Analysis of risk factors using life history
a partial membership approach to estimating the risk of liver fluke infestation.

G. T. Innocent L. Gilbert E. Jones G. Gunn G. Marion
I. J. McKendrick

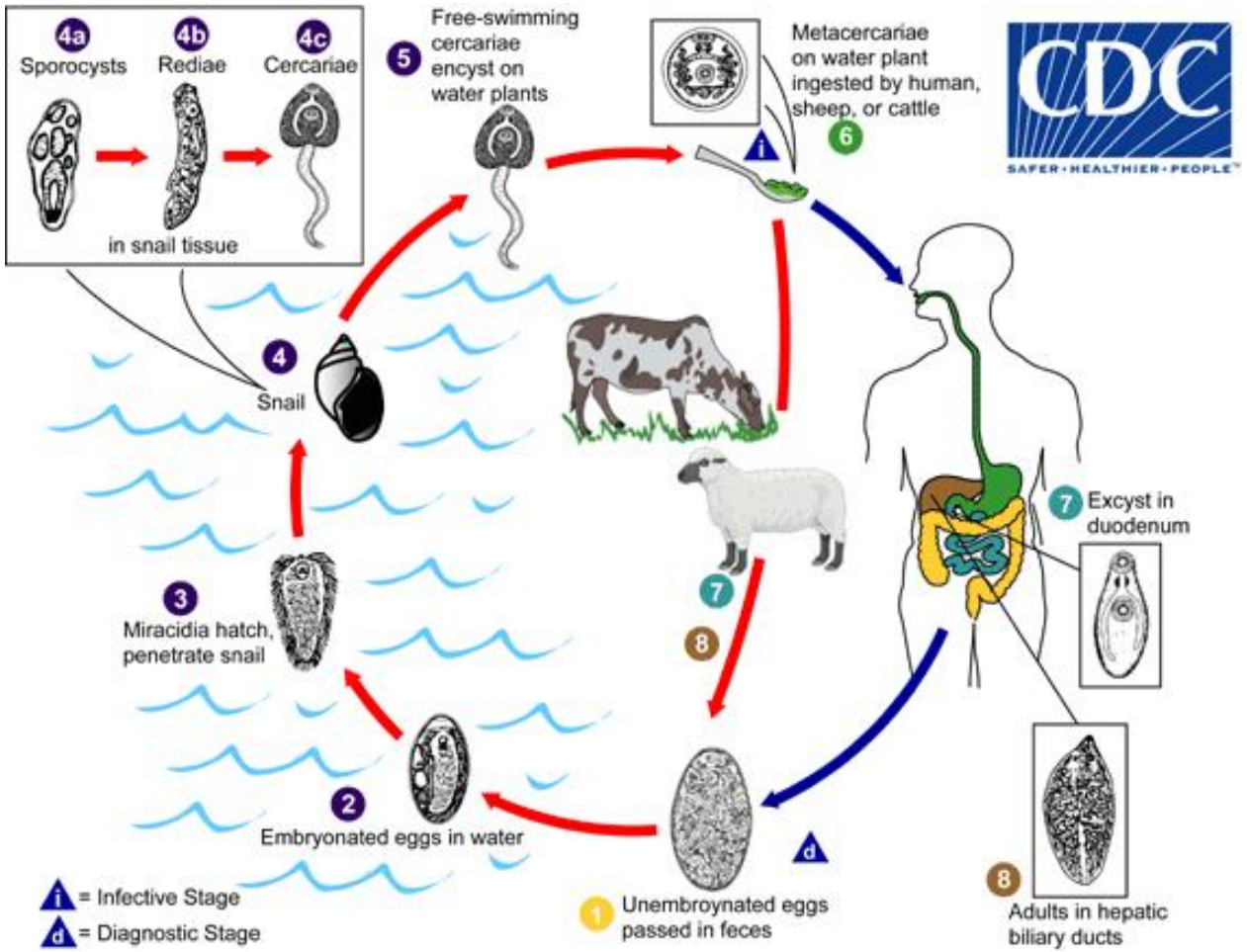
Biomathematics and Statistics Scotland
http://www.bioss.ac.uk

