Data Gaps and Needs in Model-based Risk Assessment

Lutz Edler & Annette Kopp-Schneider

Department of Biostatistics, German Cancer Research Center, Heidelberg, Germany

CONTENT

Empirical Model based risk assessment

Quantitative Methods and Modeling

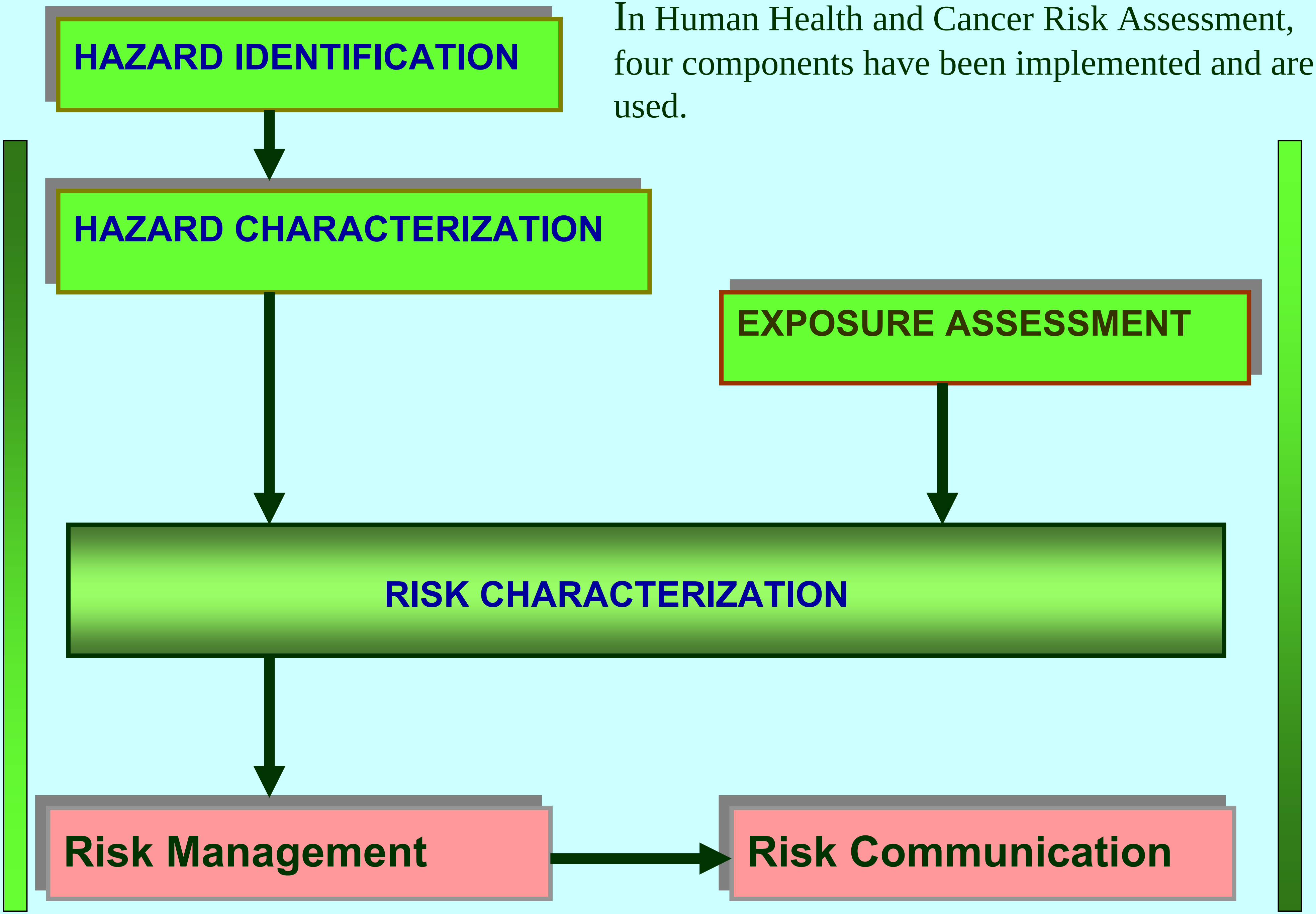
Benchmark Dose Approach

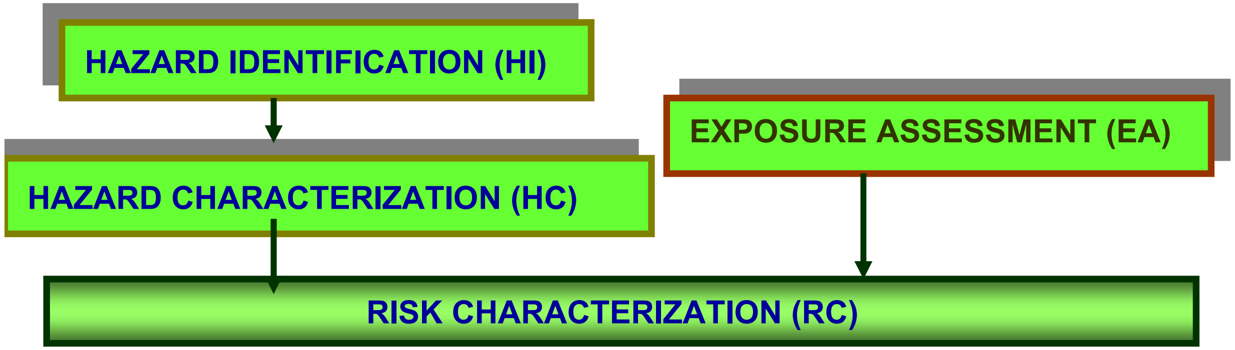
Stochastic Model based risk assessment

Testing Biological Hypotheses

Conclusion

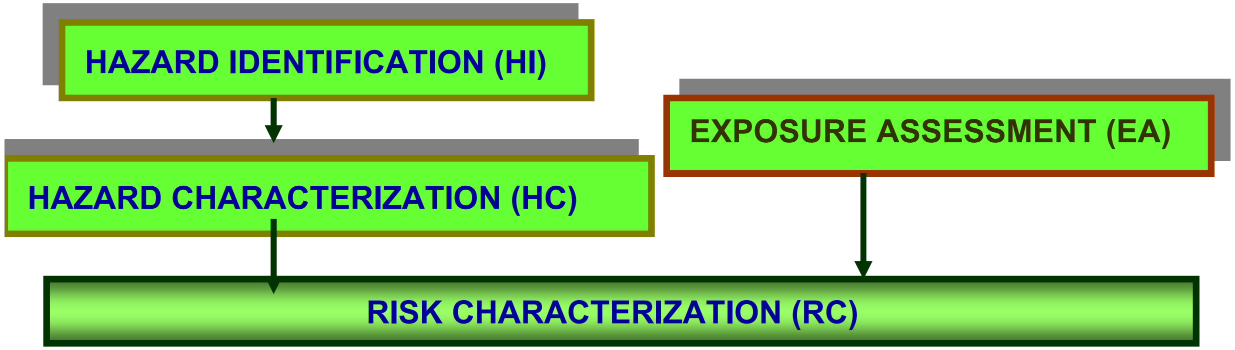
In Human Health and Cancer Risk Assessment, four components have been implemented and are used.





Statistical and Computational issues arise in all four components

Hazard identification:	traditional statistical inference, exact methods for small samples, comparison of different methods in simulation studies
Exposure assessment:	complex schemes of intake estimation applying probabilistic methods to survey data, estimation of individual exposure, developing distributions for the exposure in the population
Hazard characterization =	**non-linear regression methods,**
Dose-response assessment	**fitting stochastic models**
Risk characterisation	uncertainty analyses, probabilistic methods



First Announcement

Third International Conference on Cancer Risk Assessment
(ICCRA 3)

Mathematical, Statistical, Medical and Computational Methods

16-18 , 2009

Porto Heli, Greece

Submission of Full Paper: January 28, 2009
Submission of Abstracts to the Conference: May 1, 2009

Benchmark Dose Approach

a “simple” method to fit models to dose-response data and to estimate the response in the lower dose region together with a confidence interval.
By inverting the dose-response curve one calculates dose levels, BMD/BMDL, from which tolerable dose levels are derived.

