of late hearing loss

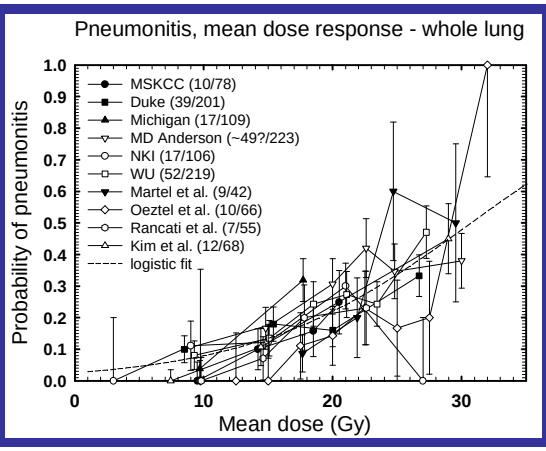
Mean dose response of pneumonitis

(A. Jackson with L. Marks/S. Kong/J.Deasy/J.Bradley/M. Martel/S. Bentzen)

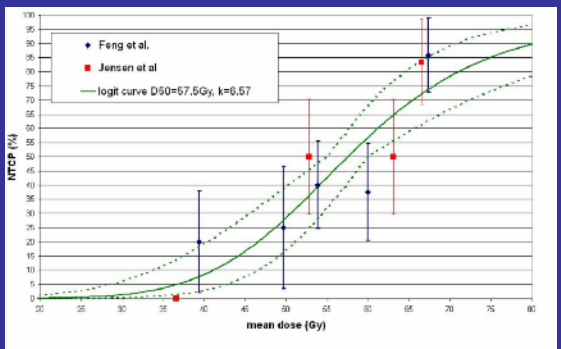
- Patients treated for NSCLC
  - Data from 9 institutions, 10 separate studies
- 1,167 patients with 222 cases of pneumonitis
- ≥ Grade 3 RTOG ~ ≥ Grade 2 SWOG
  - (requiring steroids)
  - accepted ≥ grade 1 definition if few grade 1 cases

Mean dose response of pneumonitis

- Reporting rate (and S.D.) of pneumonitis as function of mean dose to total lung
  - Numbers of pts w./w.o. pneumonitis
  - Bin locations on quartile plots
- Fit of logistic function:
  - D50 = 30.75 [29.9 – 31.7] Gy
  - γ50 = 0.907 [0.836 – 0.987]



Probability of aspiration as function of mean dose to Glottic and Supra-Glottic Larynx

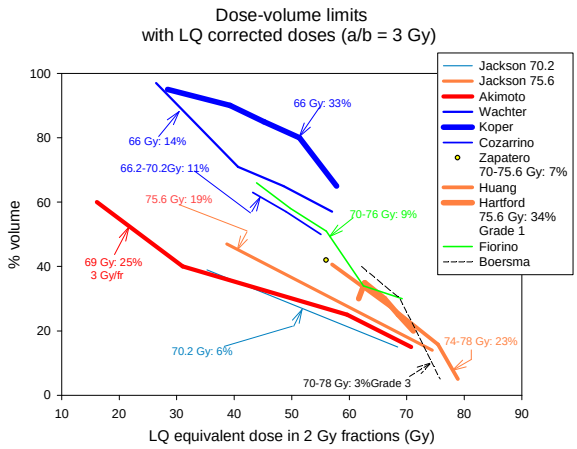


(T. Rancati, M. Schwartz, A. Allen, F. Feng, A. Popovtzer, B. Mittal, and A. Eisbruch)

Rectal dose volume limits

Jackson/Deasy/Gay/Michalski/Tucker/Zelefsky

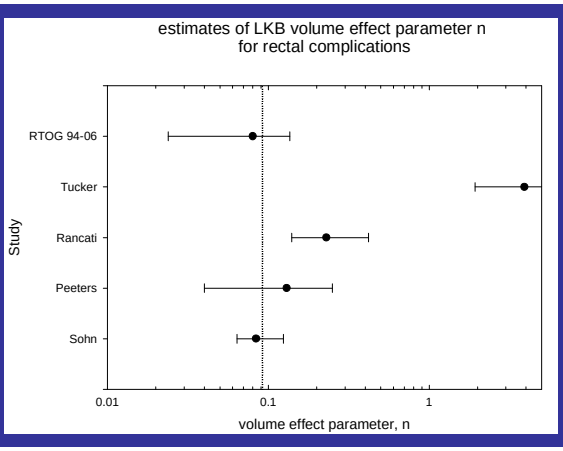
- Published limits having sig. correlation with ≥ grade 2 rectal bleeding
- Color coded to indicate prescription dose
  - Blue = 66-70 Gy
  - Red = 83 Gy (LQ equivalent dose in 2 Gy fr)
- Thickness of line indicates overall complication rate in study



