# (Malaria and TB Models) & Sensitivity/Uncertainty Analysis

SISMID 2017

Author: Andreas Handel, Department of Epidemiology and Biostatistics, University of Georgia

ahandel@uga.edu

How to use models

- So far, we ran computer code that produced dynamics (time-series).

- We changed parameter values to see how results changed.

- We often want to explore how results (as defined by us) vary with some inputs.

- Often those inputs are parameter values, corresponding to different biological conditions.

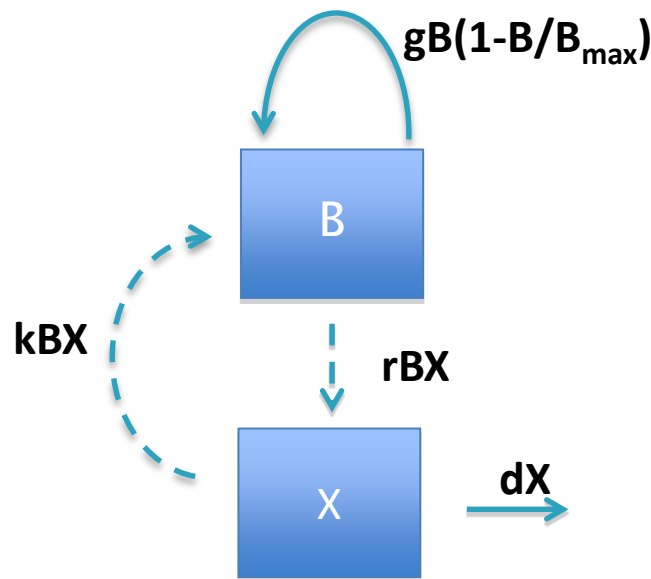- We'll briefly try ourselves, then look at such model uses for Malaria and TB.
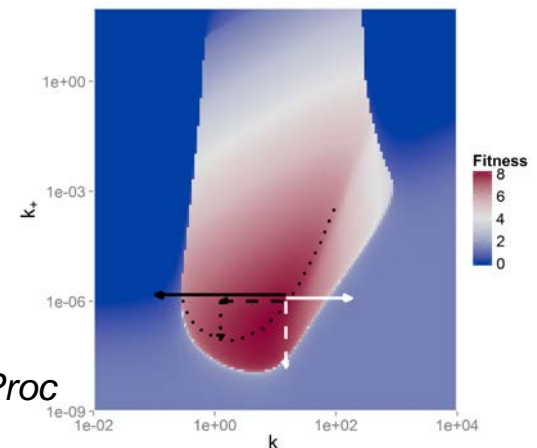
Handel et al (2014) Proc Royal Soc Interface

2

How to model - example

▸ Assume we have a simple model for a bacterial infection

$$\dot{B} = gB\left(1 - \frac{B}{B_{max}}\right) - kBX \quad \text{(bacteria)}$$

$$\dot{X} = rBX - dX \quad \text{(immune response)}$$

gB(1-B/B$_{max}$)

B

kBX

rBX

X

dX

(Earlier notation)

$$\dot{B} = gB(1 - \frac{B}{B_{max}}) - dB - kBI$$

$$\dot{I} = rBI - \delta I$$

How to model - example

▸ Let's assume our question is: How does the peak of bacteria load change as the rate of the immune response activation is changed?

▸ How do we answer this question?

$$\dot{B} = gB\left(1 - \frac{B}{B_{max}}\right) - kBX \quad \text{(bacteria)}$$

$$\dot{X} = rBX - dX \quad\quad\quad \text{(immune response)}$$

How to model - example

▸ Vary rate of the immune response activation (i.e. parameter **r** in our model).

▸ For each value of **r**, run model, record bacteria peak load B_{peak}.

▸ Plot B_{peak} as function of **r**.

$$\dot{B} = gB\left(1 - \frac{B}{B_{max}}\right) - kBX \quad \text{(bacteria)}$$

$$\dot{X} = rBX - dX \quad \text{(immune response)}$$

# How to model – R example

▸ Open SISMID-U6-model1.r

▸ Read the comments in the code to make sure you understand what goes on.

▸ Run the code, take look at the result. Make sure you understand it.

▸ Change the value of some other parameter, run again.

Malaria and TB models

- Most of the existing models for malaria and TB are rather large and complex.

- We don't have enough time to go through them in a lot of detail – it would take the rest of this module.

- I will show a few results illustrating how one can use variation of parameters to study a model/system.

- I will also give you a **very** brief glimpse of some of these models to illustrate an issue we'll then address in detail in this session.

Malaria 1
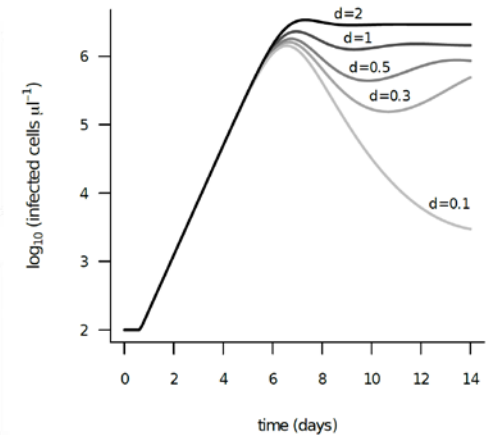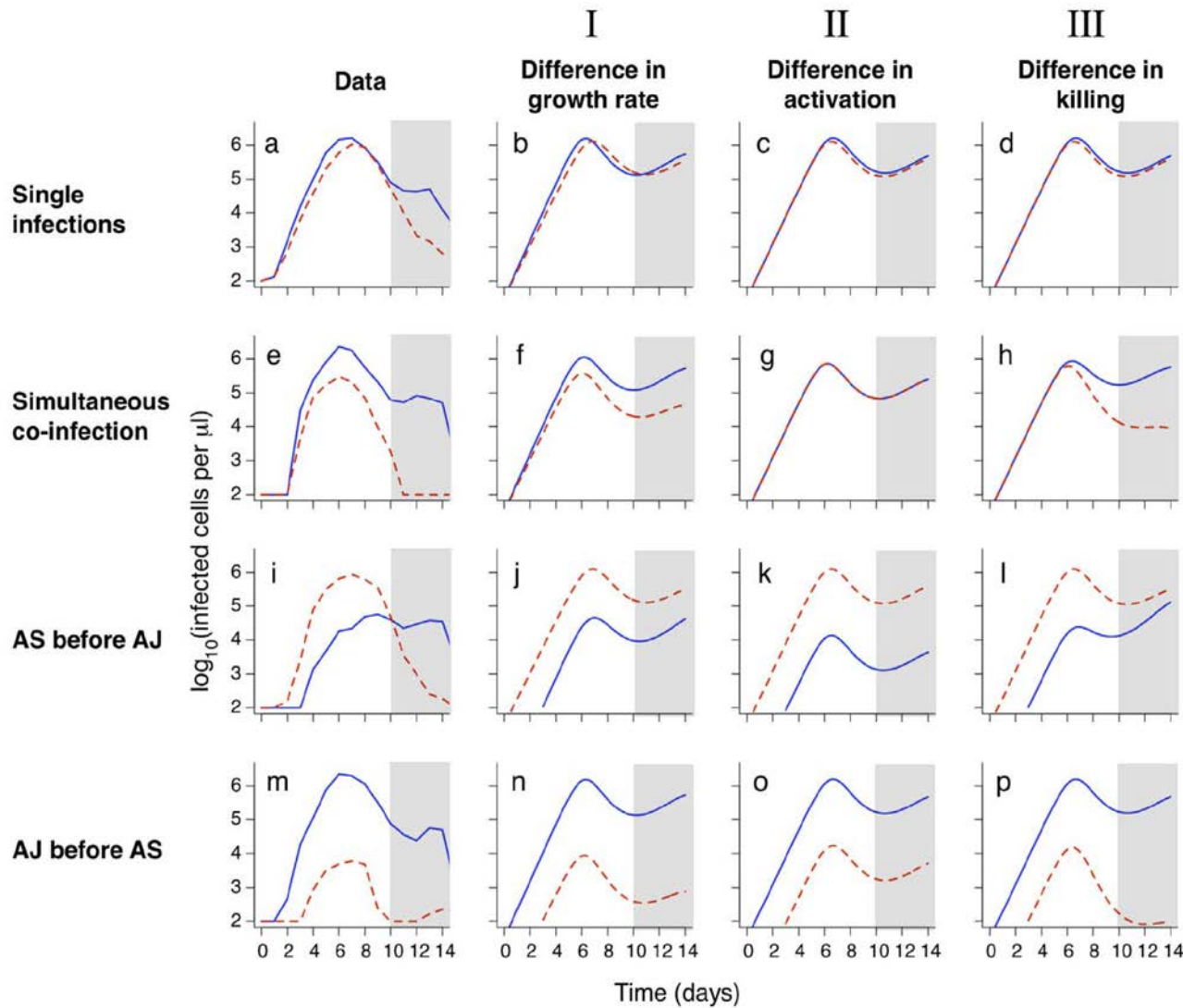
Kochin et al (2010) PLoS One

Malaria 2
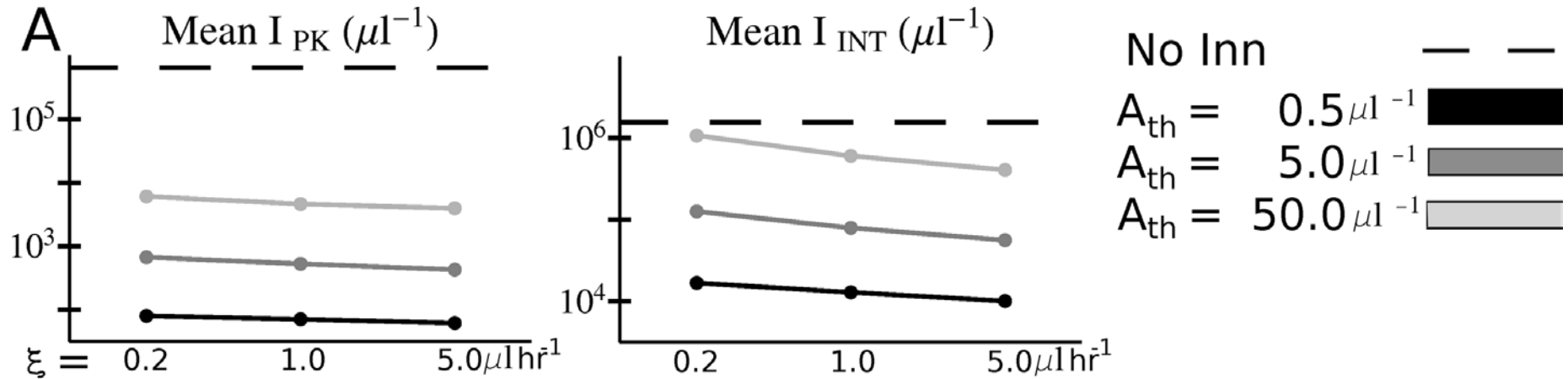
I_{PK} = peak parasite load
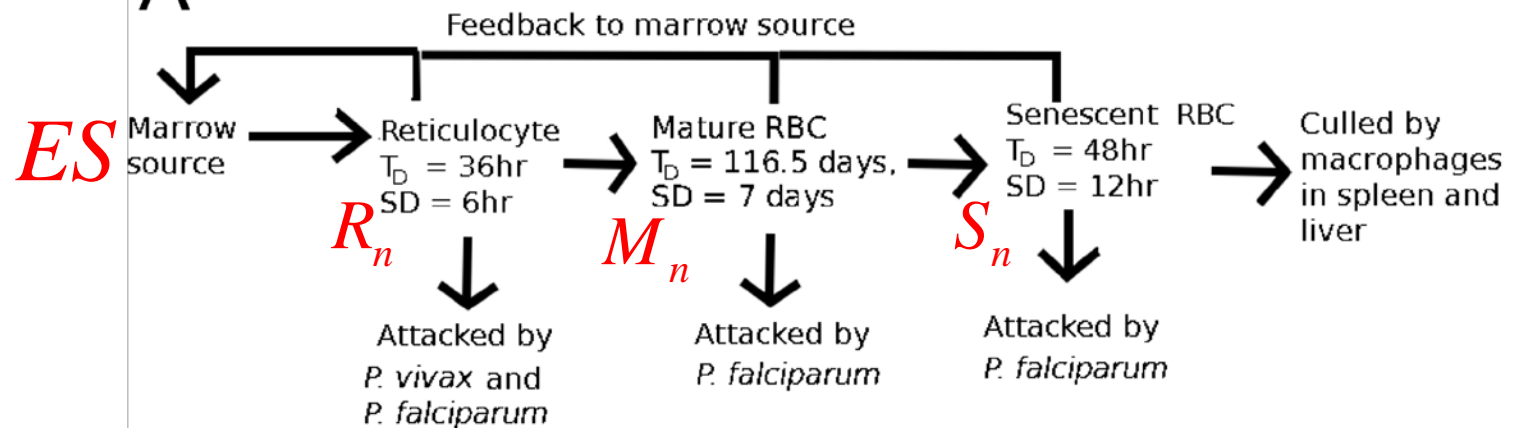I_{INT} = total parasite load
ζ = binding affinity (infection rate)
A_{th} = antigen threshold at which IR gets activated

McQueen & McKenzie (2008) PLoS Comp Bio

Malaria 2 – under the hood

A

Feedback to marrow source

ES Marrow source → Reticulocyte $T_D = 36$hr, SD = 6hr → Mature RBC $T_D = 116.5$ days, SD = 7 days → Senescent RBC $T_D = 48$hr, SD = 12hr → Culled by macrophages in spleen and liver

