

Description of the STOC free model

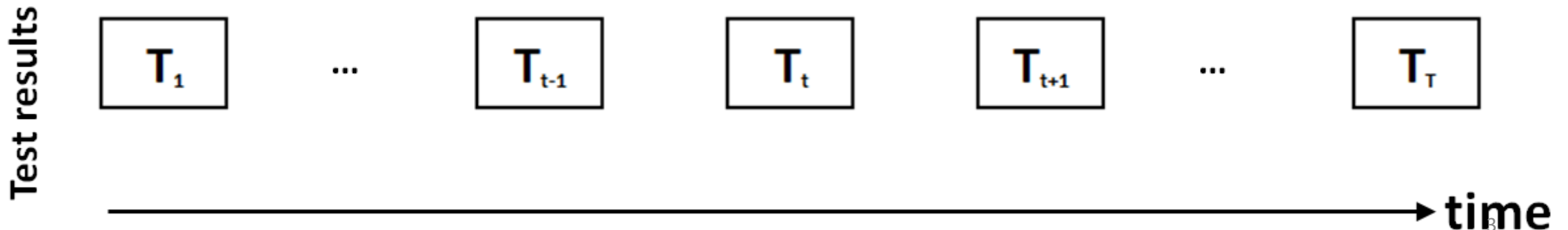
Aurélien Madouasse

11/11/2020

Surveillance programmes against cattle endemic infectious diseases

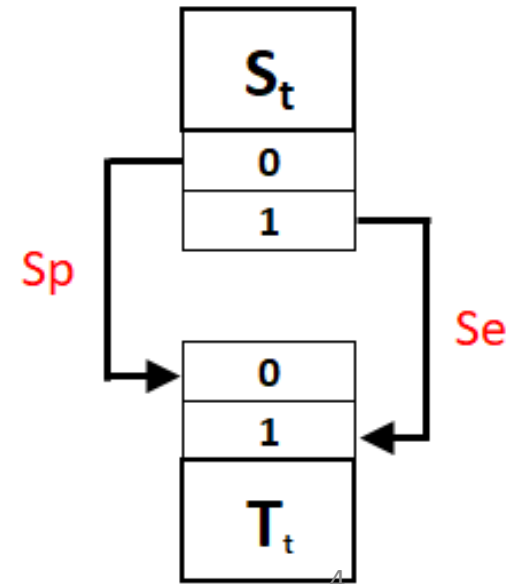
Common features among surveillance programmes

- Infection prevalence > 0
 - Objective = to control or to eradicate infection
 - Surveillance programmes designed to detect infected herds or animals
 - Control phase to eliminate infected animals or to prevent them to be sold
- Herds tested repeatedly over time
 - Longitudinal data



Common features among surveillance programmes

- Infection prevalence > 0
 - Objective = to control or to eradicate infection
 - Surveillance programmes designed to detect infected herds or animals
 - Control phase to eliminate infected animals or to prevent them to be sold
- Herds tested repeatedly over time
 - Longitudinal data
- Tests are imperfect
 - Sensitivity (Se) < 1 and specificity (Sp) < 1
 - Uncertainty in the true status (S_t) of the tested animals / herds



Modelling framework

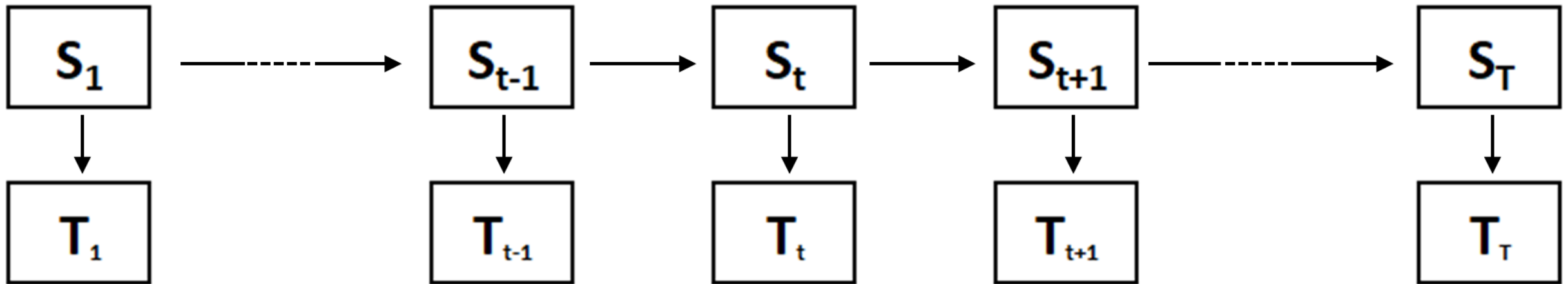
Objective

- Estimate herd level probabilities of (freedom from) infection from longitudinal test data collected as part of surveillance programmes against endemic infectious diseases of cattle
- The modelling framework should:
 - Account for imperfect test information
 - Allow the use of longitudinal data i.e. account for the fact that sequences of test results are not completely random