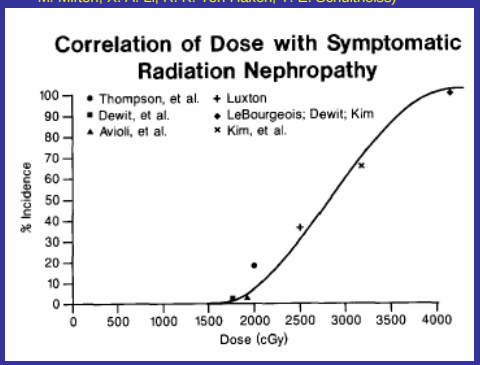
Rectal lyman model fits

- 4 published studies fitting LKB model to rectal complication data
  - Forrest plot of "n" values (n=1/a)
  - 1 S.D. indicated
  - Average value of "n" for all studies indicated
- One further study* of RTOG 94-06 included
  - n=0.08 (95% conf: 0.04-0.26)

* Tucker ASTRO 2007, IJROBP

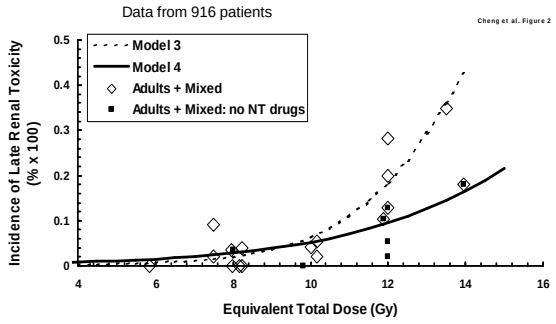


Kidney
(L. Dawson, A. Paulino, B. Kavanagh, C. Pan, S. K. Das, M. Miften, X. A. Li, R. K. Ten Haken, T. E. Schultheiss)



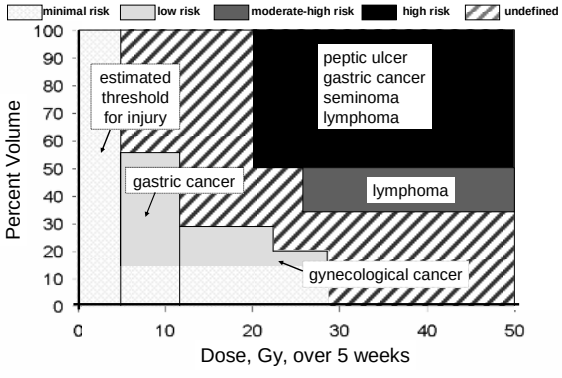
From: Cassady JR. Clinical radiation nephropathy. *IJROBP* 1995;31:1249-1256.

Bilateral Whole Kidney RT - TBI



Cheng J, Schultheiss T, Wong J. Impact of Drug Therapy, Radiation Dose and Dose Rate on Renal Toxicity following Bone Marrow Transplantation. *International Journal of Radiation Biology* 2008; In press.

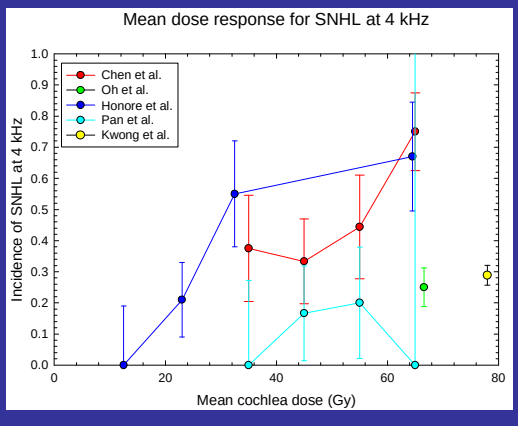
Bilateral Partial Kidney RT



Late Hearing Loss

(A. Jackson, N. Bandare, W. Mendenhall)