R_n M_n S_n

Attacked by P. vivax and P. falciparum

Attacked by P. falciparum

Attacked by P. falciparum

RBC dynamics

μ= merozoites
E_T=total RBC
T_D=average length of stage,
SD=standard deviation of stage
ζ=binding affinity

$$dR_1/dt = \mathrm{ES}(t) - \kappa_R R_1 - \zeta\mu R_1$$

$$dR_n/dt = \kappa_R(R_{n-1} - R_n) - \zeta\mu R_n,\ 1 < n < N_{cR}+1$$

For mature RBCs:

$$d\mathrm{ES}/dt = \lambda_{\mathrm{ES}}(\Phi - \mathrm{ES}(t)),\ \mathrm{ES}_{MN} < \Phi < \mathrm{ES}_{MX}$$
$$= \lambda_{\mathrm{ES}}(\mathrm{ES}_{MX} - \mathrm{ES}(t)),\ \Phi > \mathrm{ES}_{MX}$$
$$= \lambda_{\mathrm{ES}}(\mathrm{ES}_{MN} - \mathrm{ES}(t)),\ \Phi < \mathrm{ES}_{MN}$$

$$\Phi = ES_0 - \frac{dE_T}{dt} - \delta\xi\mu V$$

$$dM_1/dt = \kappa_R\ R_{NcR} - \kappa_M M_1 - \zeta\mu M_1$$

$$dM_n/dt = \kappa_M(M_{n-1} - M_n) - \zeta\mu M_n,\ 1 < n < N_{cM}+1$$

For senescent RBCs:

$$dS_1/dt = \kappa_M M_{NcM} - \kappa_S S_1 - \zeta\mu S_1$$

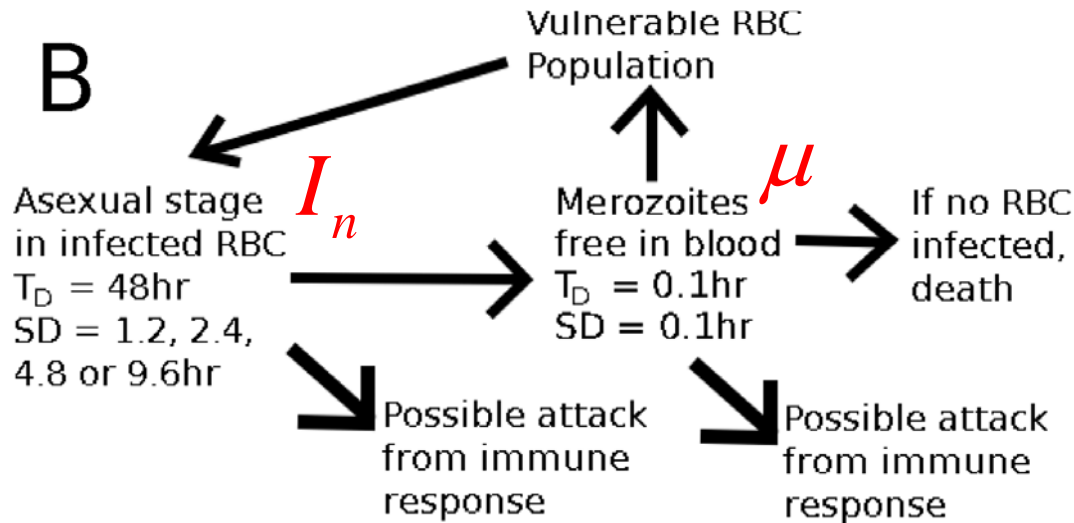$$dS_n/dt = \kappa_S(S_{n-1} - S_n) - \zeta\mu S_n,\ 1 < n < N_{cS}+1$$

Modeling Malaria – Pathogen dynamics

Vulnerable RBC Population

B

I_n

μ

Asexual stage in infected RBC
$T_D = 48hr$
$SD = 1.2, 2.4, 4.8$ or $9.6hr$

Merozoites free in blood
$T_D = 0.1hr$
$SD = 0.1hr$

If no RBC infected, death

Possible attack from immune response
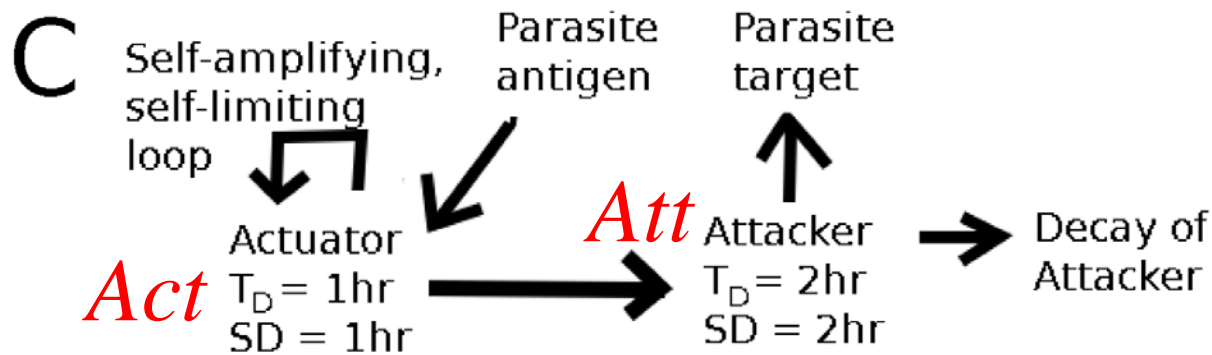
Possible attack from immune response

$$\dot{I}_1 = \xi\mu V - \kappa_I I_1 - I_1 \sum_m \text{Att}_m \xi_{m,1}$$

$$\dot{I}_n = \kappa_I I_{n-1} - \kappa_I I_n - I_n \sum_m \text{Att}_m \xi_{m,n}$$

$$\dot{\mu} = p\kappa_I I_{NcI} - \mu(\xi V - 1/T_{D\mu}) - \mu \sum_m \text{Att}_m \xi_{m,\mu} + L(t)$$

T_D=average length of stage, SD=standard deviation of stage, ζ=binding affinity, V=density of vulnerable RBC, κ=1/(T_D * SD), Att = immune response clearance, L(t) influx from liver, p=number of merozoites produced by an infected RBC

Modeling Malaria – innate IR

C Self-amplifying, self-limiting loop

Parasite antigen

Parasite target

Actuator $T_D = 1hr$ SD = 1hr

Act

Att Attacker $T_D = 2hr$ SD = 2hr

Decay of Attacker

Act=actuator level
Att=attacker level
A=parasite density

$$dAct/dt = \Theta(S_{\mathrm{Act}}) - \lambda_{\mathrm{Act}} Act$$
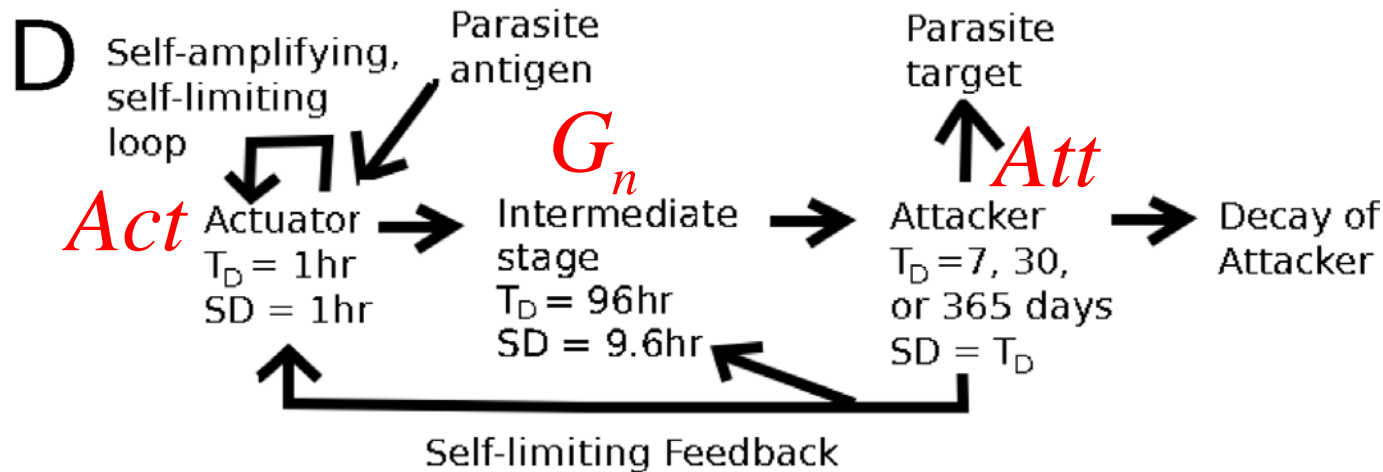
$$dAtt/dt = \Theta(S_{\mathrm{Att}}) - \lambda_{\mathrm{Att}} Att$$

where $\Theta(x) = x$ if $x > 0$, zero otherwise, and

$$S_{\mathrm{Act}} = (S_{\mathrm{Act,B}} + a\, Act\Theta(\Delta A))(1 - Act/Act_{\mathrm{Mx}})$$

$$\Delta A = A - A_{\mathrm{th}}$$

$$S_{\mathrm{Att}} = \lambda_{\mathrm{Act}} Act - S_{\mathrm{Att,th}}$$

Modeling Malaria – adaptive/Ab IR

$$\mathrm{d}Act/\mathrm{d}t = \Theta(S_{\mathrm{Act}}) - \lambda_{\mathrm{Act}}Act$$

$$\mathrm{d}G_1/\mathrm{d}t = \Theta(\lambda_{\mathrm{Act}}Act - S_{G\mathrm{th}}) - \kappa_G G_1$$

$$\mathrm{d}G_n/\mathrm{d}t = \kappa_G(G_{n-1} - G_n),\ 1 < n < N_{cG}+1$$

$$\mathrm{d}Att/\mathrm{d}t = \Theta(S_{\mathrm{G,Att}}) - \lambda_{\mathrm{Att}}Att$$

$$(10)$$

where

$$S_{\mathrm{Act}} = (S_{\mathrm{Act,B}} + aAct\Theta(\Delta A))(1 - Act/Act_{\mathrm{Mx}})(1 - Att/Att_{\mathrm{Mx}})$$

$$\Delta A = A - A_{\mathrm{th}}$$

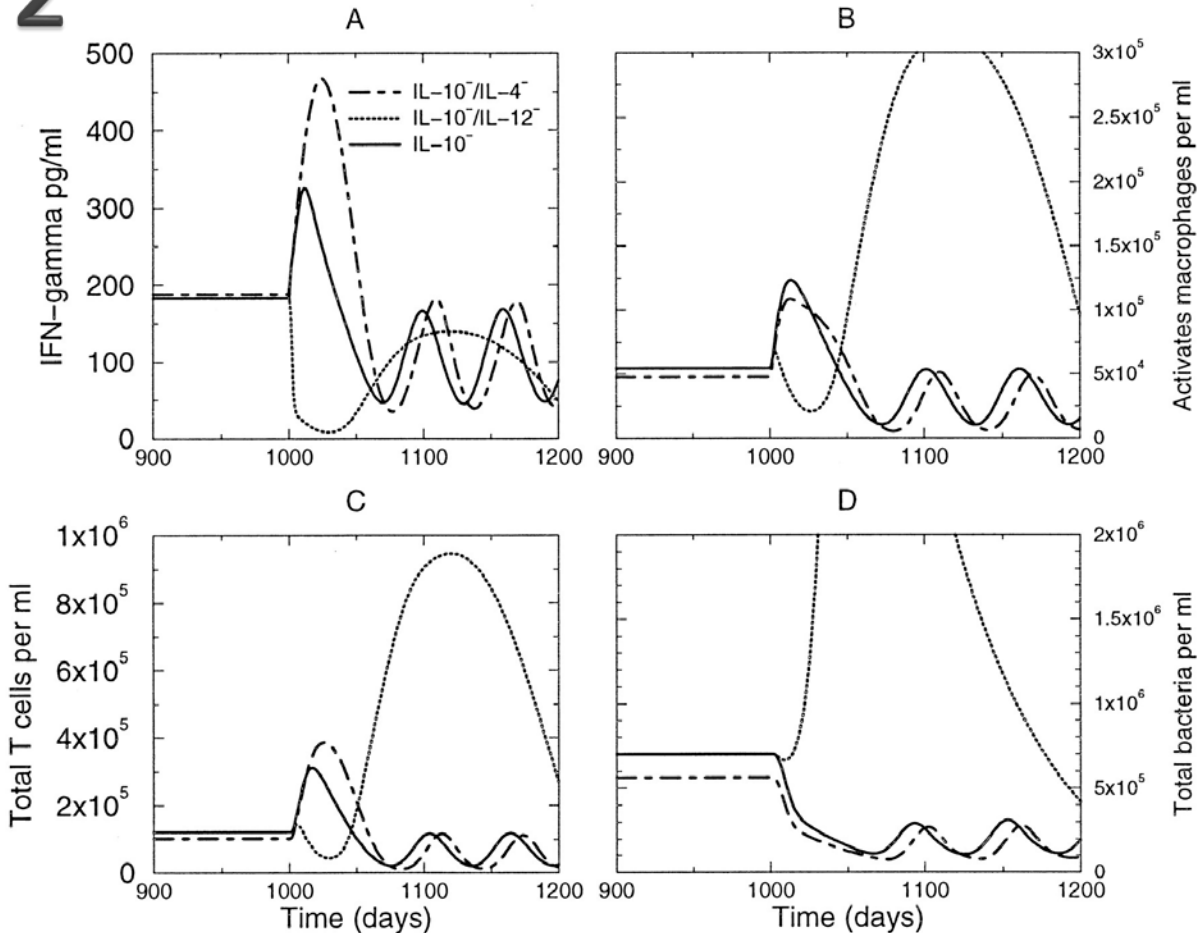$$S_{\mathrm{G,Att}} = \kappa_G G_{NcG}(1 - Att/Att_{\mathrm{Mx}})$$

