Model based risk assessment and its data needs for an improved risk management with application for food risks

Lutz Edler

Risk Analysis Research Center, Institute of Mathematical Statistics, Tokyo, Japan,
Department of Biostatistics, German Cancer Research Center, Heidelberg, Germany

Cherry Bud Workshop „Interaction Through Data“ , Yokohama, March, 2007

CONTENT

- Introduction to risk analysis and the steps of human health risk assessment
- Model based risk assessment
  - Quantitative Methods and Modeling for food risks
  - Benchmark Dose Approach
- Stochastic Carcinogenesis Models for Testing Biological Hypotheses
- Discussion

Risk Analysis

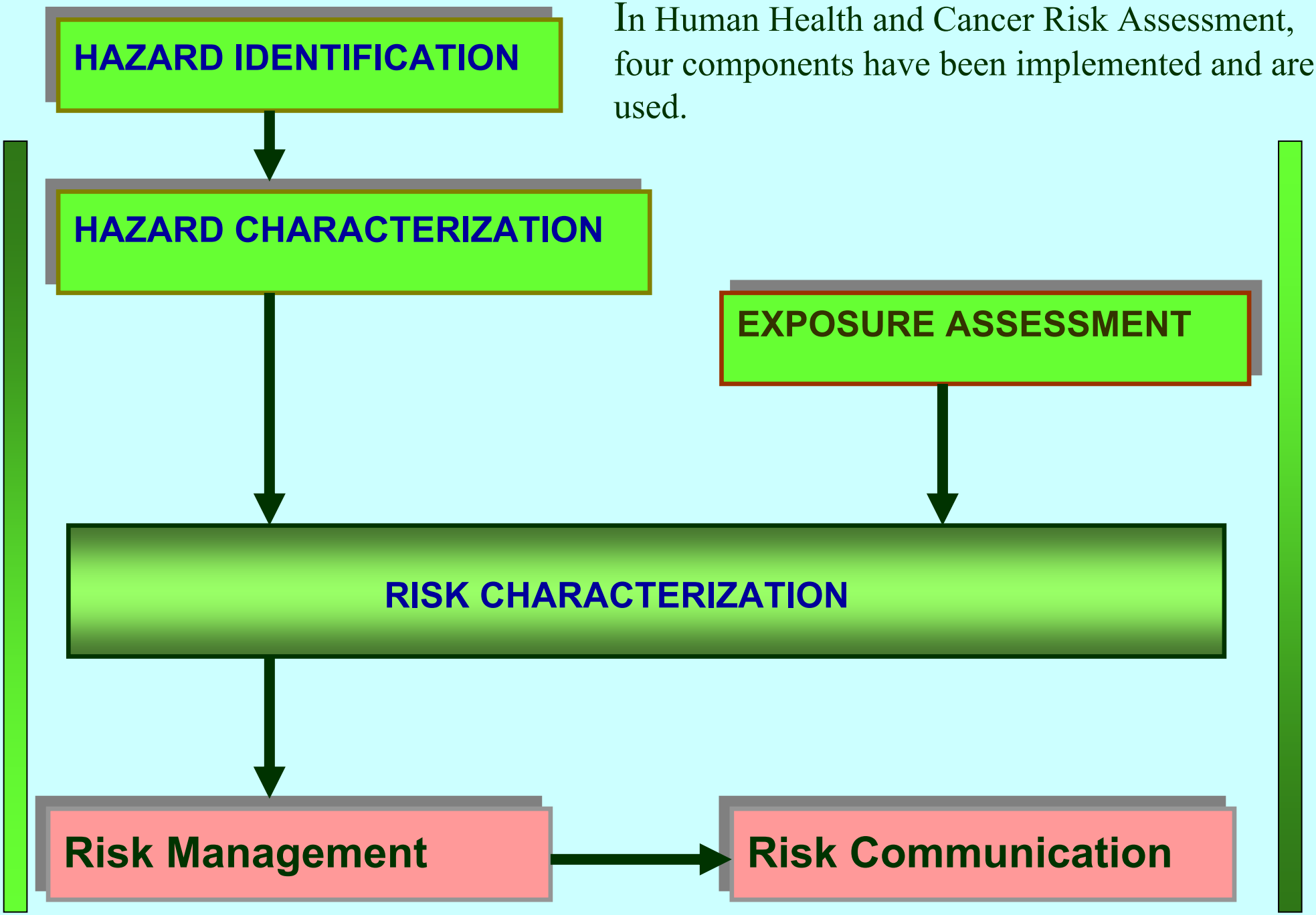
is present and relevant for many components of a society.

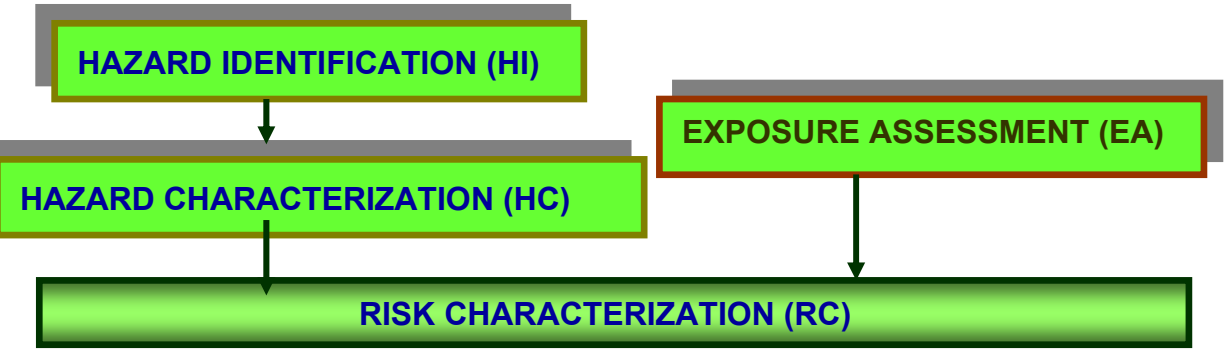
Risk Analysis in Finance
Risks of Large Corporate Units
Transportation Risk Analysis
Genetic Risks
Human Health and Cancer Risks
Food Safety and Food Risk Assessment
Risks of Illness and Injury due to Occupational Exposures
Risks of Human Activities and Interventions on Ecosystems

Risk of Poverty
Risks in Human Health Systems (Costs and Aging)
Risk of Natural Disasters
Risks of Terrorism

International Conference on Statistical Methods for Risk Analysis (ICSMRA)
Satellite Conference to the Session of the ISI
August 30-31, 2007, Lisbon, Portugal
contact: edler@dkfz.de

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 = ***Dose-response assessment***	**non-linear regression methods, fitting stochastic models**
Risk characterisation	uncertainty analyses, probabilistic methods

CONTENT

- Introduction to risk analysis and the steps of human health risk assessment
- Model based risk assessment for food and diet
  - **Quantitative Methods and Modeling**
  - Benchmark Dose Approach
- Stochastic Carcinogenesis Models for Testing Biological Hypotheses
- Discussion

Food Risk Assessment

- Food contains **natural and man made ingredients** which are toxic in principle.
- Human food is composed of **many components and ingredients** which may **accumulate** over a long time period of human life.

- Food and diet has to be **safe** for human consumption and Food producers and regulators are forced guarantee **"zero risk"**
- **Food is needed** to be daily consumed for maintenance of human functions

Goal: find an exposure level , e.g. ADI (Acceptable Daily Intake) as an "estimate of the amount of food that can be ingested daily over a lifetime without appreciable health risk".

Partners for Interaction:

- **Food Industry**
- **Regulators / Government**
- **Public** Consumer
  Press& Media

There are two different approaches

Safety assessment focuses on determining a safe dose for human exposure by identifying a dose that appears to be without adverse effects in experimental studies, usually, the so-called “No-Adverse-Effect Level” NOAEL

dividing that dose by several factors to arrive at a dose considered "safe" for human exposure, with the so-called safety/uncertainty factors UFs

Risk extrapolation focuses on estimates of risk at low doses by fitting dose-response models to data on the incidence of an adverse effect

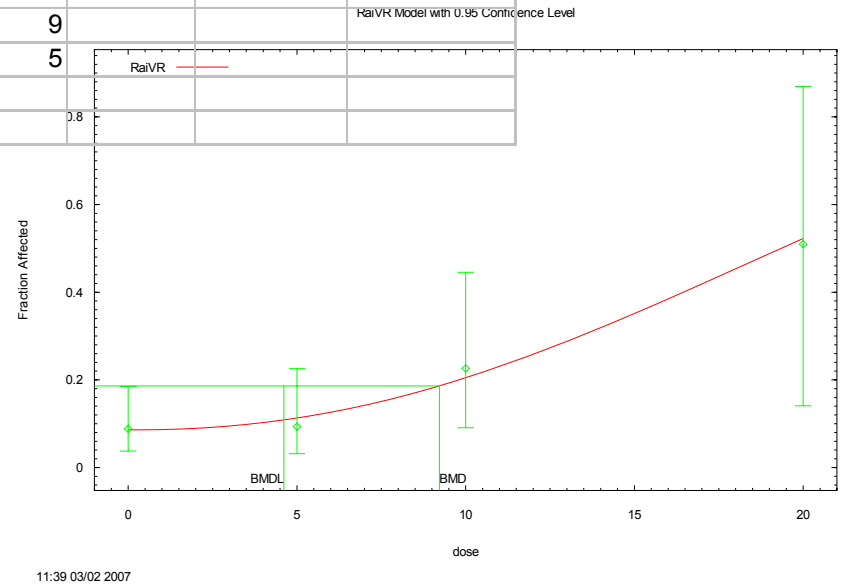
using these models to estimate the risk at lower levels of exposure.

Both Approaches make use of dose-desponse data:

n_k units examined dose levels $0=d_0< d_1< ... < d_K$, k=0,1,... K

- The NOAEL tests the dose groups against the control.
- The dose-response modeling fits a curve and estimates at low doses

Control			5 mg/kg			10 mg/kg			20 mg/kg	
Resorptions	Implantations		Resorptions	Implantations		Resorptions	Implantations		Resorptions	Implantations
1	8		3	7		0	9		0	7
0	7		0	10		1	8		5	6
3	6		1	8		0	1		0	8
0	10		0	2		1	10		8	8
0	1		0	6		7	9		10	10
0	9		0	10		0	8		0	9
2	9		1	3		0	4		3	3
1	7		0	8		2	8			
2	10		0	7		4	6			
0	7		1	9		3	6			
1	8		1	5		0	1			
0	4					1	9			
0	10					0	5			
0	10									
0	7									



Outcome of the NOAEL is design dependent

on the sample sizes and is directly related to the power of the statistical test.

NOAEL promotes low powered small studies

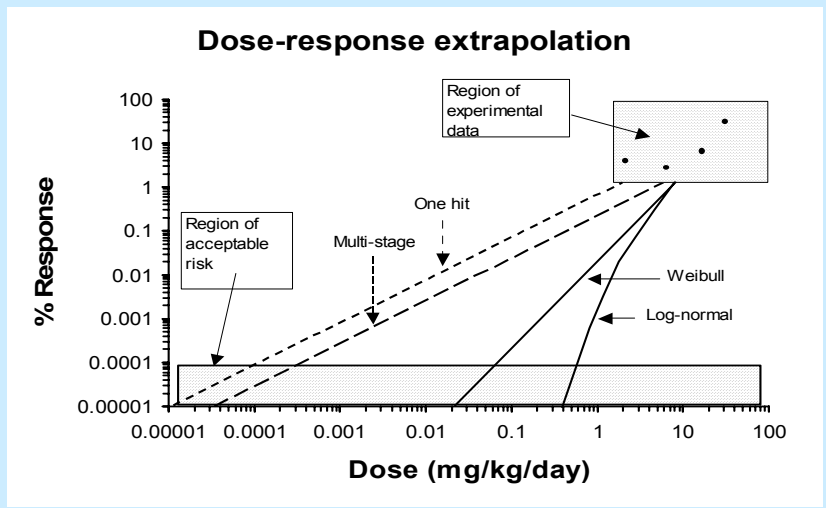
on the design of the experiment and especially on the dose-spacing. NOEAL must be one of the experimental doses.

The risk level at the NOAEL is not zero and may be higher than communicated

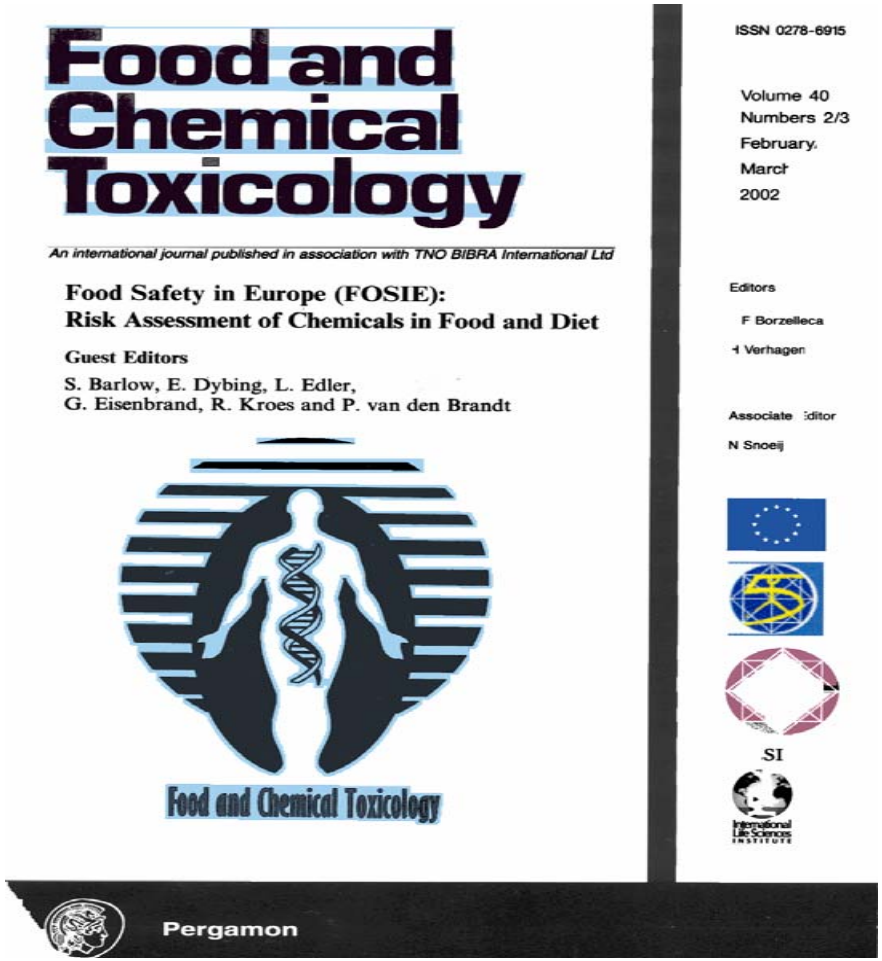
on the sensitivity of the toxicological endpoint,

sensitivity of its measurement,

Outcome of dose-response modeling is model dependent



• has been investigated in the EU Project FOSIE: Food safety in Europe: Risk Assessment of Chemicals in Food and Diet.