ScoVa16 Valencia Jan 2016

Background - liver fluke

- Serious disease of cattle and sheep throughout the world
- *Fasciola hepatica* requires damp and warm conditions
- Around 25% of cattle livers are condemned in Scottish slaughterhouses
- Imposes welfare costs on the animal and economic costs on the farmer

Life cycle


4a
Sporocysts
4b
Rediae
4c
Cercariae
in snail tissue
5
Free-swimming cercariae encyst on water plants
Metacercariae on water plant ingested by human, sheep, or cattle
6
i
CDC
SAFER · HEALTHIER · PEOPLE
4
Snail
7
Excyst in duodenum
3
Miracidia hatch, penetrate snail
7
8
2
Embryonated eggs in water
d
1
Unembroynated eggs passed in feces
8
Adults in hepatic biliary ducts
i = Infective Stage
d = Diagnostic Stage

The model

- Each animal ingests metacercariae at a (different) constant rate
- Rate depends on local environment (farm)
- Other factors may be important (age, breed etc)
- We observe absence (0) or presence (≥ 1)
- We model numbers of liver fluke as a Poisson distribution
- Identical model if we model the time to first infestation as exponential distribution

Mathematical Representation of Model

- Presence of liver fluke in animal i represents a Bernoulli trial with probability of success $p_i = 1 - e^{-\lambda_i}$
- λ_i represents the expected number of fluke in animal i
- Depends on many animal- and farm-based factors
- However, animals move from farm to farm
- Must consider the length of time that an animal has spent on each farm
- $log(RR_{ik}) = time_{ik} \times log(RR_k)$

Mathematical Representation of Model

$$log(-log(1-p_i)) = \sum^{j} \beta_j x_{ij} + \sum^{k} (t_{ik}\gamma_k + \sum^{l} t_{ik}\delta_l x_{kl})$$

- $\beta_j x_{ij}$ represents the fixed, within-animal effects
- $t_{ik}\gamma_k$ represents the random farm effects
- $t_{ik}\delta_l x_{kl}$ represents the fixed within-farm effects
- t_{ik} is the *at risk* time spent by animal i on farm k
- Fixed effect could be fitted in a standard linear model
- Farm random effects – an individual may be a member of more than one random effect group
- Cannot be easily formulated in standard statistics packages – INLA provides a mechanism

Data

- Outcome – condemnation of animal's liver – binary from abatoir
- Animal history from the Cattle Tracing System database
- Environment data
- Predictors
  - Animal specific – Age, Sex, Breed
  - Farm specific – Position (Lat, Long), Mean temperature, Mean rainfall
  - Animal history – Every farm on which an animal has lived
  - Farm-animal interplay – Number of "at risk" days spent on farm

