Modelling Epidemics

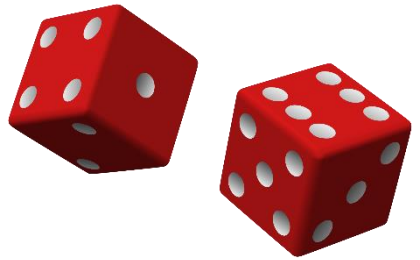
Lecture 7: Stochastic & individual based epidemiological models

Andrea Doeschl-Wilson

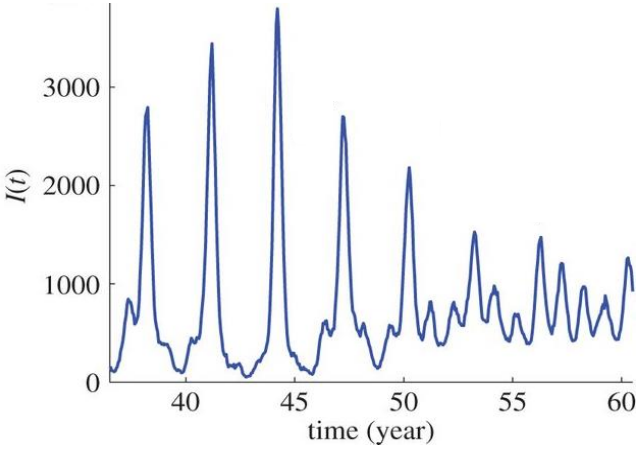
Overview

- Why and when do we need stochastic models?
- Key features of stochastic epidemiological models
- Different approaches for including stochasticity
- Implementation: Gillespie’s direct algorithm
- Analysing stochastic models
- Individual based stochastic models
- Some examples of stochastic genetic epidemiological models

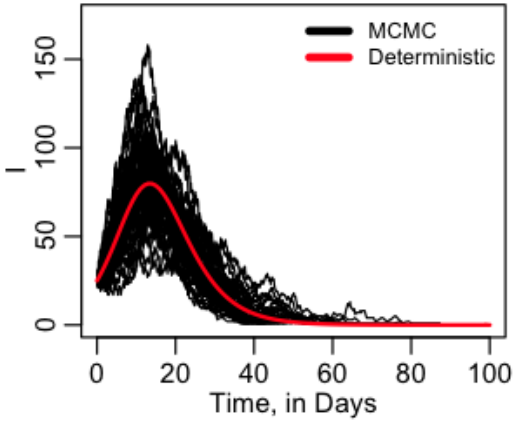
Why & when we need stochastic epidemiological models



- Infection is a chance event
  - Depends on probability of contact, transmission at contact, …
- Stochasticity is particularly important when the number of infectious individuals is small
  1. At the early stage, when disease is invading
     → Probability of an outbreak to occur
  2. During a trough phase of an epidemic cycle
     → Probability of extinction
  3. When population size is small
     →Chance fluctuations cause extinction



Distinct features of stochastic epidemiological models



1. Variability between simulations
   - Different simulations produce different outcomes
   - Precise predictions of statistical measures (e.g. mean, variance)
   - No precise prediction of disease prevalence at given point in time
2. Mean predictions of stochastic models comparable (but not equal!) to those of the equivalent deterministic model
   - Deviation from deterministic mean due to negative covariance between infectious and susceptible individuals
3. Stochastically driven extinctions, **even if R_0>1!**
   - In closed populations, chance fluctuations always result in eventual extinction. Long-term persistence requires influx of pathogens

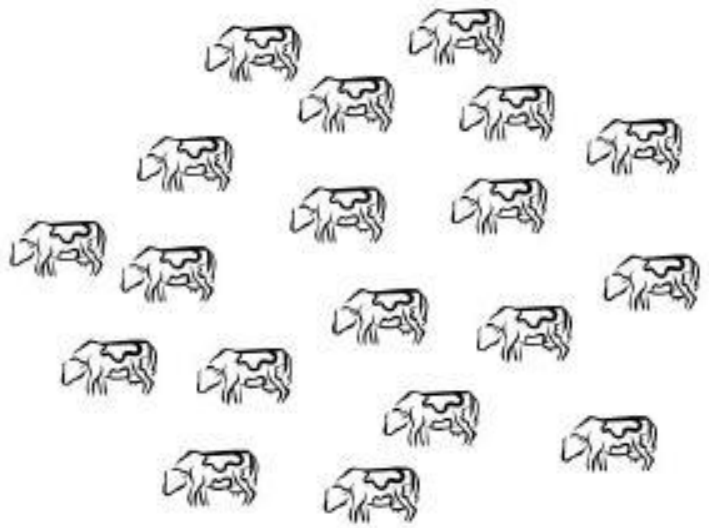
Different approaches to include stochasticity

Common for all approaches:

- **Replace proportions by integer numbers**

From now on the variables **S, I, R** denote the **number** of susceptible, infectious and recovered individuals, respectivelv

- **Implement random number generators**



Different approaches to include stochasticity

1. Stochastic differential equations (SDE)

- Add stochastic terms into the deterministic model to simulate random noise
- E.g. stochastic differential SIR model:

$$\frac{dS}{dt} = -[\beta SI/N + f_1(S,I)\omega_1]$$

$$\frac{dI}{dt} = [\beta SI/N + f_1(S,I)\omega_1] - [\gamma I + f_2(I)\omega_2]$$

$$\frac{dR}{dt} = [\gamma I + f_2(I)\omega_2]$$

Where:

- ω_1 and ω_2 are random samples from the Normal distribution N(0,1)
- $f_1(S,I)$ and $f_2(I)$ are scaling functions to scale noise with respect to variable size

Implementing different types of noise into SDEs

$$\frac{dS}{dt} = -[\beta SI/N + f(S,I)\omega_1]$$

1. Plain additive noise: f is constant
   - Independent of the population size
   - Increasing f increases the variance in the number of infected and the negative covariance between S and I, which causes changes to the mean values & variation about the mean
2. Scaled additive noise: $f = \sqrt{\beta SI/N}$
   - Variations are larger in large populations
   - For large populations, variance in the number of cases is proportional to the mean

Implementing different types of noise into SDEs

$$\frac{dS}{dt} = -[\beta SI/N + f(S,I)\omega_1]$$

1. Plain additive noise: f is constant
2. Scaled additive noise: $f = \sqrt{\beta SI/N}$

Assume that population levels are subject to random fluctuations

3. External parameter noise: $f \propto \frac{\beta SI}{N}$
   - e.g. climatic conditions may affect transmission rates
4. Heterogeneous parameter noise: $f \propto \frac{\beta S\sqrt{I}}{N}$
   - Incorporate individual variation (e.g. in contact behaviour or immune response)
   - Fluctuations depend on which individuals are infected
   - Stochasticity decreases as the number of infected individuals increases

Assume that parameters are subject to random fluctuations

Pros and cons for stochastic differential equations

Advantages:

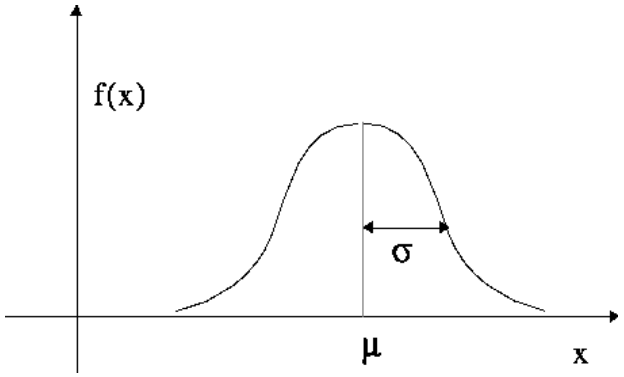
- Straight-forward to implement and low computational costs
- Similar analytical tools as for deterministic ODE models

Disadvantages:

- Account for, but do not explicitly incorporate individual differences
- Not suitable for incorporating genetics
- Not suitable when population levels (nr of infectious individuals) are small

Different approaches to include stochasticity

2. Pseudo-stochastic models



- Assume distributions for the model parameters
- Take random drawings for each parameter
  - e.g. $\beta_i \in N(\bar{\beta}, \sigma_\beta), \gamma_i \in N(\bar{\gamma}, \sigma_\gamma), \quad \mathrm{i} = 1 \ldots \mathrm{n}$
- Use deterministic model to simulate one epidemic for each parameter set (β_i, γ_i)
  - Generate many independent epidemics
- Use statistics to analyse the range of possible outputs

Pros and cons for pseudo stochastic models

Advantages:

- Represents uncertainty or fluctuations in parameter estimates
- Quick to implement and analyse
- Provides a range of possible outcomes associated with different parameter values