Hidden Markov Models (HMMs)

- HMMs model a **latent categorical** variable with a **Markovian dynamics**, whose state at a given time determines the distribution of an observed variable
 - **categorical**: the variable of interest can be in 1 of k states. In the STOC free model, $k=2$ (positive or negative status)
 - **latent**: this categorical variable is not directly observed. In the STOC free model, the latent status determines the probability of a negative or positive test result through sensitivity and specificity
 - **Markovian dynamics**: the latent status is modelled in discrete time steps. The latent status at time t only depends on the status at time $t-1$. The probabilities of transition between the k states between 2 time points are described by a $k \times k$ transition matrix.

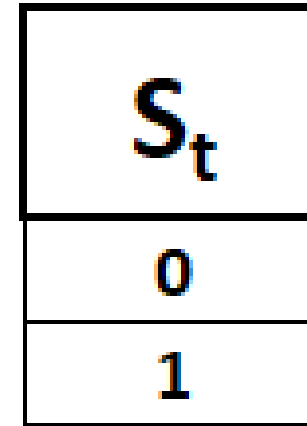
Representation of surveillance programmes as HMMs



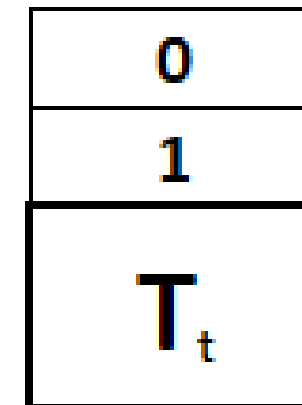
- S_t : latent status of interest at time t , from $t=1$ to $t=T$
- T_t : test result at time t

Representation of surveillance programmes as HMMs

- Herd status at time t is either negative (0) **OR** positive (1)



- Test result at time t is either negative (0) **OR** positive (1)



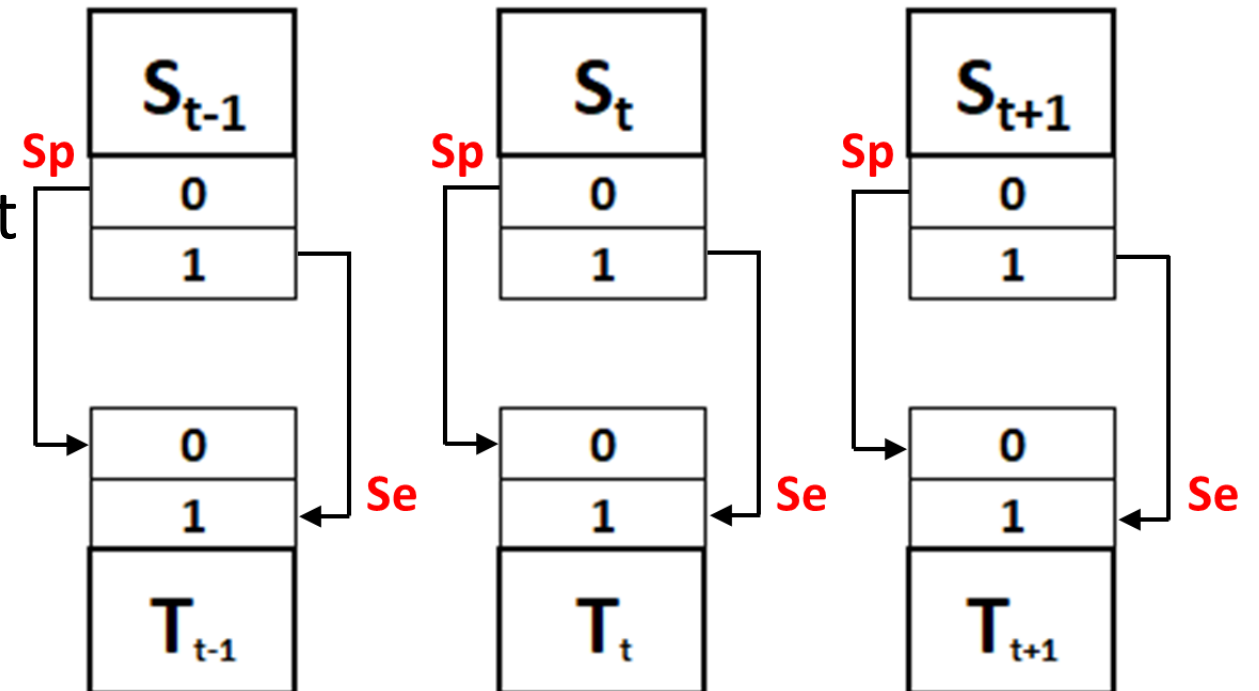
Representation of surveillance programmes as HMMs