- SAR and TTC
- THRESHOLD MODELS
- CATEGORICAL REGRESSION
- CHEMICAL-SPECIFIC ADJUSTMENT FACTORS
- NON-THRESHOLD MODEL
- **BENCHMARK DOSE**
- PROBABILISTIC RISK ASSESSMENT
- PHYSIOLOGICALLY-BASED MODELS

Alternative: TWO STEP APPROACH

- **analyse the dose-response date where effects are actually observed and determine of a point of departure**

- **do the low dose extrapolation "appropriately", usually using uncertainty factors**

this separates the regulatory risk management task of setting a reference dose better from the scientific task of estimating a dose-response curve and deriving estimates [EPA, 1996].

Example:
Benchmark Dose method (BMD)

CONTENT

- Introduction to risk analysis and the steps of human health risk assessment
- Model based risk assessment for food and diet
  - Quantitative Methods and Modeling
  - **Benchmark Dose Approach**
- Stochastic Carcinogenesis Models for Testing Biological Hypotheses
- Discussion

Benchmark Dose Approach

The Benchmark Dose approach (BMD) is 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.

$P(d|\theta)$ quantal Dose Resposne Model

p Benchmark response level (BMR)

$$p = \frac{P(d_p|\theta) - P(0|\theta)}{1 - P(0|\theta)}$$ inverse fct Q

$$d_p(\theta) = Q\left[P(0|\theta) + p(1 - P(0|\theta)\right]$$

BMD $d_p(\theta)$ to the **BMR** p

$BMDL_{95}$ is the lower one-sided confidence bound of $d_p(\theta)$ to the level 0.95 95%)

Cherry Bud W

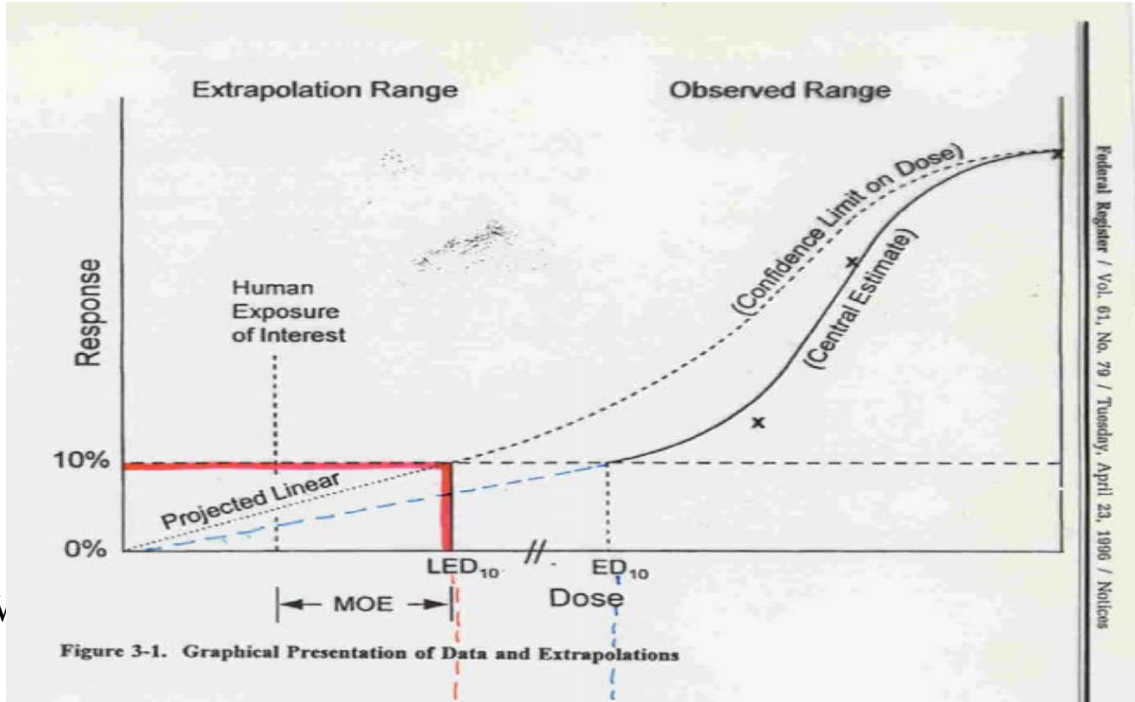
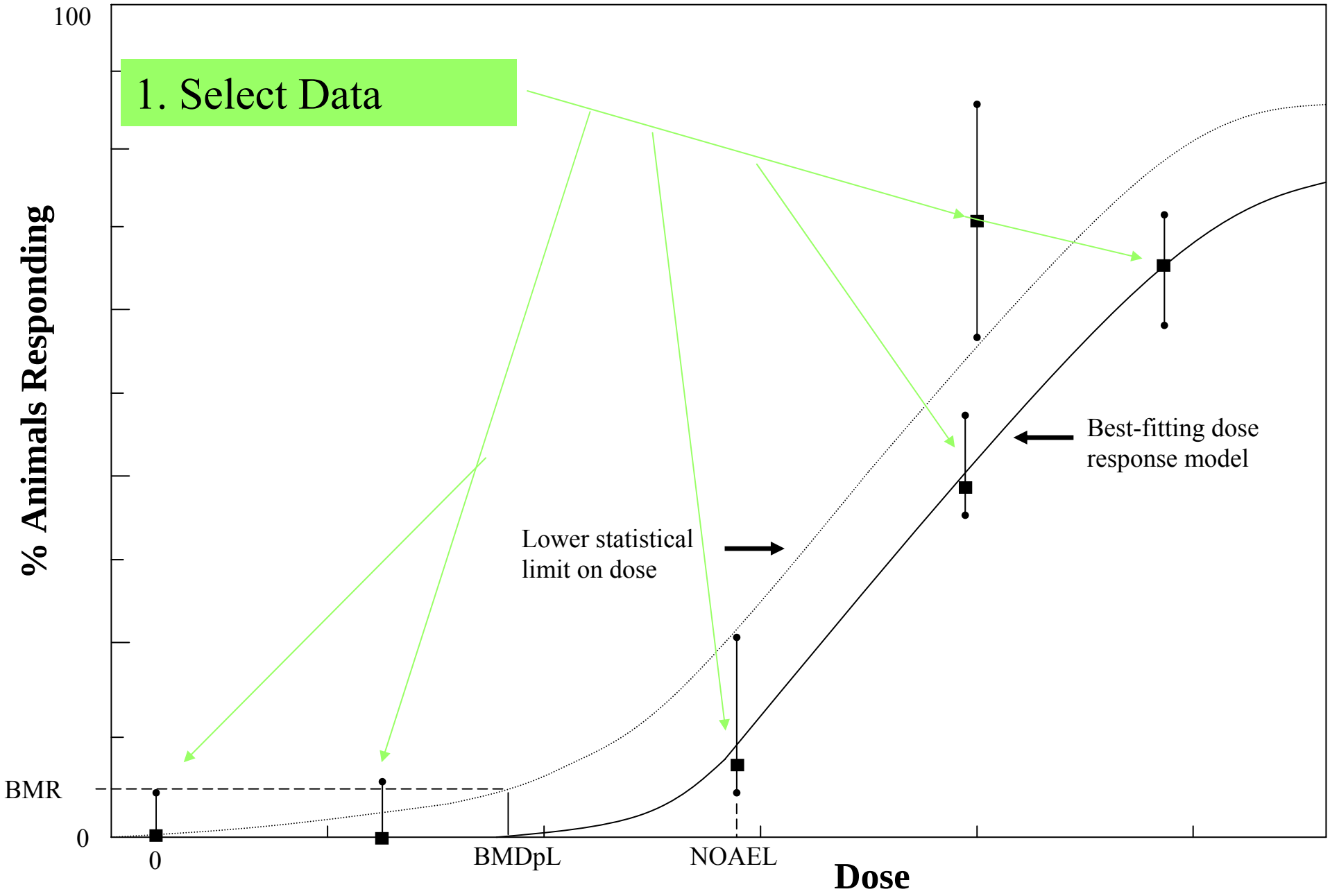
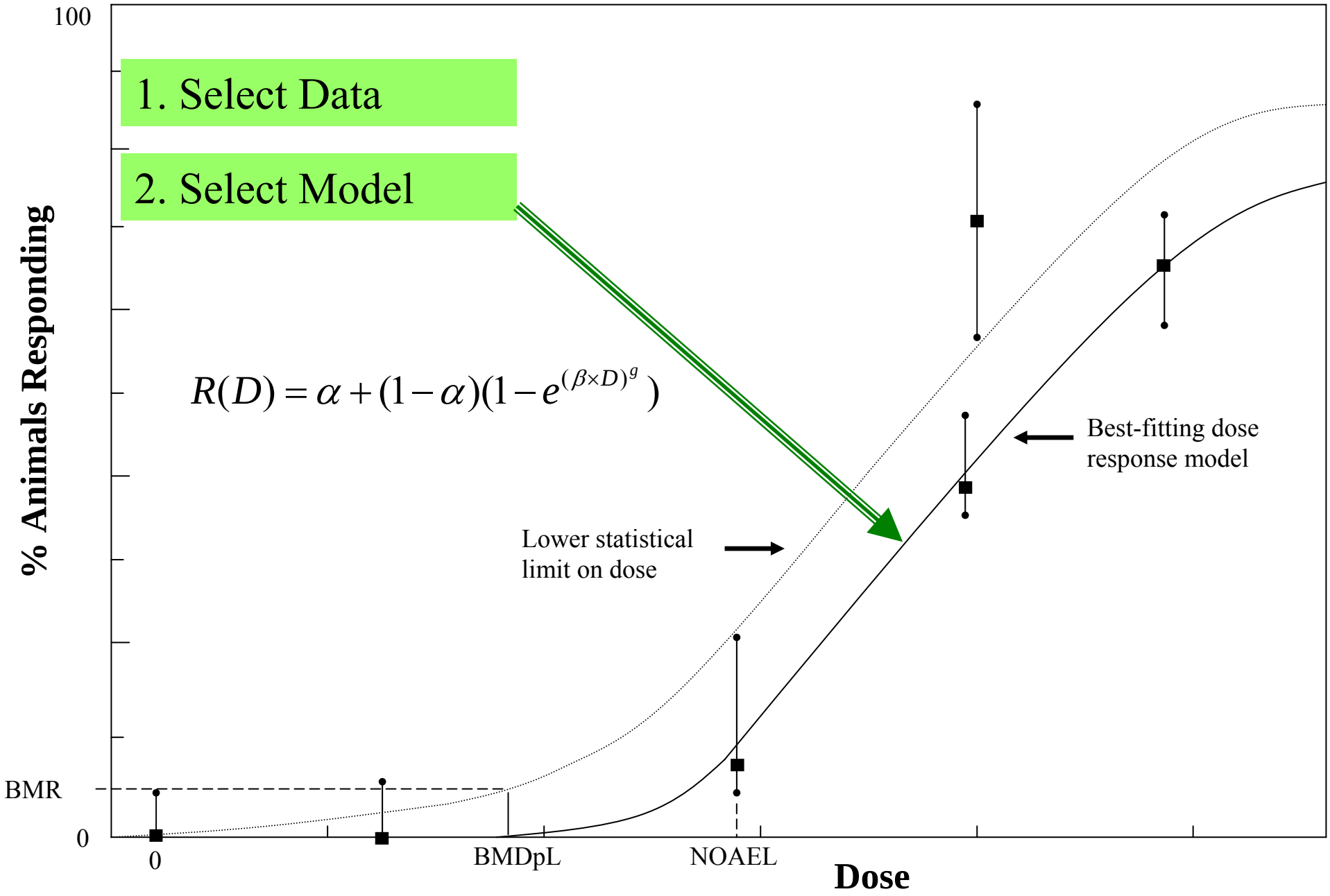


Figure 3-1. Graphical Presentation of Data and Extrapolations

How to apply the BMD?



How to apply the BMD?



Which family?

How large the family?

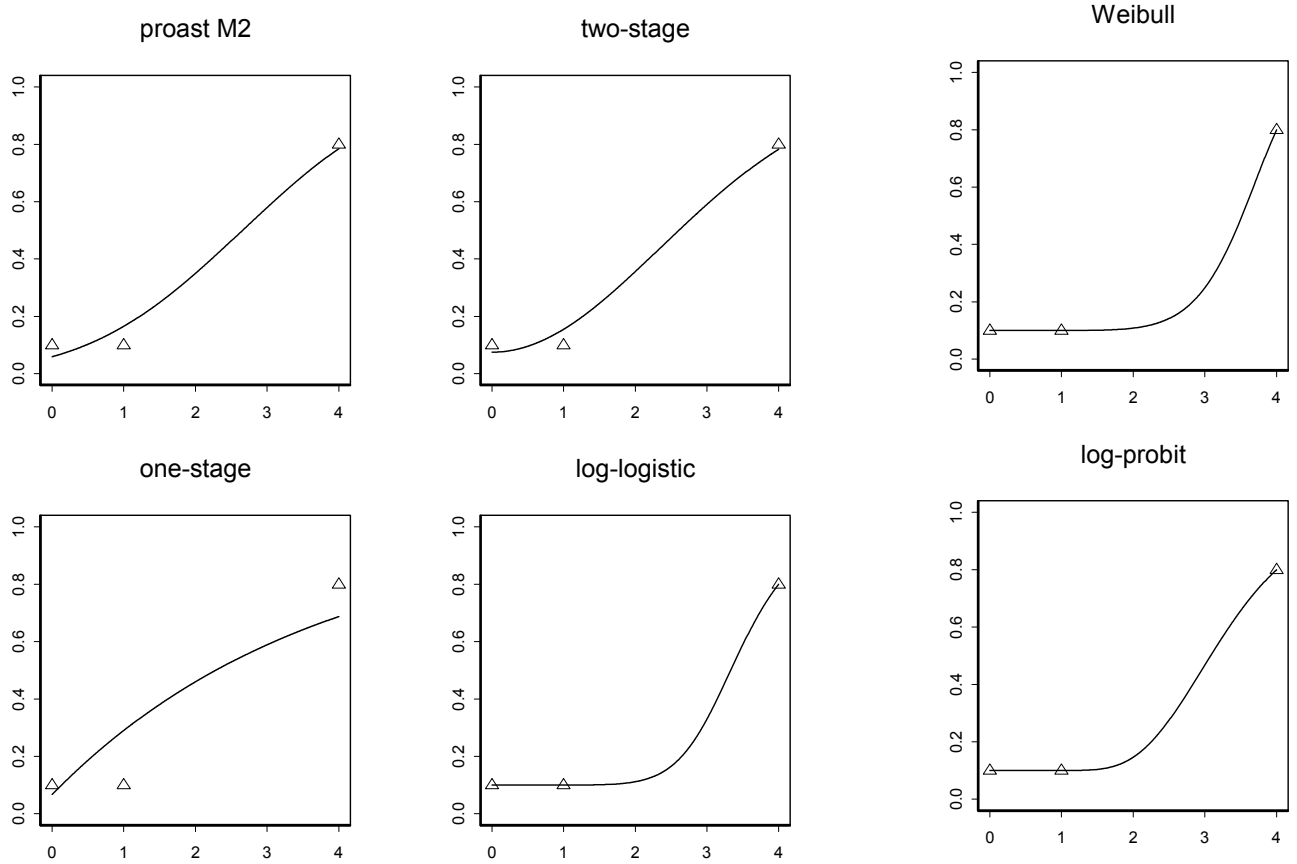
Super-families?

nested versus unnested families ?