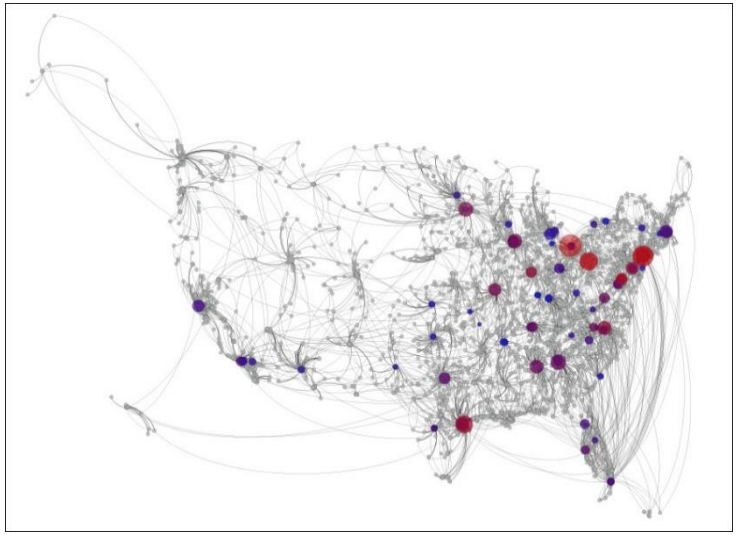
Disadvantages:

- Does not represent the stochastic nature of infection

Different approaches to include stochasticity

3. Event driven approaches

- Suitable to incorporate the random nature of events at the individual level
- Events occur at a fixed baseline probability
- But individuals experience different fates due to chance



Event driven approaches – basic methodology

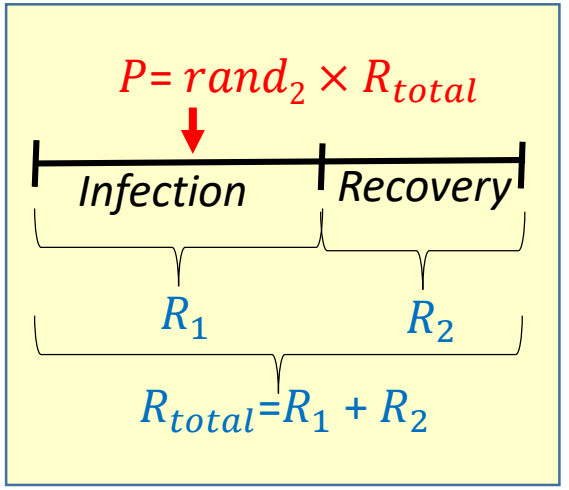
- Simulate many realizations of the same epidemic.
- Each realization is a time series of (random) events.
- Common to all realizations:

1. Define events:
   - E.g. for SIR model in closed populations the events are *infection* and *recovery*
2. Define event rates
   - E.g. for SIR model in closed populations infection occurs at a rate $\beta SI/N$ and recovery occurs at a rate γI
3. Determine the change in number of individuals in each compartment associated with each event:
   - E.g. infection: S → S-1, I → I+1
   - Recovery / Death: I→ I-1, R→R+1

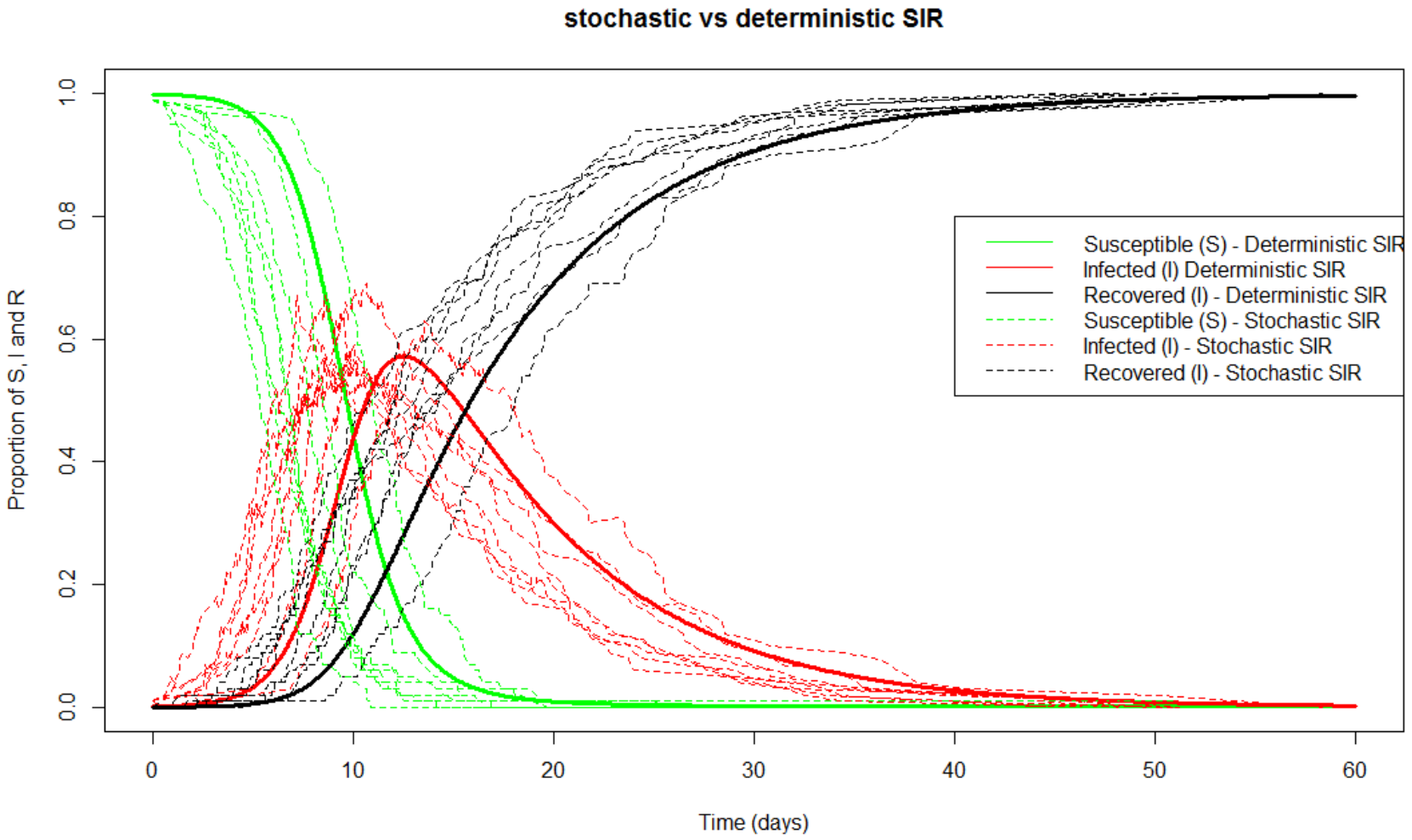
Implementing event driven approaches: Gillespie's direct algorithm

1. **Determine** all possible **events**: E_1, …, E_n
   - E.g. E_1 = infection, E_2 = recovery for SIR model
2. For each event, **determine the rate** at which it occurs, R_1, …, R_n
   - Note that the rates may vary over time (e.g. infection rate depends on nr of infected!)
3. The rate at which any event occurs is $R_{total} = \sum_{k=1}^{n} R_k$
4. Draw random number rand_1 and **determine the time until the next event**: $dt = -1/R_{total}\log(\text{rand}_1)$
5. Draw random number rand_2 and **determine the next event**:
   Set $P = rand_2 \times R_{total}$; event p occurs if $\sum_{k=1}^{p-1} R_k < P \leq \sum_{k=1}^{p} R_k$
6. **Perform event**: Update time & number of individuals in each class

REPEAT



Typical graph of a stochastic SIR model



What is the relationship between the average infection profiles from the stochastic model & those from the deterministic model?

Probability of extinction and R_0

Assume 1 infectious individual is introduced into a totally susceptible population.

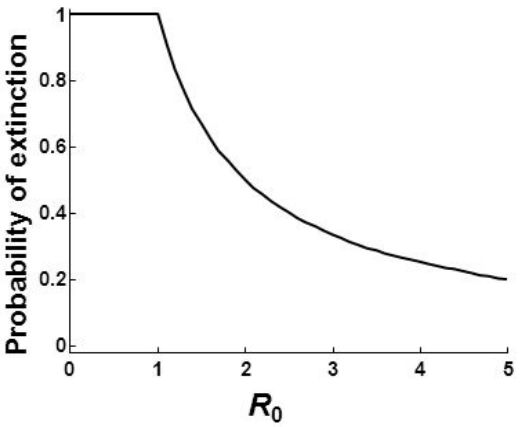
What is the probability P_{ext} that the disease goes extinct before it can cause an epidemic?

Initially 2 events can happen:

- (i) infected individual recovers → extinction
- (ii)infected individual infects another individual → 2 infected individuals
  - Probability of extinction after transmission is P_{ext}^2

Therefore:

$$\boldsymbol{P_{ext}} = \frac{\gamma}{\beta+\gamma} \times 1 + \frac{\beta}{\beta+\gamma} \times P_{ext}^2 = \frac{\gamma}{\beta} = \boldsymbol{\frac{1}{R_0}}$$



An example for a stochastic epidemiological model

JOURNAL OF ANIMAL SCIENCE
The Premier Journal and Leading Source of New Knowledge and Perspective in Animal Science

Developing stochastic epidemiological models to quantify the dynamics of infectious diseases in domestic livestock.