- Hearing loss tests from 3 studies as function of mean cochlea dose
  - (post-treatment vs pre-treatment)
- Differences in way endpoint is defined
  - Ispi- relative to contra-lateral hearing loss vs hearing loss
- Dose reconstruction
  - 1 study, doses reconstructed with surrogate CT scans
  - 1 study, ipsi- doses relative to contra-lateral



Necessity of combining data sets

- Number of complications in any given treatment series is usually low
  - False negatives
  - No statistical power to determine model parameters
- Dose-volume exposures correlated in individual series
  - Introduces phony correlations with complications (False positives)
  - Insufficient range of dose-volume combinations to determine model parameters

Problems in synthesizing data

- Endpoint definitions:
  - Need to be clinically relevant
  - Need to be specific
    - Rectal bleeding or incontinence vs grade 2 RTOG toxicity
      - Different comps. have different dose-volume effects
  - Need to be standardized

Problems in synthesizing data

- Variety of dose volume limits proposed
  - These cannot be combined
- Variety of models may be fit
  - Responses cannot be combined
    - gEUD responses with different "a" values cannot be combined

Problems in synthesizing data

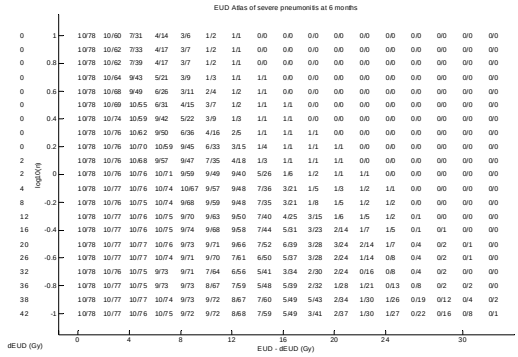
- Standard of reporting is **POOR**
  - Lack of basic statistics (numbers not stated!)
    - Schultheiss 1994: "**The information in this report would be of greater clinical use if some indication had been provided of the total number of patients from which the myelopathy cases were drawn**"
  - Locations of bins in e.g. quartile plots not given
  - Model parameters (and errors) not be stated

In other words:

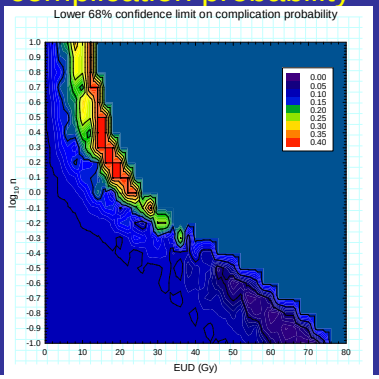
- Report the numbers of patients with complications and the number treated
  - Elementary statistics increase clinical utility
- Be comprehensive
  - Report as much about the data as possible
- How far can we take this?

Example: EUD Atlas of Pneumonitis

- Report the number of NSCLC patients whose gEUD exceeds a given level
  - Both with and without complication
- Be comprehensive:
  - Do this for each gEUD value
  - Vary the n=1/a parameter



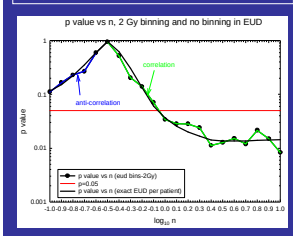
Lower 68% confidence limit on complication probability

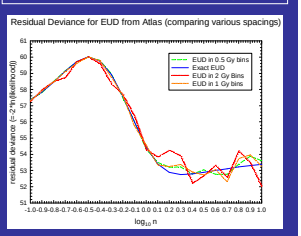


Comparison of results from atlases with 2-0.5 Gy grid spacings with those from exact EUD

For range of significant p-value, 2 Gy grid spacing in EUD is adequate

For lower 95% and 68% conf. int. on n 2 and 1 Gy spacing resp. are adequate

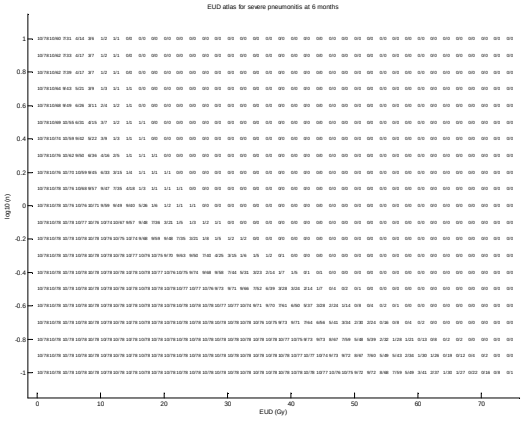




Best fit value of n sensitive to grid noise - likelihood function is flat for n > 0.1