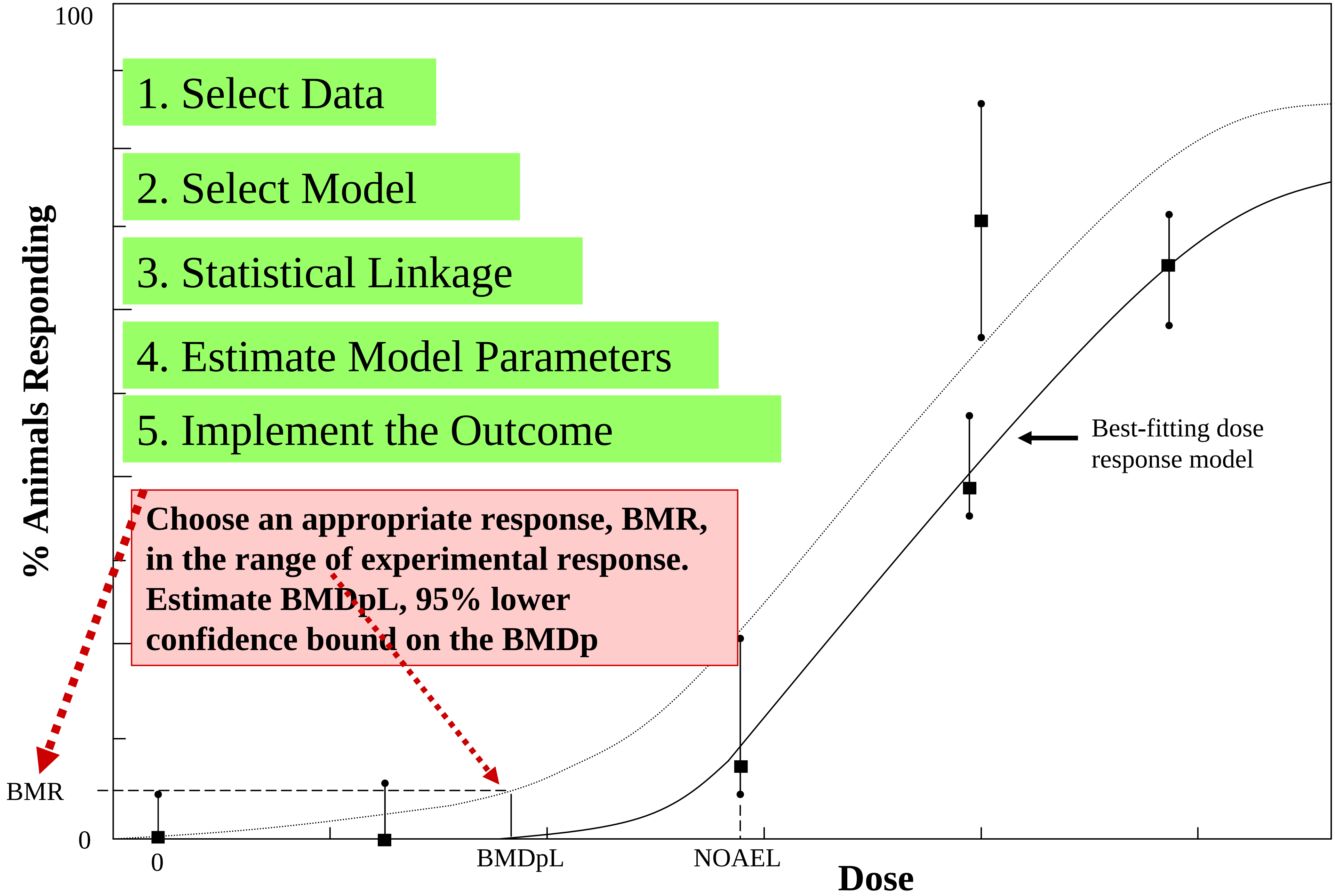
endeavors to establish the BMD in the risk management

what data
which tools

providing
the right software
to the right people
in time

BMDS package of the US EPA
PROAST package developed by RVIM, NL

How to apply the BMD?



Benchmark Dose Approach

Discussion Points between Statistician and Risk Assessor/ Risk Manager

Should the raw data be transformed before the analysis?

clarify what raw data are
clarify what is dose and what is response for the modeling
transformation of scale for graphical presentation is admitted

range and scale of dose levels

type and scale of response level **response -- effect**

Study Design

dose groups
numbers per dose group

Benchmark Dose Approach

Discussion Points between Statistician and Risk Assessor/ Risk Manager

Which model class should be chosen?

- distinguish between data type
  - quantitative, categorical, dichotomous
  - nested data
- default class of nested models
- with covariates ?

Model uncertainty --> Mat's talk

Model constraints

- **c**onstraints on parameters
- shape of curve (non-monotone?)

BMD:Model selection

Slob (2005) proposed a nested family of multiplicative nonlinear models for quantitative response

$y = a$	with $a > 0$
$y = a \exp(x/b)$	with $a > 0$
$y = a \exp((x/b)^d)$ <u>or</u> $a \exp(-(x/b)^{d)}$	with $a > 0$, $b > 0$, $d \geq 1$
$y = a [c - (c - 1) \exp(-x/b)]$	with $a > 0$, $b > 0$, $c > 0$
$y = a [c - (c - 1) \exp(-(x/b)^d)]$	with $a > 0$, $b > 0$, $c > 0$, $d \geq 1$

Benchmark Dose Approach: Covariates

dose response data stratified e.g. by sex of the animal)

When choosing the best model for each of the two strata

and then run through the possible scenarios of equality or non-equality of each parameter,

one may miss the best model-model parameter combination,

since there could be model-parameter scenarios in other combinations of models leading to a better fit.

Benchmark Dose Approach

Discussion Points between Statistician and Risk Assessor/ Risk Manager

Model fit

- search algorithm
- criterion for the selection of admissible models
    - nested: log-lik
    - non-nested: AIC
    - compare with saturated