K MacKenzie and S C Bishop

J ANIM SCI 2001, 79:2047-2056.

- Model transmission dynamics for microparasitic infections in a typical pig farm
- Predict the impact of altering the farm structure on the epidemiology

MacKenzie & Bishop, JAS 2001 (1)

JOURNAL OF ANIMAL SCIENCE
The Premier Journal and Leading Source of New Knowledge and Perspective in Animal Science

Utilizing stochastic genetic epidemiological models to quantify the impact of selection for resistance to infectious diseases in domestic livestock

K. MacKenzie and S. C. Bishop

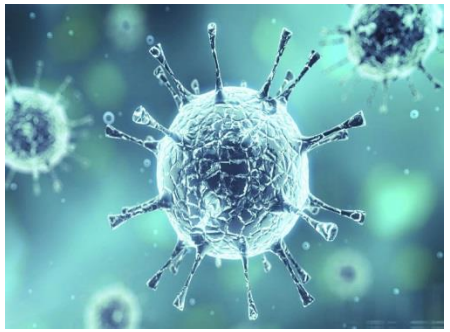
J ANIM SCI 2001, 79:2057-2065.

- Incorporate genetics into the previous model
- Predict the impact of genetic selection on the epidemiology

MacKenzie & Bishop, JAS 2001 (2)

Transmissible Gastroenteritis (TGE)

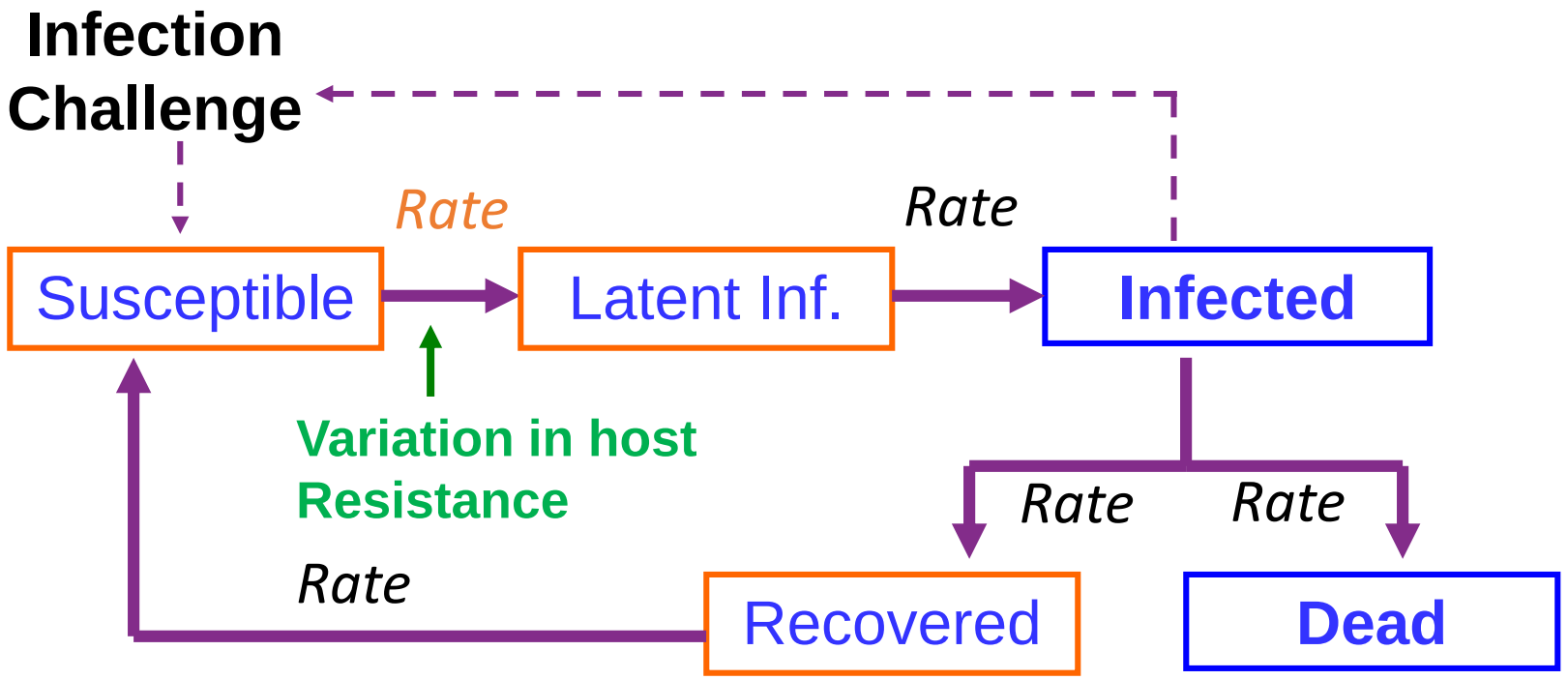
- Gastro-intestinal disease in pigs caused by coronavirus
- Acute, rapidly spreading
- Causes diarrhea and vomiting



Transmissible Gastroenteritis (TGE)

- Susceptible pigs
  become latent with a probability
- Latent pigs
  become infectious at a rate
- Infected piglets
  die at a given rate
- Infected pigs
  recover at a given rate
- Recovered pigs
  lose immunity at a given rate

Epidemiological TGE model



Incorporate the structure of a typical pig farm

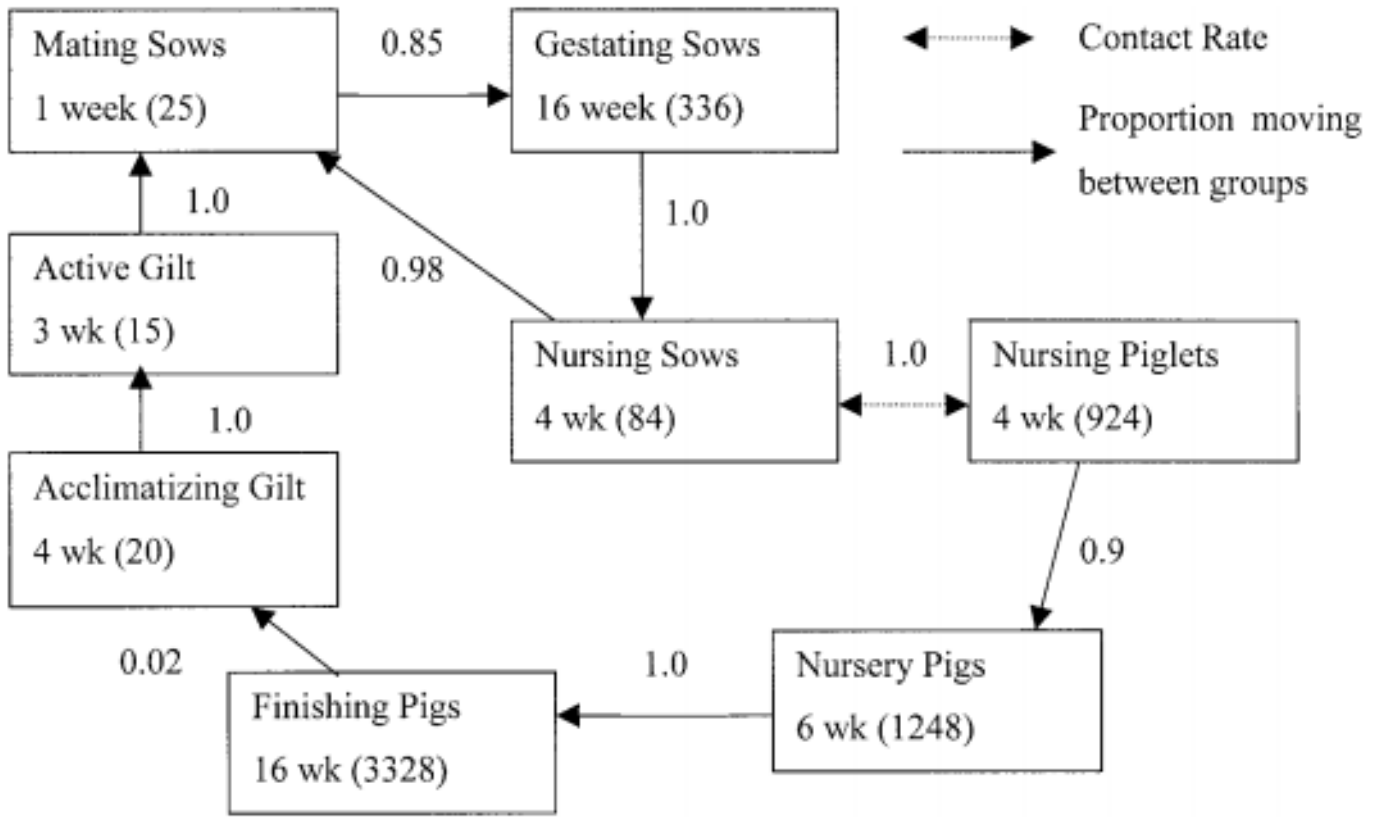


Figure 1. Pig farm structure.