Conclusions

- QUANTEC is:
  - Updating our clinical understanding of normal tissue tolerances
  - Providing clinical guidelines where possible
    - With appropriate caveats
  - Defining areas of our ignorance
    - recommend studies to remedy this
  - Investigating future directions:
    - Reporting standards
    - Clinically relevant but specific endpoint definitions
    - Inter-institutional data synthesis (atlases or pooling)



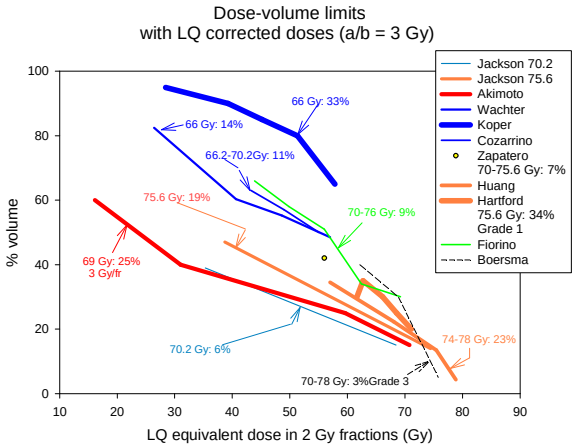
EUD AND DOSE-VOLUME BASED ATLASES OF COMPLICATION INCIDENCE

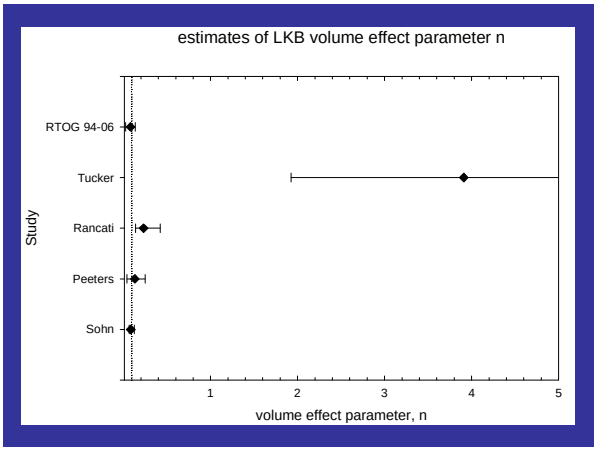
- A PROPOSAL FOR NEW STANDARDS IN REPORTING RESULTS OF TREAMENT PROTOCOLS.

Andrew Jackson, Ellen D. Yorke,
Kenneth E. Rosenzweig,
Ennapadam Venkatraman,
and C. Clifton Ling

Memorial Sloan-Kettering Cancer Center

1) MSKCC, Yorke et al. IJROBP 63 2005: 672-682, from Fig 4a) (≥RTOG grade 3, 6 months)
2) Duke, Hernando et al. IJROBP 51 2001: 650-659, from Table 4 (≥CTC grade 1, 6 months)
3) Michigan, Kong et al. IJROBP 65 2006: 1075-1086, from Table 4 and Fig 2a) (≥SWOG grade 2, 6 months) – bin location and time from authors.
4) MD Anderson, Wang et al. IJROBP 66 2006: 1399-1407, from Fig 2 (≥CTC grade 3, 1 year actuarial – includes concurrent chemo patients)
5) NKI, Seppenwoolde et al. IJROBP 60 2004: 748-78, from Fig 3a) (≥SWOG grade 2, 6 months)
6) WU, Hope et al. IJROBP 65 2006: 112-124, from Fig 9c) (≥SWOG grade 2 – no time limit) with bin locations from authors, increased by 11% to ~account for inhomogeneity corrections.
7) Michigan, Martel et al. IJROBP 28 1994: 575-581, from Table 1 (≥SWOG grade 1) with mean doses calculated from relationship between EUD (n=0.87) and mean dose from Kwa et al., Radiotherapy and Oncology 48 1998: 61-69 Fig 2a).
8) Heidelberg, Oeztel et al. IJROBP 33 1995: 455-460, from Fig 2 and text (≥RTOG acute grade 1).
9) Milan, Rancati et al. Radiother. and Oncol. 67 2003: 275-283, from Fig 3 (≥SWOG Grade 2 – no time limit, patients without COPD – includes induction chemo patients).
10) Gyeonggi, Kim et al. Radiology 235 2005: 208-215, from Table 5 (≥RTOG grade 3, 6 months – includes concurrent chemo patients) – median values of mean dose in each bin provided by the authors.
11) Logistic fit: data fit to the form (f/(1+f)), where f=exp(b0+b1*dmean). Best fit values [95% confidence intervals] are b0 = -3.63 [-3.53–3.74], b1 = 0.118 [0.109-0.128], corresponding to D_{50} = 30.75 [29.9 – 31.7] Gy and γ_{50} = 0.907 [0.836 – 0.987].