When test results at the herd level

- Test results depend on the latent status through sensitivity and specificity

$$Se = p(T=1 \mid S = 1)$$

$$Sp = p(T=0 \mid S = 0)$$



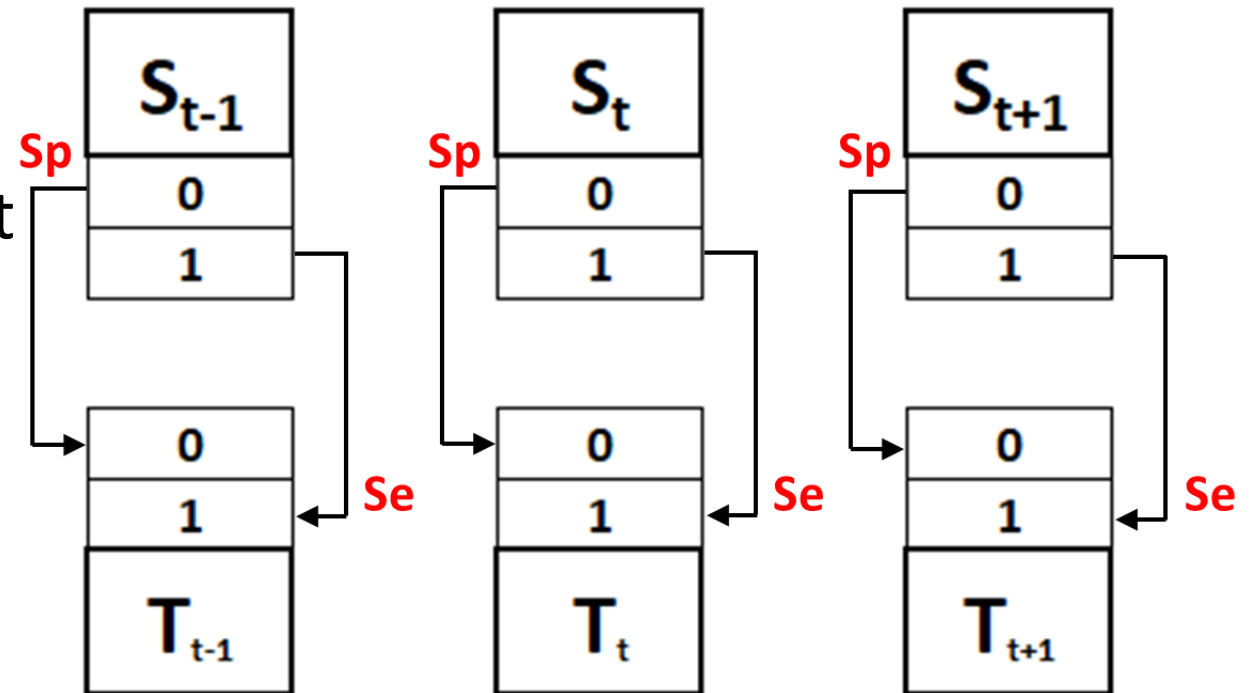
Representation of surveillance programmes as HMMs