multiple BMDs/BMDLs from multiple models

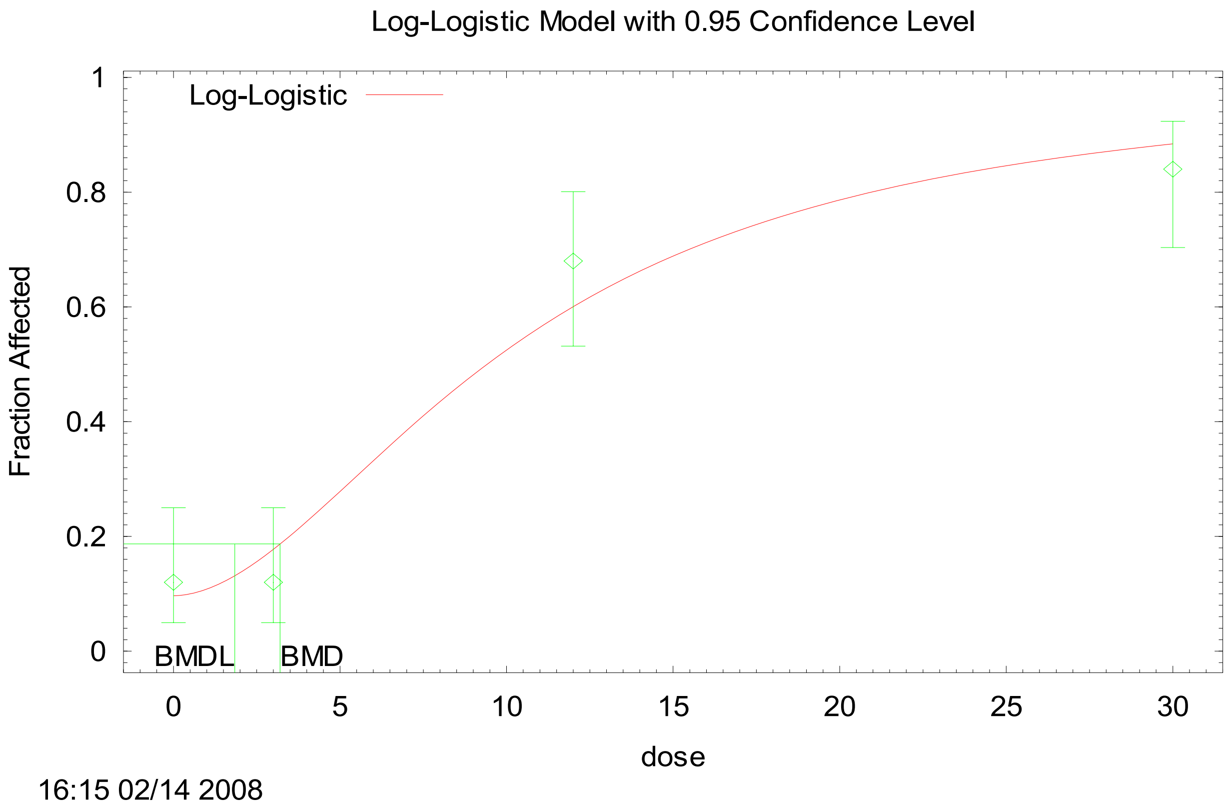
- Bayes averaging

software bugs

etc

Benchmark Dose Approach : BMDS

choice of color rather odd



CONTENT

Empirical Model based risk assessment

Quantitative Methods and Modeling

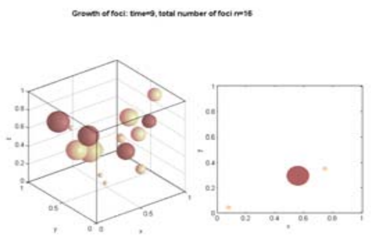
Benchmark Dose Approach

Stochastic Model based risk assessment

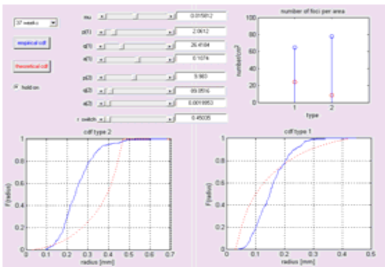
Testing Biological Hypotheses

Conclusion

Visualization to:

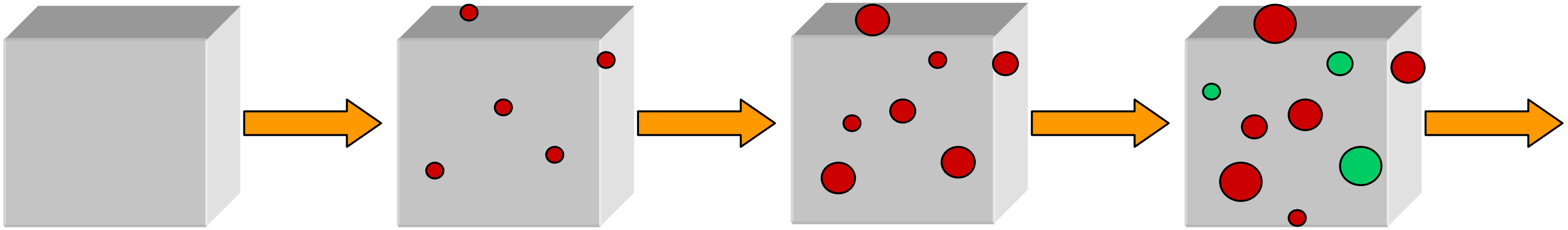


demonstrate general aspects of modeling



observe the influence of model parameters.

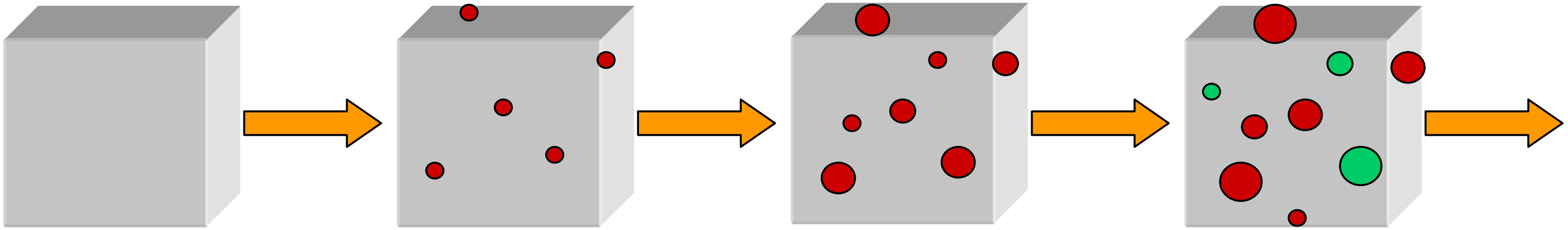
The color-shift model



- 3D-Model for formation and growth of intermediate lesions (‚foci‘)
- Centers of foci generated by Poisson process
- Starting radius r_0
- Foci growth according to exponential law
- ‚Color‘ change after exponential waiting time

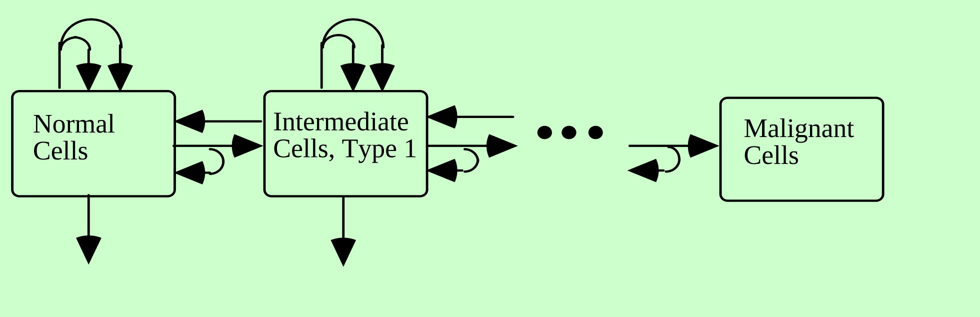
Kopp-Schneider, Portier, Bannasch (1998)
Geisler (2001)

The color-shift model



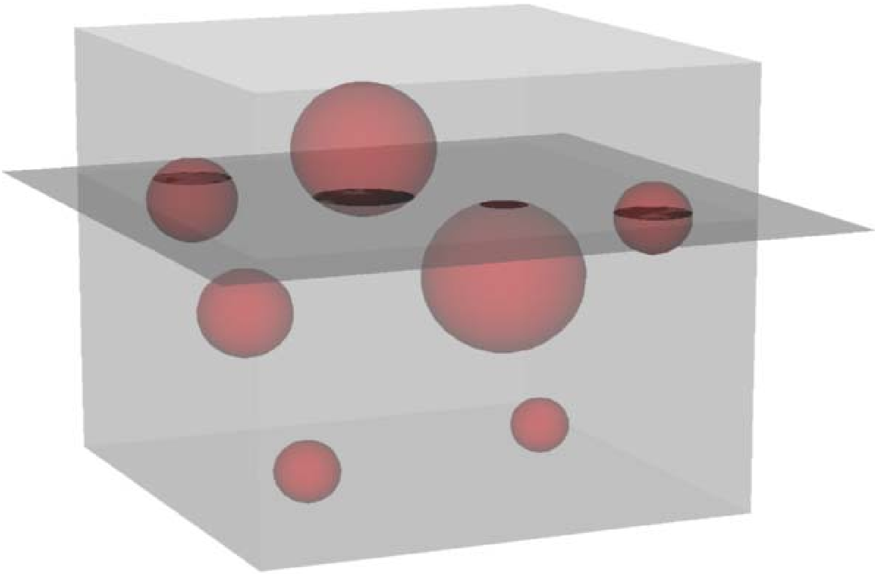
- 3D-Model for formation and growth of intermediate lesions (‚foci‘)
- Centers of foci generated by Poisson process
- Starting radius r_0
- Foci growth according to exponential law
- ‚Color‘ change after exponential waiting time

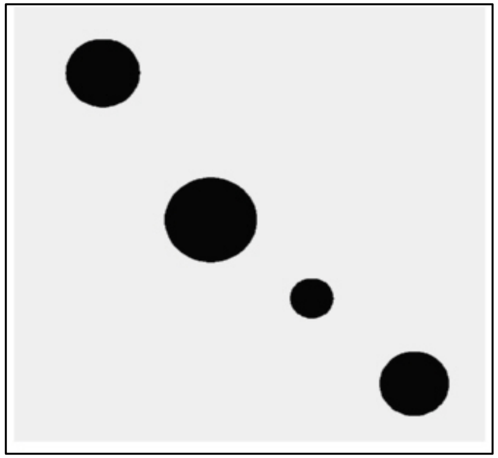
General multistage model with clonal expansion



Kopp-Schneider, Portier, Bannasch (1998)
Geisler (2001)

