Policlinico Agostino Gemelli
Università Cattolica del Sacro Cuore

Gemelli

Realization and implementation of a dose-response curves analysis tool integrated in a free statistical software (R)

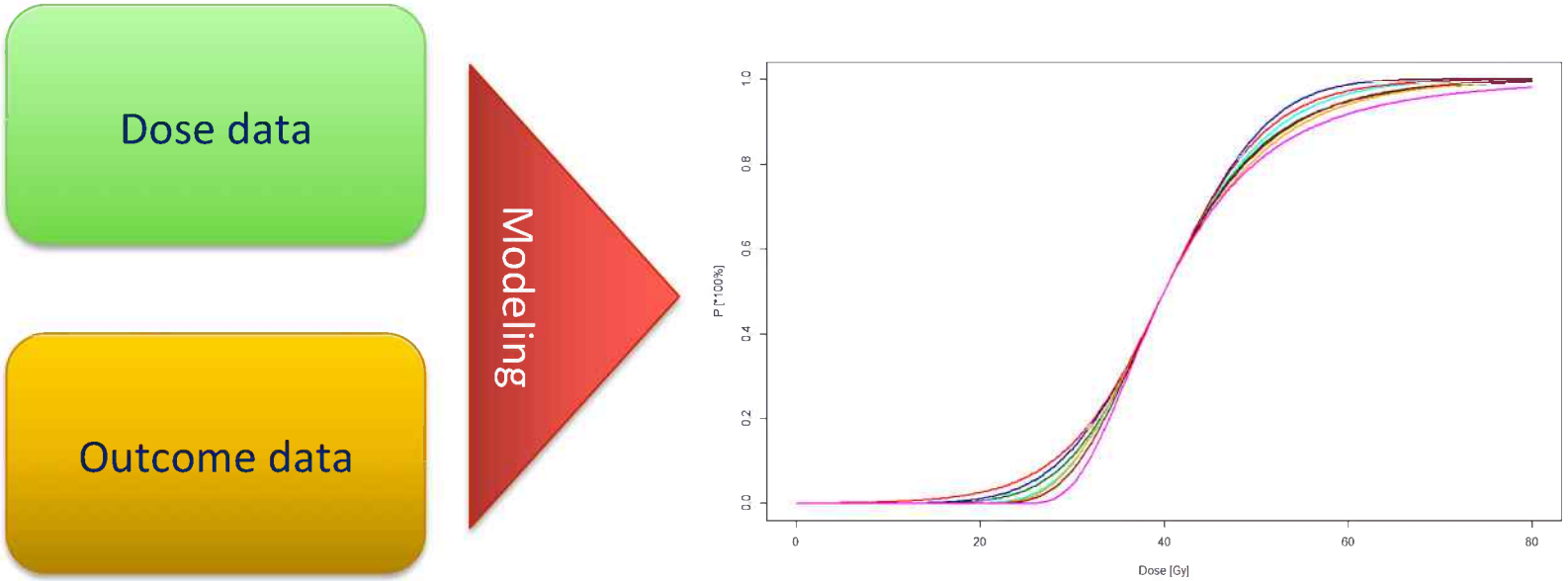
N. Dinapoli, R. Gatta, S. Chiesa, A. Damiani, V. Valentini

Datamining
Radiomics
KBO Labs
Knowledge Based Oncology Labs

Gemelli
ADVANCED RADIATION THERAPY

BACKGROUND

Radiobiological modeling is a complex statistical process based mainly on the implementation of some inferential statistical techniques that allow to fit dose, or dose-distribution data, to outcome detected according clinical protocols on specific oncological situations



DATA EXTRACTION

Different sources can be used as dataset for achieving dose and dose-distribution data:

1) **DICOM-RT files** (RT Structures, RT Plans and RT Dose)
2) (Varian) **Text format** DVH files
3) **Legacy dataset** (xls, csv…)

DATA MANIPULATION

Why "R", what is "R"?