When test results at the herd level

- Test results depend on the latent status through sensitivity and specificity

$$T_t \sim \text{Bernoulli}(p(T_t^+))$$

$$p(T_t^+) = S_t Se + (1 - S_t)(1 - Sp)$$



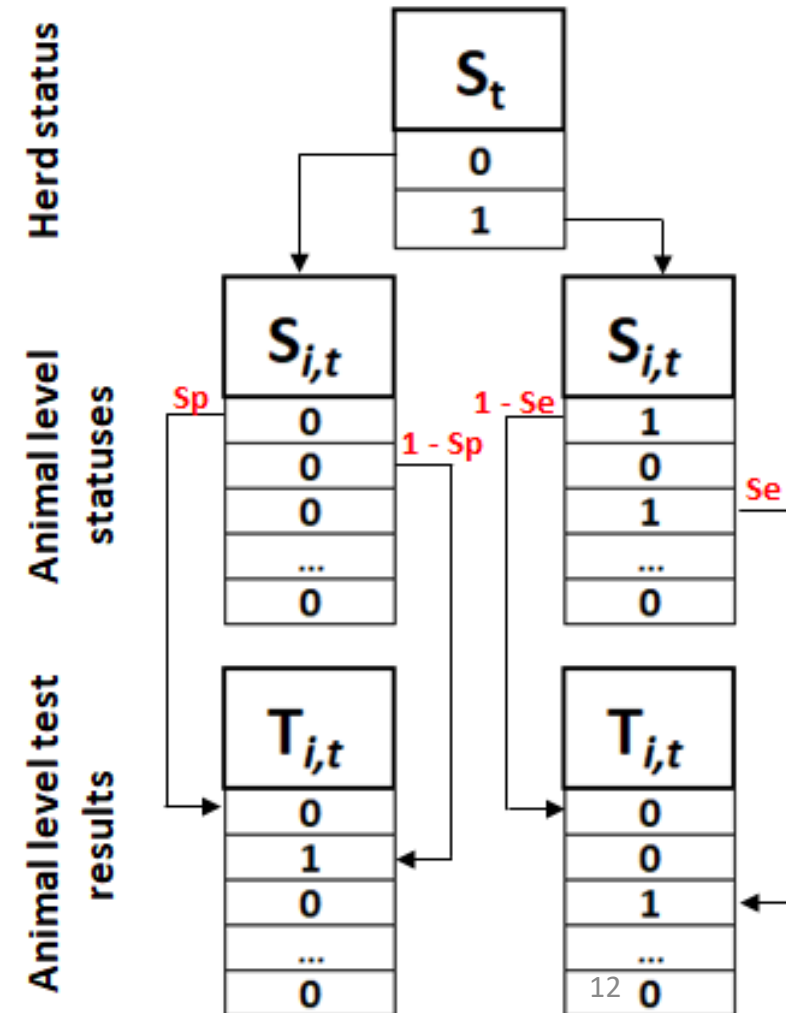
Representation of surveillance programmes as HMMs

When test results at the animal level

- Status defined at both herd and animal levels
 - Status negative herds have an animal prevalence of 0
 - Status positive herds have an animal level prevalence of π_{within}

In the Figure, the 2 herd statuses should be read as negative OR positive while individual animal statuses and test results represent possible values for the i^{th} animal

- Test results depend on individual animal statuses through sensitivity and specificity
- Status dynamics modelled at the herd level