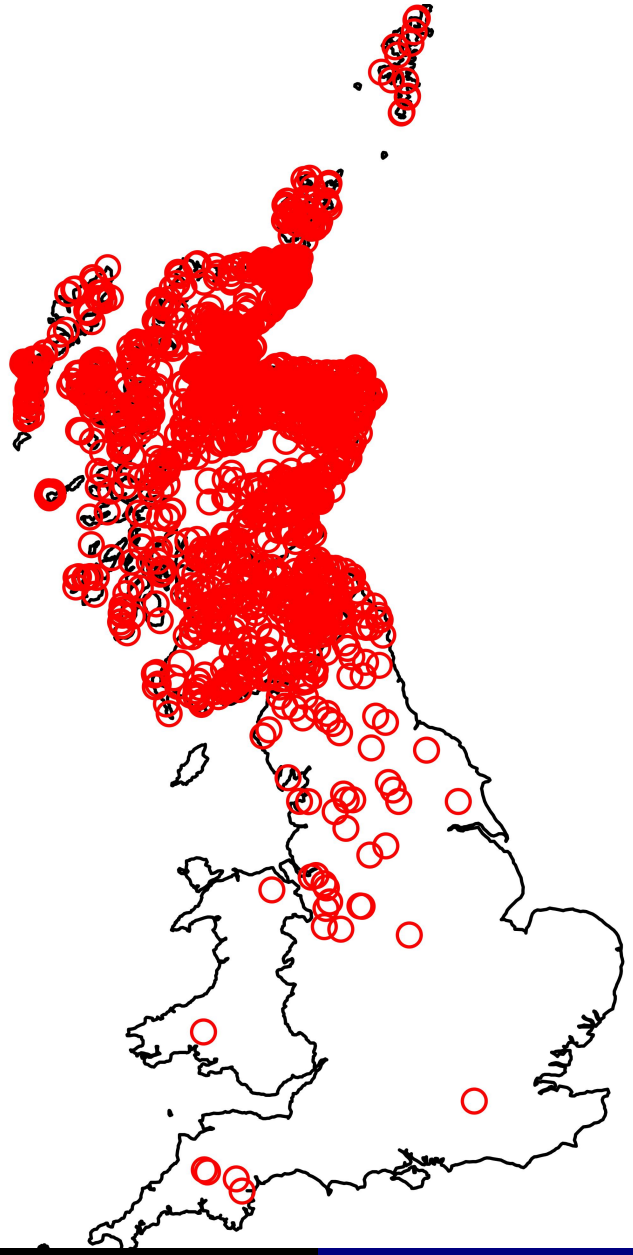
Data description

- Total of 2925 animals: 1075 farms
- 337 animal stayed on a farm less than 50 days, 43 less than 10 days
- 1185 animals stayed on only 1 farm, none stayed on more than 5 farms
- 524 farms are only represented by one animal
- Only 82 farms were represented by more than 10 animals
- I.E. sparse data