Source: MacKenzie & Bishop, JAS 2001 (1)

Model outputs: Probability of a minor / major epidemic

- For fixed model parameter values, produce many realizations of an epidemic
- Define no/ minor / major epidemics
- Count proportion of realizations within each epidemic category

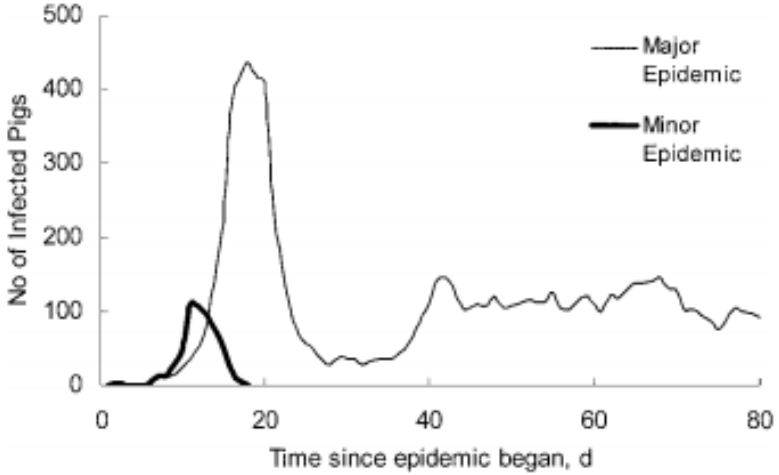


Figure 2. Example of a minor and a major epidemic ($\beta = 0.0005$, $\gamma = 0.05$).

Table 1. Probability[a] of no epidemic or a minor or major epidemic for varying transmission coefficients (β) and recovery rates (γ)

		Probability of		
β	γ	No epidemic	Minor epidemic	Major epidemic
0.0001	0.01	0.51	0.13	0.35
0.0001	0.05	0.88	0.12	0.00
0.0001	0.10	0.95	0.05	0.00
0.0005	0.01	0.15	0.01	0.85
0.0005	0.05	0.44	0.10	0.46
0.0005	0.10	0.64	0.29	0.07
0.001	0.01	0.08	0.00	0.92
0.001	0.05	0.27	0.02	0.71
0.001	0.10	0.44	0.08	0.48

[a]The failure of a given row of probabilities to sum to 1.0 is attributable to rounding error.

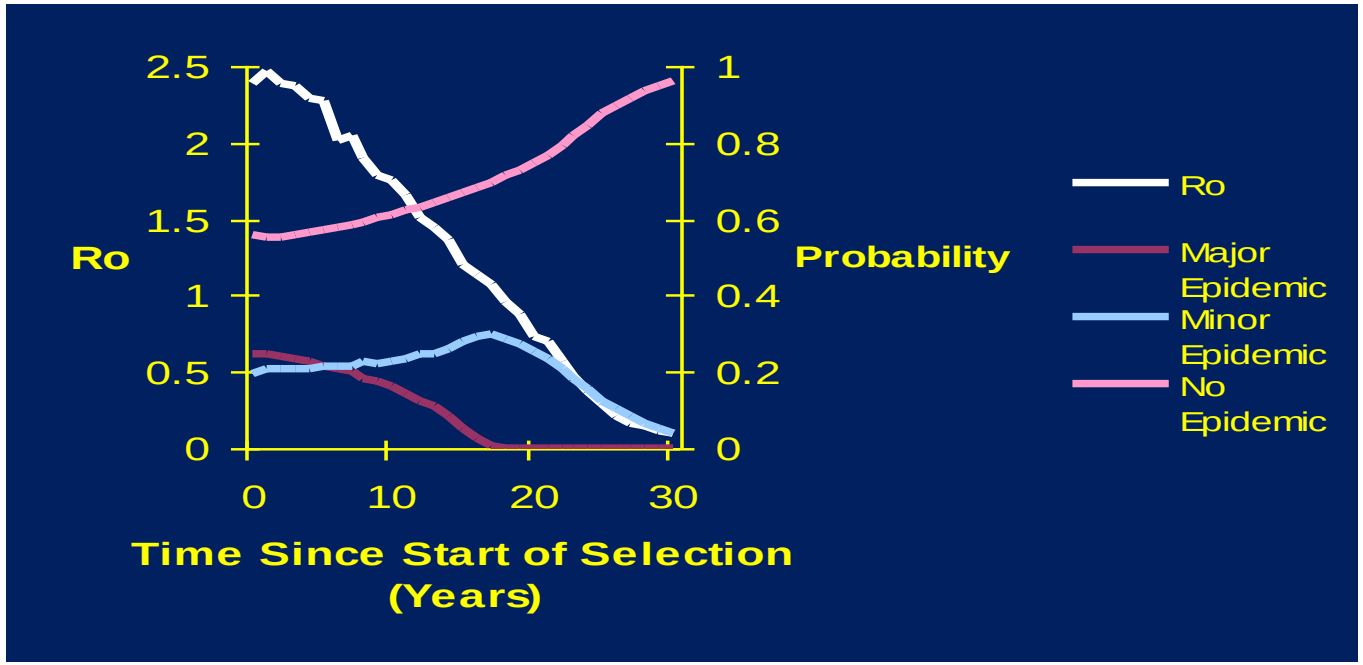
Source: MacKenzie & Bishop, JAS 2001 (1)

Use the model to predict the impact of selection

- Very simplistic implementation of genetic selection:

  Assume a genetic improvement in the transmission rate of ΔG

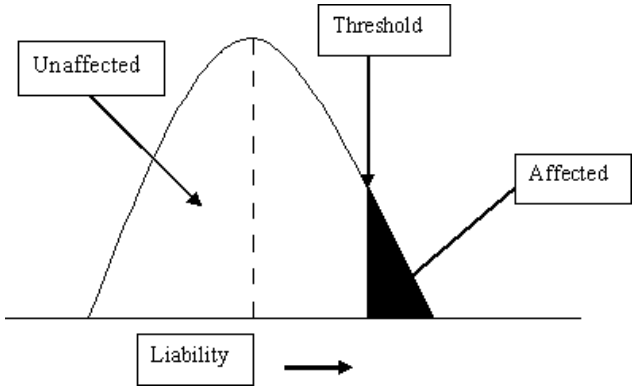
$$\beta_{year\ k+1} = \beta_{year\ k} - \Delta G \beta_{initial}$$



Note that the model does not include an explicit expression for genetic variation in resistance!

Source: MacKenzie & Bishop, JAS 2001(2)

Individual based stochastic models

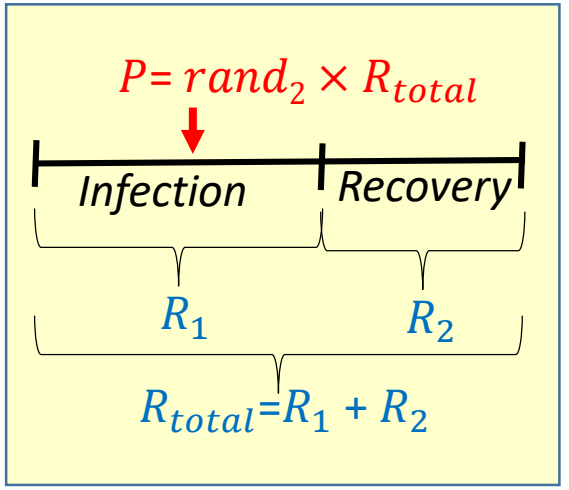


- **Model the infection status of each individual over time**
- **Allow to assign different parameter values to different individuals**
  - E.g. individuals differing in susceptibility have different transmission rates β_i
  - These can be continuous traits, i.e. drawn from distributions
  - In line with quantitative genetics concept of 'polygenic effects'
  - Can incorporate genetic structure, i.e. β_i = mean + genetic effect + environmental effect
- **Easy to implement** by straight-forward extension of event-based approach
- **Extremely flexible** (can implement many types of heterogeneity)
  - Easy to get carried away with making the model unnecessarily complex and untractable!
- **But can be very, very slow**

Implementing individual based models with Gillespie's direct algorithm

REPEAT

1. **Determine** all possible **events**: E_1, …, E_n
   - E.g. E_1 = infection, E_2 = recovery for SIR model
2. For each event, **determine the rate** at which it occurs, R_1, …, R_n
   - Event rates are now the sum of rates specific to each individual eligible for the event
3. The rate at which any event occurs is $R_{total} = \sum_{k=1}^{n} R_k$
4. Draw random number rand_1 and **determine the time until the next event**: $dt = -\frac{\log(rand1)}{R_{total}}$
5. Draw random number rand_2 and **determine the next event**: Set $P = rand_2 \times R_{total}$; event p occurs if $\sum_{k=1}^{p-1} R_k < P \leq \sum_{k=1}^{p} R_k$
6. **Determine the individual** to which the event happens
   - Use similar approach as in 5: draw randomly according to individual rate
7. **Perform event**: Update time & number of individuals in each class



An individual based stochastic genetic-epidemiological model for bovine Tuberculosis in cattle