Representation of surveillance programmes as HMMs

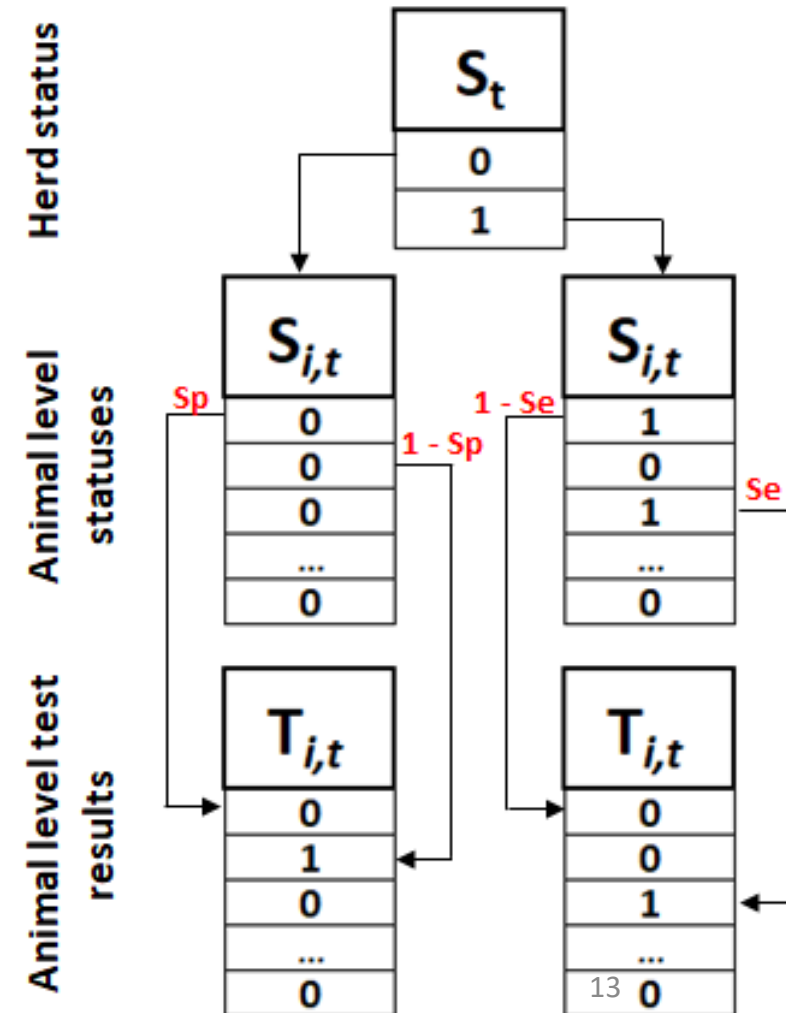
When test results at the animal level

$$T_t \sim \text{Binomial}(n, p(T_t^+))$$

$$p(T^+) = S_t \pi_{\text{within}} Se + [(1 - S_t) + S_t (1 - \pi_{\text{within}})](1 - Sp)$$

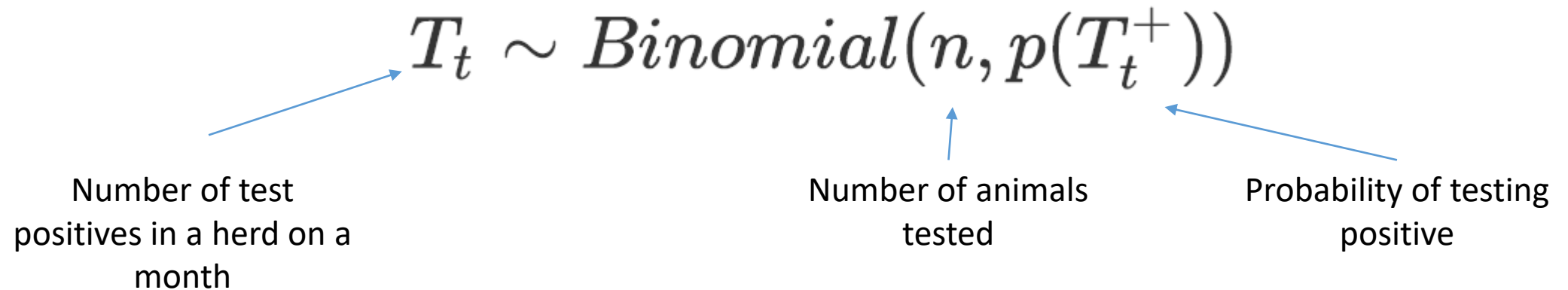
$$Se = p(T_{i,t} = 1 | S_{i,t} = 1)$$

$$Sp = p(T_{i,t} = 0 | S_{i,t} = 0)$$



Current questions on modelling of animal level tests

- Initially, we assumed the denominator of the Binomial distribution to be the number of animals tested

$$T_t \sim \text{Binomial}(n, p(T_t^+))$$


Number of test positives in a herd on a month

Number of animals tested

Probability of testing positive

- Leads the model to overestimate the probability of infection in some cases

e.g. when testing every calf borne, when obtaining 3 negative test results out 3 calves tested, the model assumes that there is an infinite population of untested calves that could be infected. If π_{within} is low, chances of missing the infection are assumed to be high.