13

TB 1
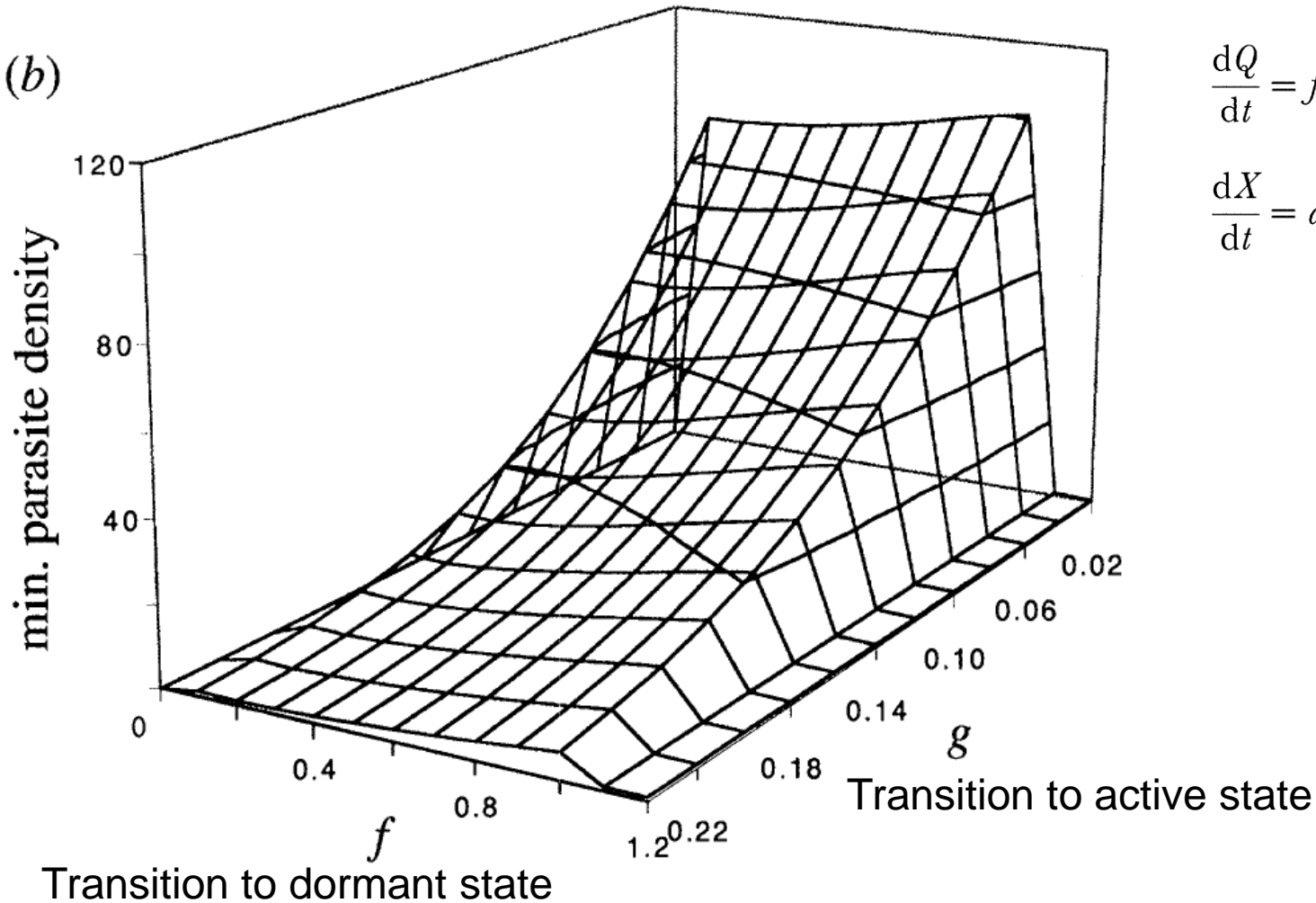
$$\frac{dP}{dt} = rP - hPX - fP + gQ,$$

$$\frac{dQ}{dt} = fP - gQ,$$

$$\frac{dX}{dt} = a + sX\left(\frac{P}{k+P}\right) - dX.$$

(b)

min. parasite density

Transition to active state

g

Transition to dormant state

f

*Antia et al (1996), Proc Royal Soc B*

14

TB 2

Wigginton and Kirschner (2001), Journal of Immunology.

IL-10 depletion "experiments". A) Effects on IFN-γ when IL-10 only (solid line), IL-10 and IL-4 (dashed line), and IL-10 and IL-12 (dotted line) are each depleted; B) effects on the activated macrophage population; C) effects on the total T lymphocytes; D) Total bacterial population.

TB 2 – Under the hood

Macrophage Component
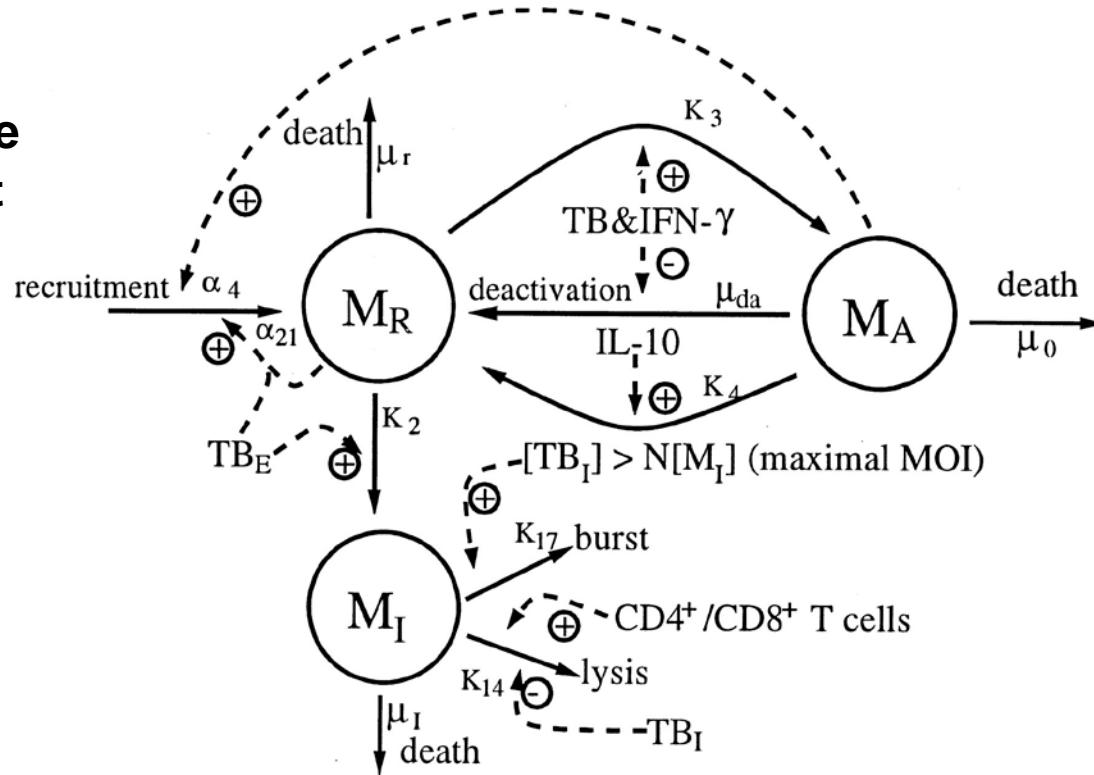
Three populations of macrophages are included in the model: resting, activated and chronically infected macrophages, denoted, respectively, by $M_R(t)$, $M_A(t)$, and $M_I(t)$.

Wigginton and Kirschner (2001), Journal of Immunology

Macrophage Equations

$$\frac{dM_R}{dt} = s_M + \alpha_4(M_A + wM_I) + \alpha_{21}M_R\left(\frac{B_T}{B_T + c_{28}}\right) + k_4 M_A\left(\frac{I_{10}}{I_{10} + s_8}\right)$$

$$- k_2 M_R\left(\frac{B_E}{B_E + c_9}\right) - k_3 M_R\left(\frac{I_\gamma}{I_\gamma + s_3}\right)\left(\frac{B_T}{B_T + c_8}\right) + \mu_{da}M_A\left(\frac{s_3}{I_\gamma + s_3}\right)\left(\frac{c_8}{B_T + c_8}\right) - \mu_r M_R$$
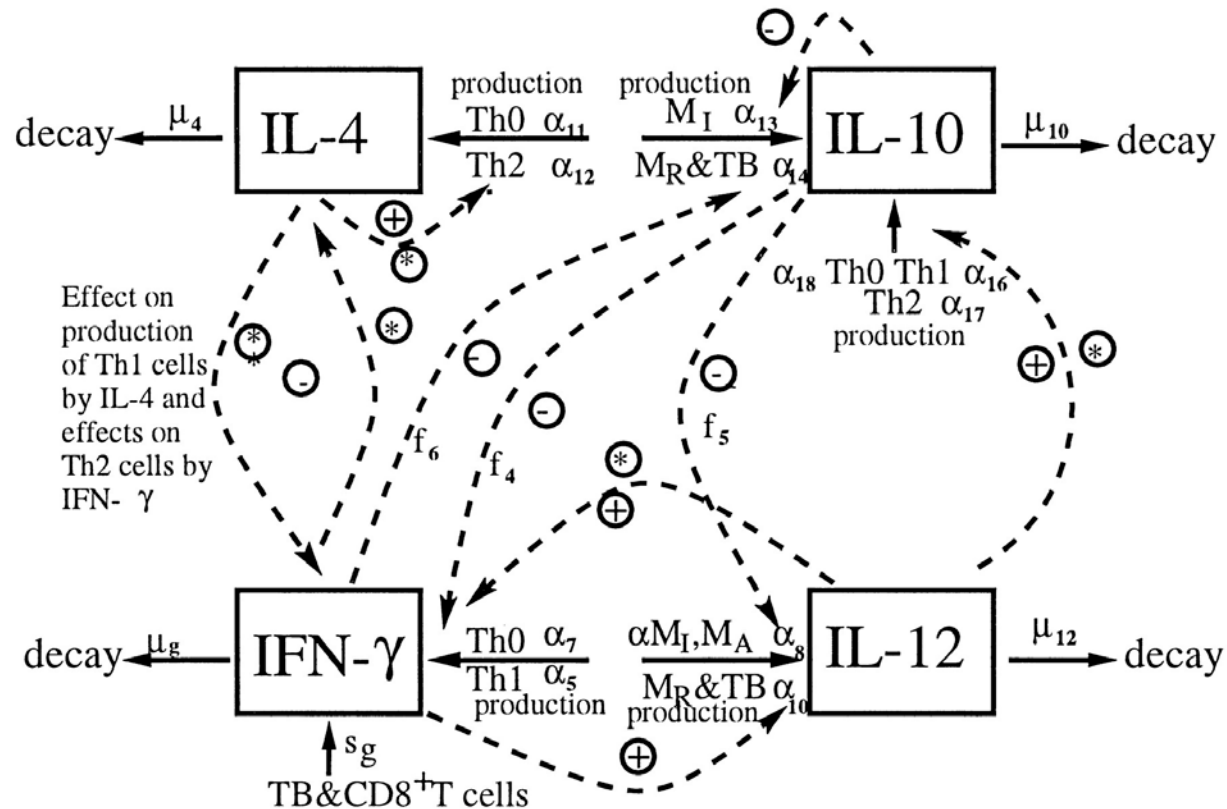
$$\frac{dM_A}{dt} = k_3 M_R\left(\frac{I_\gamma}{I_\gamma + s_3}\right)\left(\frac{B_T}{B_T + c_8}\right) - k_4 M_A\left(\frac{I_{10}}{I_{10} + s_8}\right) - \mu_{da}M_A\left(\frac{s_3}{I_\gamma + s_3}\right)\left(\frac{c_8}{B_T + c_8}\right) - \mu_a M_A$$

$$\frac{dM_I}{dt} = k_2 M_R\left(\frac{B_E}{B_E + c_9}\right) - k_{17}M_I\left(\frac{B_I^m}{B_I^m + (N\ M_I)^m + \epsilon}\right) - k_{14}M_I\left(\frac{T_T/M_I}{(T_T/M_I) + c_4}\right)\left(1 - p\frac{B_I}{B_I + N\ M_I + \epsilon}\right) - \mu_I M_I$$

Three differential equations for resting, activated and chronically infected macrophages, denoted, respectively, by $M_R(t)$, $M_A(t)$, and $M_I(t)$.

Wigginton and Kirschner (2001), Journal of Immunology

Cytokine Component

Four cytokines are tracked in the model: IL-12, IL-10, IL-4, and IFN-γ.