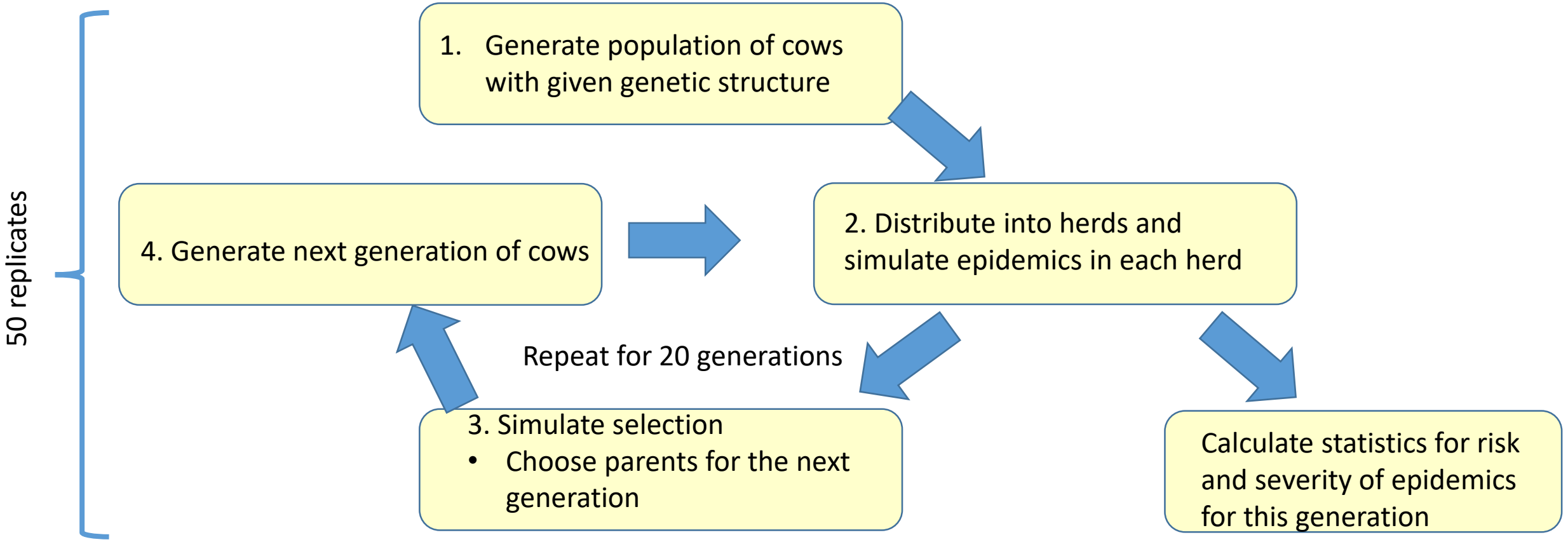
To what extent can genetic selection help eradicate bovine TB in cattle within the next 20 years?

Aim:

Develop a genetic-epidemiological model to predict the impact of selection on bTB risk & prevalence

- **Calibrate the model with field data**

Simulation process



Step 1: Generate a large population of cows with genetic variation in susceptibility

- Assume polygenic genetic variation in susceptibility (i.e. susceptibility controlled by many genes, each with very small effects)
- Apply standard genetic model for susceptibility g_i of each individual cow i:

$$\boldsymbol{g_i = mean + additive\ genetic\ effect\ A_i + environmental\ effect\ E_i}$$

Where E_i is a random drawing from the normal distribution, $E_i \sim Normal(0, \sigma_e^2)$

And genetic effects A_i are generated using

$$\boldsymbol{A_i = \frac{1}{2}(A_{sire_i} + A_{dam_i}) + Mend_i}$$

with A_{sire} / $A_{dam} \sim Normal(0, \sigma_a^2)$

and Mendelian sampling term $Mend_i \sim N(0, 0.5(1 - \bar{F})\sigma_a^2)$

Use this to assign individual susceptibility values to a population of N = 20,000 cows, which are offspring of e.g. 200 bulls, mated to 50 cows

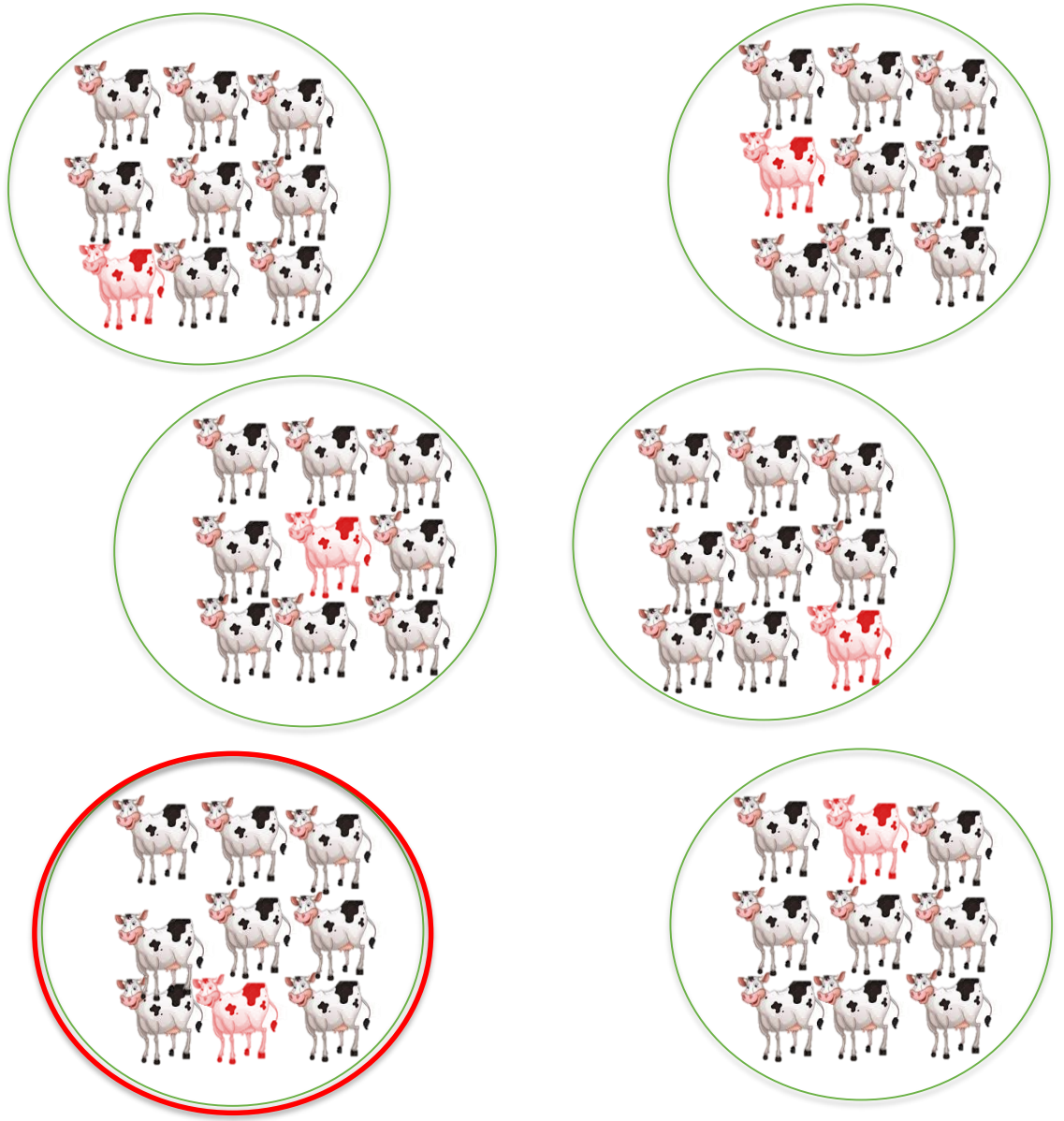
Step 2: Allocate cows into individual herds & simulate epidemics

- E.g. random allocation of individuals into herds of size 100
- Introduce bTB into each herd by one infected cow (chosen at random)