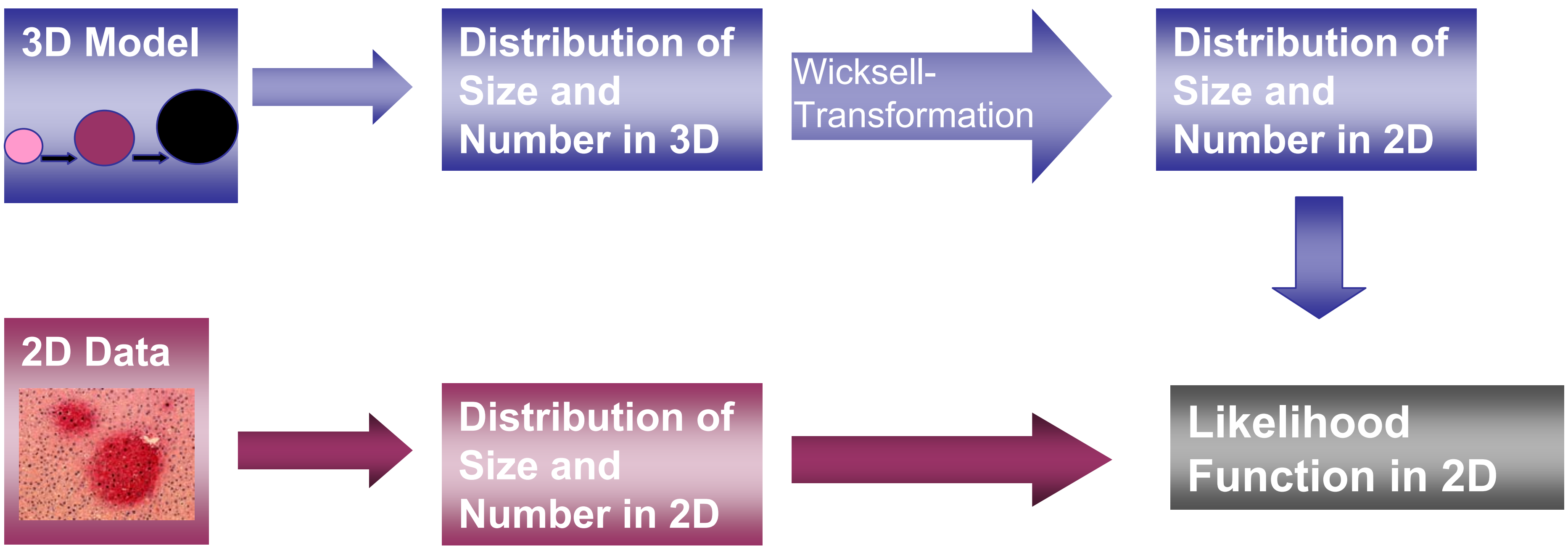
The stereological problem



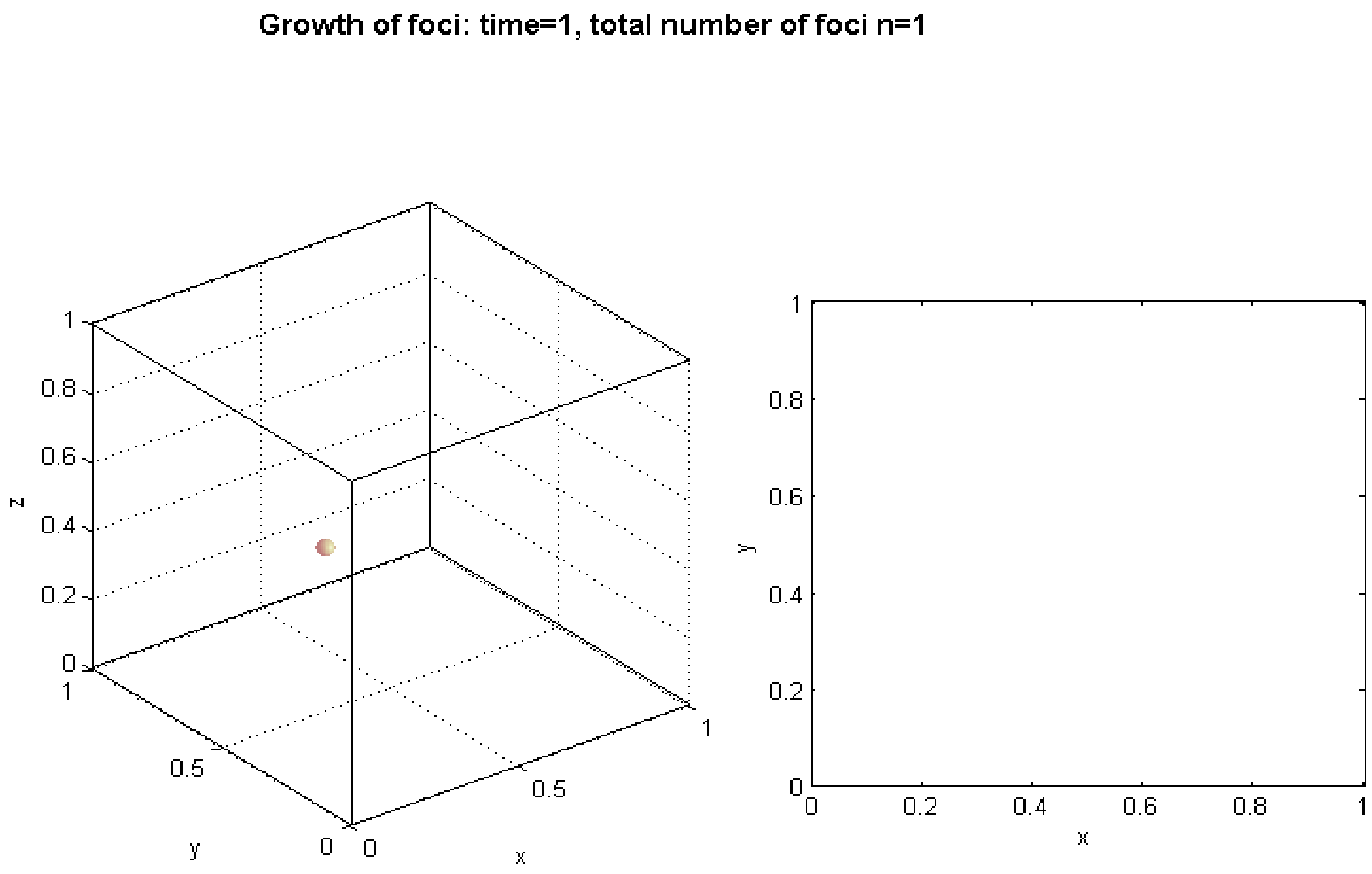


Computing the Likelihood Function:

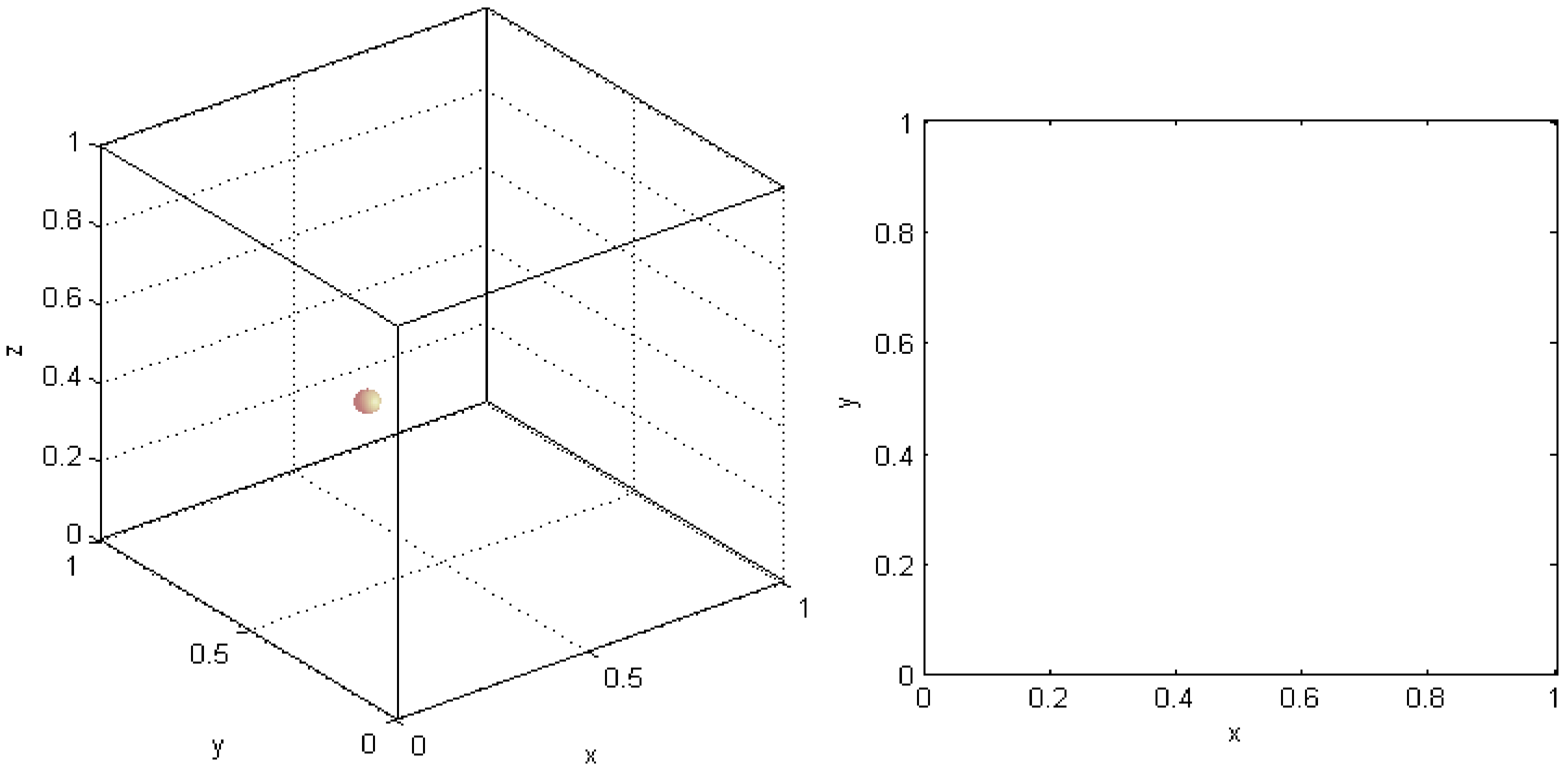
Model



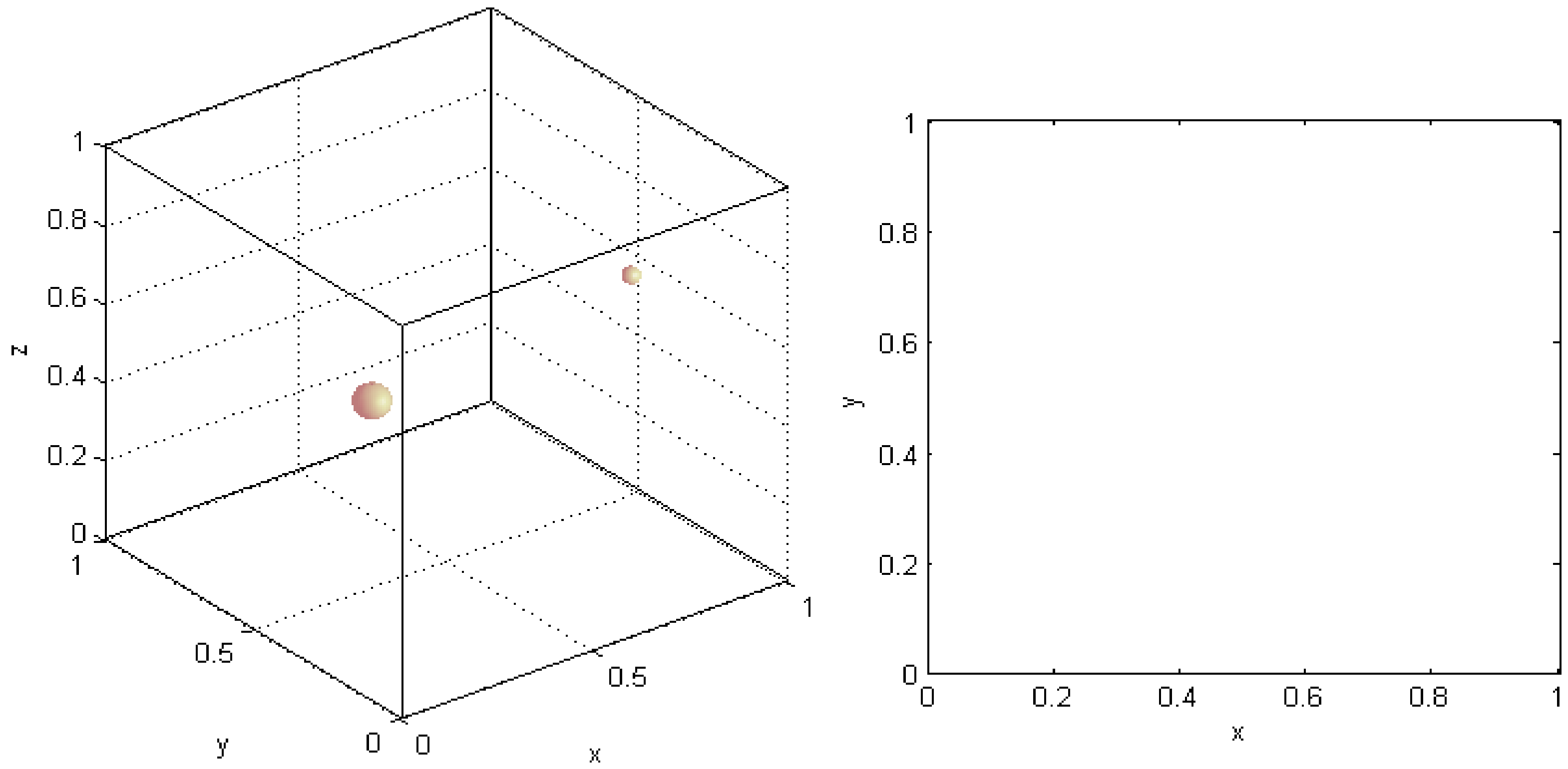
Data


Growth of foci: time=1, total number of foci n=1
z
y
x
0
0.2
0.4
0.6
0.8
1
0.5

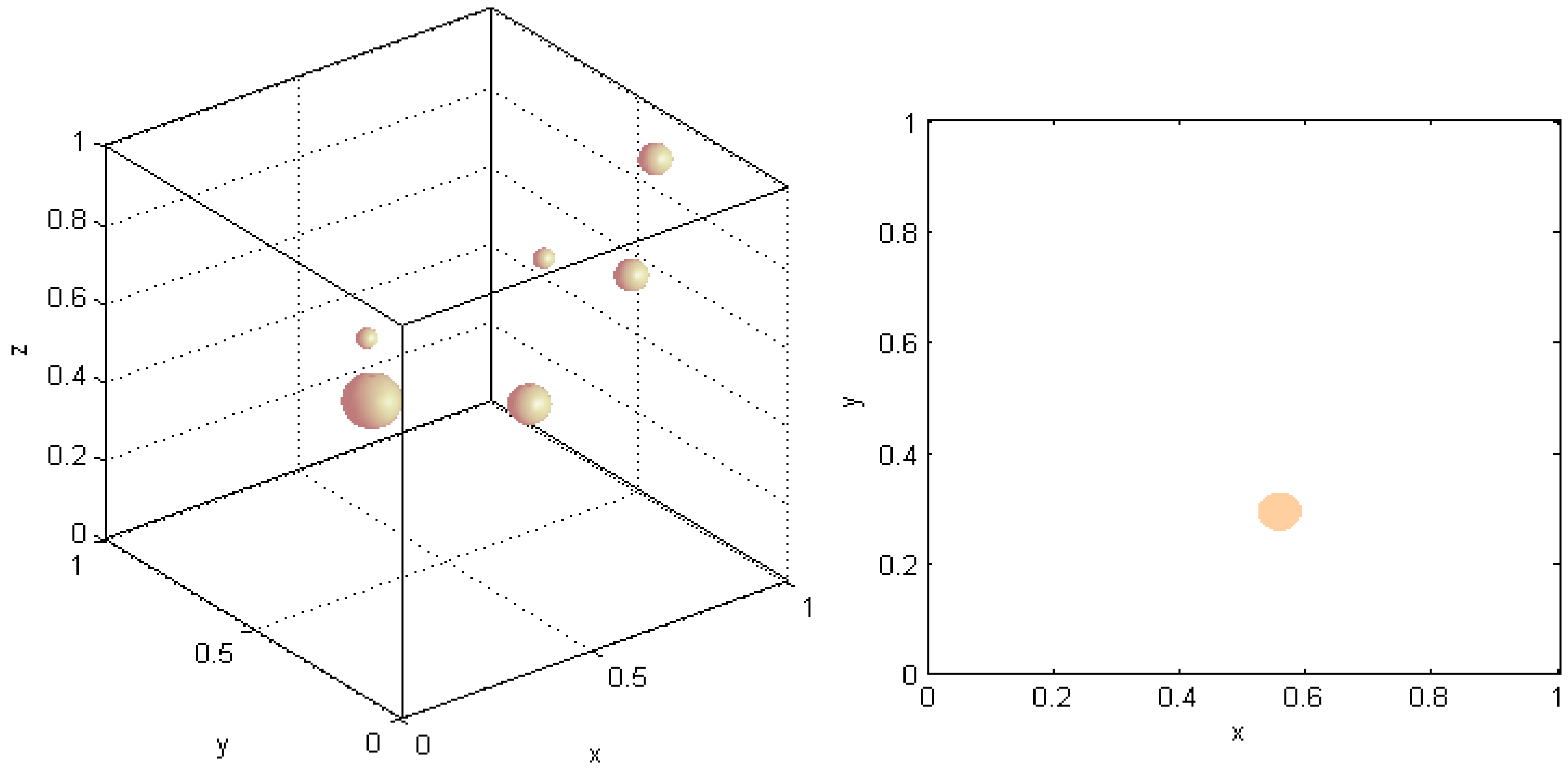
Growth of foci: time=2, total number of foci n=1