Wigginton and Kirschner (2001), Journal of Immunology

Cytokine Equations

$$\frac{dI_\gamma}{dt} = s_9 \left(\frac{B_T}{B_T+c_{10}}\right) \left(\frac{I_{12}}{I_{12}+s_7}\right) + \alpha_5 T_1 \left(\frac{M_A}{M_A+c_{14}}\right) + \alpha_7 \left(\frac{I_{12}}{I_{12}+f_4 I_{10}+s_4}\right) \left(\frac{M_A}{M_A+c_{14}}\right) T_0 - \mu_9 I_\gamma$$

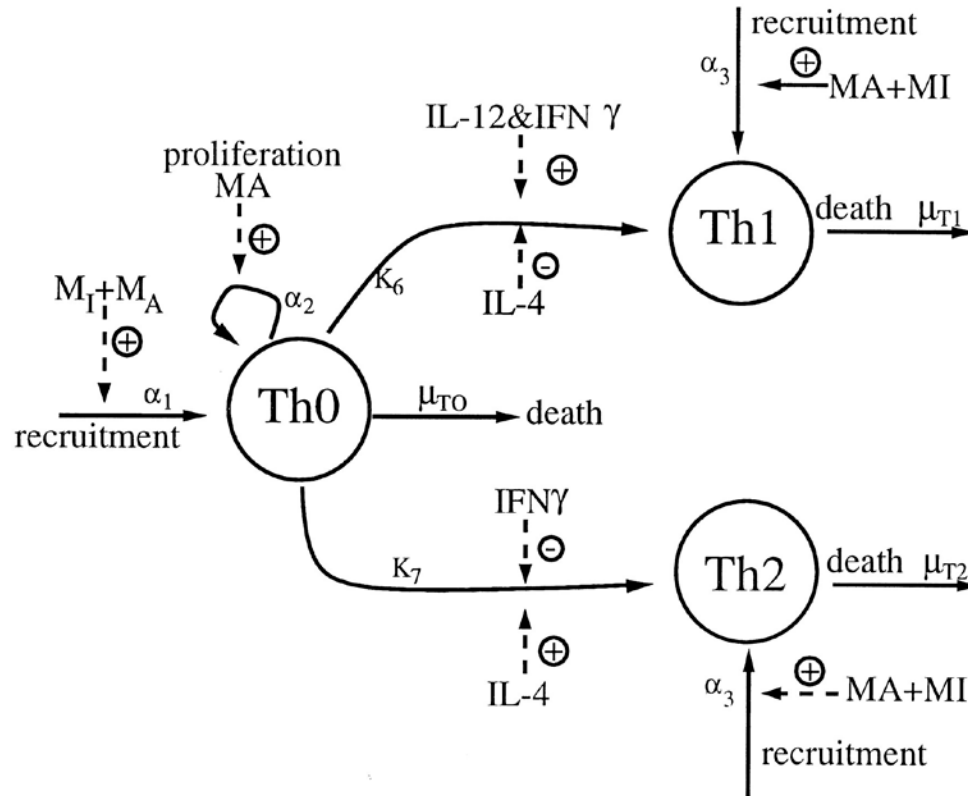$$\frac{dI_4}{dt} = \alpha_{11} T_0 + \alpha_{12} T_2 - \mu_{I4} I_4$$

$$\frac{dI_{10}}{dt} = \left(\alpha_{13} M_R \left(\frac{B_T}{B_T+c_{12}}\right) + \alpha_{14} M_A\right) \left(\frac{s_6}{I_{10}+f_6 I_\gamma+s_6}\right) + \alpha_{16} T_1 + \alpha_{17} T_2 + \alpha_{18} T_0 \left(\frac{I_{12}}{I_{12}+s_9}\right) - \mu_{I10} I_{10}$$

$$\frac{dI_{12}}{dt} = \alpha_8 M_A + \alpha_{22} M_I + \alpha_{10} \left(\frac{I_\gamma}{I_\gamma+f_5 I_{10}+s_5}\right) \left(\frac{B_T}{B_T+c_{18}}\right) M_R - \mu_{I12} I_{12}$$

Four differential equations for the four cytokines: IFN-γ, IL-4, IL-10, IL-12.

Wigginton and Kirschner (2001), Journal of Immunology

T-cell component

Three T cell populations are modeled: T_0, T_1, and T_2 cells.

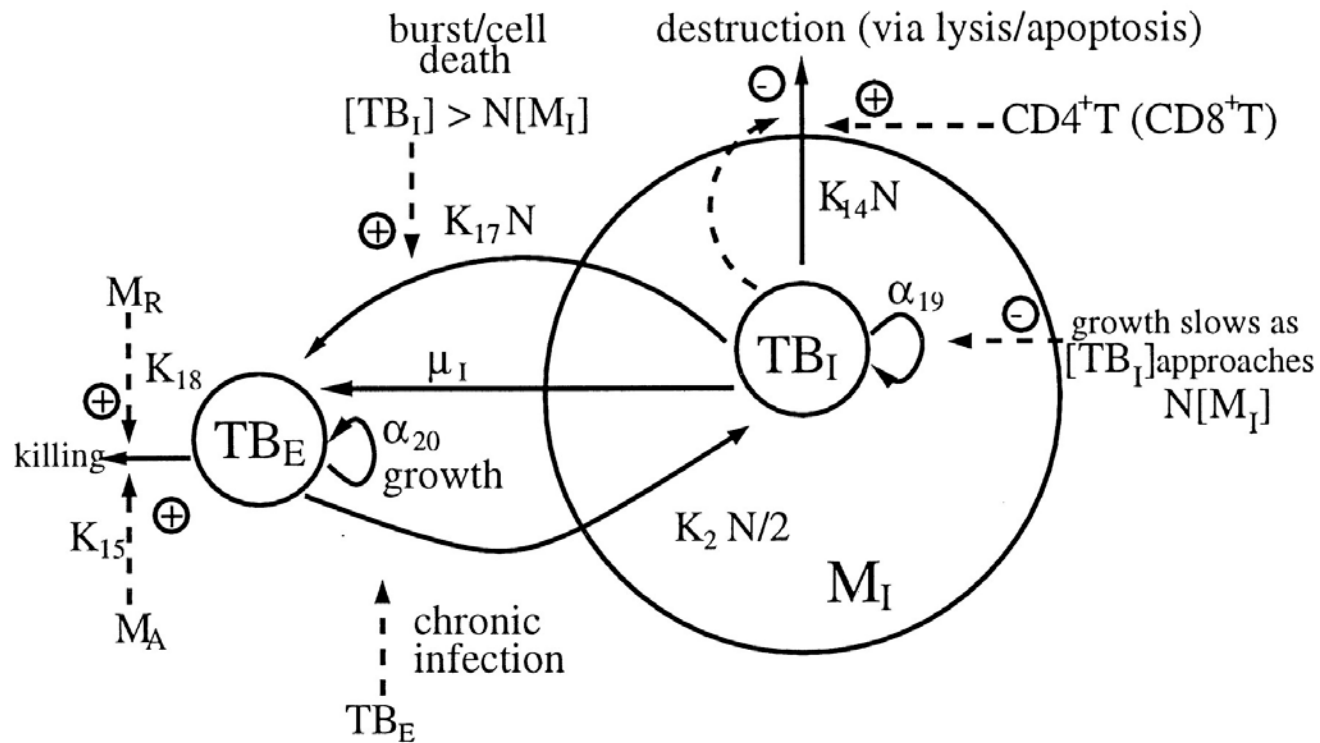*Wigginton and Kirschner (2001), Journal of Immunology*

T-cell equations

$$\frac{dT_0}{dt} = \alpha_1(M_A + wM_I) + \alpha_2 T_0\left(\frac{M_A}{M_A + c_{15}}\right) - \mu_{T0}T_0$$

$$- k_6\left(\frac{I_\gamma}{I_\gamma + f_1 I_4 + s_1}\right)I_{12}T_0 - k_7\left(\frac{I_4}{I_4 + f_2 I_\gamma + s_2}\right)T_0 \tag{8}$$

$$\frac{dT_1}{dt} = \alpha_3(M_A + wM_I) + k_6\left(\frac{I_\gamma}{I_\gamma + f_1 I_4 + s_1}\right)I_{12}T_0 - \mu_{T1}T_1 \tag{9}$$

$$\frac{dT_2}{dt} = \alpha_3(M_A + wM_I) + k_7\left(\frac{I_4}{I_4 + f_2 I_\gamma + s_2}\right)T_0 - \mu_{T2}T_2 \tag{10}$$

Wigginton and Kirschner (2001), Journal of Immunology

Bacteria component

Two bacterial populations are modeled: intracellular bacteria (B_I) and extracellular bacteria (B_E).

Wigginton and Kirschner (2001), Journal of Immunology

Bacteria equations

$$\frac{dB_E}{dt} = \alpha_{20}B_E - k_{15}M_AB_E - k_{18}M_RB_E + \mu_IB_I$$

$$+ k_{17}N\,M_I\left(\frac{B_I^m}{B_I^m + (N\,M_I)^m + \epsilon}\right) - k_2\left(\frac{N}{2}\right)M_R\left(\frac{B_E}{B_E + c_9}\right) \qquad (11)$$

$$\frac{dB_I}{dt} = \alpha_{19}B_I\left(1 - \frac{B_I^m}{B_I^m + (N\,M_I)^m + \epsilon}\right)$$

$$- k_{17}N\,M_I\left(\frac{B_I^m}{B_I^m + (N\,M_I)^m + \epsilon}\right) + k_2\left(\frac{N}{2}\right)M_R\left(\frac{B_E}{B_E + c_9}\right)$$

$$- k_{14}N\,M_I\left(\frac{T_T/M_I}{(T_T/M_I) + c_4}\right)\left(1 - p\,\frac{B_I}{B_I + N\,M_I + \epsilon}\right) - \mu_IB_I \qquad (12)$$

Wigginton and Kirschner (2001), Journal of Immunology

Table 4 – Cytokine parameters

Name	Range	Reference	Definition	Units
α_5	(0.02, 0.066)	50, 51, 56	IFN- by T_1	pg/T_1 day
α_7	(0.02, 0.066)	50, 51, 56	IFN- by T_0	pg/T_0 day
α_8	8.0×10^{-5}	60, 79, 111	IL-12 by M_A	pg/M_A day
α_{22}	8.0×10^{-5}	60, 79, 111	IL-12 by M_I	pg/M_I day
α_{11}	$(2.8 \times 10^{-3}, 9.12 \times 10^{-3})$	16	IL-4 by T_0	pg/T_0 day
α_{12}	$(2.18 \times 10^{-2}, 9.12 \times 10^{-2})$	16	IL-4 by T_2	pg/T_2 day
α_{16}	$(2.0 \times 10^{-4}, 1.0 \times 10^{-3})$	53, 55	IL-10 by T_1	pg/T_1 day
α_{17}	$(2.0 \times 10^{-4}, 6.0 \times 10^{-3})$	53, 55	IL-10 by T_2	pg/T_2 day
α_{10}	$(2.75 \times 10^{-5}, 2.75 \times 10^{-4})$	56, 56, 60, 111	Max[1] IL-12 from M_R	pg/M_R day
α_{18}	$(2.0 \times 10^{-3}, 6 \times 10^{-3})$	53, 55	Max IL-10 by T_0	pg/T_0 day
α_{13}	$(1.1 \times 10^{-3}, 1.25 \times 10^{-3})$	51, 56, 59	Max IL-10 from M_R	pg/M_R day
α_{14}	$(1.1 \times 10^{-3}, 1.25 \times 10^{-3})$	51, 56, 59	Max IL-10 from M_A	pg/M_A day
s_g	(360, 730)	16, 51, 129	Max IFN from CD8[+] and NK	pg/ml day
s_1	(50, 110)	Estimated	Half-sat, IFN- on T_0 to T_1	pg/ml
s_2	(1, 2)	Estimated	Half-sat, IL-4 on T_0 to T_2	pg/ml
s_3	(50, 110)	Estimated	Half-sat, IFN-g on M activation	pg/ml
s_4	(50, 100)	64, 112	Half-sat, IL-12 on IFN-g	pg/ml
s_5	(100, 500)	60	Half-sat, IFN-g on IL-12	pg/ml
s_6	(51, 58)	130, estimated	Half-sat, IL-10, IFN- on IL-10	pg/ml
s_7	(5, 100)	64	Half-sat, IL-12 on NK, CD8[+] IFN-g	pg/ml
s_8	(100, 500)	20	Half-sat, IL-10 on M_A deactivation	pg/ml
s_9	100	53, 54	Half-sat, IL-12 on IL-10	pg/ml
f_1	(2.9, 410)	79	Adjustment, IL-4/IFN-g on T_0 to T_1	Scalar
f_2	(0.0012, 0.16)	79	Adjustment, IFN-g/IL-4 on T_0 to T_2	Scalar
f_4	(0.76, 3.2)	52, 79	Adjustment, IL-10/IL-12 on IFN-g	Scalar
f_5	(4.8, 65)	59, 60, 79, 112	Adjustment, IL-10/IFN-g on IL-12	Scalar
f_6	(0.025, 0.053)	59, 79, 130	Adjustment, IFN-g, IL-10 on IL-10	Scalar

Wigginton and Kirschner (2001), Journal of Immunology

Table 6 – T-cell parameters

Name	Range	Reference	Definition	Units
α_1	0.05	116, estimated	Max[1] T_0 cell recruitment	T_0/M day
α_2	(1.4, 2.8)	42	Max growth rate for T_0	1/day
α_3	0.0528	116, estimated	Max T_1/T_2 recruitment	T/M day
α_4	(0.03, 0.05)	Estimated	M_R recruitment by M_A, M_I	1/day
α_{21}	(0.01, 0.07)	131, estimated	Max M_R recruitment by M_R	1/day
s_M	(330, 430)	67–70, 117	M_R source	M/ml day
μ_a	0.011	132	Death rate, M_A	1/day
μ_r	0.011	132	Death rate, M_R	1/day
μ_I	0.011	132	Death rate, M_I	1/day
μ_{da}	(0.3, 2.0)	Estimated	Deactivation rate, M_A	1/day
μ_{T0}	0.3333	114	Death rate, T_0	1/day
μ_{T1}	0.3333	114	Death rate, T_1	1/day
μ_{T2}	0.3333	114	Death rate, T_2	1/day
k_6	$(2.9\times 10^{-4}, 1.0\times 10^{-3})$	8, 62	Max T_0 to T_1 rate	ml/pg day
k_7	(0.02, 0.7)	8, 62	Max T_0 to T_2 rate	1/day
k_2	(0.2, 0.4)	Estimated	Chronic infection rate	1/day
k_3	(0.2, 0.4)	Estimated	M_R activation rate	1/day
k_4	(0.36, 0.4)	20	M_A deactivation by IL-10	1/day

Wigginton and Kirschner (2001), Journal of Immunology

Further reading

▸ Malaria: Mideo et al. "Modeling malaria pathogenesis" (2008) Cellular Microbiology and references therein.