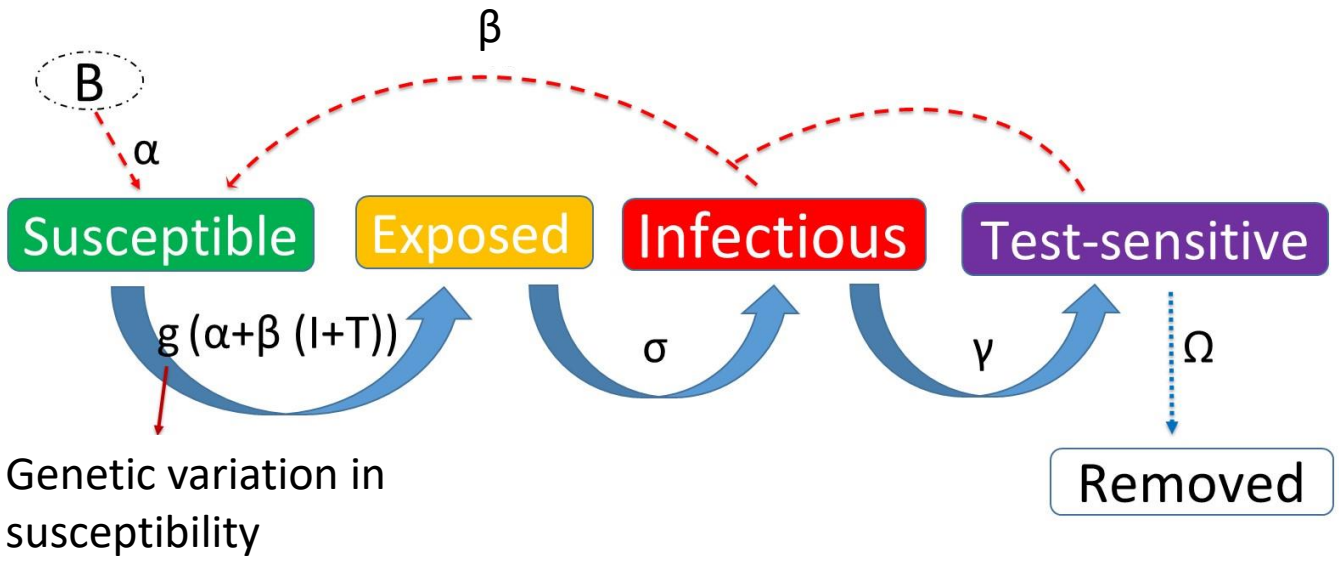
➢Simulate epidemic in each herd

Assume transmission only within herds, but not between herds

- VERY strong assumption!
- Justified as detection of one bTB case leads to herd closure)

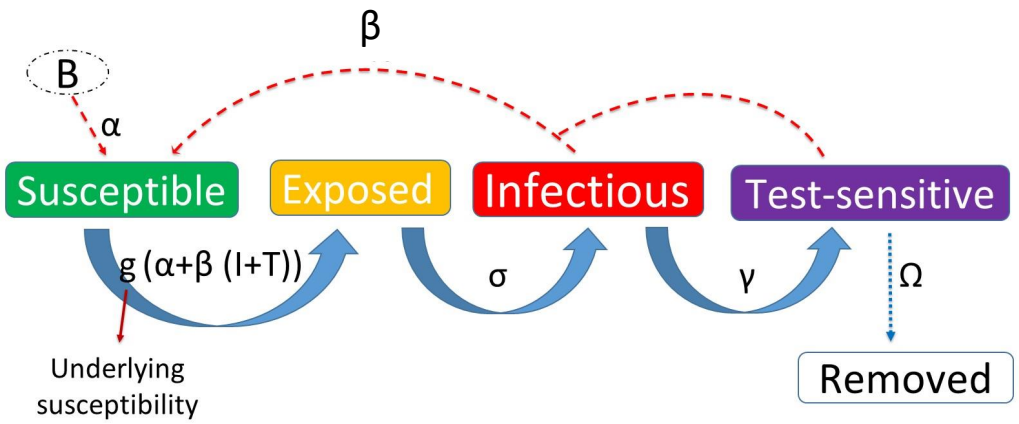


The genetic-epidemiological bTB model



- **Susceptible (S):** uninfected, but could become infected if exposed
- **Exposed (E):** infected but without clinical signs; unable to infect others
- **Infectious (I):** infected and able to transmit the infection
- **Test-sensitive (T):** possibility to be detected by a diagnostic test
- **Reactor (R):** diagnosed animal, removed from herd
- **Background infection (B):** external infection (wildlife, neighbouring cattle)

The genetic-epidemiological bTB model



Parameter	Value
β (avg. transmission rate)	0.012 (days^{-1})
α (external force of infection)	5x10^{-7}
*σ (*Rate from exposed to infectious state)	0.04 (days^{-1})
*γ (*Rate from infectious to test-sensitive state)	0.5 (days^{-1})
Ω (Test sensitivity; probability of detection & removal)	0.60
σ^2_A, σ^2_E (Genetic / environmental variance for susceptibility g)	0.3

Input values obtained from literature and from calibrating model predictions to field data

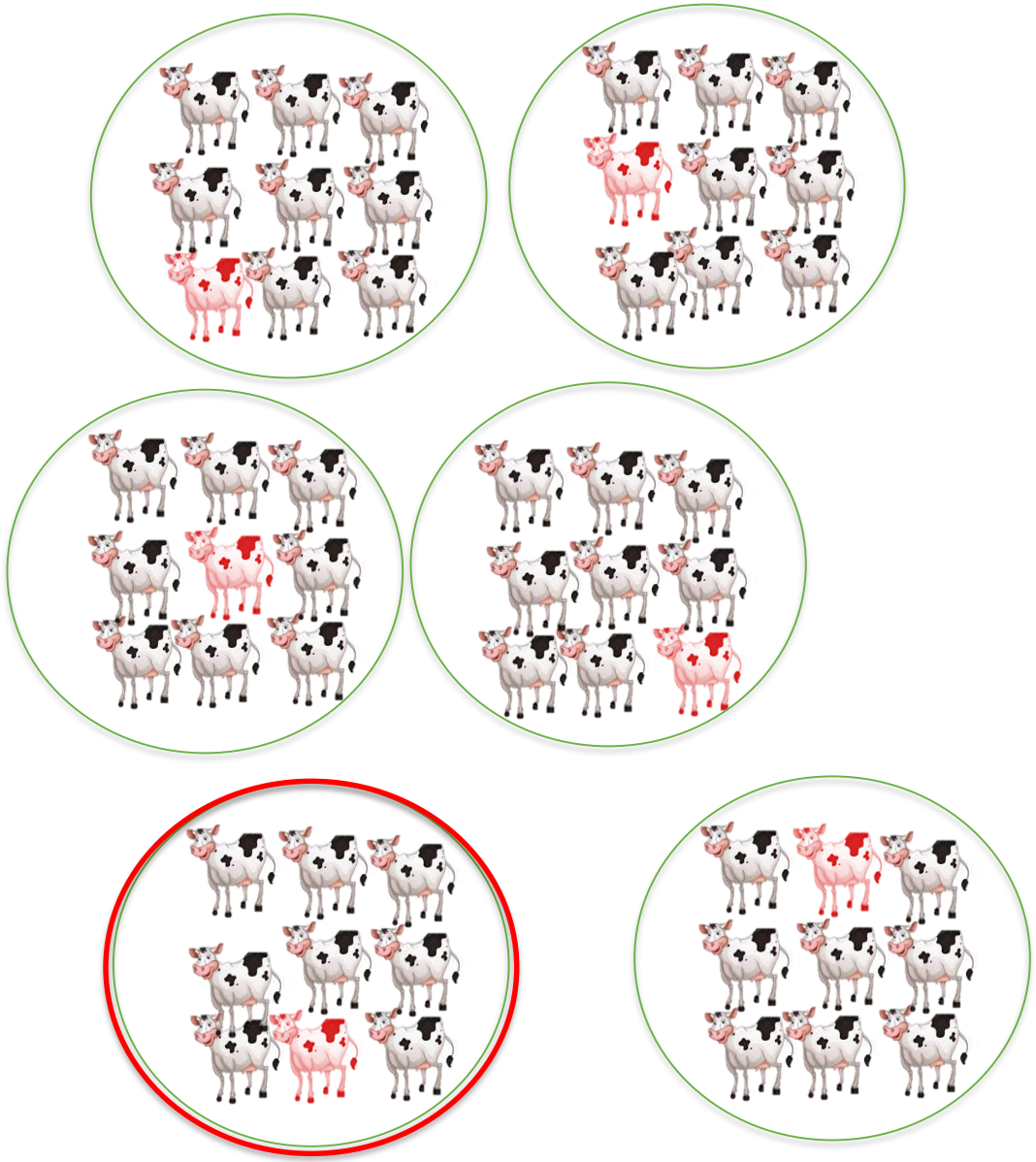
Simulation process

Apply stochastic model with Gillespie Algorithm

2 types of results:

1. Infection status of every individual cow over time

2. Summary statistics (epidemic characteristics):
   - Risk of bTB spread per herd
   - Mean percentage of reactors per herd
   - Mean duration of epidemic
   - Genetic parameters of susceptibility to bTB in the observed scale