R is "GNU S", a freely available language and environment for statistical computing and graphics which provides a wide variety of statistical and graphical techniques: linear and nonlinear modelling, statistical tests, time series analysis, classification, clustering, etc. Please consult the R project homepage for further information.

http://http://cran.r-project.org/

MODEL SELECTION

Several well known dose-response models have been chosen for being implemented into our package: the names of the dose-response functions correspond to the name of the main author of the published model:

1. ***Bentzen***, *S M, and S L Tucker. 1997. "Quantifying the Position and Steepness of Radiation Dose-Response Curves." International Journal of Radiation Biology 71 (5) (May): 531–42.*
2. ***Burman***, *C, G J Kutcher, B Emami, and M Goitein. 1991. "Fitting of Normal Tissue Tolerance Data to an Analytic Function." International Journal of Radiation Oncology, Biology, Physics 21 (1) (May 15): 123–35.*
3. ***Goitein***, *Michael. 1979. "The Utility of Computed Tomography in Radiation Therapy: An Estimate of Outcome." International Journal of Radiation Oncology, Biology, Physics 5 (10) (October): 1799–807.*
4. ***Munro***, *T R, and C W Gilbert. 1961. "The Relation between Tumour Lethal Doses and the Radiosensitivity of Tumour Cells." The British Journal of Radiology 34 (400) (April): 246–51.*
5. ***Okunieff***, *P, David Morgan, A Niemierko, and Herman D Suit. 1995. "Radiation Dose-Response of Human Tumors." International Journal of Radiation Oncology, Biology, Physics 32 (4) (July 15): 1227–37.*
6. ***Suit***, *H, S Skates, A Taghian, P Okunieff, and J T Efird. 1992. "Clinical Implications of Heterogeneity of Tumor Response to Radiation Therapy." Radiotherapy and Oncology : Journal of the European Society for Therapeutic Radiology and Oncology 25 (4) (December): 251–60.*
7. ***Warkentin***, *Brad, Pavel Stavrev, Nadia Stavreva, Colin Field, and B Gino Fallone. 2004. "A TCP-NTCP Estimation Module Using DVHs and Known Radiobiological*

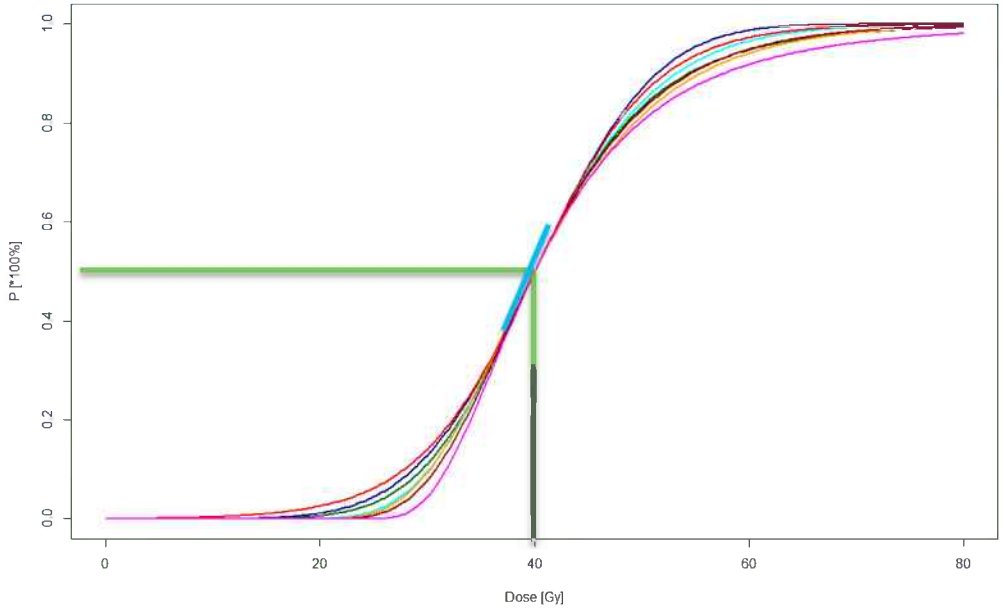
MODEL SELECTION

Several well known dose-response models have been chosen for being implemented into our package: the names of the dose-response functions correspond to the name of the main author of the published model:

1. ***Bentzen***, *S M, and S L Tucker. 1997. "Quantifying the Position and Steepness of Radiation Dose-Response Curves." International Journal of Radiation Biology 71 (5) (May): 531–42.*
2. ***Burman***, *C, G J Kutcher, B Emami, and M Goitein. 1991. "Fitting of Normal Tissue Tolerance Data to an Analytic Function." International Journal of Radiation Oncology, Biology, Physics 21 (1) (May 15): 123–35.*
3. ***Goitein***, *Michael. 1979. "The Utility of Computed Tomography in Radiation Therapy: An Estimate of Outcome." International Journal of Radiation Oncology, Biology, Physics 5 (10) (October): 1799–807.*
4. ***Munro***, *T R, and C W Gilbert. 1961. "The Relation between Tumour Lethal Doses and the Radiosensitivity of Tumour Cells." The British Journal of Radiology 34 (400) (April): 246–51.*
5. ***Okunieff***, *P, David Morgan, A Niemierko, and Herman D Suit. 1995. "Radiation Dose-Response of Human Tumors." International Journal of Radiation Oncology, Biology, Physics 32 (4) (July 15): 1227–37.*
6. ***Suit***, *H, S Skates, A Taghian, P Okunieff, and J T Efird. 1992. "Clinical Implications of Heterogeneity of Tumor Response to Radiation Therapy." Radiotherapy and Oncology : Journal of the European Society for Therapeutic Radiology and Oncology 25 (4) (December): 251–60.*
7. ***Warkentin***, *Brad, Pavel Stavrev, Nadia Stavreva, Colin Field, and B Gino Fallone. 2004. "A TCP-NTCP Estimation Module Using DVHs and Known Radiobiological*

Parameters

- TD50
- Gamma50
- EUD
- «a»



MODEL SELECTION

Bentzen (Log-Poisson): $$P = 0.5^{\left(\frac{D_{50}}{x}\right)^{\frac{2\gamma_{50}}{\log 2}}}$$

Lyman (Probit): $$P = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{\frac{-z^2}{2}} dz \qquad t = \frac{x - D_{50}}{D_{50}} \cdot \gamma_{50} \sqrt{2\pi}$$

Goitein (Log-Probit): $$P = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{\frac{-z^2}{2}} dz \qquad t = \frac{\log x - D_{50}}{D_{50}} \cdot \gamma_{50} \sqrt{2\pi}$$

Munro (Poisson appr.): $$P = 2^{-e^{e\gamma_{50}\left(1-\frac{x}{D_{50}}\right)}}$$

MODEL SELECTION

Okunieff (Logit):

$$P = \frac{1}{1 + e^{4\gamma_{50}\left(1 - \frac{x}{D_{50}}\right)}}$$

Suit (Log-Logit):

$$P = \frac{1}{1 + \left(\frac{D_{50}}{x}\right)^{4\gamma_{50}}}$$

Warkentin (Poisson):

$$P = 0.5^{e^{\frac{2\gamma_{50}}{\log 2}\left(1 - \frac{x}{D_{50}}\right)}}$$

Niemierko gEUD:

$$gEUD = \left(\frac{1}{N}\sum_{j=1}^{N}\left(v_j D_j^a\right)\right)^{\frac{1}{a}}$$