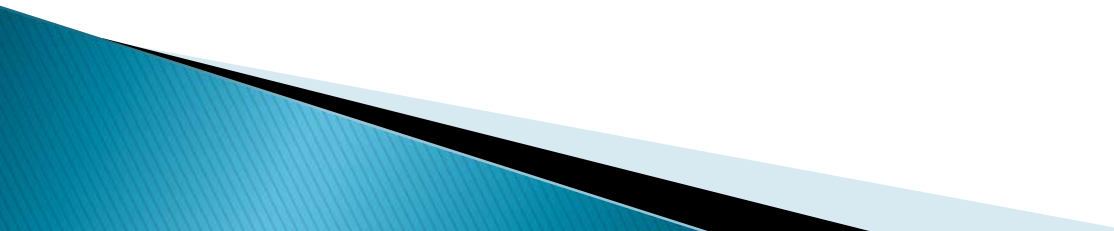▸ TB: The models by Denise Kirschner's group are a good starting point.

Model building

- Models are (should be!) built based on known biology.

- Every model is an approximation of the real system. The trick is to capture the salient features of what's going on in the system without the model becoming unwieldy.

- If one wants to include a lot of realism, the models will inevitably become rather large.

- Large models tend to contain lots of parameters.

Model Parameters

- Sometimes one can obtain estimates for parameter values needed for the model from direct fitting to data.

- Data is often not available and fitting only works for simple models.

- Parameters are usually obtained from the literature.

- Problem: Many parameters are not well known, which leads to model uncertainty.

Model Parameters

▸ Sometimes, we might be mainly interested in how results change as we vary one parameter, but we also want to know how all the other less important parameters affect our outcome.

▸ Other times, we might want to use our model to make predictions for a given scenario. But usually, we do not know the values for all parameters precisely. We need to figure out how uncertainty in parameters affects our outcome.

▸ In both cases, we need to study the impact of variations/uncertainty in multiple parameters on the output/result.

Uncertainty/Sensitivity Analysis

Uncertainty & Sensitivity Analysis

▶ Varying multiple inputs/parameters over a usually broad range is called a (global) uncertainty & sensitivity analysis.

◦ Uncertainty Analysis: Given uncertainty in the inputs, how much uncertainty is there in the outputs/results?

◦ Sensitivity Analysis: How much do individual inputs contribute to the uncertainty in outputs/results?

▶ Notes:

◦ Varying one input/parameter is sometimes called a local sensitivity analysis, since it answers the question: How sensitive is the result to changes in one input?

◦ Sometimes, the term "local sensitivity analysis" is also used with respect to changing one/multiple parameter(s) a little bit around a given value – instead of varying it over a large range.
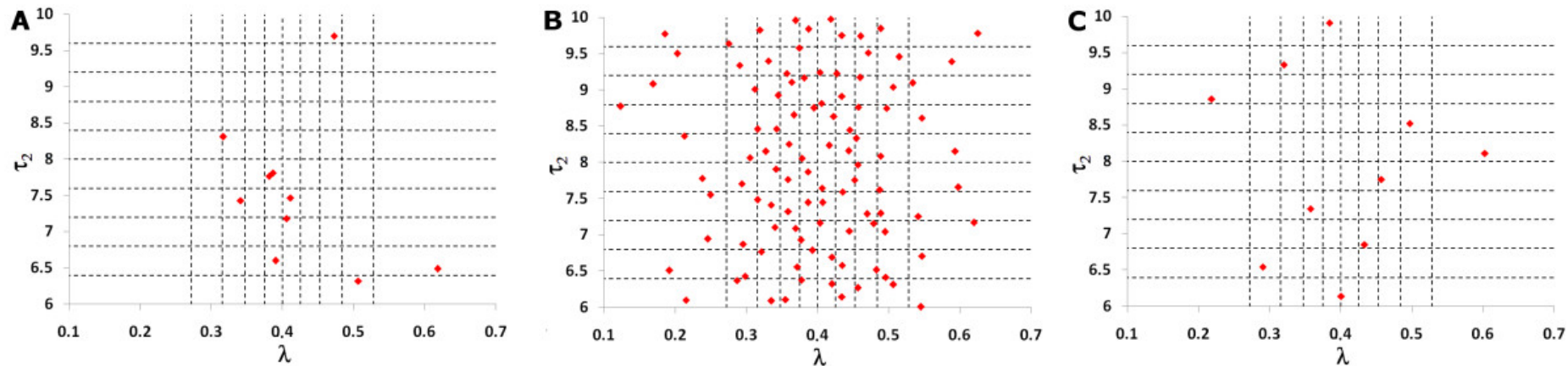
Sampling parameter values

- We want to run the model for as many different parameter values as possible.

- Usually, there are many parameters we don't know exactly. We need to vary them all and see how results change.

- If there are many parameters, simple sampling becomes too computationally intensive: If we want to fully sample 66 parameters with 10 samples for each parameter, we would need to run 10^{66} simulations (the universe is $<10^{18}$ seconds old).

- Smart ways of sampling exist, a commonly used method is Latin Hypercube Sampling (LHS).

Sampling parameter values

- First, we need to specify the distribution/range of uncertainty of the parameters.

Hoare et al. (2008), Theoretical Biology & Medical Modeling
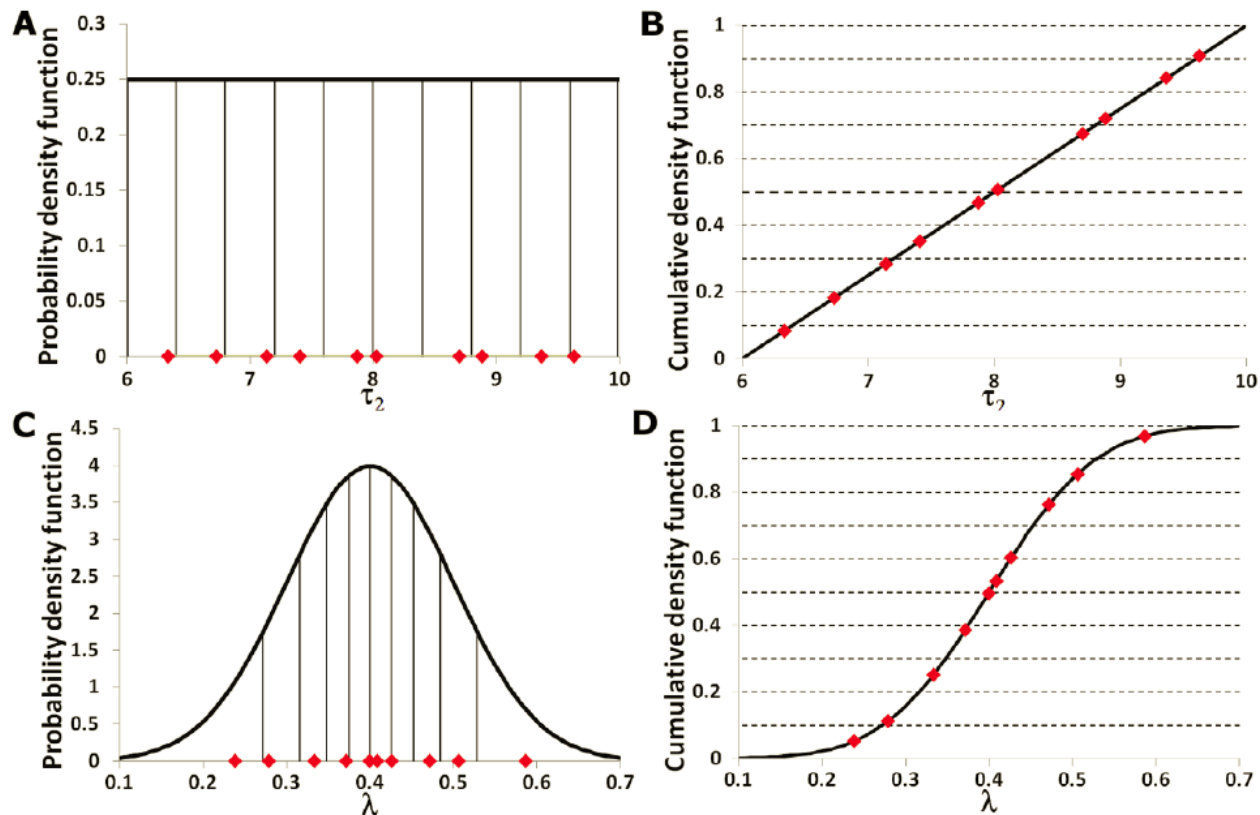
Latin Hypercube Sampling (LHS)

Example: 2 parameters. **A)** In random sampling, there are regions of the parameter space that are not sampled and other regions that are heavily sampled; **B)** in full factorial sampling, a random value is chosen in each interval for each parameter and every possible combination of parameter values is chosen; **C)** in Latin Hypercube Sampling, a value is chosen once and only once from every interval of every parameter (it is efficient and adequately samples the entire parameter space).

Hoare et al. (2008), Theoretical Biology & Medical Modeling

LHS – a few comments

- If parameters are uncertain within several orders of magnitude, probability sampling in log-space might be suitable to ensure full coverage of the whole range.

- The minimum number of samples needed for LHS is # of parameters + 1. Usually, many more should be done. For most situations, it's not too computationally challenging to do 1000s of simulations.

Uncertainty analysis

▸ For each parameter sample, we run the simulation and record the result.

▸ We can then see how uncertainty in inputs affects the results.

▸ A convenient way to represent the results is by using boxplots.
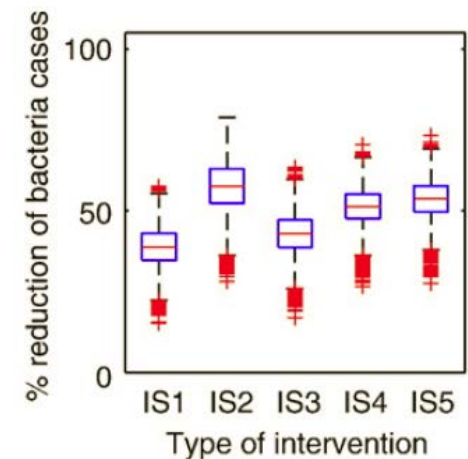
Handel et al. (2009), Epidemics

A simple worked example

▸ Simple model as an example ("TB light")

$$\dot{B} = gB\left(1 - \frac{B}{B_{max}}\right) - kBX \qquad \text{(TB bacteria)}$$

$$\dot{X} = rBX - dX \qquad\qquad\qquad \text{(immune response)}$$

▸ Model has 5 parameters: **g, k, r, d, B_{max}**
  ◦ If we don't know the initial conditions (B_0 and X_0), we can consider them model parameters. For this example, we assume we know them.
  ◦ Assume we know **k, B_{max}** and **r** exactly, but we only know that **g** and **d** are in a certain range.

▸ We want to know how "the result" depends on the values for the parameters **g** and **d.**

▸ We need to define "the result", i.e. the model output that we are interested in studying.

A simple worked example

- Let's decide that we are interested in the peak value for the bacteria. This is our output/result.

- Our inputs are the parameters **g** and **d**.

- Question: How does uncertainty in the parameters affect uncertainty in the result?

- To answer this, we produce samples for the parameters, run the model for each parameter sample and record B_{peak}. We than plot all those values.
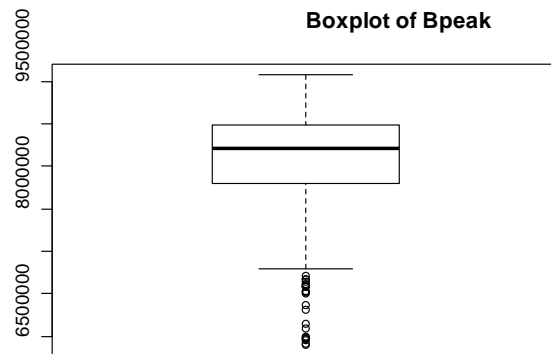
A simple worked example

- First, we need to define the ranges for the two parameters *g* and *d*
  - For *g*, assume that we know it's close to 0.5. We can choose a normal distribution with mean 0.5 and SD 0.1 (check negative values!)
  - For *d*, assume that we know nothing about *d* other than that it is between 0.1 and 0.4. Then we choose a uniform distribution.
- Next, we specify the number of LHS samples we want to run. Usually, one should choose many 1000s, but that slows down the simulation. So we pick 100.
- Now we can implement and run the whole thing.

R example

- ▶ If you haven't already, install the R packages "lhs" and "sensitivity" (the 2nd one we need later)

- ▶ Open SISMID-U6-us1.r

- ▶ Read the comments in the code to figure out what goes on
  - ◦ The script creates 100 LHS for the parameters
  - ◦ It then simulates the ODEs for each of the parameter samples and records the maximum number of bacteria
  - ◦ At the end, it plots all 100 results for Bpeak in a boxplot

- ▶ Run the code, make sure you understand what the results mean.