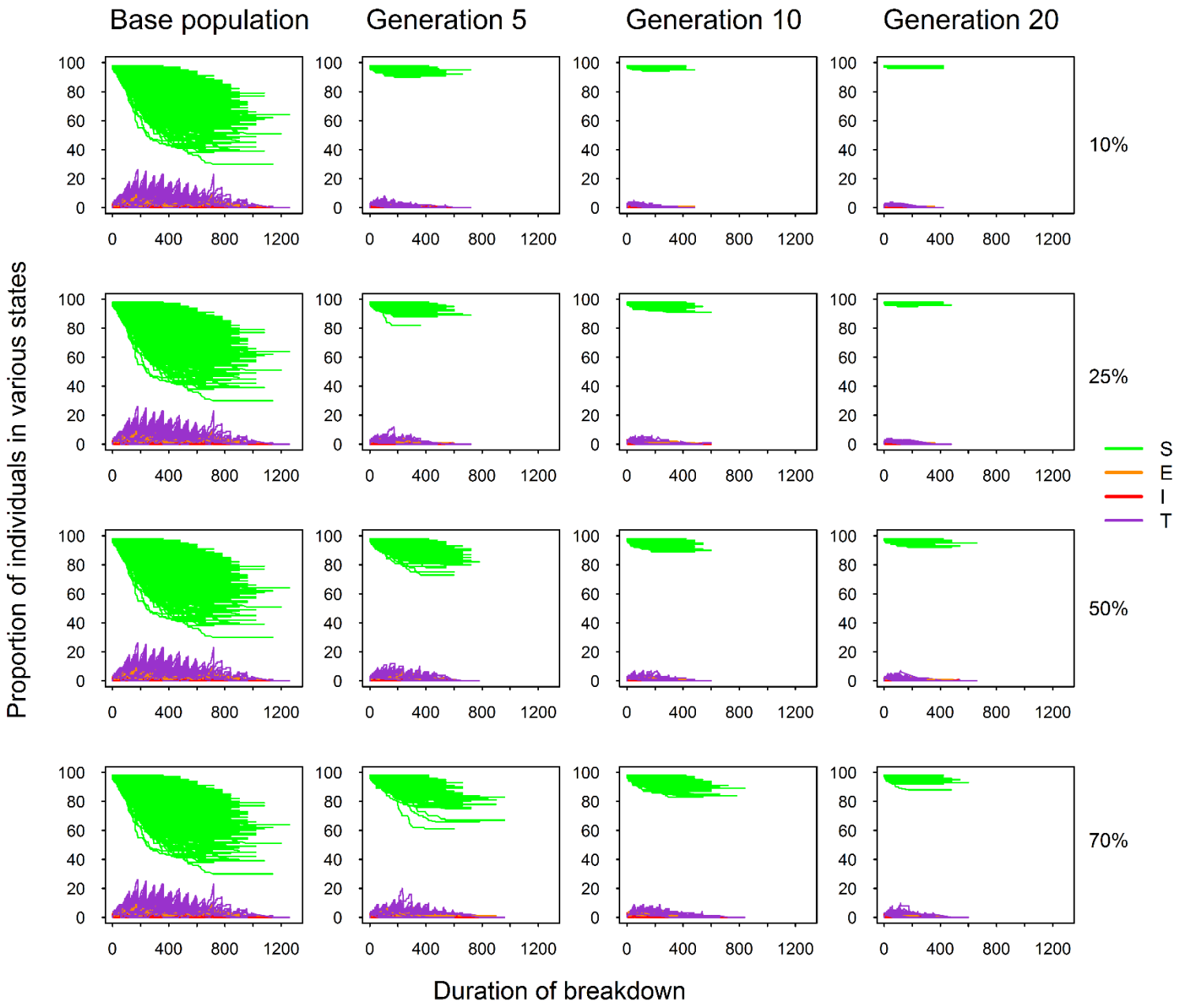
Step 3: Simulate selection

- **Simulate a selection scheme at the population (not herd) level:**
- **Identify genetically best individuals based on simulated binary disease phenotype (reactor / non-reactor) from epidemiological model**
  - **Apply genetic evaluation to simulated data:**
  - **→ Estimate Breeding value (EBV) for susceptibility for each individual sire**
  - **Select the best x% of individuals as parent for the next generation**
    - **Various selection intensities explored (best 10-70% of sires)**

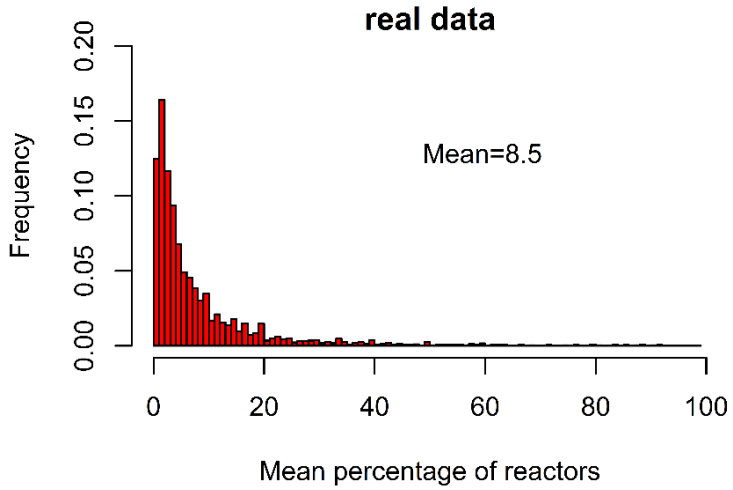
Model results:

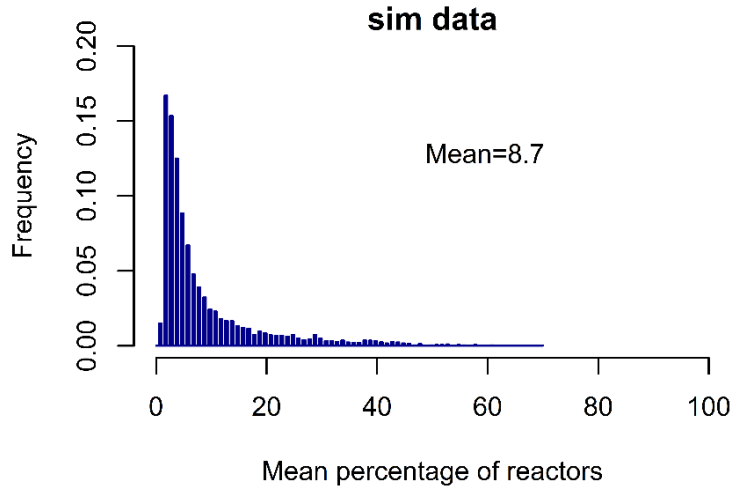
1. Ensemble of epidemic profiles per generation