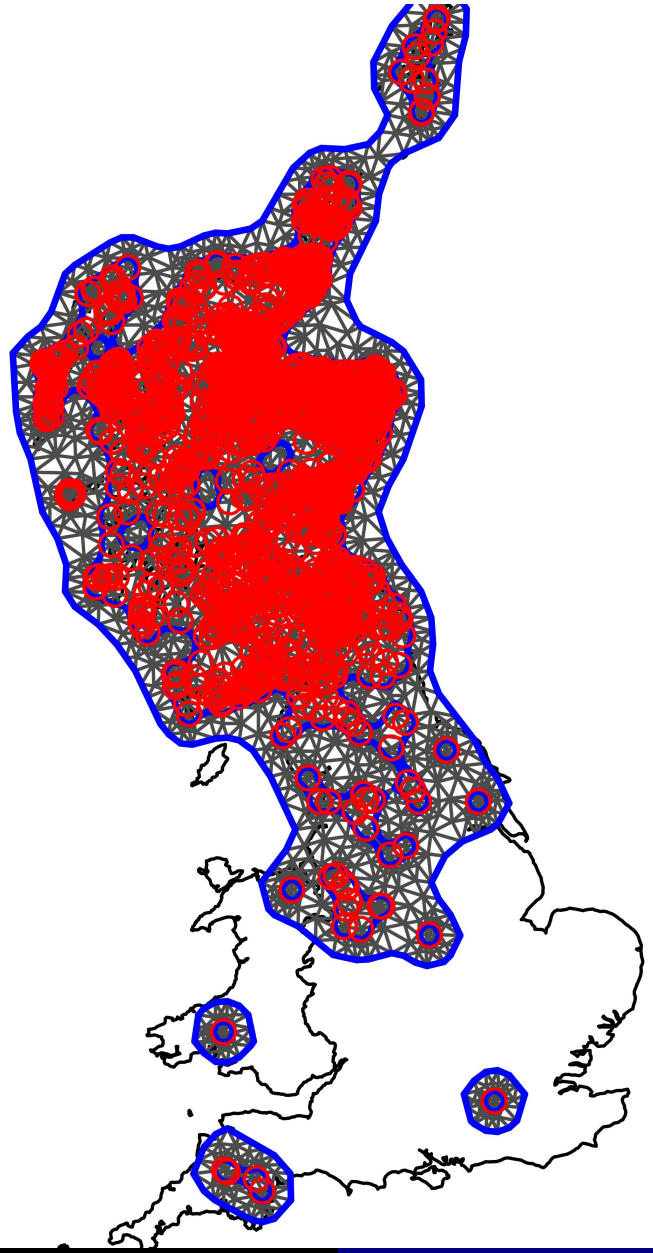
Random effects model - farm effects

- We might expect farms close in space to be similar in risk
- Partial answer: farm co-ordinates as fixed effects
- This allows for clines, but not local effects
- We use stochastic partial diferential equation (spde)
- We also allow for a non-spatial farm random effect
- Fitted using INLA
- As animals stay on up to 5 farms need to repeat random effect using “copy”
- Need to “copy” both spatial and non-spatial random effect

Position of farms in data



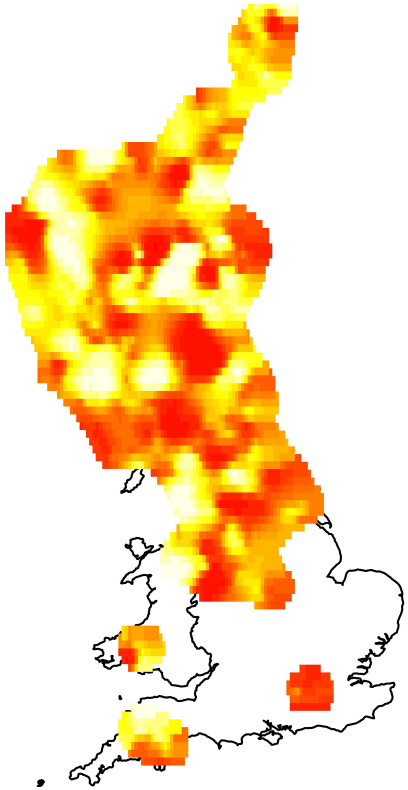
Mesh used for spde model



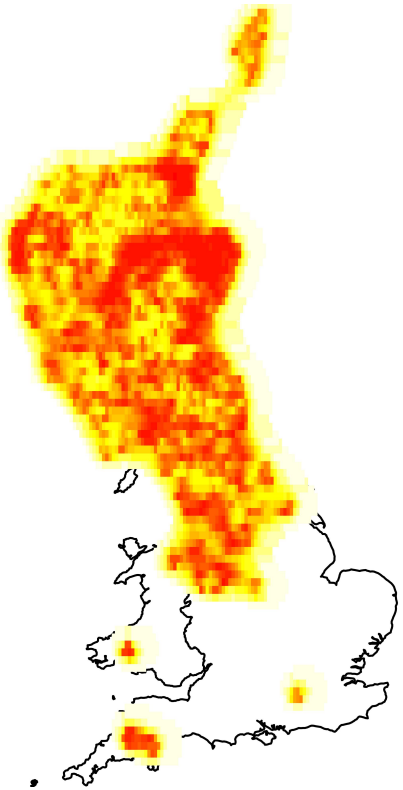
Fixed effects model

Parameter	Coefficient	95% CI
age	0.0026	0.0017 – 0.0035
east	-0.0293	-0.0360 – -0.0228
north	0.0002	-0.0028 – 0.0032
rain	0.0011	0.0008 – 0.0015
temp	0.0003	0.0001 – 0.0006