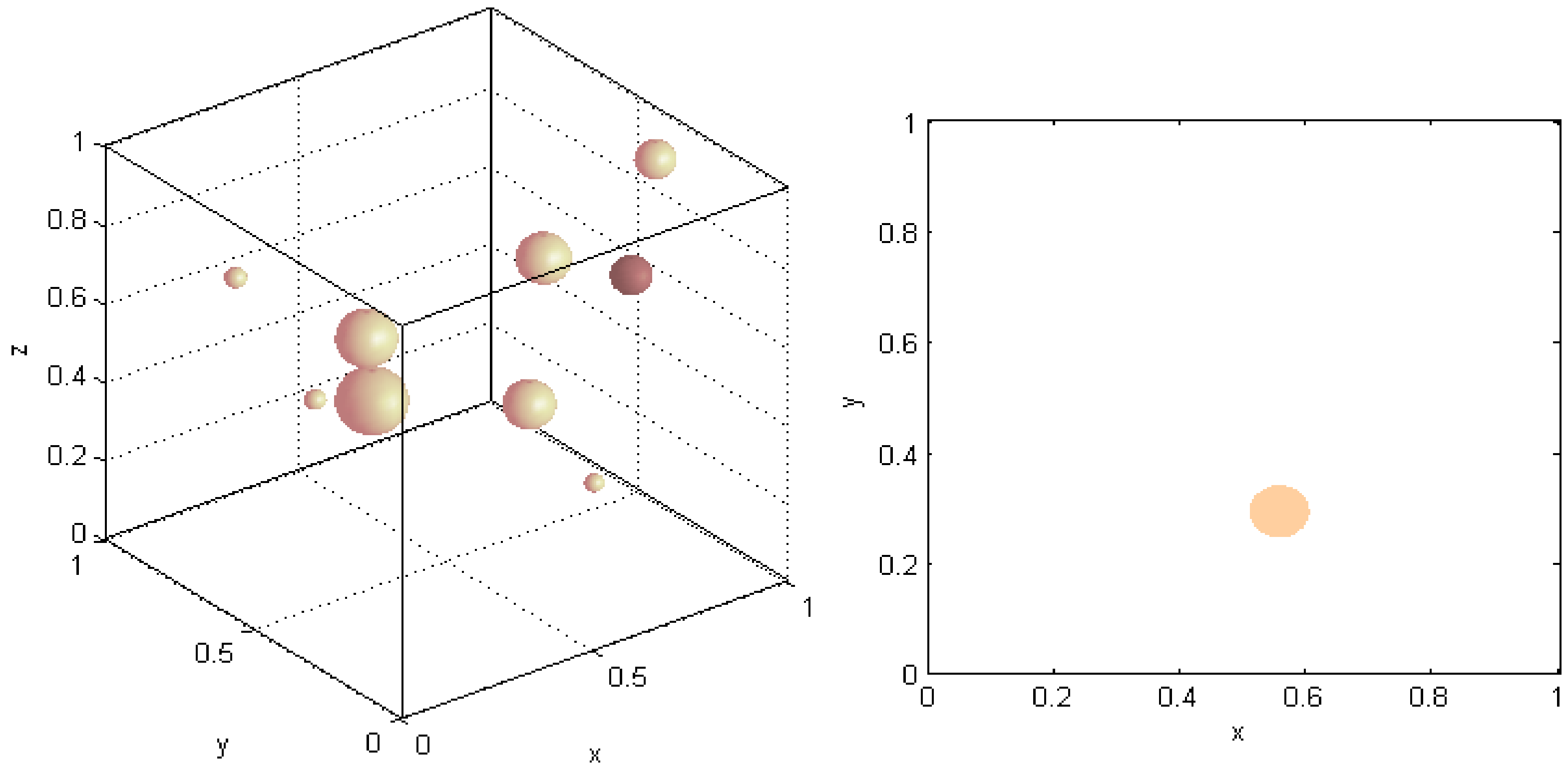
Growth of foci: time=4, total number of foci n=2



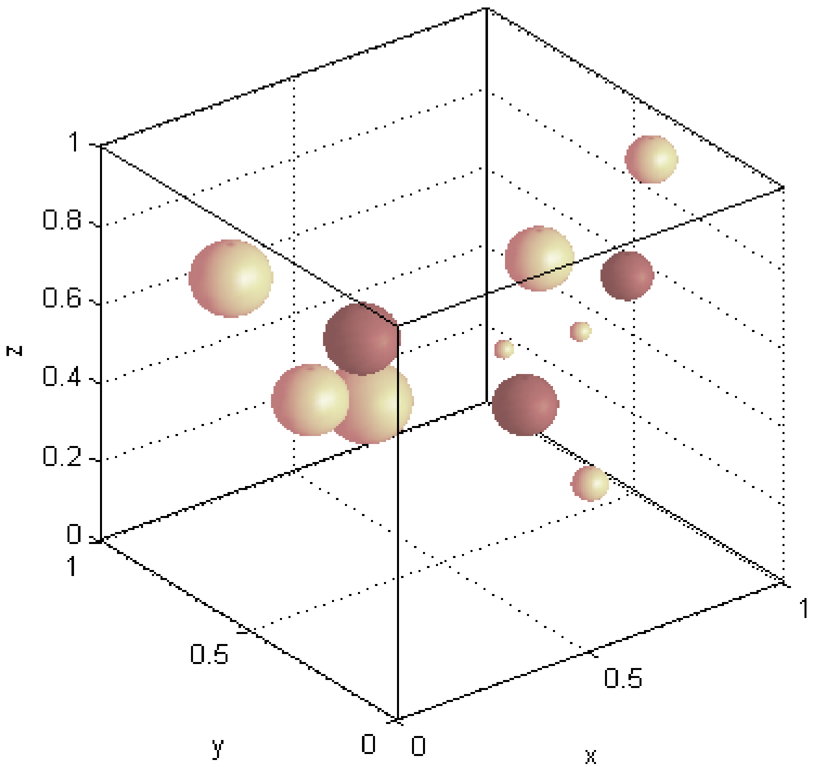
Growth of foci: time=6, total number of foci n=6