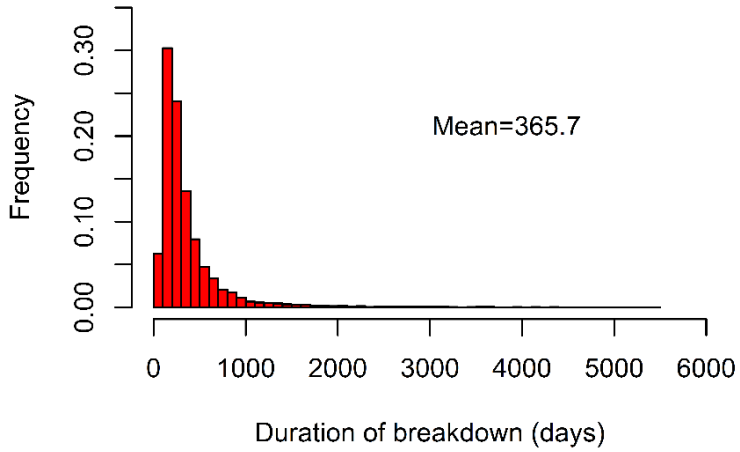
--> useful for qualitative assessment

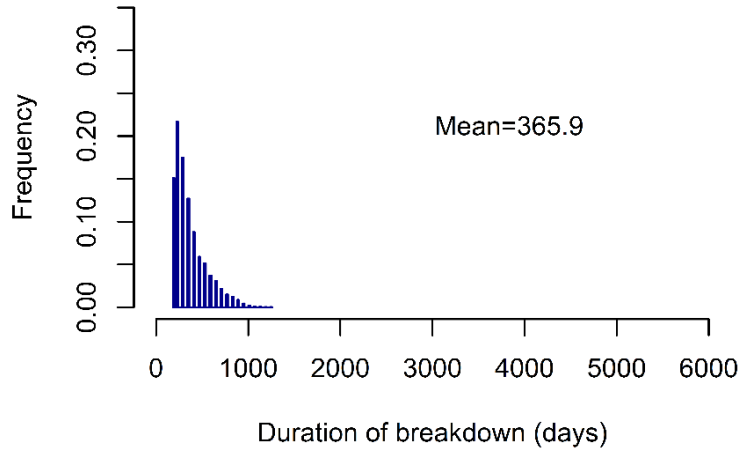


Model results 2:
Summary statistics for comparison with real data





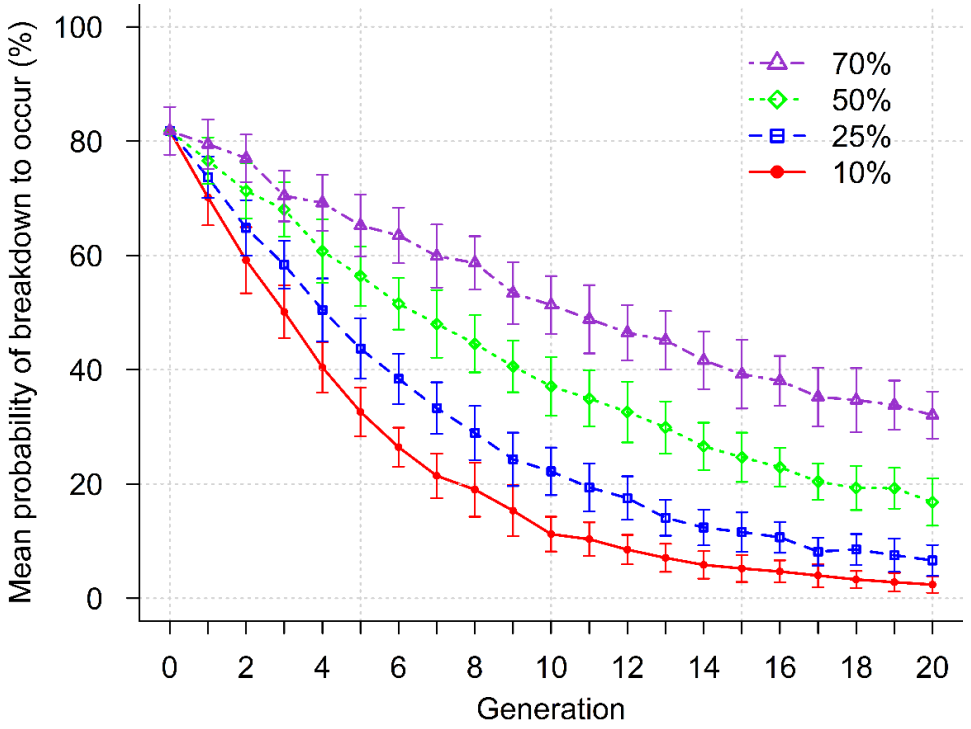




→ The model produces some of the characteristics for bTB

→ Important to show for credibility

Model results 3: Predictions for impact of selection



Different colours refer to different selection intensities

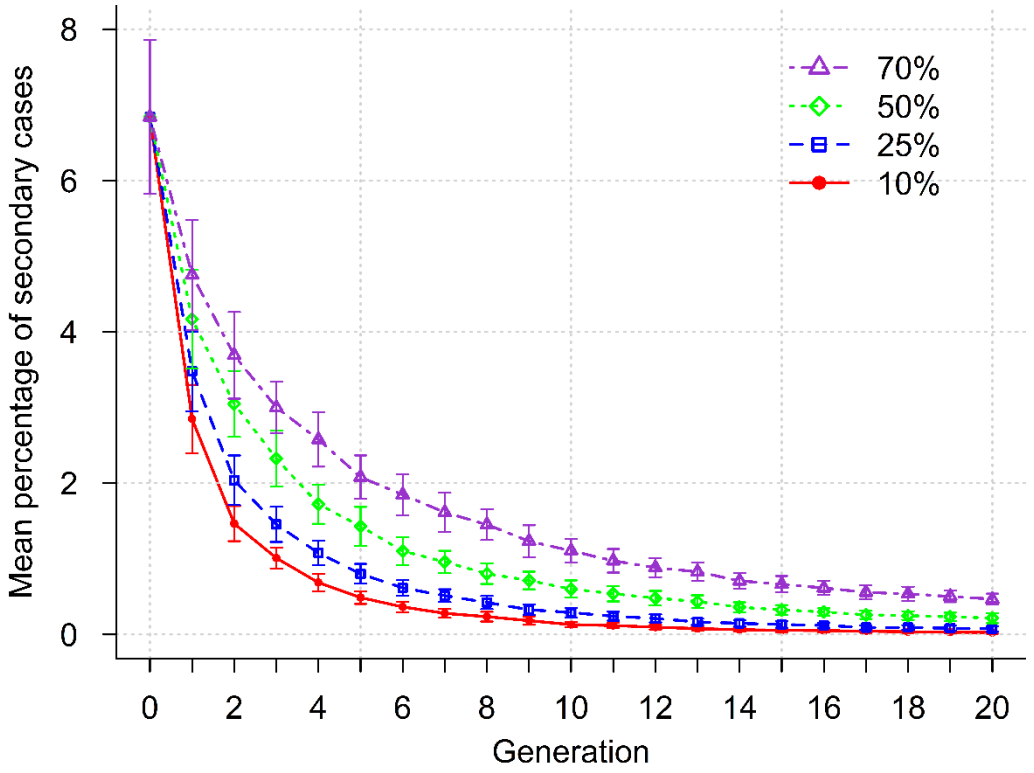
Risk of epidemic

- Before selection = 81.8%
- Reduced by half
  - 4 generations

    (select best 10% of sires)
  - 15 generation

    (select best 70% of sires)

Model results 3: Predictions for impact of selection



Different colours refer to different selection intensities

Severity of epidemic

- Before selection = 6.8% secondary cases
- On average 1% of secondary cases
  - 4 generations (best 10% of sires)
  - 11 generations (best 70% of sires)
- On average 0% of secondary cases
  - 12 generations (best 10% of sires)
  - >>20 generations (best 70% of sires)

bTB model conclusions

- **Breeding for bTB resistance can substantially reduce the risk and severity of bTB epidemics**
- **Selection benefits arise within 5-15 generations**
- **Selection is a viable complementary long-term strategy to existing control measures**

Kethusigele et al., 2018

Summary

- Stochasticity is an intrinsic property of infectious disease
- Stochasticity matters most when nr. of infected is small (prob. of extinction)
- Stochasticity can be implemented in various ways
  - For large populations, where the individual nature is unimportant, stochasticity can be mimicked by adding noise to the differential equations
  - Event driven approaches more flexible and powerful, but less elegant
- Stochastic models are comparable to deterministic models
- Event driven approaches can be easily extended to individual based models
- Stochastic models can be a powerful tool for assessing the impact of diverse disease control strategies (under different environmental conditions)

Some final remarks: The bigger picture – what he have not covered here

- Most of the methods shown here apply to microparasitic infections
- Modelling macroparasites requires modelling the life cycle of the parasite as well as transmission dynamics (e.g. Laurenson et al., 2012; Gharbi et al. 2015)
- There are other approaches for modelling epidemics:
  - **Branching processes:** Markov process that models a population in which each individual in generation n produces some random number of individuals in the next generation according to a probability distribution (e.g. Lloyd-Smith et al., 2005 used this to show that the superspreading is a common phenomenon)
  - **Agent based models:** Computational model for simulating (inter) actions of autonomous agents with a view to assessing their effect on the system as a whole (see e.g. Perez & Dragicevic 2009)
  - **Network models:** Implement network theory; useful to model spatial structure and heterogeneity (see e.g. Danone et al. 2011)

Where next?

We need to link the model world with the real world

Statistical inference

See lectures on Thursday

- Good understanding of underlying factors for epidemiological characteristics
- Predictions for effect of control strategies on epidemics
- But simplified problem

- Real observations
- Real problem

References

- Keeling, Matt J., and Pejman Rohani. *Modeling infectious diseases in humans and animals*. Princeton University Press, 2008.
- Allen, Edward J., et al. "Construction of equivalent stochastic differential equation models." *Stochastic analysis and applications* 26.2 (2008): 274-297.
- Gillespie, Daniel T. "Exact stochastic simulation of coupled chemical reactions." *The journal of physical chemistry* 81.25 (1977): 2340-2361.
- MacKenzie, K., and S. C. Bishop. "Utilizing stochastic genetic epidemiological models to quantify the impact of selection for resistance to infectious diseases in domestic livestock." *Journal of animal science* 79.8 (2001): 2057-2065.
- MacKenzie, K., and S. C. Bishop. "Developing stochastic epidemiological models to quantify the dynamics of infectious diseases in domestic livestock." *Journal of animal science* 79.8 (2001): 2047-2056.
- Nath, M., J. A. Woolliams, and S. C. Bishop. "Identifying critical parameters in the dynamics and control of microparasite infection using a stochastic epidemiological model." *Journal of animal science* 82.2 (2004): 384-396.
- Nath, M., J. A. Woolliams, and S. C. Bishop. "Assessment of the dynamics of microparasite infections in genetically homogeneous and heterogeneous populations using a stochastic epidemic model." *Journal of animal science*86.8 (2008): 1747-1757.
- Doeschl-Wilson, Andrea B., et al. "Implications of host genetic variation on the risk and prevalence of infectious diseases transmitted through the environment." *Genetics* 188.3 (2011): 683-693.
- Laurenson, Yan Christian Stephen Mountfort, et al. "Exploration of the epidemiological consequences of resistance to gastro-intestinal parasitism and grazing management of sheep through a mathematical model."*Veterinary parasitology* 189.2 (2012): 238-249.
- Gharbi, Karim, et al. "The control of sea lice in Atlantic salmon by selective breeding." *Journal of The Royal Society Interface* 12.110 (2015): 20150574.
- Lloyd-Smith, James O., et al. "Superspreading and the effect of individual variation on disease emergence." *Nature* 438.7066 (2005): 355-359.
- Keeling, Matt J., and Bryan T. Grenfell. "Individual-based perspectives on R 0." *Journal of Theoretical Biology* 203.1 (2000): 51-61.
- Perez, Liliana, and Suzana Dragicevic. "An agent-based approach for modeling dynamics of contagious disease spread." *International journal of health geographics* 8.1 (2009): 1.
- Danon, Leon, et al. "Networks and the epidemiology of infectious disease."*Interdisciplinary perspectives on infectious diseases* 2011 (2011).