Continuous data		
Exponential family	no.param.	function
Model 2	2	$y = a \exp(bx)$
Model 3	3	$y = a \exp(bx^d)$
Model 4	3	$y = a [c – (c-1)\exp(bx)]$
Model 5	4	$y = a [c – (c-1)\exp(bx^d)]$
Hill family		
Model 2	2	a [1 - x/(b+x)]
Model 3	3	a [1 - x^d/(b^d+x^d)]
Model 4	3	a [1 + (c-1)x/(b+x)]
Model 5	4	a [1 + (c-1)x^d/(b^d+x^d)]
Power model[1)]	3	$y = a + bx^d$
Quantal data		
Log-probit	3	$y = a + (1-a)\ \Phi\ (\log(x/b) / c)$
Log-logistic	3	……
Weibull	3	$y = a + (1-a) \exp((x/b)^c)$
Multistage		
One-stage	2	$y = \exp(-a – bx)$
Two-stage	3	$y = \exp(-a – bx – cx^2)$
Three-stage	4	$y = \exp(-a – bx – cx^2 – dx^3)$

form EFSA BMD.WG

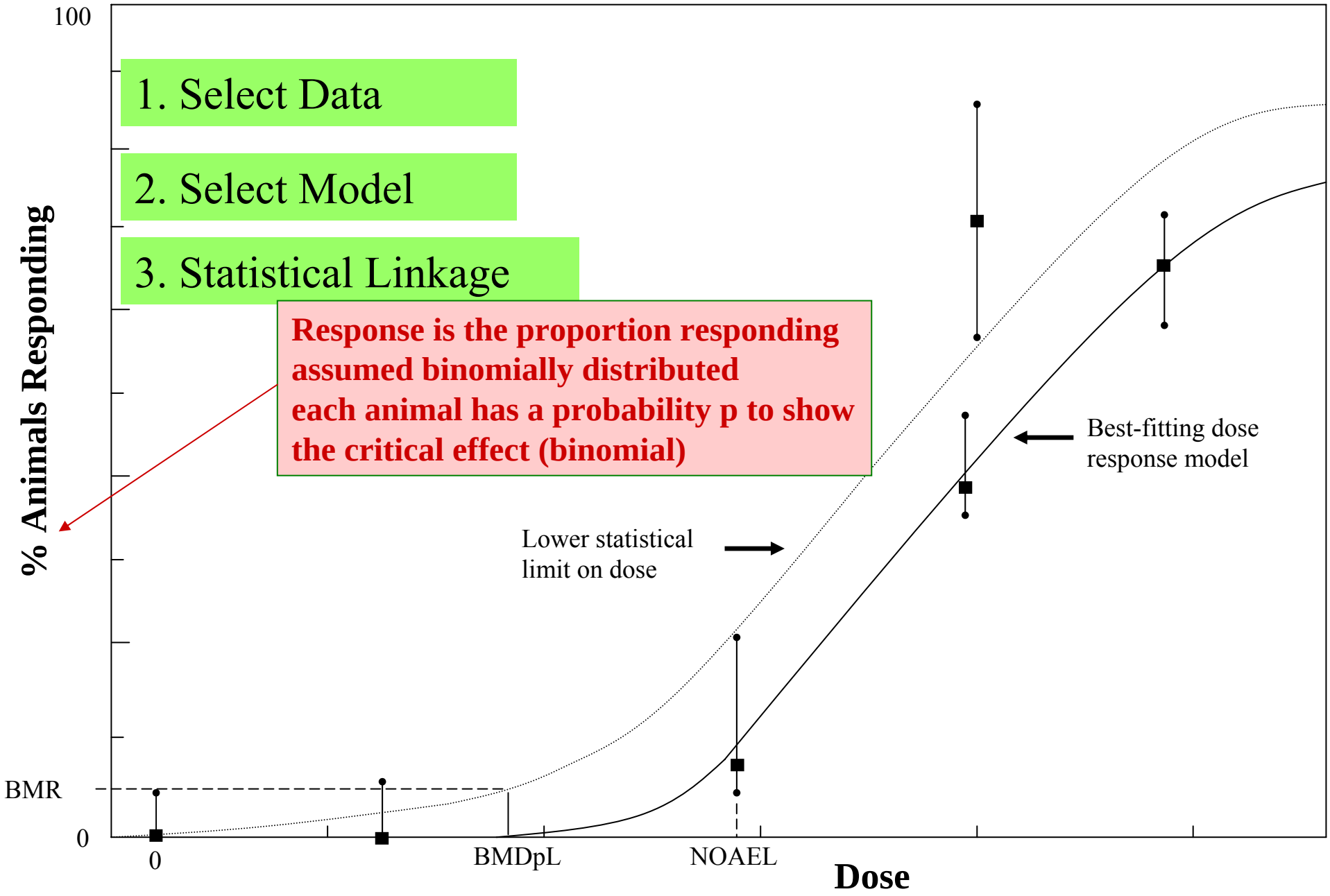
Limitations of "data poor" experiments: model dependency



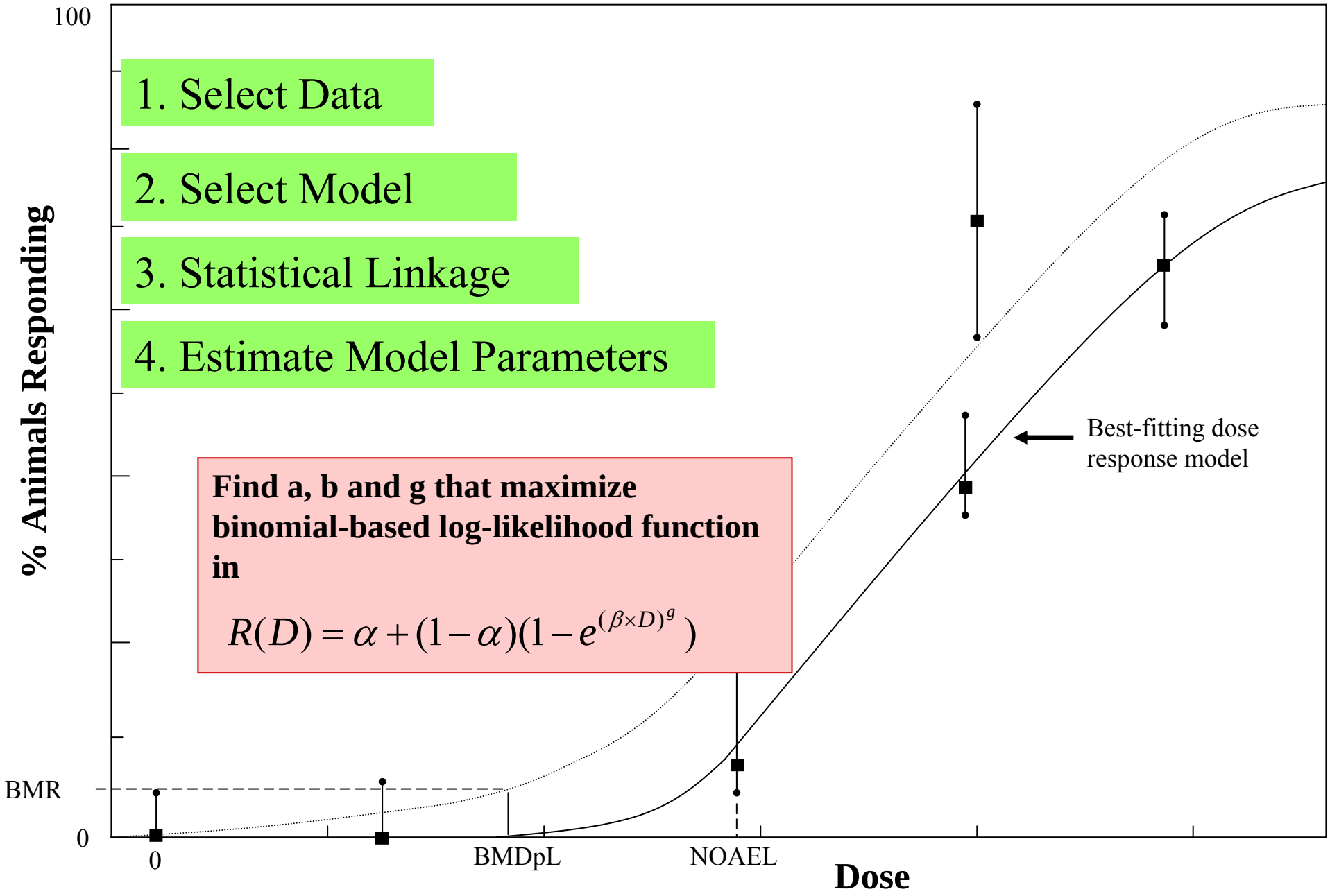
example form EFSA BMD.WG, W. Slob

This needs a strong interaction between statisticians and risk managers

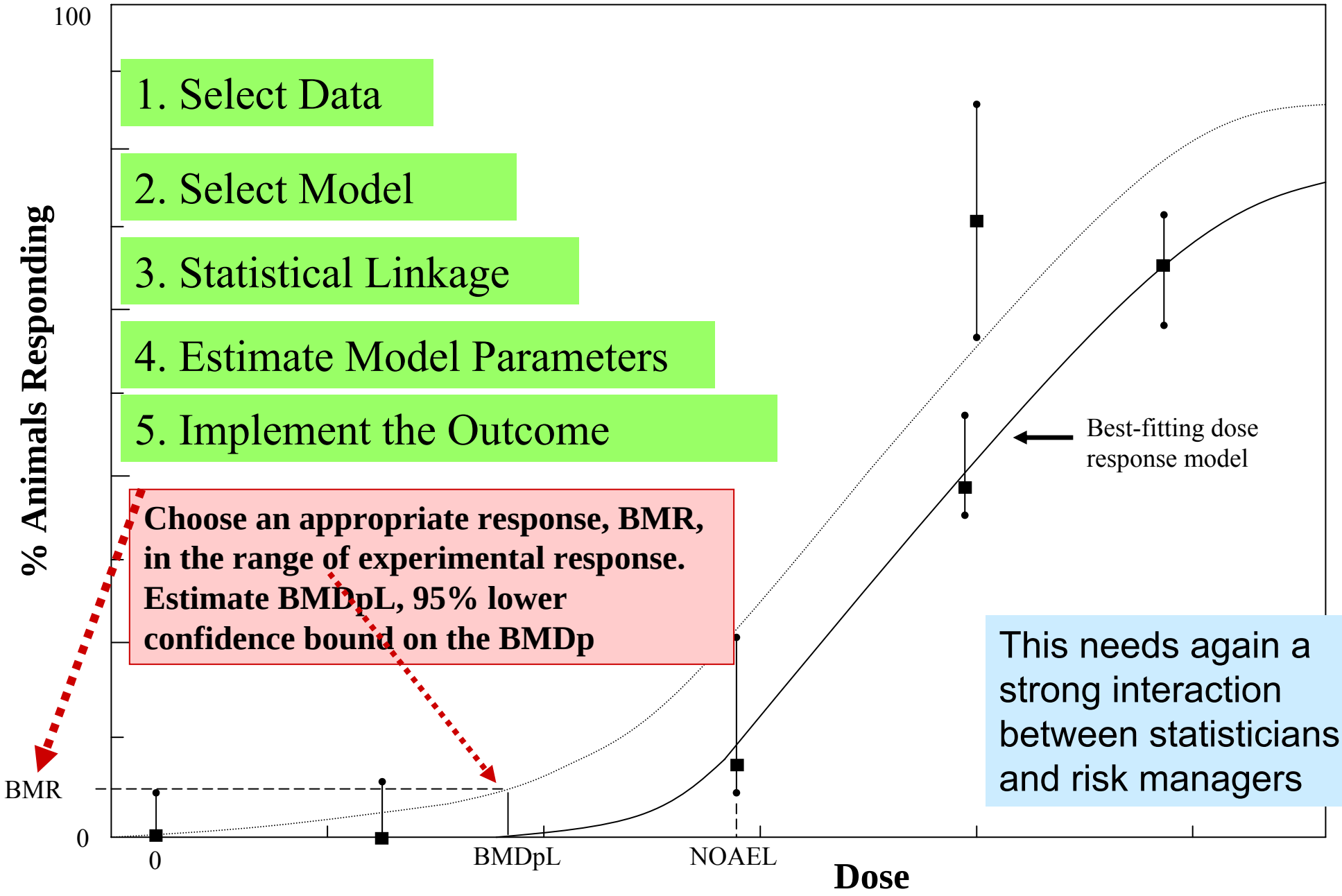
How to apply the BMD?



How to apply the BMD?



How to apply the BMD?