Data Set

- 78 patients from a Phase I dose escalation study of 3D-CRT of non-small cell lung cancer*
  - Treated doses 57.6-90 Gy
- Endpoint: ≥ RTOG Grade 3 radiation pneumonitis (G3RP)
  - G3RP = steroids and/or oxygen
  - Pts either developed G3RP within 6 months following treatment or survived that time without it
  - There were 10 instances of G3RP

* Rosenzweig et al, Cancer 103:2118-27, 2005; Yorke et al, IJROBP 63:672-682, 2005

Dose-Volume and EUD Atlases

- We previously introduced the dose-volume atlas of complication incidence*
  - analyzed data from the lung protocol
  - does not account for overall shape of DVH
- Here we introduce the EUD atlas
  - Accounts for overall shape of the DVH
  - Atlases can be based on other models

* Jackson et al, Sem in Rad Onc 16:260-268, 2006

Calculation of EUD

- EUD values were generated for total lung
  - Doses biologically corrected to l.q. equivalent doses delivered in 2 Gy fractions using α/β=3Gy
  - EUD calculated for each of a grid of n values (log10(n) varying from -1 to + 1 steps of 0.1)

$$EUD(n_j) = \left[\sum_k v_k (d_k)^{1/n_j} \right]^{n_j}$$

Additive property of the atlas

- Provided dosimetric and endpoint information is compatible
  - Doses defined in same way
  - Endpoint defined in same way
- Data from different atlases (A and B) can be added, grid point by grid point:

$$N_c(A,B) / N_p(A,B) = [N_c(A) + N_c(B)] / [N_p(A) + N_p(B)]$$

Additive property of the atlas

- Facilitates meta-analysis of data from different institutions
  - Low numbers of complications from individual protocols can be combined
  - Potentially synergistic effect if adopted as a standard of reporting

Inefficient use of space

- Many repeated numbers where no patient treatments occur
- Solution:
  - Shift the range to eliminate this dead space
  - Record to shift so that correct position can be reconstructed

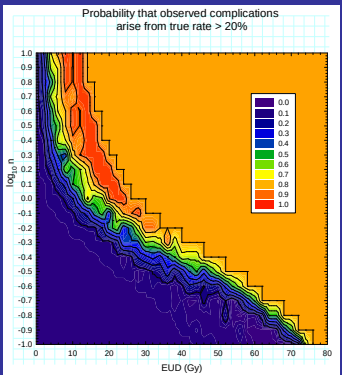
Calculations based on Atlas data

- At each grid point, can calculate e.g.:
  - Confidence limits on complication probability for patients with EUD > grid value
  - Probability that complication rate in patients with EUD > grid value is greater than tolerance

Calculations based on Atlas data

- For each value of n, can calculate e.g.:
  - Logistic regression of EUD with endpoint
    - p-value for correlation of EUD with endpoint
    - range of n values for significant correlation with endpoint
    - Likelihood of fit of EUD response
    - Find max likelihood value and confidence intervals for n

Probability that G3PR rate is > 20%



Atlas grid spacing introduces statistical noise

- Need to determine effect of grid spacing noise on
  - Range of significant correlation
  - Best fit value of n parameter
  - confidence interval on fitted value of n

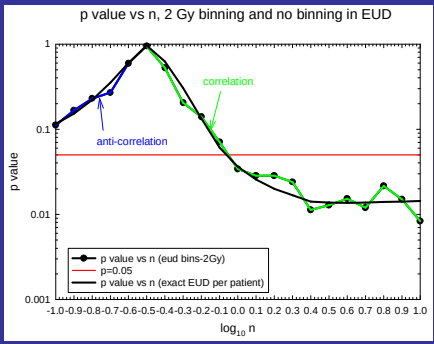
Conclusions

- Demonstrated utility of EUD and dose volume atlases using our lung data
- Fitting n values from EUD atlas data
  - At n ≥ 1 use 0.5 Gy grid
  - At n < 1 use 2 Gy grid
  - Noise should decrease as # pts and comps increase
  - Reformatting atlas (shifting range) for each value of n:
    - better use of atlas space
    - better grid resolution