- ▶ Change the range of d to be between 1 and 4. Rerun the code, again make sure you understand the result.
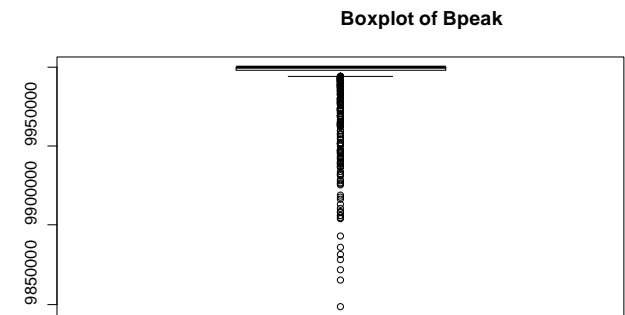
Uncertainty analysis

▸ The boxplot and CDF plot show the variation in the result as the parameters vary.
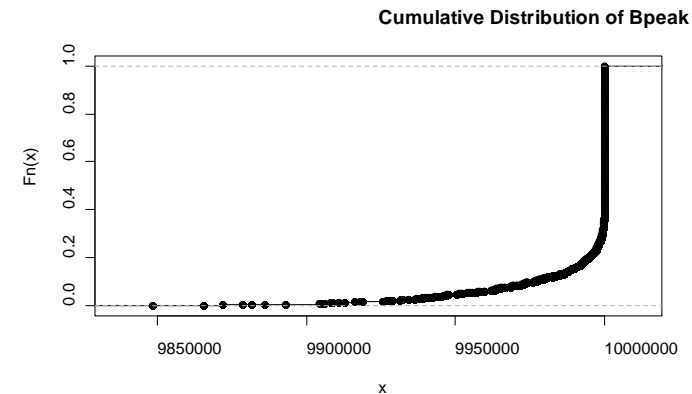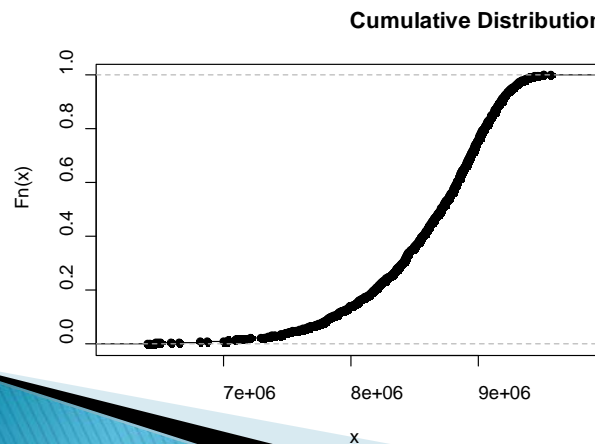
Boxplot of Bpeak

Boxplot of Bpeak

d=0.1-0.4

$$\dot{B} = gB\left(1 - \frac{B}{B_{max}}\right) - kBX$$

d=1-4

Cumulative Distribution of B

$$\dot{X} = rBX - dX$$

Cumulative Distribution of Bpeak

Sensitivity analysis

▸ Instead of just recording the amount of uncertainty in the result for a given level of uncertainty in the inputs, we can go further.

▸ We can study how much impact each parameter has on the result, i.e. how sensitive the result is to variations in each parameter/input.

Measuring Sensitivity - Correlation Coefficients

▸ Correlation Coefficients (CC) indicate how correlated a given output is with a given input.

▸ CC are between -1 and 1. Large CC means strong (negative) correlation, CC ≈ 0 means no correlation.

▸ CC assume monotonicity between input and output. That can be checked with scatterplots.
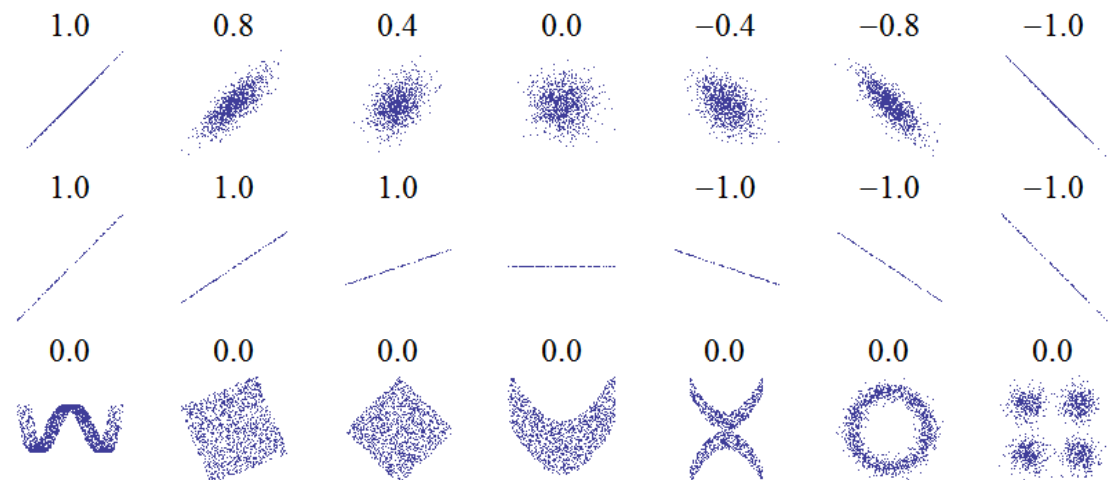
Source: Wikipedia

Rank Correlation Coefficients (RCC)

- Input-output relation is often nonlinear, therefore a Rank CC is usually more suitable.

- Example (adopted from Wikipedia): Consider the following four input/output pairs (x, y): (0, 1), (10, 100), (101, 500), (102, 2000).
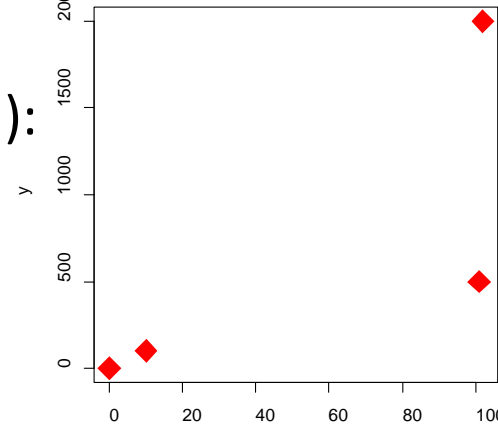
  For each pair, as x increases, so does y. But the increase is not linear. Therefore, a standard (Pearson) CC would be <1.

- If we rank-transform the values, we get (1,1), (2,2), (3,3), (4,4), i.e. a perfect correlation and a RCC would return 1.

- We usually don't know if the input/output relation is linear and we are often only interested in knowing if an output changes as function of input, but we don't care if that change is linear or not.

Partial Rank Correlation Coefficients

- RCC assume one input and one output.

- Usually, we have more than one input. We want to know the impact of a given input on the output while controlling for the other inputs.

- That can be done with Partial Rank Correlation Coefficients (PRCC).

- The package "sensitivity" in R can do PRCC. We will compute both "normal" (i.e. Spearman) RCC and PRCC.

Back to the example

▸ Let's decide that we are now interested in a different result/outcome, namely the steady state of the model

$$\dot{B} = gB\left(1 - \frac{B}{B_{\max}}\right) - kBX, \qquad \dot{X} = rBX - dX$$

$$B_s = ?, \qquad X_s = ?$$

Sensitivity Analysis Example

▸ Let's decide that we are interested in the steady state of the model

$$\dot{B} = gB\left(1 - \frac{B}{B_{max}}\right) - kBX, \qquad \dot{X} = rBX - dX$$

steady state $(\dot{X} = 0, \ \dot{B} = 0):$

$$B_s = \frac{d}{r}, \qquad X_s = \frac{g}{k}\left(1 - \frac{d}{rB_{max}}\right)$$

Sensitivity Analysis Example

▸ Our outputs/results is are the steady state values for B and X, **B$_s$** and **X$_s$**.

▸ Our inputs are still the parameters *g* and *d.*

▸ Question: How does variation in the parameters affect the results?

$$B_s = \frac{d}{r}, \quad X_s = \frac{g}{k}\left(1 - \frac{d}{rB_{\max}}\right)$$

Sensitivity Analysis Example

▸ Since we have the equations for the steady state, we know immediately how **B$_s$** and **X$_s$** depend on *g* and *d*:

  ◦ *B$_s$* increases linearly with *d* and is independent of *g*
  ◦ *X$_s$* decreases with *d* and increases linearly with *g*

▸ For more complicated models/outputs (e.g. peak bacteria in the previous example), we usually can't write down equations and therefore need to rely on simulations.
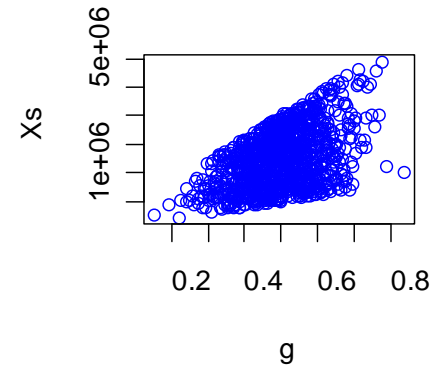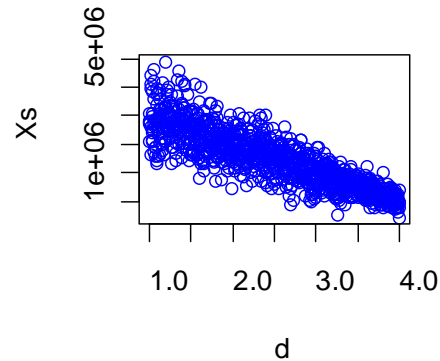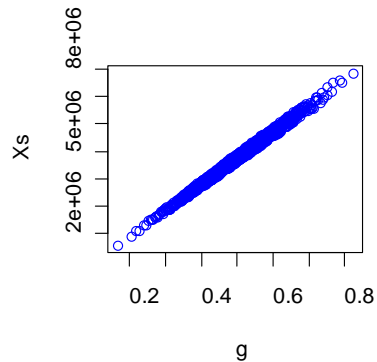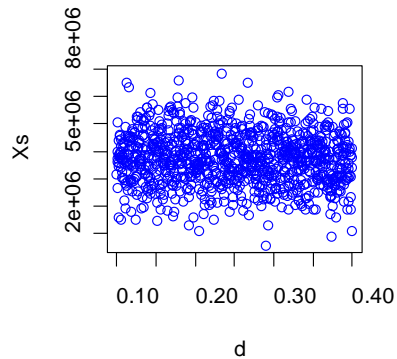
$$B_s = \frac{d}{r}, \quad X_s = \frac{g}{k}\left(1 - \frac{d}{rB_{max}}\right)$$

Sensitivity Analysis Example - R

▸ Let's see if a sensitivity analysis agrees with our analytical results (it has to!)

▸ Open SISMID-U6-us2.r, read the comments in the code to make sure you understand what goes on

▸ Run the code. Compare your plots with the equations.

▸ Look at the printout produced by the program, try to understand it.
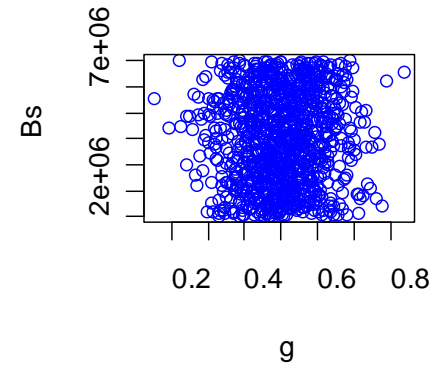
▸ Change **d** to be between 1 and 4, run the code again.

$$B_s = \frac{d}{r}, \quad X_s = \frac{g}{k}\left(1 - \frac{d}{rB_{max}}\right)$$

Scatterplots

d=0.1-0.4

d=1 - 4

$$B_s = \frac{\color{red}{d}}{r}$$
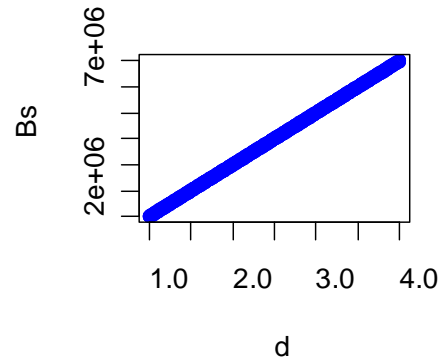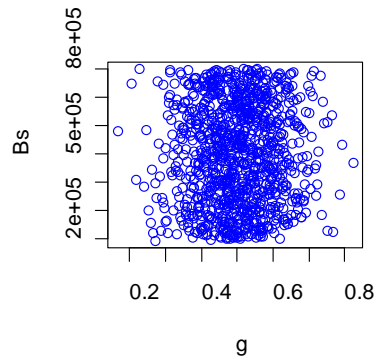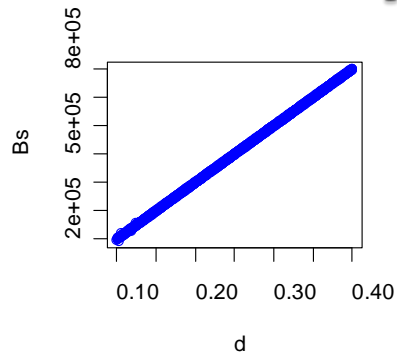
$$X_s = \frac{\color{red}{g}}{k}\left(1 - \frac{\color{red}{d}}{rB_{\max}}\right)$$

Correlation Coefficients

▸ The R script also computes and shows Spearman Rank Correlation Coefficients (SRCC) and PRCC, which indicate the importance/correlation of a parameter and the output:

◦ "SRCC for Bs and d: 1.000000, p-value: 0.000000e+00"
◦ "SRCC for Bs and g: 0.021889, p-value: 4.892352e-01"
◦ "SRCC for Xs and d: -0.877692, p-value: 0.000000e+00"
◦ "SRCC for Xs and g: 0.422837, p-value: 0.000000e+00"
◦ "PRCC for Bs and d: 1.000000"
◦ "PRCC for Bs and g: -0.041859"
◦ "PRCC for Xs and d: -0.978984"
◦ "PRCC for Xs and g: 0.922647"

Sensitivity analysis – further comments

▸ There are more ways to do uncertainty/sensitivity analysis and present results

- ◦ Graphical: Boxplots, scatterplots, …
- ◦ Numerical: Spearman rank correlation coefficient, Partial rank correlation coefficient, regression coefficients, variance decomposition…

▸ See references for more.

Discussion

▸ Uncertainty/sensitivity analysis helps understand how changes in parameter values affect the results.

▸ Especially useful for large models where analytical approaches don't work anymore.

▸ Potentially numerically challenging (takes long).

▸ Results (plots, correlation coefficients) need to be interpreted properly.

▸ For a big model "information overload" can occur.

Discussion

▸ What should one do with all that information?

- ◦ If a parameter is found to play no role on the result, maybe one can reduce/simplify the model? It could suggest that this parameter is also biologically not important.

- ◦ If a parameter is very influential, it suggests that one should try to measure/determine it in more detail.

- ◦ If one wants to affect the results (e.g. implement treatment strategies to reduce infection), one should find methods to target those influential parameters.