Both Approaches make use of Dose-Response Data:

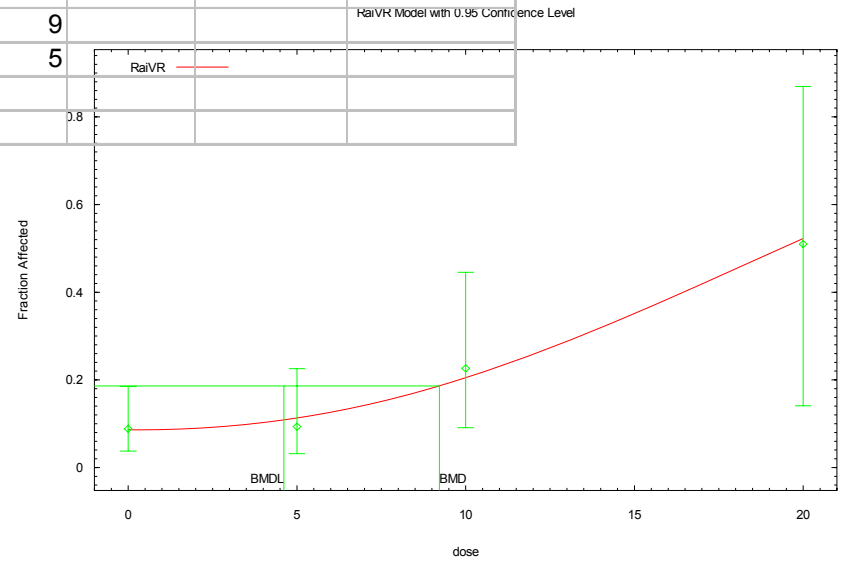
n_k units examined dose levels $0=d_0< d_1< \ldots < d_K$, k=0,1,.... K

- The NOAEL tests the dose groups against the control.
- The dose-response modeling fits a curve and estimates at low doses

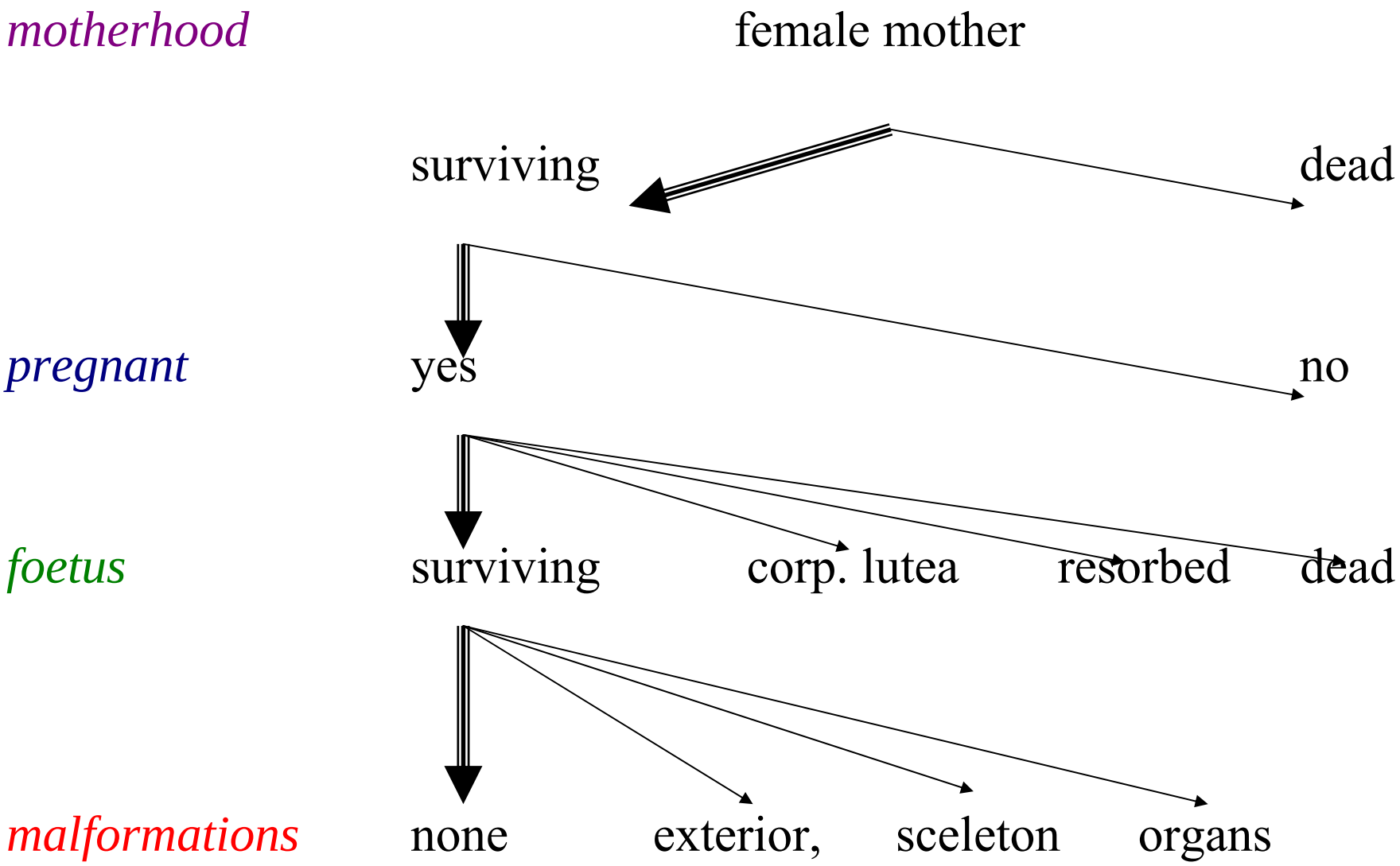
Control			5 mg/kg			10 mg/kg			20 mg/kg	
Resorptions	Implantations		Resorptions	Implantations		Resorptions	Implantations		Resorptions	Implantations
1	8		3	7		0	9		0	7
0	7		0	10		1	8		5	6
3	6		1	8		0	1		0	8
0	10		0	2		1	10		8	8
0	1		0	6		7	9		10	10
0	9		0	10		0	8		0	9
2	9		1	3		0	4		3	3
1	7		0	8		2	8			
2	10		0	7		4	6			
0	7		1	9		3	6			
1	8		1	5		0	1			
0	4					1	9			
0	10					0	5			
0	10									
0	7									

not using litter specific covariates,
but estimating intralitter correlations

Model	*BMD*	*BMDL*
NLogistic	9.2	4.7
NCTR	9.2	4.6
RaiVR	9.2	4.6 (identical to NCTR)



The evolvement of the experimental unit reproductive data



CONTENT

- Introduction to risk analysis and the steps of human health risk assessment
- Model based risk assessment for food and diet
  - Quantitative Methods and Modeling
  - Benchmark Dose Approach
- **Stochastic Carcinogenesis Models for Testing Biological Hypotheses**

- Discussion

Visualization of Parametric Carcinogenesis Models – Interaction with biologists

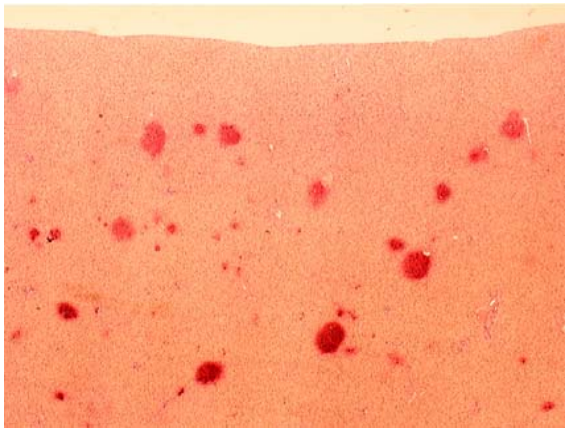
Stochastic carcinogenesis models is a class of mathematically complex models developed for hypotheses generation and falsification in cancer research. They have the potential to base the risk assessment on biological mechanisms.

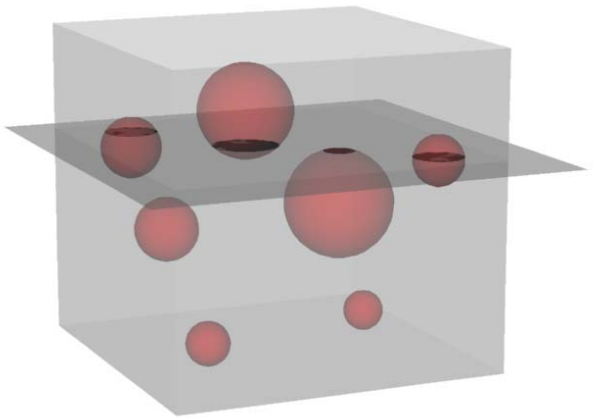
- analytical limited tractability and use of numerical methods

- visualization of analytical and computational components and issues

Jutta Groos, Iris Burkholder and Annette Kopp-Schneider

(Dept. Biostatistics, German Cancer Research Center, Heidelberg)





Animal Carcinogenesis -Experiment

(Weber and Bannasch 1994)

Liver Section

11 weeks

15 weeks

20 weeks

27 weeks

37 weeks

6 Groups of 5 Rats

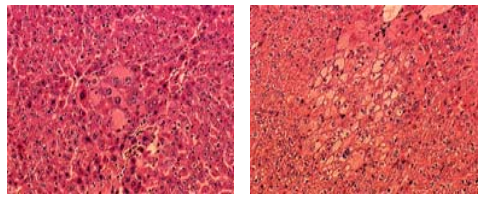
5 5 5 5 5 5

NNM: Continuous,

6 mg/kg Body Weight

Marker H&E

2Types of Foci



Data :

- Area of Liver Section
- Type and Area of every Focus

Data

1,0,82.12248,51.75546,.3930646,23.10435
2,2,5.051971E-02,.87343,.8321747,.32771
3,3,5.514535E-02,.8382204,.9862847,.2743699
4,3,2.695747E-02,.592346,.9654685,.2024913
5,3,4.596029E-02,.8893687,.7301784,.3339285
6,2,7.091237E-03,.3317455,.8096957,.1149725
7,3,3.641833E-02,.7545831,.8037404,.3033004
8,2,.0163475,.5372459,.7117294,.1865115
9,3,2.224992E-02,.6544041,.6528994,.246189
10,3,9.72255E-03,.3561169,.9633963,.1192311
11,2,1.010202E-02,.4423806,.6486731,.1795579
12,10,1.099347E-02,.3878037,.9185884,.1281707
13,3,3.285237E-02,.6981657,.8469536,.2553731
14,3,3.315317E-02,.766953,.7082677,.2945103
15,3,1.975396E-02,.6487775,.5897562,.2199199
16,1,1.052525E-02,.4066853,.7996969,.1587129
17,3,.0113858,.4371457,.7487229,.1747762
18,3,1.806738E-02,.5111027,.8691379,.1752452

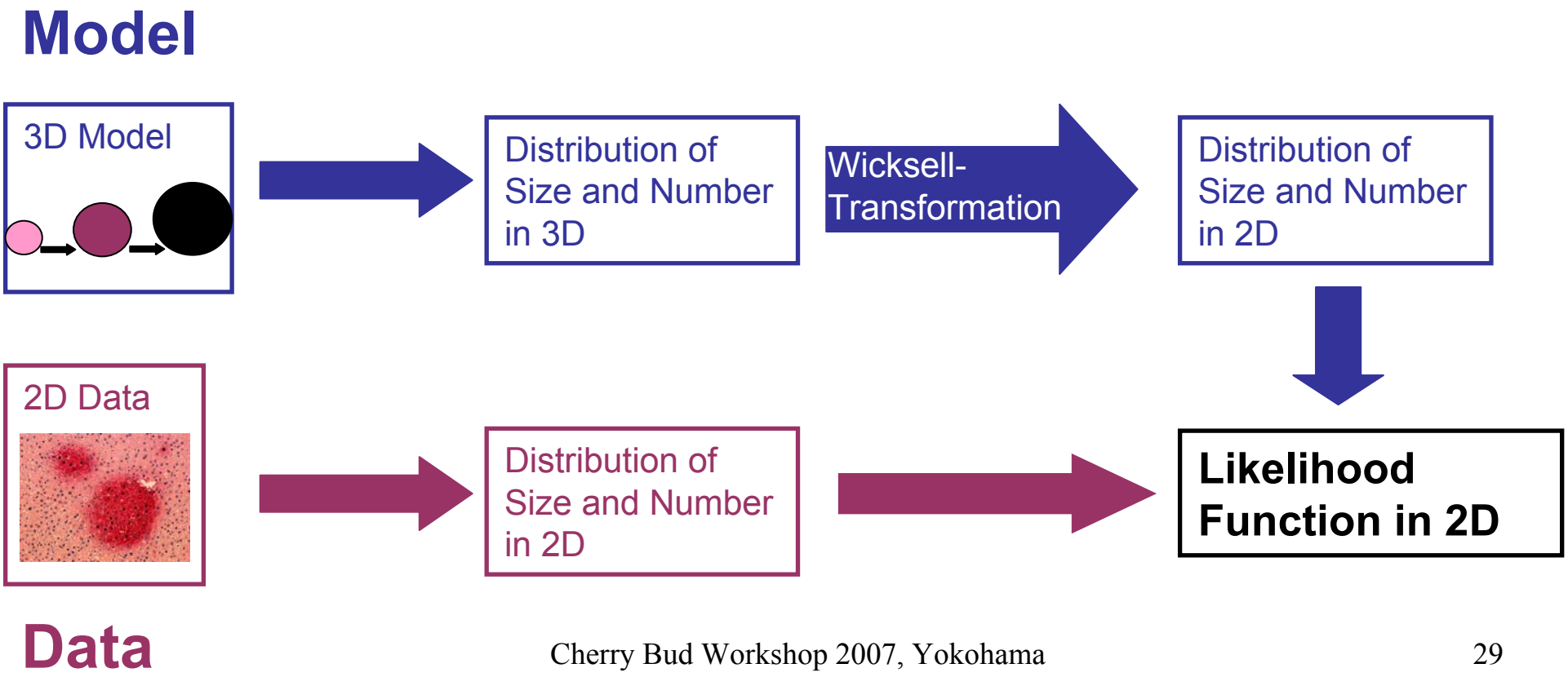
ID, type of focus, area (mm^2), circumvent (mm), shape factor, maximum diameter (mm)

NNM-Data: number of foci (6mg/kgBW/d NNM)

Time	number of animals	Total area (in mm^2)	# Typ 1 foci	# Typ 2 foci	# Typ 3 foci
11	5	490	85	1	-
15	5	367	117	3	1
20	5	406	234	18	-
27	5	543	317	151	3
37	5	455	296	354	13

BUT...