Spatial random effect

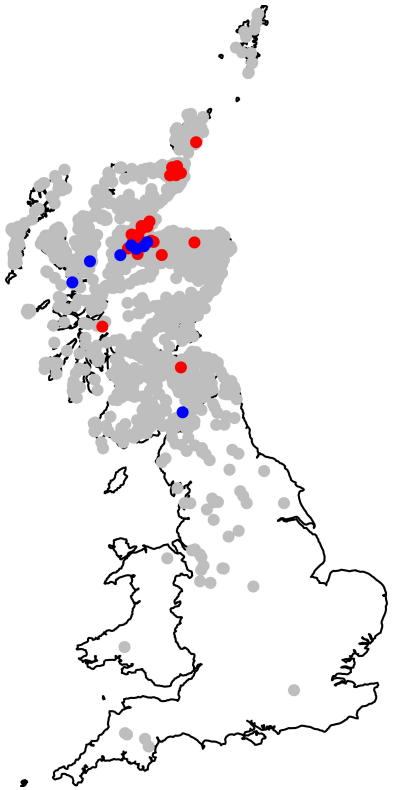


Posterior mean



Posterior SD

Non-spatial random effect



Conclusions

- Increase risk with age *after inclusion of times on farm*
- Cline East → West, but not North → South
- Increased risk with rainfall and temperature
- No evidence of problem areas
- Residual farm effects important

Discussion

- Even in INLA not a quick analysis
- Increases in complexity the further back in an animal's history you go
- On my (old) machine couldn't use a more regular mesh, not based on the observed farms as ran out of memory

Other relevant work

- Design of a field trial for a bovine TB DIVA test
- Analyis of S-I-S data

Design of a field trial for a bovine TB DIVA test

- 2012-13 England cost £25m
- Defra/Welsh Government commissioned project
- Develop protocols to validate vaccine and DIVA test
- Determine the likely level of buy-in by farming community
- Determine costs and benefits of implementation
- BioSS involvement: DIVA test
- Distinguishes between Infected and Vaccinated Animals

The Consortium Partners

- HEIs
  - University of Cambridge
  - University of Aberystwyth
- BioSS
- Veterinary advisers
- CRO: Triveritas Ltd.

Our work

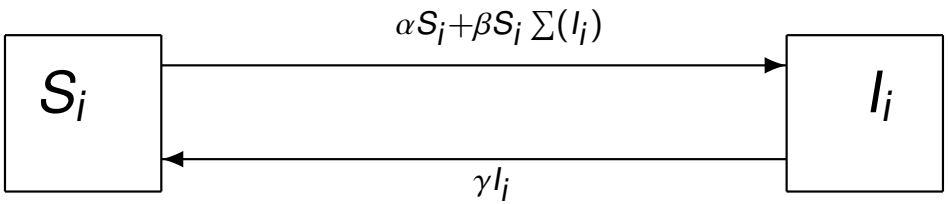
- Examine power to assess the properties of DIVA test under different designs
- “Gold standard” exists - post mortem examination and culture
- Our proposal – two-tier approach
  - Use “Gold standard” approach to get a reasonably accurate estimate of sensitivity and specificity
  - Latent class analysis to further refine accuracy of our estimates
- Target sufficient precision to make vaccination/DIVA no more costly that test-and-cull
- Effectively specificity ≥ 0.9985 I.E. no more than 15 false positives out of 10,000 negative animals

Modelling S-I-S data

- S-I-S models complex
- Uncertain number of transitions between states
- Observations prone to error (Se/Sp)
- Cross-sectional studies can estimate transmission parameters
- Identifiability problems when we try to estimate transmission and test parameters
- Chapman-Kolmogorov equations