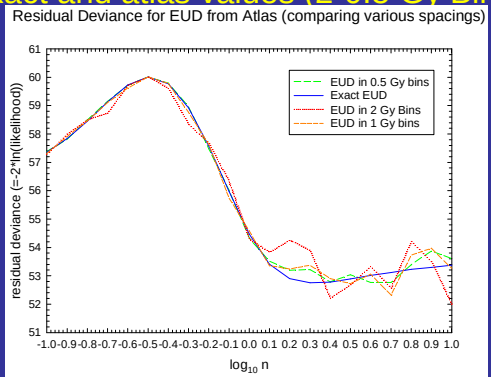
Conclusions

- Publication of atlases facilitates in depth meta-analysis of dependence of outcome data on atlas variables
- Atlases allow for useful publication of treatment series with few to no comps
- Atlases facilitate the quantitative assessment of potential risks and benefits of future treatments

P value, correlation with severe pneumonitis – exact vs atlas (2Gy bins)



Residual deviance - exact and atlas values (2-0.5 Gy Bins)



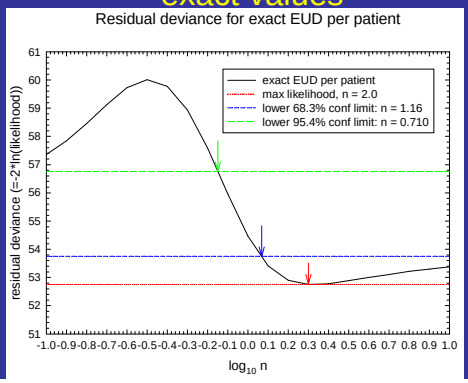
Determining grid spacing

- Noise increases at n>1
  - Range of EUD values ↓, grid size gets cruder
- For p value range, and 95% confidence interval on best fit value of n,
  - 2 Gy grid spacing in EUD is adequate
- For lower limit of 68% confidence interval on n,
  - 1 Gy grid spacing in EUD is adequate
- Best fit value of n sensitive to grid noise
  - likelihood is a flat function of n
  - low statistics

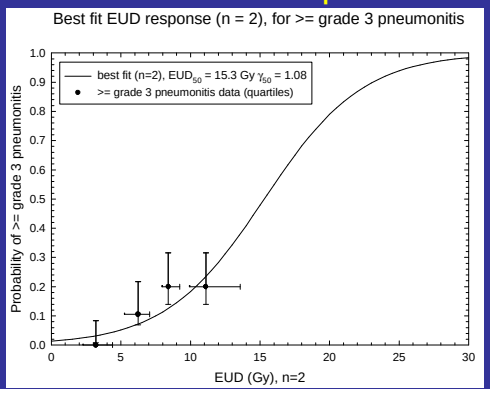
Basic Atlas Data

- 2 dimensional grid in EUD and log10(n)
- At each grid point (EUDi, log10(nj)), record
  - # patients with G3RP, (N_ci,j), with EUD(nj) > EUDi
  - total # patients, (N_pi,j), with EUD(nj) > EUDi.
- Display as explicit ratio:
  $N_c$ij/$N_p$ij

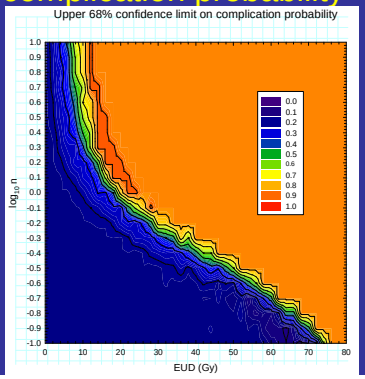
Residual deviance - exact values



Best fit EUD response

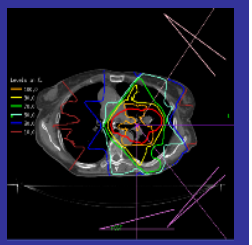


Upper 68% confidence limit on complication probability



Data Set

- treatments based on a single CT data set
- treatment plans restorable from electronic archive



Liver

- Charlie Pan, Brian D Kavanagh, Laura A. Dawson, X. Allen Li, Shiva K Das, Moyed Miften, Randall K Ten Haken