Computing the Likelihood Function:



Wicksell-Transformation

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

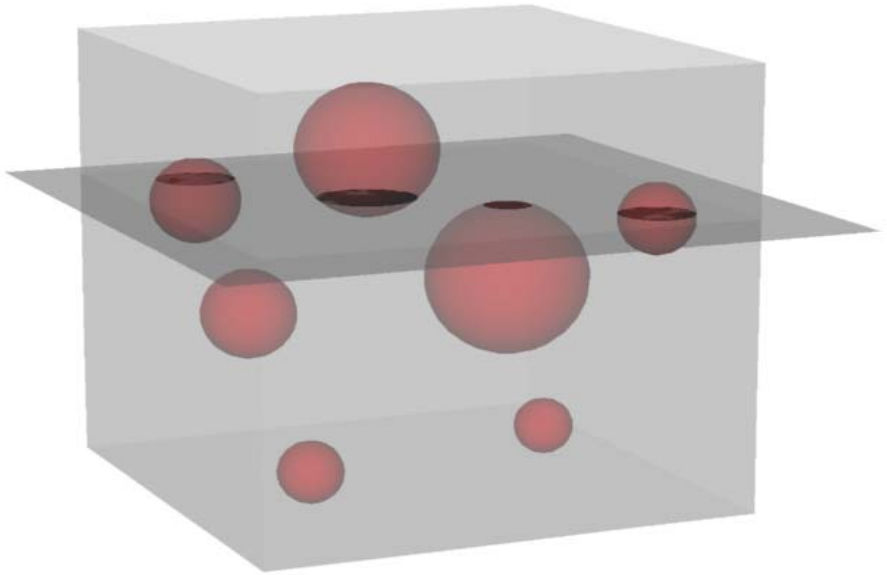
$$G_2^{\varepsilon}(y) = 1 - \frac{1}{\mu_{\varepsilon}} \int_y^{\infty} \sqrt{r^2 - y^2}\, g_3(r)\, dr$$ Size Distribution in 2D

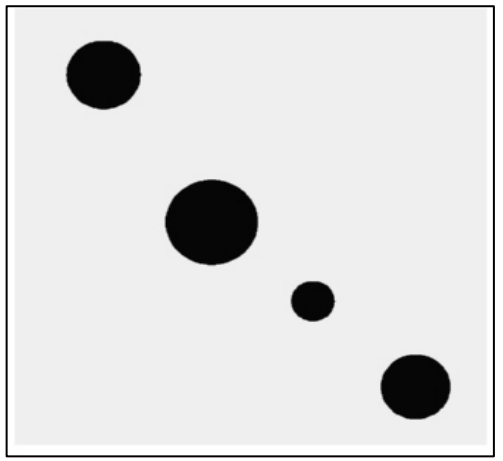
$$g_2^{\varepsilon}(y) = \frac{y}{\mu_{\varepsilon}} \int_y^{\infty} \frac{1}{\sqrt{r^2 - y^2}}\, g_3(r)\, dr$$ Density in 2D

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

$g_3(r)$ Density in 3D

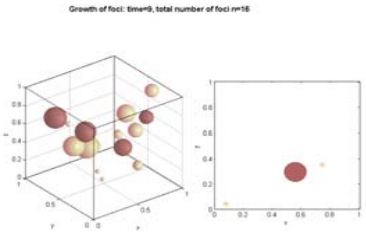
The stereological problem



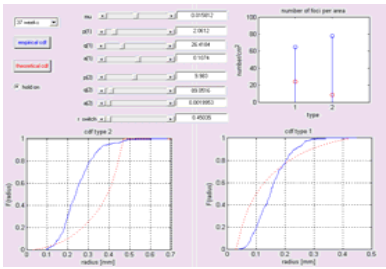


- Probability(focus is cut) is proportional to the size(focus)
- #focal transections is proportional to the #foci and size(foci)

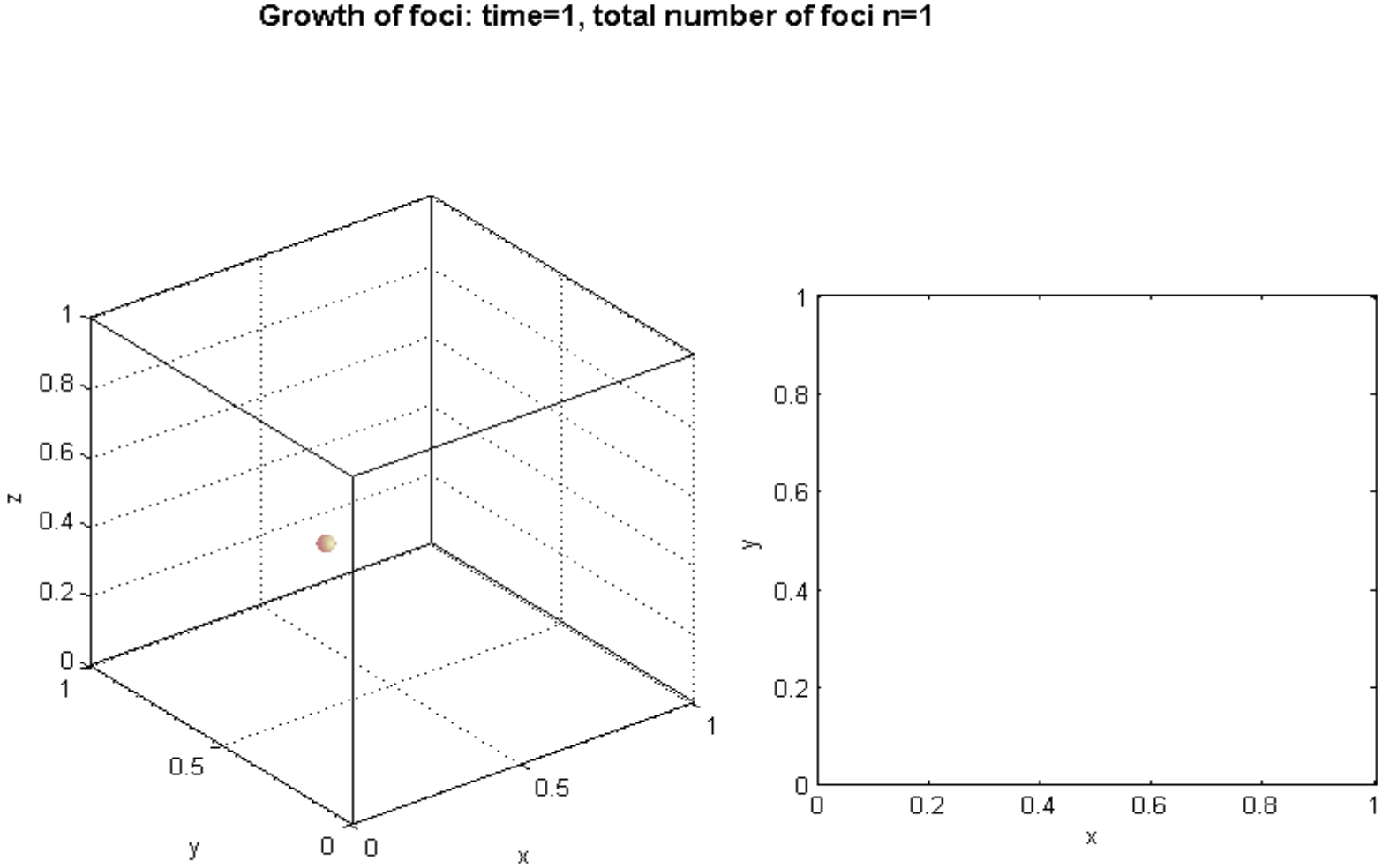
Visualization to:

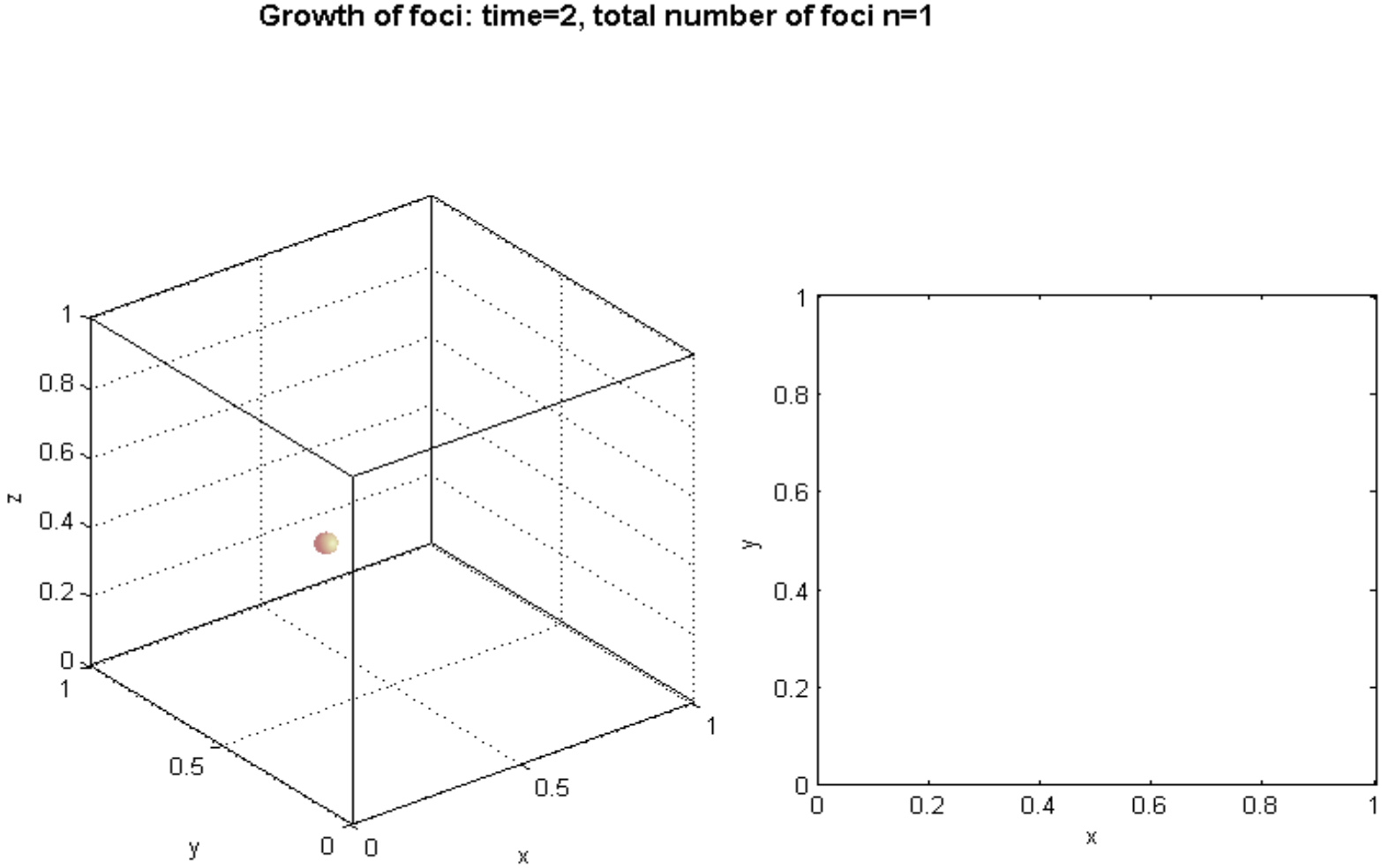


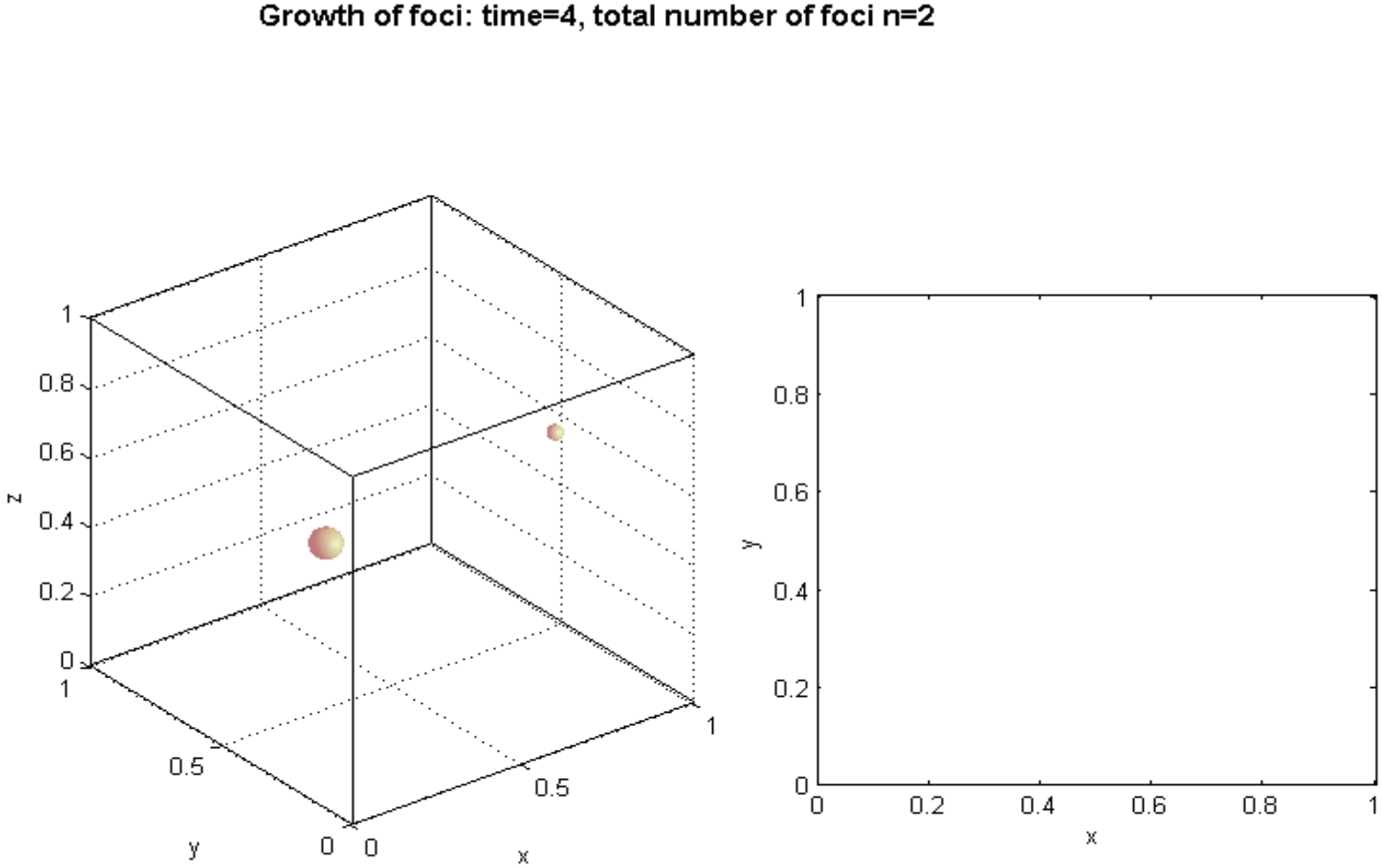
demonstrate general aspects of modeling

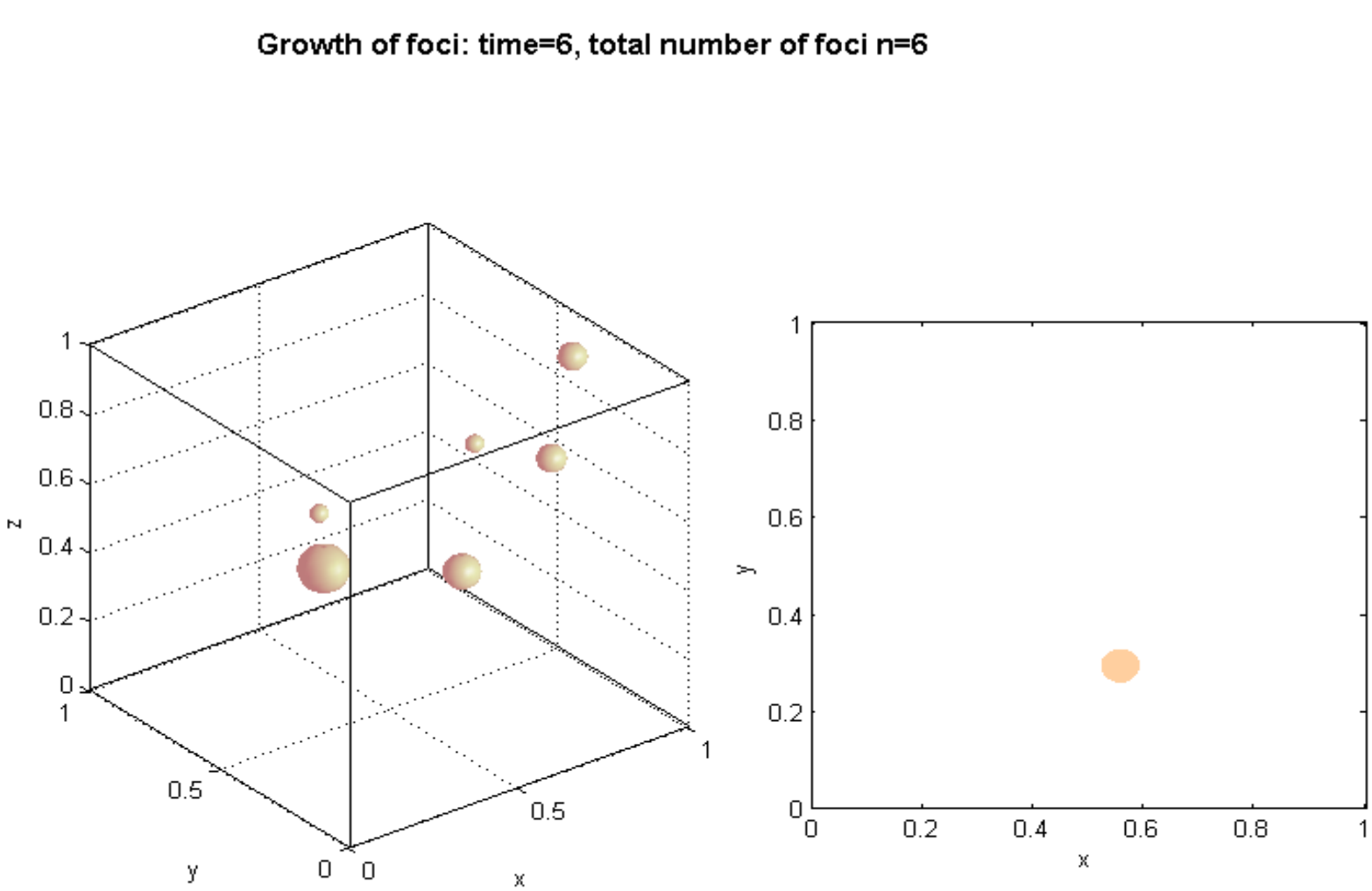


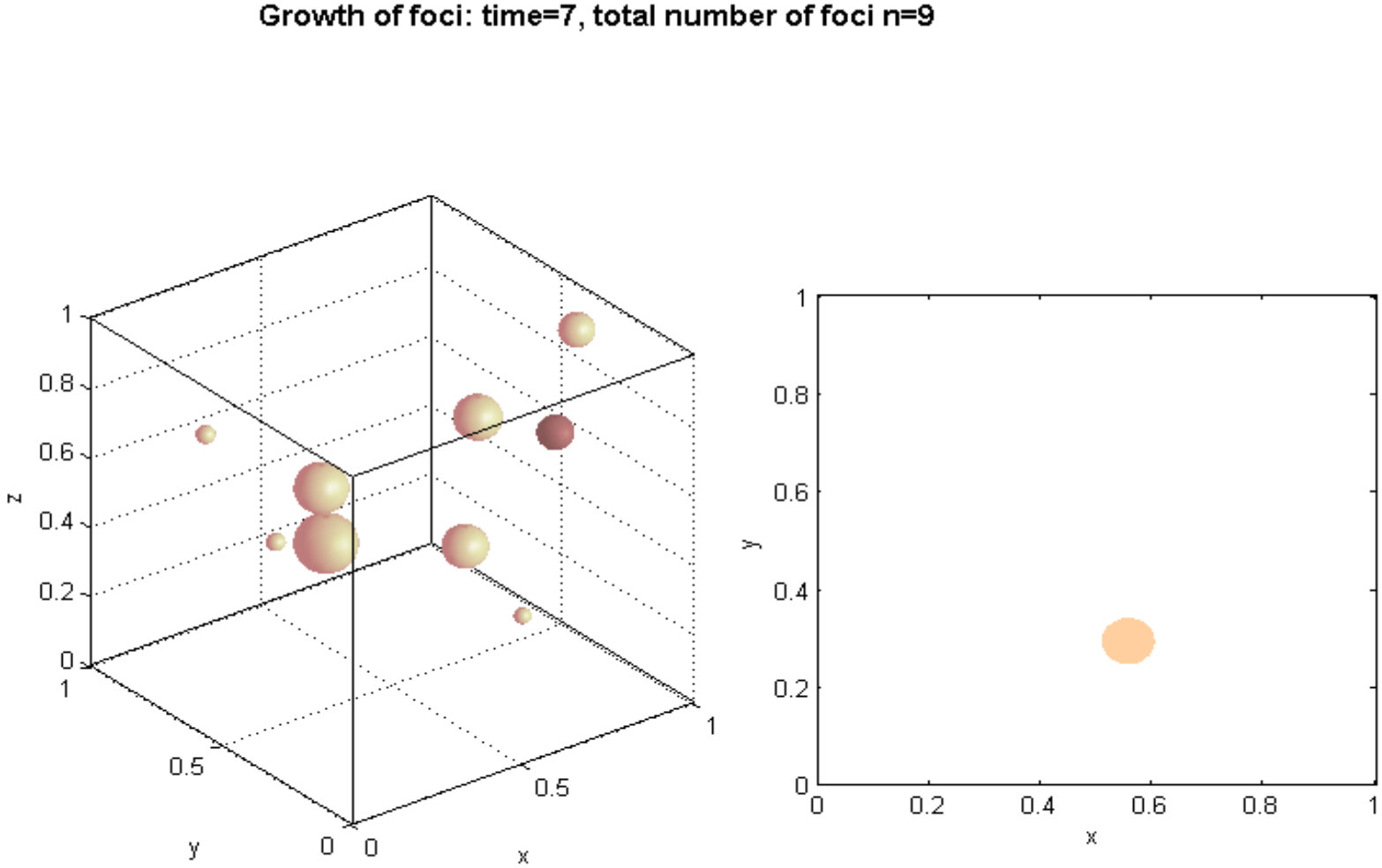
observe the influence of model parameters.

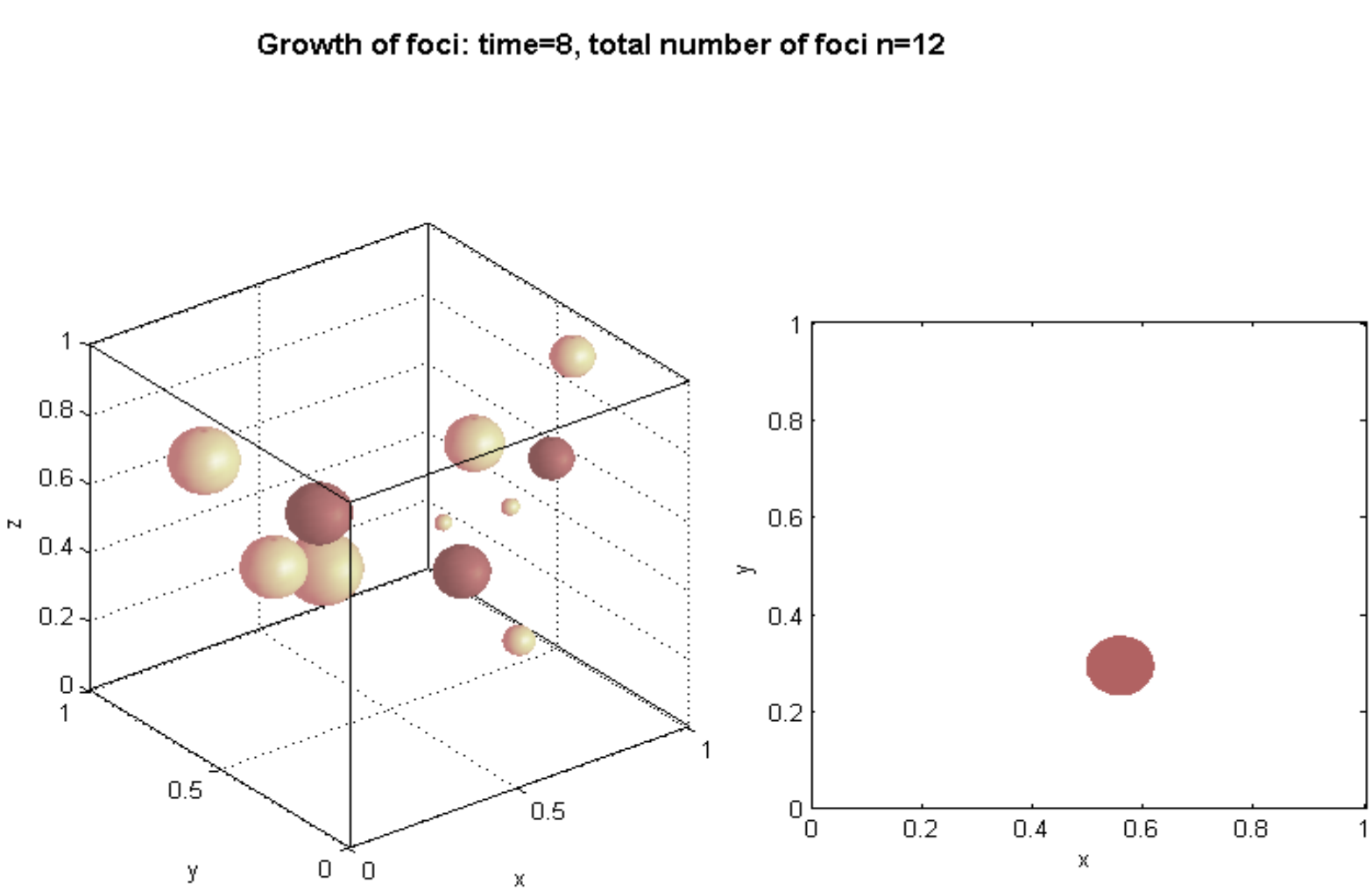

Growth of foci: time=1, total number of foci n=1
z
y
x
y
0
0.2
0.4
0.6
0.8
1
0.5

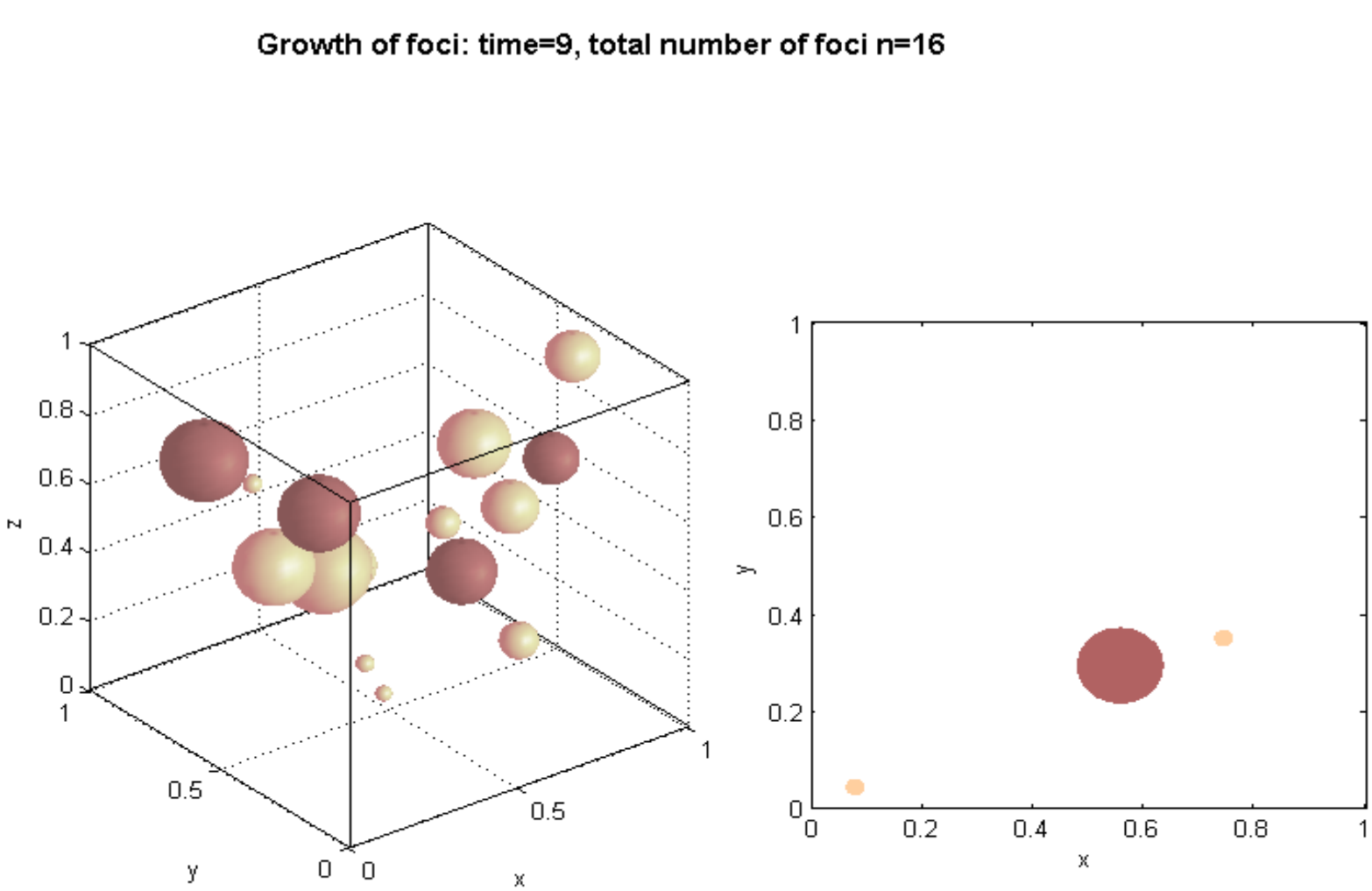

Growth of foci: time=2, total number of foci n=1
z
y
x
y
x


Growth of foci: time=4, total number of foci n=2
z
1
0.8
0.6
0.4
0.2
0
1
0.5
y
0
0
0.5
1
x
y
1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1
x