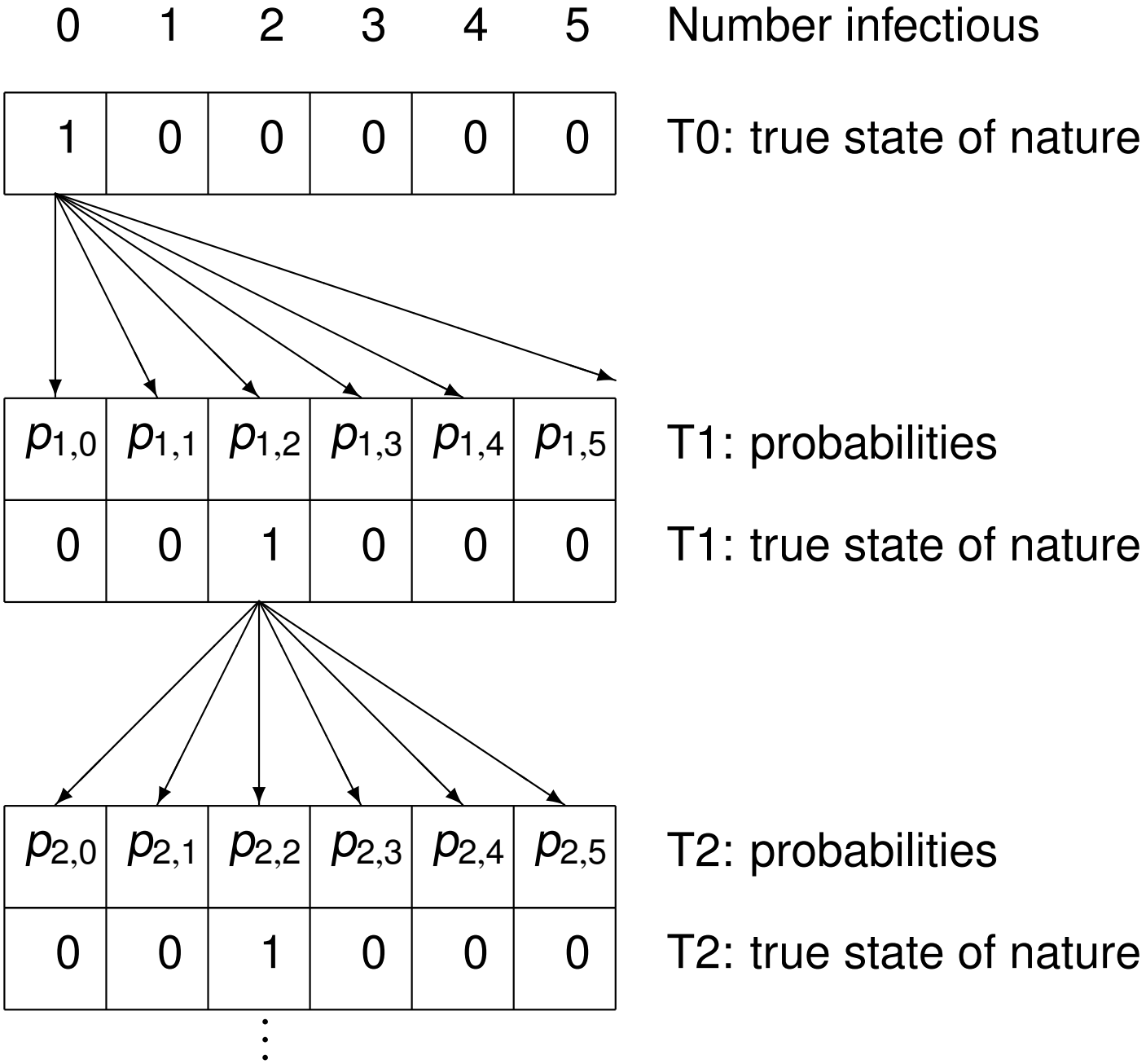
Model

α constant environmental/external infection rate
β transmission infection rate
γ recovery rate



Model

$$\begin{aligned}\frac{dP(I=m)}{dt} =& (\alpha(N-m+1)+\beta(m-1)(N-m+1))P(I=m-1) \\ & -\gamma m P(I=m) \\ & -(\alpha(N-m)+\beta m(N-m))P(I=m) \\ & +\gamma(m+1)P(I=m+1)\end{aligned}$$


0 1 2 3 4 5
Number infectious
1 0 0 0 0 0
T0: true state of nature
$p_{1,0}$ $p_{1,1}$ $p_{1,2}$ $p_{1,3}$ $p_{1,4}$ $p_{1,5}$
T1: probabilities
0 0 1 0 0 0
T1: true state of nature
$p_{2,0}$ $p_{2,1}$ $p_{2,2}$ $p_{2,3}$ $p_{2,4}$ $p_{2,5}$
T2: probabilities
0 0 1 0 0 0
T2: true state of nature