- ◦ Reporting the whole uncertainty/sensitivity results (or at least important aspects) allows the reader to better decide how much trust they put into the result.

A word of caution

- Extensive uncertainty/sensitivity analysis can give a false sense of accuracy.

- You specify the range of the parameters. That range could be completely wrong.

- Any other external influences (noise, unaccounted changes, etc.) are ignored.

- **Main problem: Sensitivity analysis does not help if your underlying model is structurally wrong!**

- All models are wrong. Some are decent approximations, but it's hard to say which ones are.

A word of caution

▸ So: If you learn that a model and an extensive uncertainty/sensitivity analysis predicts that a median of 31,247 people will get the flu and be hospitalized this coming flu season, with a 95% confidence interval of (12,451 – 43,887) how should you react?

A word of caution

▸ The correct response (though maybe not good to say that out loud): Whatever, that's BS.

▸ A bit more professional: Take a close look at the model.

  ◦ You might decide that it's a decent model, so you trust the ballpark figures, i.e. that the cases will be somewhere between 10K and 100K. Forget about that pretty 12,451 – 43,887 CI.

  ◦ Or you decide that the model is junk and that you can't trust any of it.

Further reading

- Good recent articles:
    - Hoare et al. (2008), Theoretical Biology & Medical Modeling
    - Marino et al. (2008) Journal of Theoretical Biology
- Application to epidemiological models: Blower and Dowlatabadi (1994) International Statistical Review
- Books:
    - W.J. Conover (1999) "Practical Nonparametric statistics", Wiley
    - A. Saltelli (2004) "Sensitivity Analysis in Practice: A guide to Assessing Scientific Models", Wiley
- Original papers on LHS: R.L. Iman & colleagues (see references in the above papers).
- More on sensitivity: Work by Saltelli, Helton (see refs in Marino et al.).
- Online tutorial: http://sensitivity-analysis.jrc.ec.europa.eu/tutorial/index.asp
- Application to an ID model: Handel et al. (2009) Epidemics

Additional topics

▸ Epidemiological model of influenza and bacteria co-infection and the effects of antiviral and antibacterial treatment.
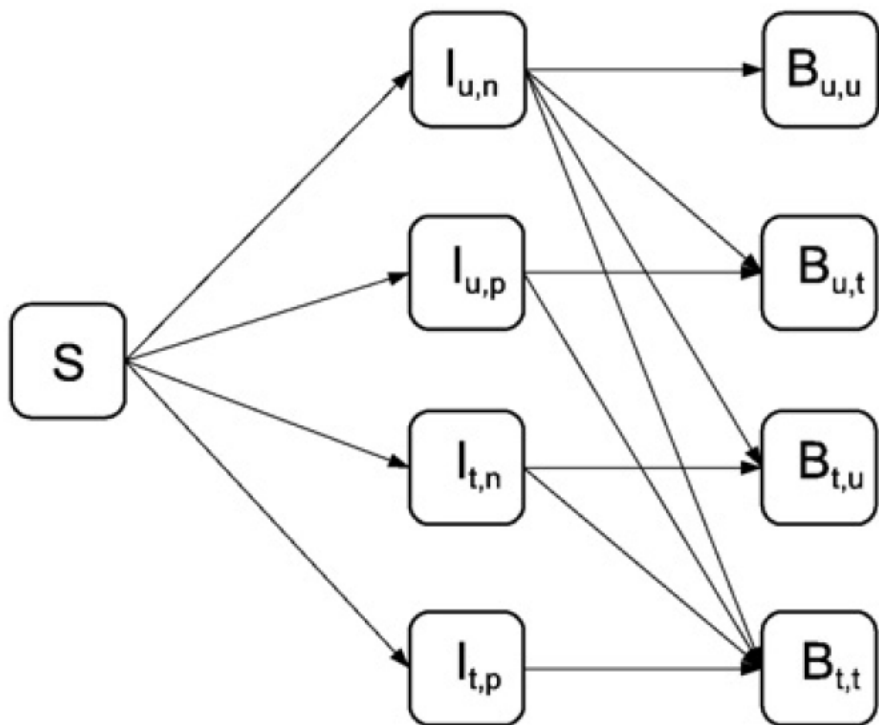
Handel et al. (2009), Epidemics

$$\lambda_{ii} = \beta_{u,n}I_{u,n} + \beta_{u,p}I_{u,p} + \beta_{t,n}I_{t,n} + \beta_{t,p}I_{t,p}$$
$$\lambda_{bi} = \alpha_{u,u}B_{u,u} + \alpha_{u,t}B_{u,t} + \alpha_{t,u}B_{t,u} + \alpha_{t,t}B_{t,t}$$
$$\lambda_1 = \lambda_{ii} + \lambda_{bi}$$
$$\lambda_{bb} = \gamma_{u,u}B_{u,u} + \gamma_{u,t}B_{u,t} + \gamma_{t,u}B_{t,u} + \gamma_{t,t}B_{t,t}$$
$$\lambda_{ib} = \kappa_{u,n}I_{u,n} + \kappa_{u,p}I_{u,p} + \kappa_{t,n}I_{t,n} + \kappa_{t,p}I_{t,p}$$
$$\lambda_2 = \lambda_{bb} + \lambda_{ib}$$
$$\dot{S} = -\lambda_1(1 - e_p f_p)S$$
$$\dot{I}_{u,n} = (1 - g_p)(1 - f_t)(1 - f_p)\lambda_1 S - v_{u,n}I_{u,n} - k_{u,n}\lambda_2 I_{u,n}$$
$$\dot{I}_{u,p} = g_p(1 - f_t)(1 - f_p)\lambda_1 S - v_{u,p}I_{u,p} - k_{u,p}\lambda_2 I_{u,p}$$
$$\dot{I}_{t,n} = (1 - g_p)(f_t(1 - f_p) + f_p(1 - e_p))\lambda_1 S - v_{t,n}I_{t,n} - k_{t,n}\lambda_2 I_{t,n}$$
$$\dot{I}_{t,p} = g_p(f_t(1 - f_p) + f_p(1 - e_p))\lambda_1 S - v_{t,p}I_{t,p} - k_{t,p}\lambda_2 I_{t,p}$$
$$\dot{B}_{u,u} = (1 - f_t)(1 - g_t)(v_{u,n}c_{u,n} + k_{u,n}\lambda_2)I_{u,n} - \delta_{u,u}B_{u,u}$$
$$\dot{B}_{u,t} = (1 - f_t)g_t(v_{u,n}c_{u,n} + k_{u,n}\lambda_2)I_{u,n}$$
$$+ (1 - f_t)(v_{u,p}c_{u,p} + k_{u,p}\lambda_2)I_{u,p} - \delta_{u,t}B_{u,t}$$
$$\dot{B}_{t,u} = f_t(1 - g_t)(v_{u,n}c_{u,n} + k_{u,n}\lambda_2)I_{u,n}$$
$$+ (1 - g_t)(v_{t,n}c_{t,n} + k_{t,n}\lambda_2)I_{t,n} - \delta_{t,u}B_{t,u}$$
$$\dot{B}_{t,t} = f_t g_t(v_{u,n}c_{u,n} + k_{u,n}\lambda_2)I_{u,n} + f_t(v_{u,p}c_{u,p} + k_{u,p}\lambda_2)I_{u,p}$$
$$+ g_t(v_{t,n}c_{t,n} + k_{t,n}\lambda_2)I_{t,n} + (v_{t,p}c_{t,p} + k_{t,p}\lambda_2)I_{t,p} - \delta_{t,t}B_{t,t}$$

U&S analysis example – parameter ranges

Symbol	Meaning	Value/range	Assumptions/references
N_0	Population size	3×10^8	U.S. population
e_p	Efficacy of AV prophylaxis	0.8	Based on Refs. (Yang et al., 2006; Halloran et al., 2007)
$v_{u,n}$	Clearance rate (1/mean duration of infection) of influenza infected hosts: AV untreated, AB non-prophylaxed	1/4.8	Based on Ref. (Carrat et al., 2008)
$v_{u,p}$	AV untreated, AB prophylaxed	$v_{u,n}$	AB prophylaxis does not alter duration of influenza infection (Louria et al., 1959; Maeda et al., 1999; Carrat et al., 2004)
$v_{t,n}$	AV treated, AB non-prophylaxed	1/3.4	Reduction of infectious period by $\approx 30\%$ (Hayden et al., 1999; Treanor et al., 2000; Whitley et al., 2001; Aoki et al., 2003)
$v_{t,p}$	AV treated, AB prophylaxed	$v_{t,n}$	AB prophylaxis does not alter duration of influenza infection
$\delta_{u,u}$	Clearance rate (1/mean duration of infection) of bacteria infected hosts: AV untreated, AB untreated	1/5–1/15	based on (Brundage and Shanks, 2008; Klugman et al., 2009)
$\delta_{u,t}$	AV untreated, AB treated	$(1.5-2)\,\delta_{u,u}$	AB treatment leads to faster recovery (Kaiser et al., 1996; Kaiser et al., 2001)
$\delta_{t,u}$	AV treated, AB untreated	$(1-1.25)\,\delta_{u,u}$	AV treatment leads to somewhat faster recovery (McCullers, 2004)
$\delta_{t,t}$	AV treated, AB treated	$(1.75-2)\delta_{u,u}$	Combined AB and AV treatment are synergistic
$c_{u,n}$	Fraction of influenza infected hosts who become bacteria infected: AV untreated, AB non-prophylaxed	0.05–0.2	Based on Refs. (Kaiser et al., 2000; Kaiser et al., 2003; Brundage, 2006)
$c_{u,p}$	AV untreated, AB prophylaxed	$(0.25-0.75)\,c_{u,n}$	AB prophylaxis lowers the probability of bacterial infection (Madhi et al., 2004; Moberley et al., 2008)
$c_{t,n}$	AV treated, AB non-prophylaxed	$(0.75-1)\,c_{u,n}$	AV treatment lowers the probability of bacterial infection (Kaiser et al., 2000; Kaiser et al., 2003; Whitley et al., 2001; Cole et al., 2002)
$c_{t,p}$	AV treated, AB prophylaxed	$(0.1-0.5)\,c_{u,n}$	Combined AV treatment and AB prophylaxis are synergistic
$d_{u,n}$	Fraction of influenza infected hosts who die: AV untreated, AB non-prophylaxed	$(1-5) \times 10^{-3}$	$\approx 1\%-5\%$ deaths for 1918-like, most deaths due to bacterial infections (Soper, 1918; Mills et al., 2004; Brundage, 2006; Bonten and Prins, 2006; Brundage and Shanks, 2008; Morens et al., 2008); $\approx 0.1\%-0.5\%$ deaths for 2009-like, few deaths due to bacterial infections (Fraser et al., 2009; for Disease Control, C., Prevention, 2009; Jamieson et al., 2009; Garske et al., 2009)
$d_{u,p}$	AV untreated, AB prophylaxed	$d_{u,n}$	AB prophylaxis does not affect death due to virus (Louria et al., 1959; Maeda et al., 1999; Carrat et al., 2004)
$d_{t,n}$	AV treated, AB non-prophylaxed	$(0.25-0.75)\,d_{u,n}$	AV treatment reduces mortality
$d_{t,p}$	AV treated, AB prophylaxed	$d_{t,n}$	AB prophylaxis does not affect death due to virus
$\varepsilon_{u,u}$	Fraction of bacteria infected hosts who die: AV untreated, AB untreated, 1918-like	$(1-2.5) \times 10^{-1}$	$\approx 1\%-5\%$ deaths for 1918-like, most deaths due to bacterial infections (Soper, 1918; Mills et al., 2004; Brundage, 2006; Bonten and Prins, 2006; Brundage and Shanks, 2008; Morens et al., 2008)
	2009-like	$(1-2.5) \times 10^{-3}$	$\approx 0.1\%-0.5\%$ deaths for 2009-like, few deaths due to bacterial infections (Fraser et al., 2009; for Disease Control, C., Prevention, 2009; Jamieson et al., 2009; Garske et al., 2009)
$\varepsilon_{u,t}$	AV untreated, AB treated	$(0.1-0.5)\,\varepsilon_{u,u}$	AB treatment reduces death due to bacterial infection (Louria et al., 1959)
$\varepsilon_{t,u}$	AV treated, AB untreated	$(0.5-1)\,\varepsilon_{u,u}$	AV treatment reduces death due to bacterial infection (McCullers, 2004; McCullers and Bartmess, 2003)
$\varepsilon_{t,t}$	AV treated, AB treated	$(0.01-0.2)\,\varepsilon_{u,u}$	Combined AB and AV treatment are synergistic (McCullers, 2004)
f_p	Fraction of uninfecteds receiving AV prophylaxis	0–1	Varied for different scenarios
f_t	Fraction of influenza infecteds receiving AV treatment	0–1	Varied for different scenarios
g_p	Fraction of influenza infecteds receiving AB prophylaxis	0–1	Varied for different scenarios
g_t	Fraction of co-infecteds receiving AB treatment	0–1	Varied for different scenarios