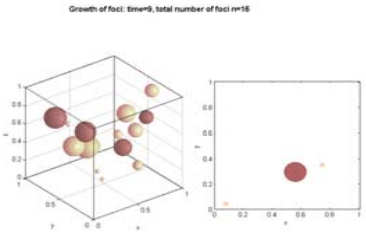

Growth of foci: time=6, total number of foci n=6
z
y
x
y
x


Growth of foci: time=7, total number of foci n=9
z
y
x
0
0.2
0.4
0.6
0.8
1
0.5

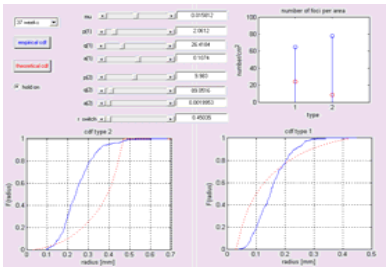

Growth of foci: time=8, total number of foci n=12
z
y
x
y
x


Growth of foci: time=9, total number of foci n=16
z
y
x
0
0.2
0.4
0.6
0.8
1
0.5

Visualization to:

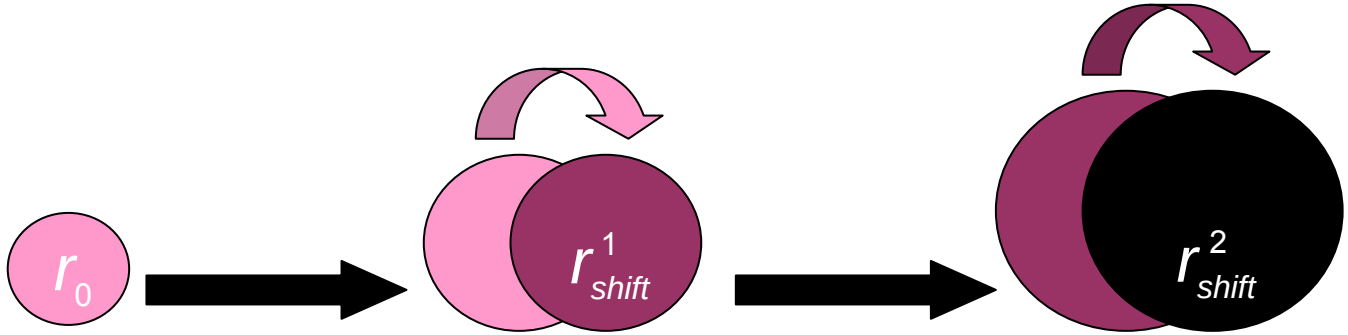


demonstrate general aspects of modeling



observe the influence of model parameters.

ColorShiftModel (CSM) with Stochastic Growth and Deterministic Shift-Radius



Stochastic exponential growth with type-dependent growth rates:

$$r(t,c) = r^c_{shift} \exp\left(B_c\,(t - \tau_c)\right)$$

radius of focus of type c at time t

B_c positive random variable.

size distribution: 3D-Density

Formation of foci:
Poisson-process parameter μ.

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

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

Color 1

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

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 :

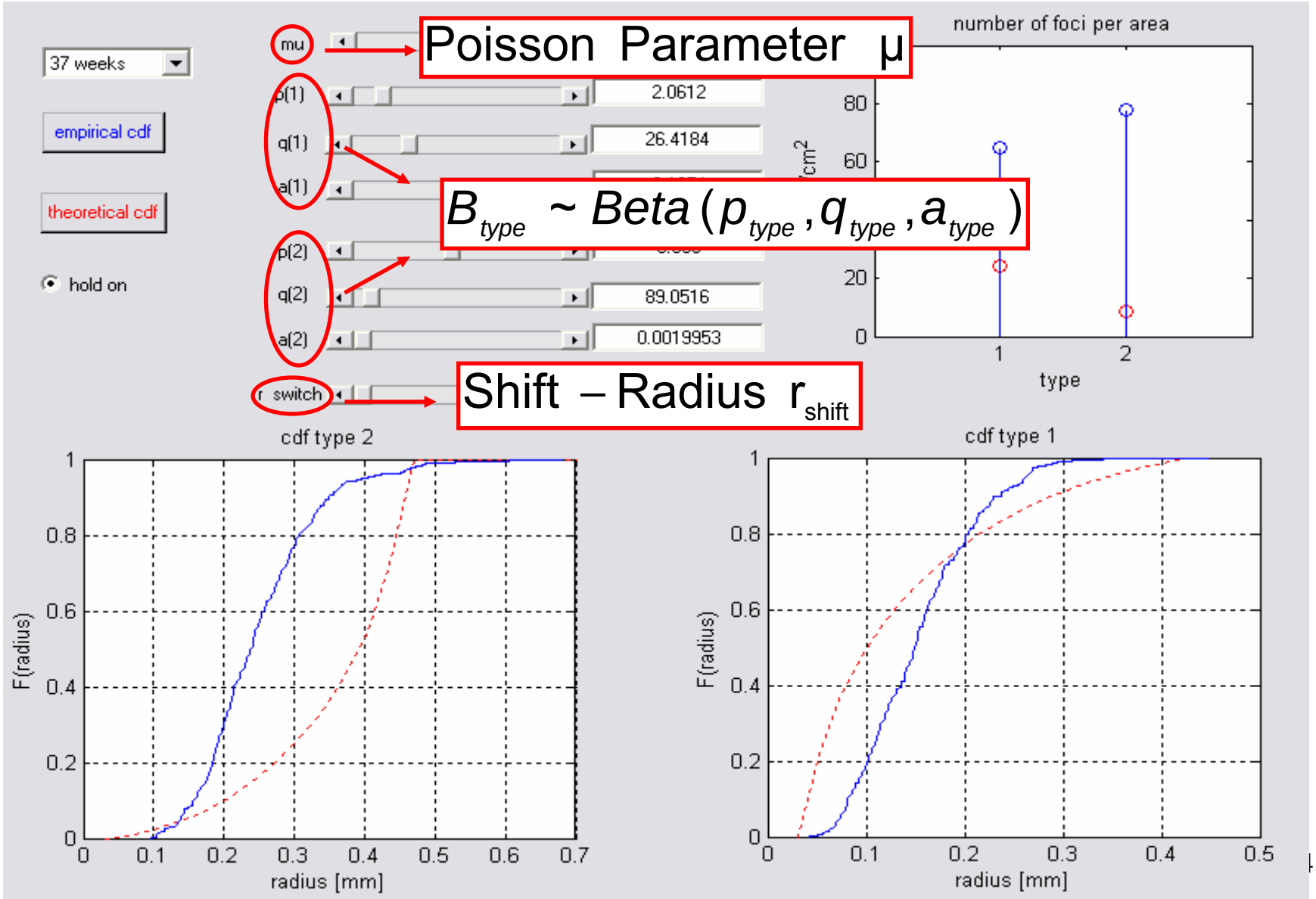
1. Develop analytical expression for distribution of size and number of foci in 2D and derive log-likelihood-function .

2. Find starting parameters for optimization.

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

4. Compare results with data.

MATLAB GUI


37 weeks
empirical cdf
theoretical cdf
hold on
mu
p(1)
q(1)
a(1)
2.0612
26.4184
p(2)
q(2)
a(2)
89.0516
0.0019953
r switch
Poisson Parameter μ
$B_{type} \sim Beta(p_{type}, q_{type}, a_{type})$
Shift – Radius r_{shift}
number of foci per area
/cm2
type
cdf type 2
F(radius)
radius [mm]
cdf type 1
F(radius)
radius [mm]

CONTENT

- Introduction to risk analysis and the steps of human health risk assessment
- Model based risk assessment
  - Quantitative Methods and Modeling for food risks
  - Benchmark Dose Approach
- Stochastic Carcinogenesis Models for Testing Biological Hypotheses
- **Discussion**

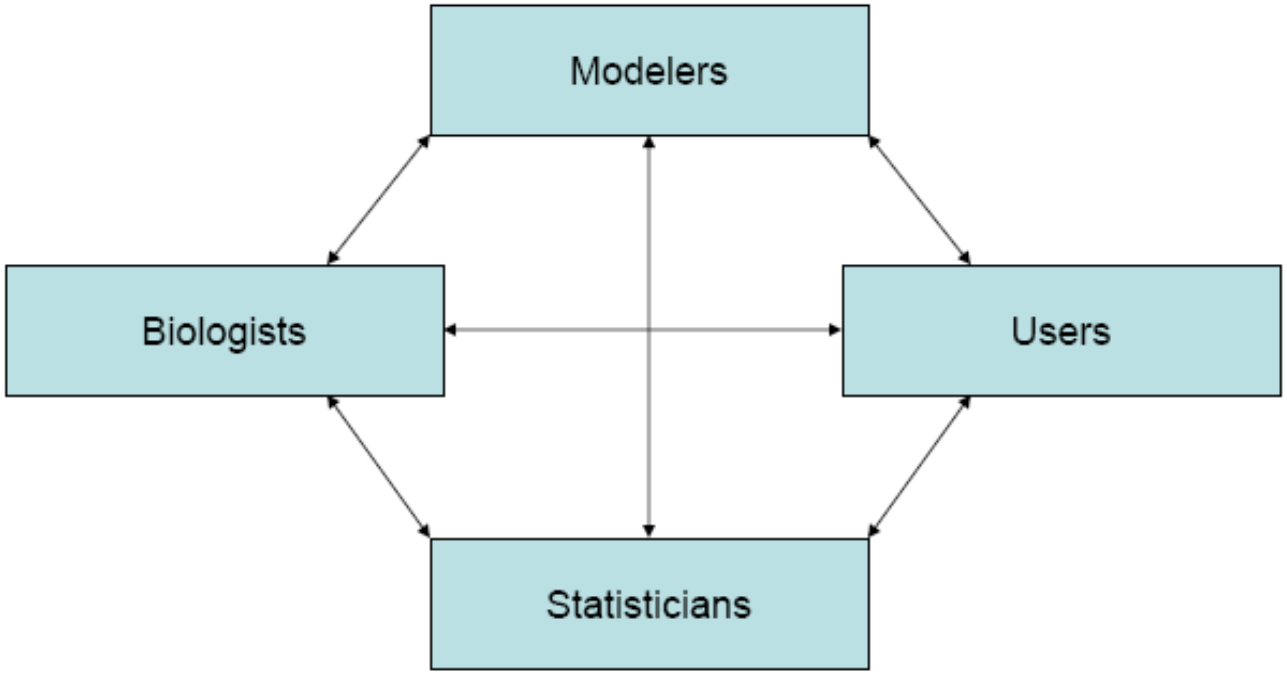
Statistical methods in the risk assessment and risk management process

- **are critical for science based decisions because of the large amount of uncertainty at various steps of risk analysis and since conclusions are made on populations**

- **have somehow lost of its impact and influence on risk assessed and management**
  - failure in the low dose extrapolation
  - faster development in the applied sciences than statistical modeling could keep pace, with generation of a large amount of amount of quantitative and qualitative data
  - statistical analysis methods published in non-statistical are often behind statistical methods but used in applications
  - development of evaluation software by agencies and companies aside of the statistical computing community e.g., GLOBAL, Crystal Ball, @RISK etc

- **need to get closer to the risk managers and interact also with risk communicators besides improved modeling**
  - review of methods and their applicability and use
  - more use of data and data bases
  - interaction with all players

The Interaction Network in Risk Assessment



Biologists: Biologists, Toxicologists, Physiologists

Modelers: Modelers & Mathematicians

Users: Risk Assessors & Decision Makers

Benchmark Dose Approach is an example for this interaction:

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

Which model class should be chosen?

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

Model constraints

- parameters
- shape of curve (non-monotone?)

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

Study Design

- dose groups
- numbers per dose group

Thank you!

Discussion slides

Problems with threshold models

- **we observe harmful effects at high doses**
- **we observe ubiquitous presence of the substance in the human environment at very low doses**

Do threshold doses exist below which no effects happen?