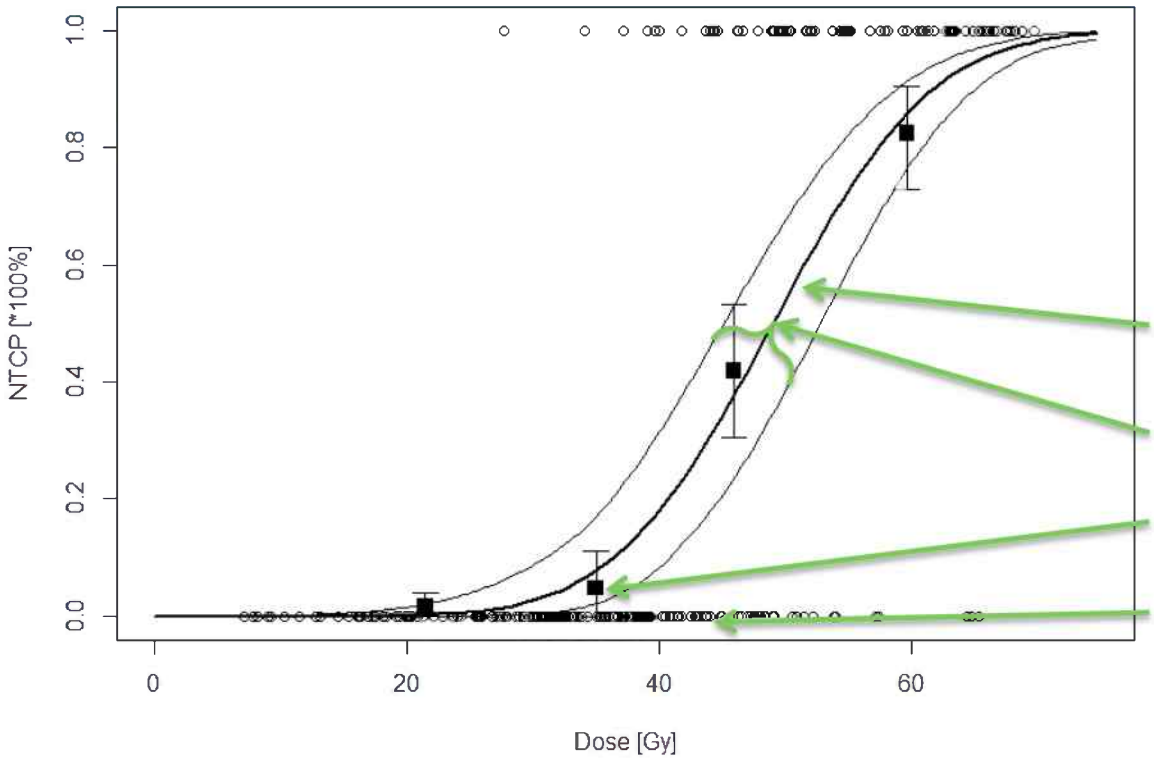
MODELING PROCEDURE

Modeling procedure is designed by using the **Maximum Likelihood Estimation (MLE)** technique, that aloows to detect the most likely parameters that fit a given dose-outcome series. Modeling can be performed both on three parameters (D_{50}, γ_{50} and a), taking into account the need of definition of parallel/serial organization of the structure, and on two parameters (D_{50}, γ_{50}) when modeling on nominal dose series has to be performed.
The MLE procedure finds the maximum of the **log-Likelihood function (log-L)** related to each dose-response function.
It is possible to compare the performance of different models fitted on case series by using the p-value of the z-test of each variable and the Akaike Information Criterion.

$$\log L = \sum_{i=1}^{n} y_i \log P_i + (1 - y_i)\log(1 - P_i)$$

***Bentzen**, S M, and S L Tucker. 1997. "Quantifying the Position and Steepness o*
Dose-Response Curves". International Journal of Radiation Biology 71 (5) (May