U&S analysis example – outcome uncertainty
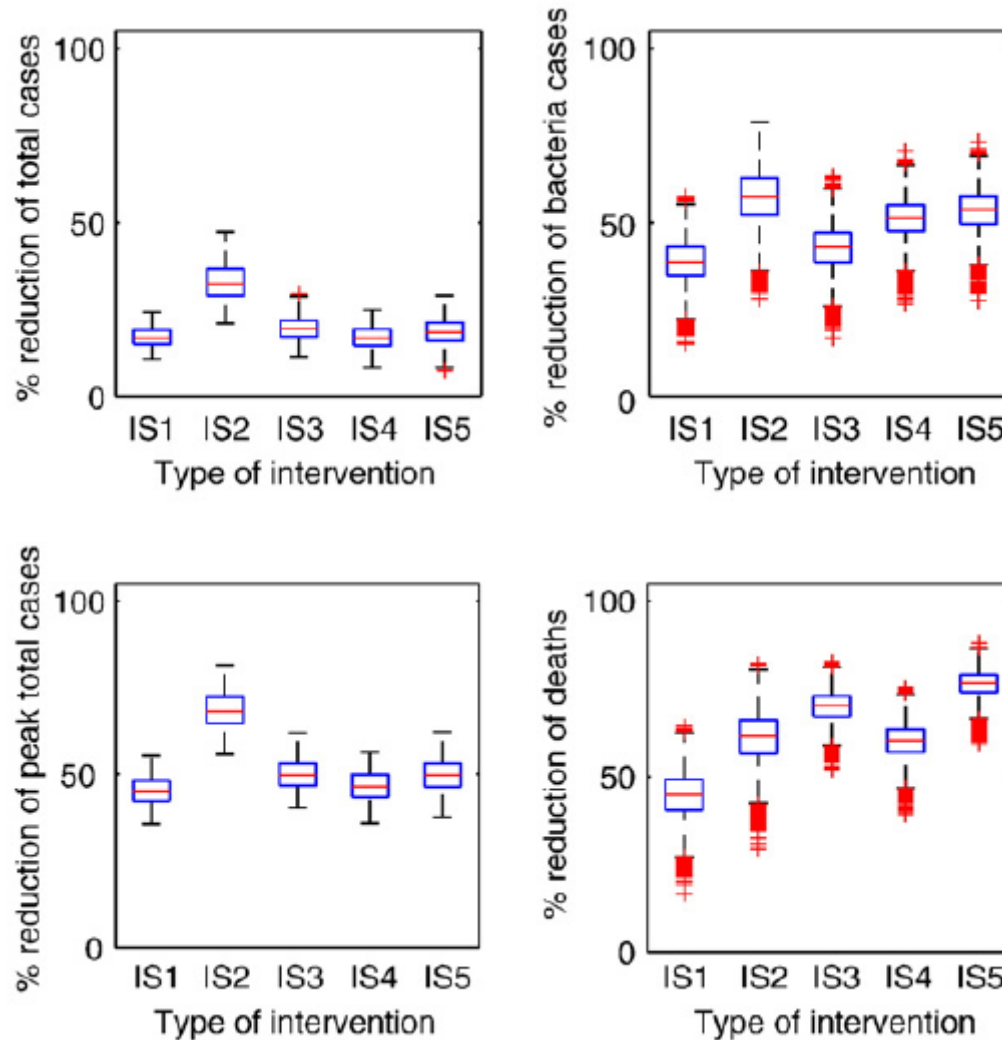
U&S analysis example – correlation coefficients

Parameter	Intervention strategy				
	IS1	IS2	IS3	IS4	IS5
Reduction of total cases					
$\beta_{u,n}$	−0.96 (1)	−0.96 (1)	−0.95 (1)	−0.95 (1)	−0.93 (1)
$c_{u,n}$	−0.15 (3)	−0.18 (2)	−0.01 (19)	−0.02 (15)	0.04 (12)
$\alpha_{u,u}$	−0.11 (5)	−0.16 (3)	0.29 (3)	0.13 (3)	0.28 (2)
$\alpha_{u,t}$	0.002 (29)	−0.006 (18)	−0.33 (2)	−0.08 (10)	−0.23 (3)
$\delta_{t,u}$	0.22 (2)	0.15 (4)	0.06 (11)	0.11 (4)	0.02 (17)
$k_{u,p}$	−0.007 (19)	−0.001 (29)	−0.02 (15)	0.27 (2)	0.18 (4)
Reduction of bacteria cases					
$\beta_{u,n}$	−0.65 (1)	−0.73 (1)	−0.64 (1)	−0.58 (1)	−0.56 (1)
$c_{u,p}$	−0.01 (13)	−0.02 (14)	−0.02 (19)	−0.51 (2)	−0.48 (2)
$c_{t,n}$	−0.42 (2)	−0.33 (2)	−0.37 (2)	−0.30 (3)	−0.27 (3)
$\kappa_{u,n}$	0.16 (3)	0.14 (3)	0.16 (4)	0.13 (6)	0.12 (8)
$\gamma_{u,u}$	0.15 (4)	0.11 (7)	0.16 (3)	0.12 (7)	0.13 (7)

Sensitivity analysis - regression

▸ Instead of using CC, another approach is to do sensitivity analysis with regression.

▸ Regression analysis is well developed in statistics.

▸ Simple linear regression example: $Y = b_0 + \sum_i b_i X_i$

  ◦ Y is the outcome (every outcome gets its own equation).
  ◦ X_i are the parameters/inputs.
  ◦ b_i are the regression coefficients (RC).

▸ The RC tell us how a change in a parameter X_i affects the result Y.

Sensitivity analysis - regression

▸ Usually we want to do generalized linear regression, e.g. by including quadratic terms.

$$Y = b_0 + \sum_i b_i X_i + \sum_i b_{ii} X_i^2 + \sum_{i,j} b_{ij} X_i X_j$$

▸ Note: Even though the model is nonlinear in the X_i terms, statisticians call this a linear model, because the RC (b_i) appear only in linear form. This can be confusing since mathematicians (or physicists/engineers) would call this a nonlinear model/equation.

Sensitivity analysis - regression

- The regression model/equation is a surrogate for our real (dynamic, ODE) model.

- Only if the regression model "works well", i.e. can explain the data well, can we consider it a good proxy.

- To measure quality of the regression model, we can use the (adjusted) coefficient of determination (R^2).

- R^2 is the is the proportion of variability that is accounted for by the regression model. Only if this value is close to 1 can we trust the regression analysis.

- See "Coefficient of determination" on Wikipedia for more.

Sensitivity analysis - regression

- Since the parameters/inputs are of different magnitude, so are the RC. It is therefore difficult to compare and rank them.

- One can normalize/standardize the coefficients by multiplying with SD_x/SD_y

$$Y = b_0 + \sum_i b_i X_i \qquad b_i^* = \frac{SD_{X_i}}{SD_Y} b_i$$

- The SRC (b_i^*) are comparable because they all refer to a 1 standard deviation change in their respective independent variables. A standard deviation change in X_i results in b_i^* standard deviation increase in Y.

- This standardization doesn't mean they have to be between -1 and 1, though they often are.

Sensitivity analysis - regression

- The same issue with regard to linearity between input and output that applies to CC also applies to RC.

- Since we often don't know/care about linearity between input and output, we can again use rank transformation to obtain standardized rank regression coefficients (SRRC).

Regression analysis - Example

- Let's revisit our familiar model.

- Our input is the same as before, the parameters *d* and *g*.

- We now want to use B_s and X_s as well as B_{peak} as output.

- We need to define a regression model. Let's choose

$$Y = b_1 X_1 + b_2 X_2 + b_3 X_1 X_2$$

- The X_i are our inputs/parameters, **Y** are our outputs (one equation for each of the 3 outputs) and the b_i are the RC.

Regression analysis - R

- Open SISMID-U6-us3.r

- Read the code, make sure you understand it. Run it, read/understand the output.

- Run the code to perform regression and compare the results from the models.

- Change **d** to be between 1 and 4, run the code again.

$$B_s = \frac{d}{r}, \quad X_s = \frac{g}{k}\left(1 - \frac{d}{rB_{\max}}\right)$$

Regression analysis – R Results
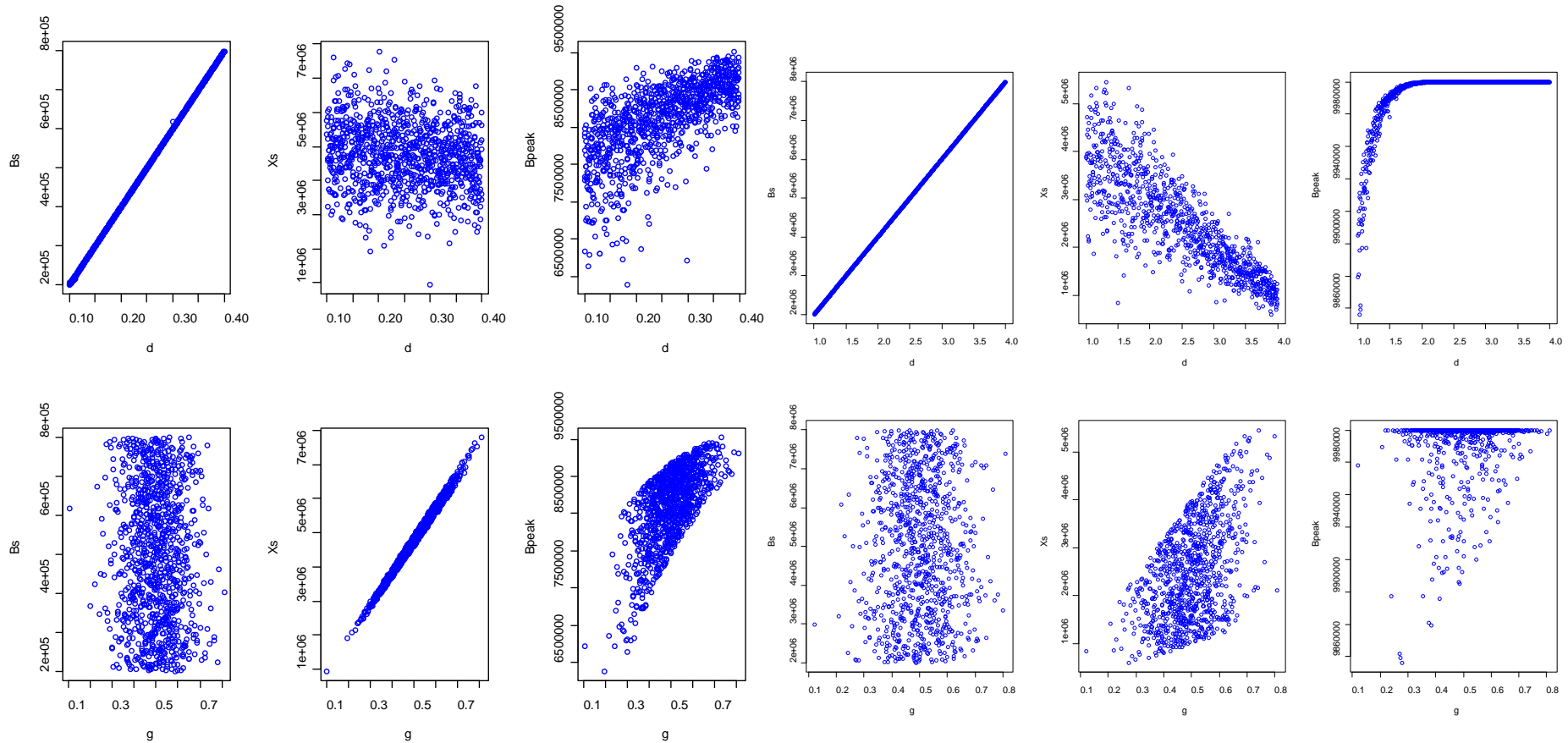
d between 0.1 and 0.4

d between 1 and 4

$$B_s = \frac{d}{r}, \quad X_s = \frac{g}{k}\left(1 - \frac{d}{rB_{\max}}\right)$$

72

Regression analysis – R Results

- Adjusted R-squared for Bs, Xs, Bpeak - first regression model: 1, 1, 0.39035
- Adjusted R-squared for Bs, Xs, Bpeak - simple regression model: 1, 0.9688, 0.3466
- [1] "PRCC for Bs and d: 1.000000"
- [1] "SRRC for Bs and d: 1.000000"
- [1] "SRRC2 for Bs and d: 1.000000"
- [1] "PRCC for Bs and g: -0.093637"
- [1] "SRRC for Bs and g: -0.000000"
- [1] "SRRC2 for Bs and g: -0.000000"
- [1] "PRCC for Xs and d: -0.971894"
- [1] "SRRC for Xs and d: -0.980639"
- [1] "SRRC2 for Xs and d: -0.894087"
- [1] "PRCC for Xs and g: 0.910344"
- [1] "SRRC for Xs and g: 0.430890"
- [1] "SRRC2 for Xs and g: 0.476393"
- [1] "PRCC for Bpeak and d: 0.999078"
- [1] "SRRC for Bpeak and d: 1.034723"
- [1] "SRRC2 for Bpeak and d: 0.997386"
- [1] "PRCC for Bpeak and g: 0.387463"
- [1] "SRRC for Bpeak and g: 0.037640"
- [1] "SRRC2 for Bpeak and g: 0.018011"

Regression – Interpreting the results

- The equation for the steady state are

$$B_s = \frac{d}{r}, \quad X_s = \frac{g}{k}\left(1 - \frac{d}{rB_{max}}\right)$$

- Our regression models are

$$Y = b_1 X_1 + b_2 X_2 + b_3 X_1 X_2 \qquad Y = b_1 X_1 + b_2 X_2$$

- With $X_1 = d$, $X_2 = g$, we get

$$Y = b_1 d + b_2 g + b_3 gd \qquad Y = b_1 d + b_2 g$$

- Both regression models perfectly describe how B_s depends on the parameters (i.e. $b_1 = 1/r$, $b_2 = b_3 = 0$). Therefore, we expect $R^2 = 1$.

- Regression model 1 can perfectly describe how X_s depends on g, d (i.e. $b_1 = 0$, $b_2 = 1/k$, $b_3 = -1/krB_{max}$), model 2 can not. Therefore, we expect $R^2 = 1$ for the 1st model, $R^2 < 1$ for the second.

Regression - Interpreting the results

▶ Since we don't have an analytic expression for B_{max}, we can't check that one easily.

▶ We find that the quality of regression differs depending on the range of **d.** Range of 1-4:

◦ Adjusted R-squared for Bpeak - first regression model: 0.390359
◦ Adjusted R-squared for Bpeak - simple regression model: 0.34665

▶ Range 0.1-0.4:

◦ Adjusted R-squared for Bpeak - first regression model: 0.972277"
◦ Adjusted R-squared for Bpeak - simple regression model: 0.94579

▶ Why is that?