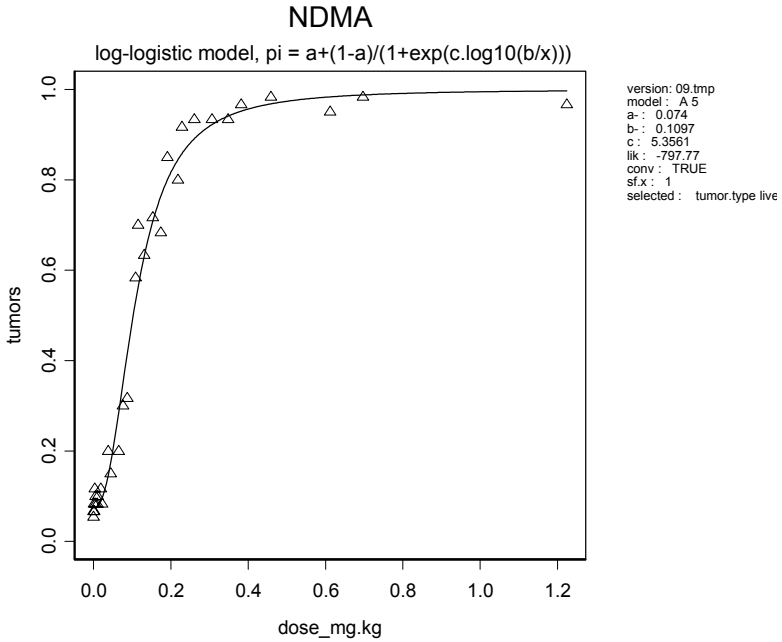
No Consensus ! refer to W. Slob Int. J Tox (1999) 259-68

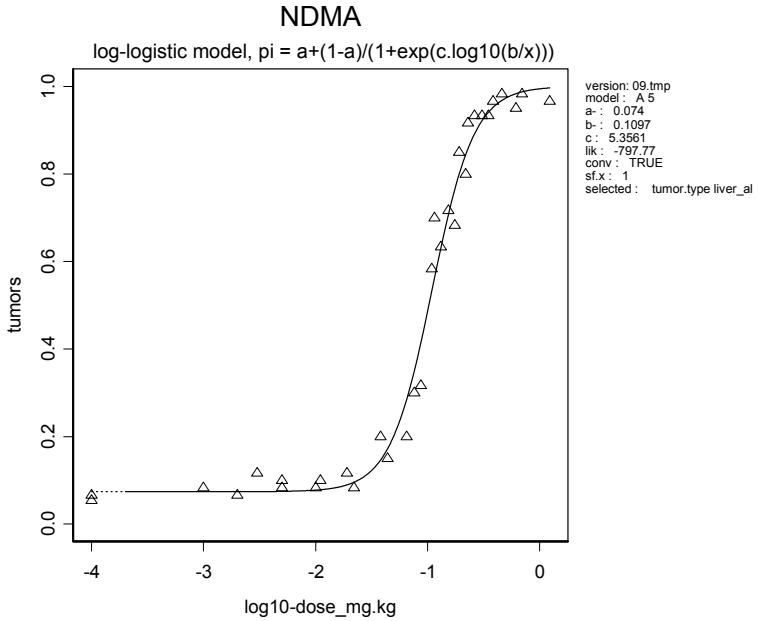
Biological
The dose below which the organism does not suffer from an adverse effect

Experimental
The dose below which no effects are observed (NOAEL)

Mathematical
The dose below which the true response is zero and above which it is non-zero

This is the same fitted curve only plotted at different scales





Plots from W. Slob made with the RIVM software PROAST

BMD: Response functions model selection

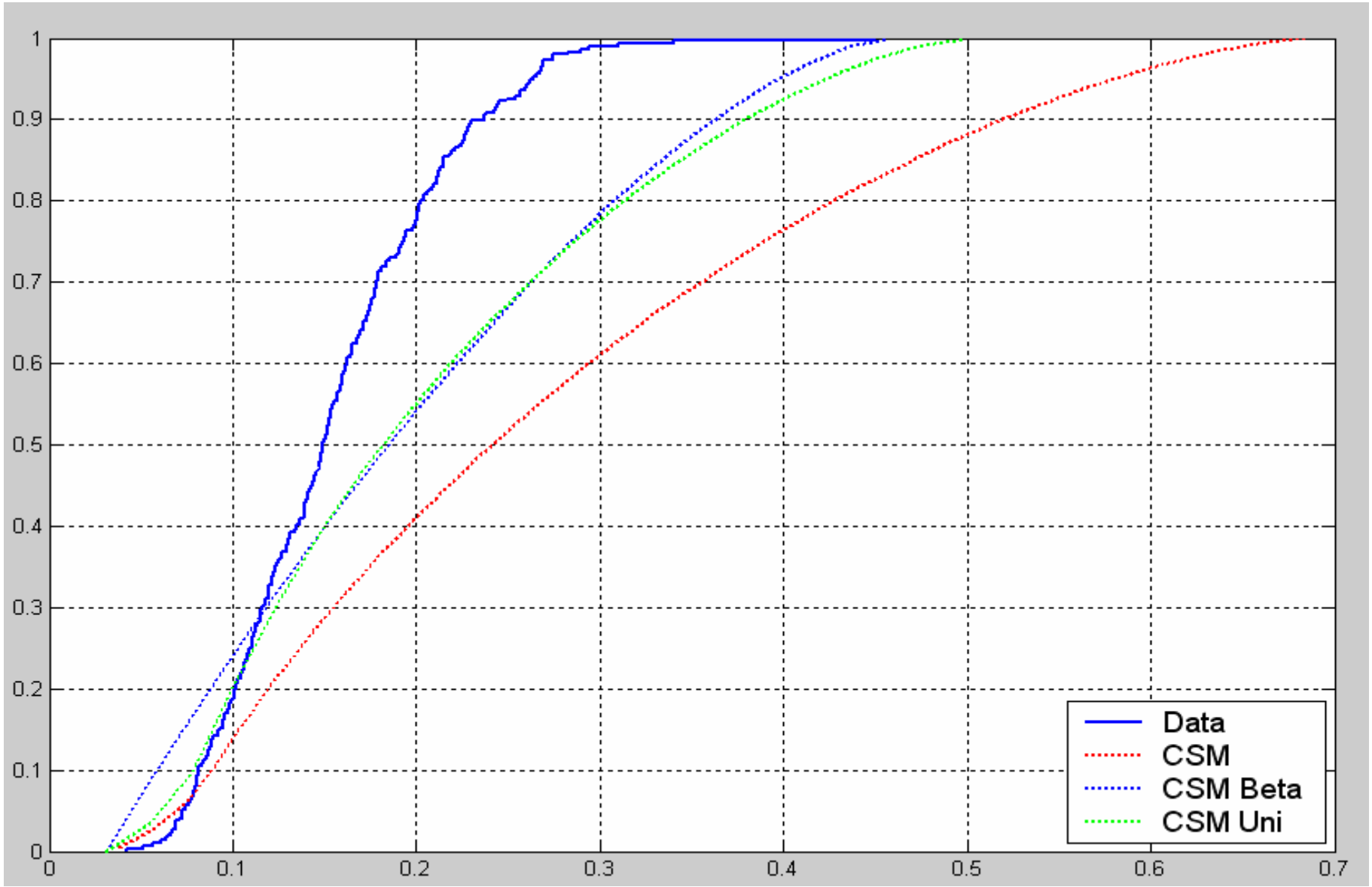
Slob (2005) proposed a nested family of multiplicative nonlinear models for quantitative response

$y = a$	with $a > 0$
$y = a \exp(x/b)$	with $a > 0$
$y = a \exp((x/b)d)$ or $a \exp(-(x/b)d)$	with $a > 0, b > 0, d \neq 1$
$y = a [c - (c - 1) \exp(-x/b)]$	with $a > 0, b > 0, c > 0$
$y = a [c - (c - 1) \exp(-(x/b)d)]$	with $a > 0, b > 0, c > 0, d \geq 1$

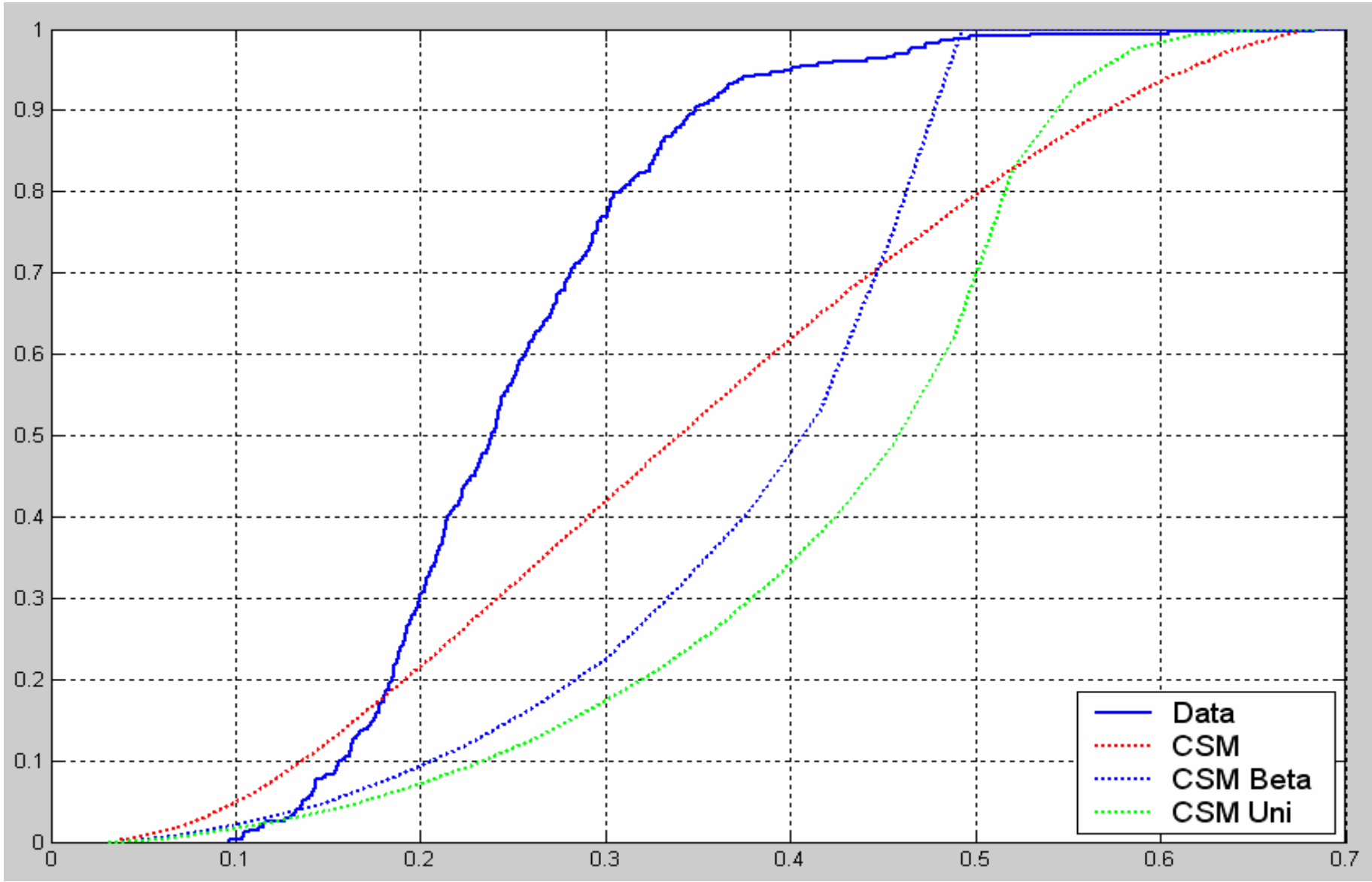
and performed three simulation studies each with a large number of designs.

- the performance of a design is determined by the total sample size
- distribution of the animals over the dose groups does not lead to poorer performance, except when the variation of the endpoint is high (6 dose groups may be used)
- dose placement can affect performance and the risk of misplacement is lower when distributing the dose groups
- high level dose groups help to estimate at low doses, high doses should not be deleted

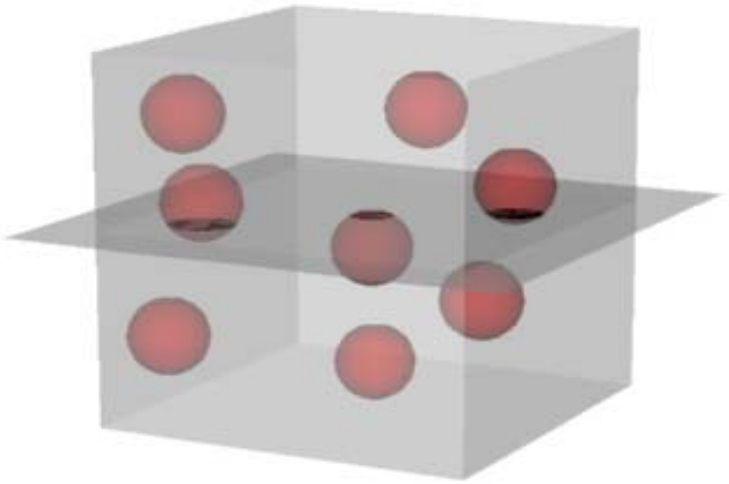
Size Distribution of Type 1 Foci

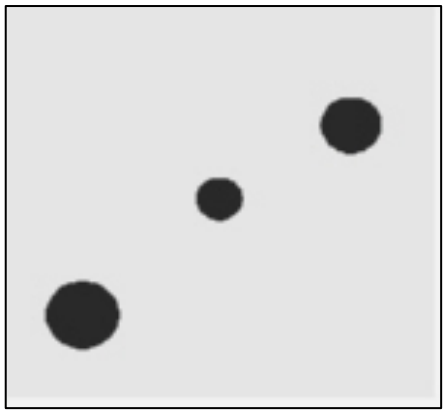


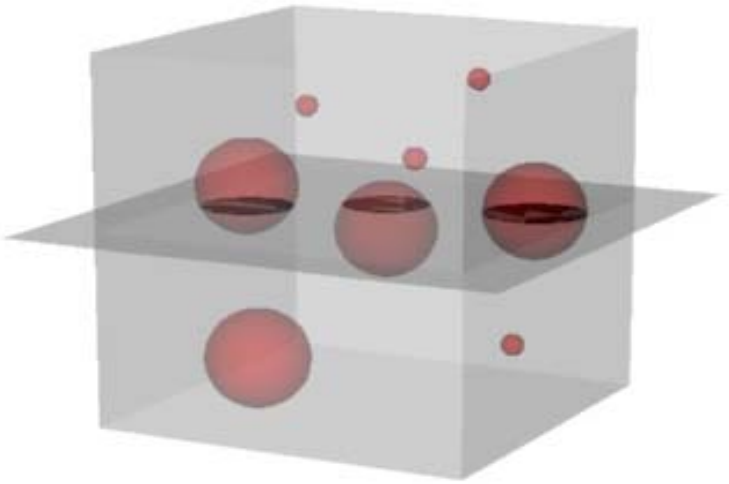
Size Distribution of Type 2 Foci

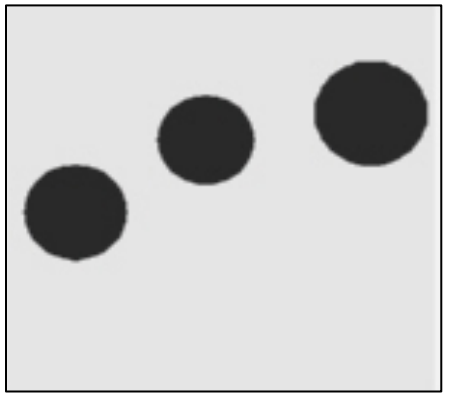


Paradoxa in stereology









Mean sphere diameter = 3 > Mean sphere diameter = 2.5

Mean transection diameter = 2.4 < Mean transection diameter = 2.7