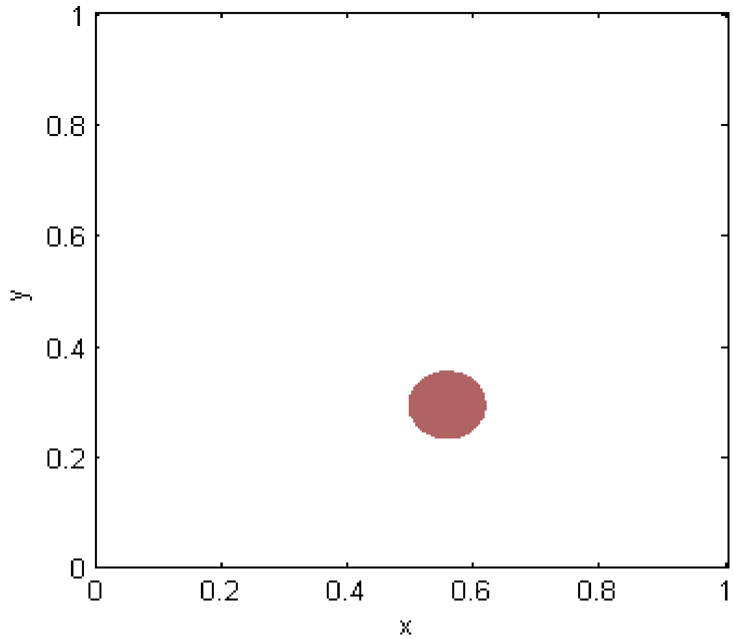


Growth of foci: time=7, total number of foci n=9

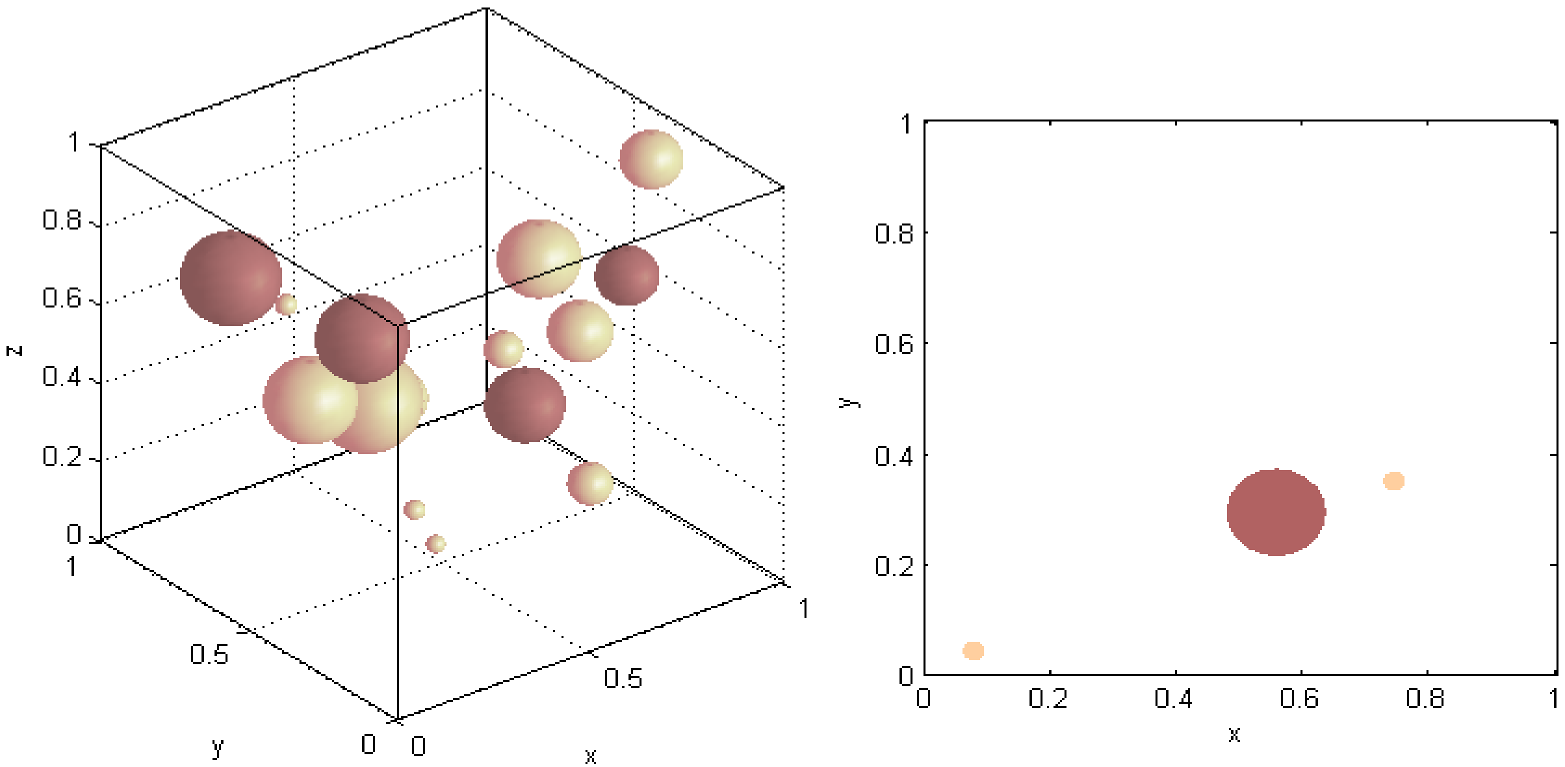


Growth of foci: time=8, total number of foci n=12





Growth of foci: time=9, total number of foci n=16



size distribution: 3D-Density

Formation of foci:
Poisson-process parameter μ.

Growth of foci: $B_1 \sim Beta(p_1,q_1,a_1)$
$B_2 \sim Beta(p_2,q_2,a_2)$

Color-Shift: At radius r_{shift}
$\Rightarrow$ **8 parameters**

Color 1

$$g_3(r) = \frac{p_1+q_1-1}{rta_1(p_1-1)}\left(I_{1-\frac{\ln(\frac{r}{r_0})}{a_1 \cdot t}}(q_1,(p_1-1))\right)$$

Color 2

$$g_3(r) = \frac{p_2+q_2-1}{r(p_2-1)B(p_1,q_1)a_2}\int_{\frac{\ln(\frac{r_{switch}}{r_0})}{a_1 t}}^{1} b_1^{(p_1-1)}(1-b_1)^{(q_1-1)}\left(I_{1-\underline{b_2}(b_1,r)}(q_2,(p_2-1))\right)db_1$$

with:

$$\underline{b_2}(b_1,r) := \min\left\{\left(\frac{\ln(\frac{r}{r_{switch}})}{a_2\left(t-\frac{\ln(\frac{r_{switch}}{r_0})}{a_1 b_1}\right)}\right),1\right\}$$

$B(p,q)$: Beta-function with parameters p,q .
$I_x(p,q)$: Incomplete Beta-function with parameters p,q,x .

Steps to find maximum-likelihood-parameters :