EXAMPLE OF GRAPHICAL OUTPUT



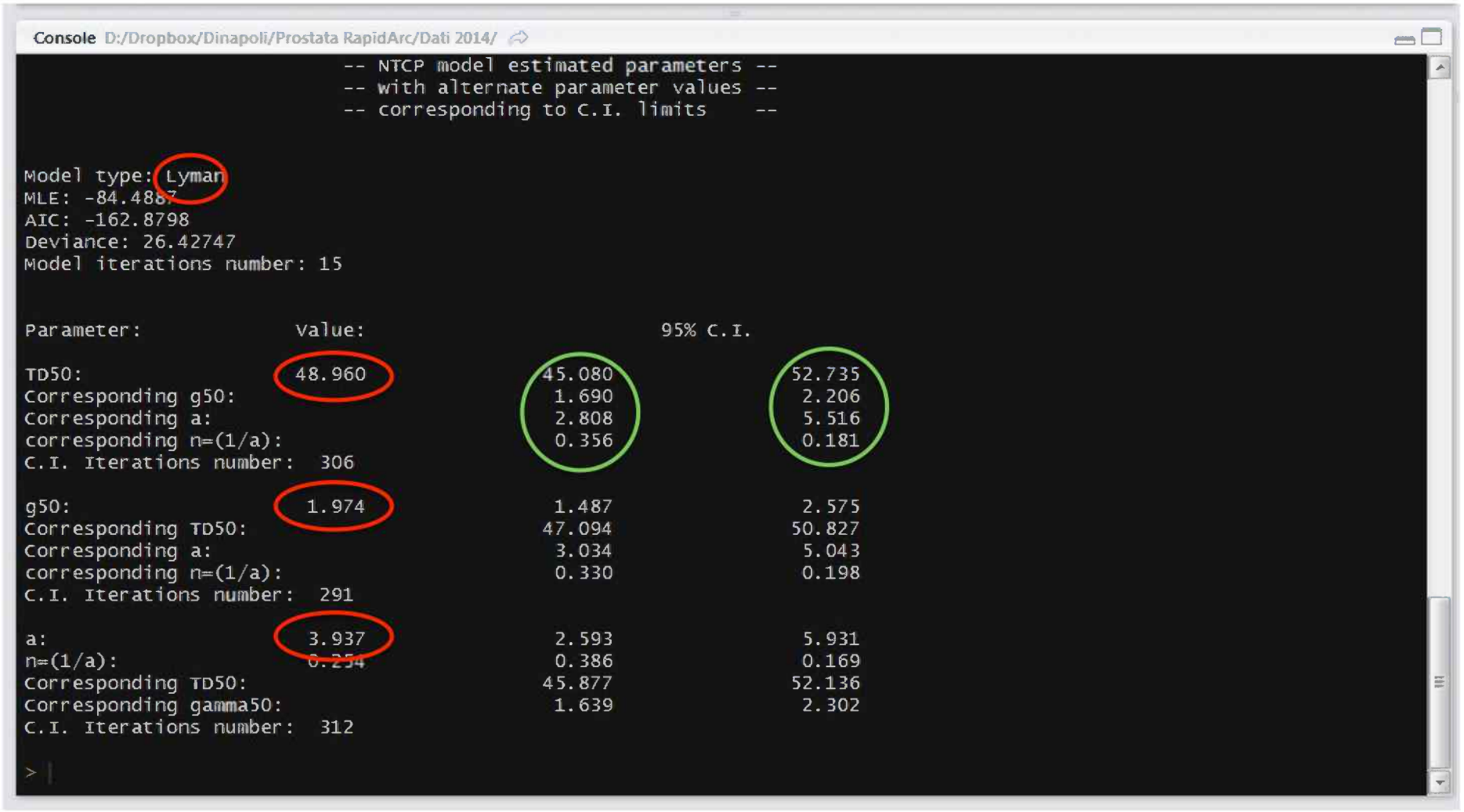
Dose response curve

Confidence interval

Quantiles of cases

Cases

EXAMPLE OF MODELING OUTPUT



```
Console D:/Dropbox/Dinapoli/Prostata RapidArc/Dati 2014/

                        -- NTCP model estimated parameters --
                        -- with alternate parameter values --
                        -- corresponding to C.I. limits    --


Model type: Lyman
MLE: -84.4887
AIC: -162.8798
Deviance: 26.42747
Model iterations number: 15


Parameter:              Value:                    95% C.I.

TD50:                   48.960                    45.080          52.735
Corresponding g50:                                 1.690           2.206
Corresponding a:                                   2.808           5.516
corresponding n=(1/a):                             0.356           0.181
C.I. Iterations number:  306

g50:                    1.974                      1.487           2.575
Corresponding TD50:                               47.094          50.827
Corresponding a:                                   3.034           5.043
corresponding n=(1/a):                             0.330           0.198
C.I. Iterations number:  291

a:                      3.937                      2.593           5.931
n=(1/a):                0.254                      0.386           0.169
Corresponding TD50:                               45.877          52.136
Corresponding gamma50:                             1.639           2.302
C.I. Iterations number:  312

>
```

EXAMPLE OF MODELING OUTPUT

Console D:/Dropbox/Dinapoli/Prostata RapidArc/Dati 2014/

```
> summary(model$model)
Maximum likelihood estimation

Call:
mle2(minuslogl = negLogLik, start = list(TD50 = start.TD50, gamma50 = start.gamma50,
    aa = start.a))

Coefficients:
        Estimate Std. Error z value   Pr(z)
TD50    66.49909    6.84338  9.7173 < 2e-16 ***
gamma50  1.46643    0.59305  2.4727 0.01341 *
aa       3.84089    2.10253  1.8268 0.06773 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

-2 log L: 128.3828
> |
```

FURTHER DEVELOPMENTS

1. Publication of the package in the CRAN depository to be downloaded by everybody

2. Inclusion of multivariate modeling procedures to find factors that can affect the values of parameters themselves in model output

3. Definition of error analysis in the DVH extraction function for calculating DVHs as much as possible similar (if not equal) to the ones provided by TPS workstation

http://www.rstudio.com/
http://cran.r-project.org/