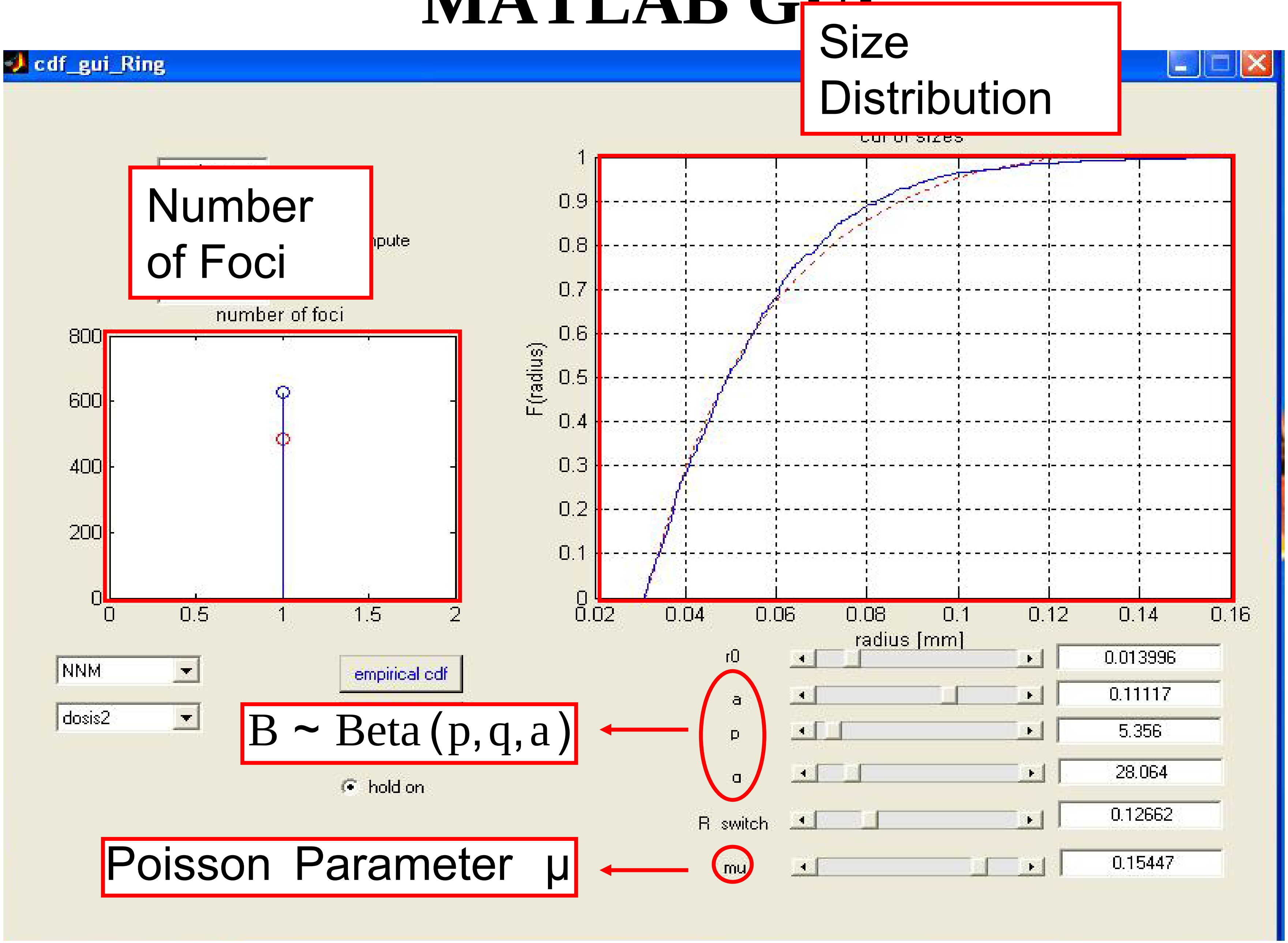
1. Develop analytical expression for distribution of size and number of foci in 2D and derive log-likelihood-function .

2. Find starting parameters for optimization.

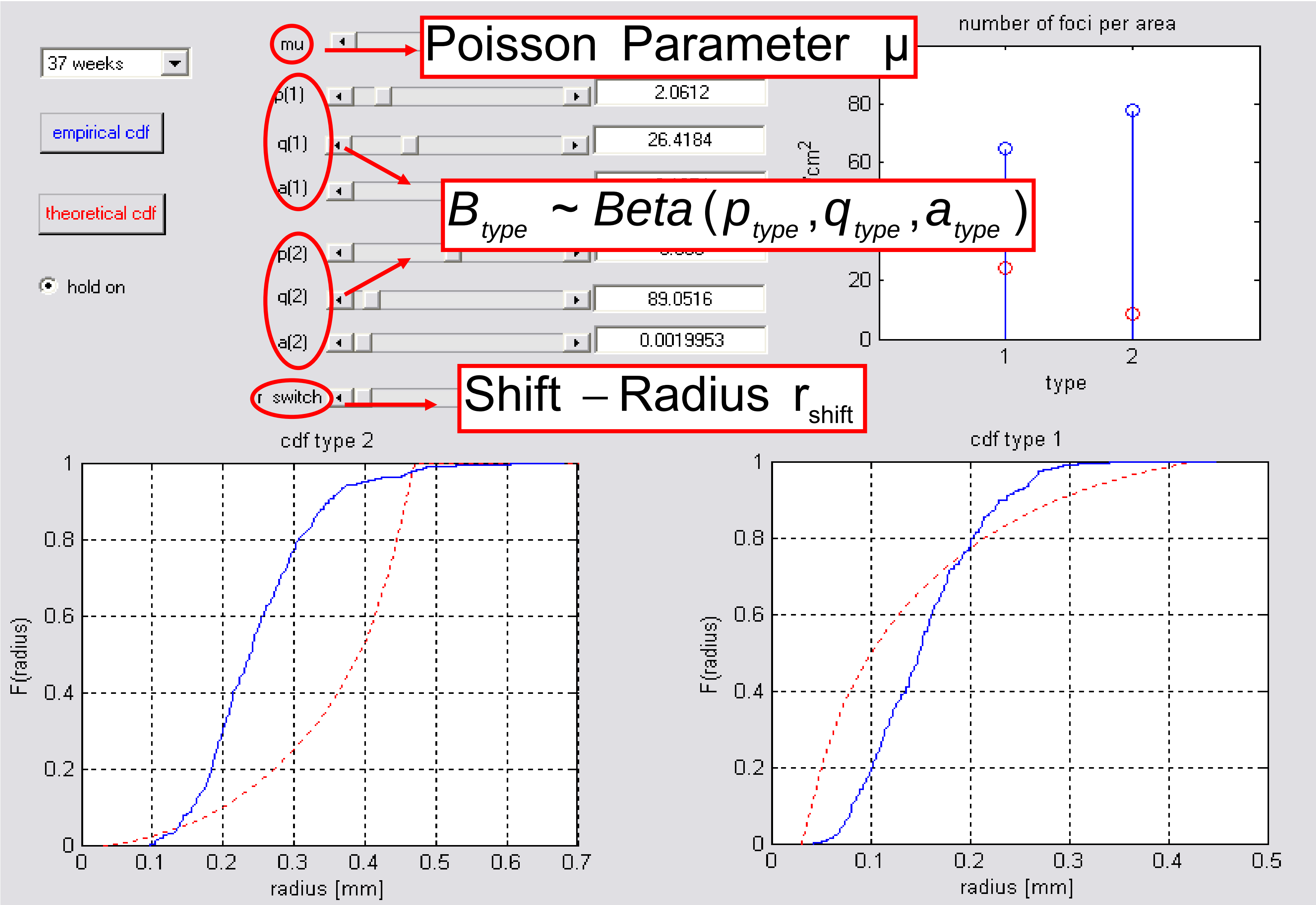
3. Maximize log-likelihood-function using *fmincon* in *MATLAB*.

4. Compare results with data.

MATLAB GUI


cdf_gui_Ring
Size Distribution
cdf of sizes
Number of Foci
number of foci
F(radius)
radius [mm]
NNM
dosis2
empirical cdf
B ~ Beta (p, q, a)
hold on
r0
a
p
q
R switch
mu
0.013996
0.11117
5.356
28.064
0.12662
0.15447
Poisson Parameter μ

MATLAB GUI



37 weeks
empirical cdf
theoretical cdf
hold on
mu
p(1)
q(1)
a(1)
p(2)
q(2)
a(2)
r switch
2.0612
26.4184
89.0516
0.0019953
Poisson Parameter μ
$B_{type} \sim Beta(p_{type}, q_{type}, a_{type})$
Shift – Radius r_{shift}
number of foci per area
type
cdf type 2
cdf type 1
F(radius)
radius [mm]

Objectives were

demonstrate developments in statistical science which can be applied for the risk assessment and should be recognized by risk assessors in regulation and industry

draw the attention of the computational statistics community to an important field of application of statistical computing techniques widely not or no more realized

In accordance with the mission of IASC (Int. Assoc. Statistical Computing)

fostering evaluations of statistical computing techniques and programs

training individuals in sound and useful statistical computing procedures

Thank You!

Wicksell-Transformation

Size Distribution in 2D (given radius > ε) :

$$G_2^{\varepsilon}(y) = 1 - \frac{1}{\mu_\varepsilon}\int_y^\infty \sqrt{r^2 - y^2}\, g_3(r)\,dr$$ Size Distribution in 2D

$$g_2^{\varepsilon}(y) = \frac{y}{\mu_\varepsilon}\int_y^\infty \frac{1}{\sqrt{r^2 - y^2}}\, g_3(r)\,dr$$ Density in 2D

$$\mu_\varepsilon = \int_\varepsilon^\infty \sqrt{r^2 - \varepsilon^2}\, g_3(r)\,dr$$ Mean size in 3D

$g_3(r)$ Density in 3D

Overview of current carcinogenesis models