Current questions on modelling of animal level tests

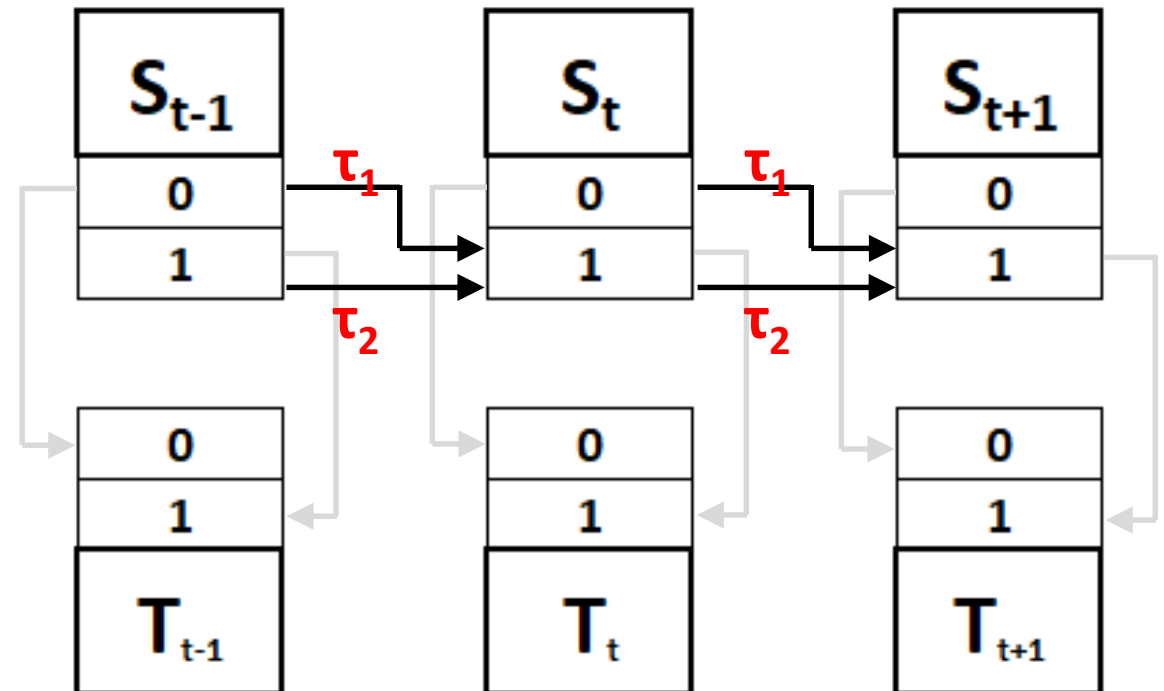
- Proposed solution:
 - Use herd size as the denominator of the Binomial distribution
- Other possibilities
 - In probability, this is a case of sampling without replacement which can be modelled with a hypergeometric distribution
 - Cannot be implemented easily in JAGS

Representation of surveillance programmes as HMMs

- Time steps of equal duration
(discrete-time model)
- Herd status at time t depends on herd status at time $t-1$
(Markovian property)

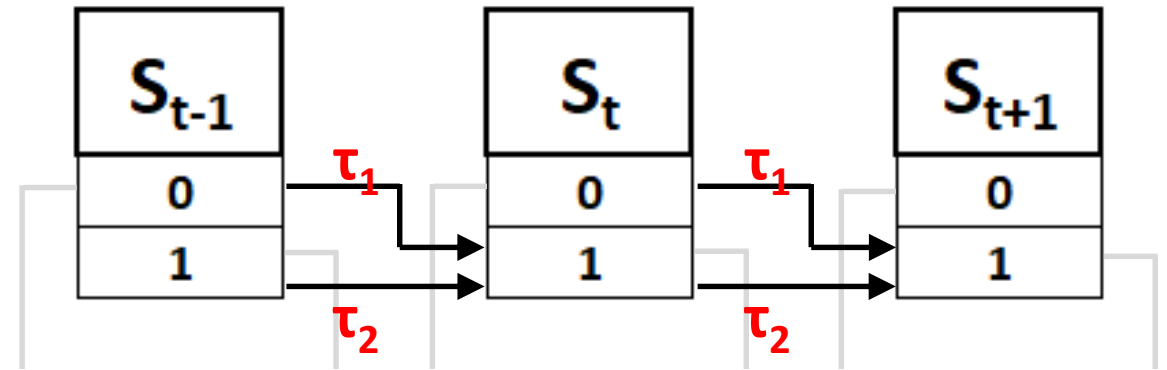
$$\tau_1 = p(T_t=1 \mid T_{t-1} = 0)$$

$$\tau_2 = p(T_t=1 \mid T_{t-1} = 1)$$



Representation of surveillance programmes as HMMs

- Time steps of equal duration
(discrete-time model)
- Herd status at time t depends on herd status at time $t-1$
(Markovian property)



$$S_t \sim \text{Bernoulli}(p(S_t))$$

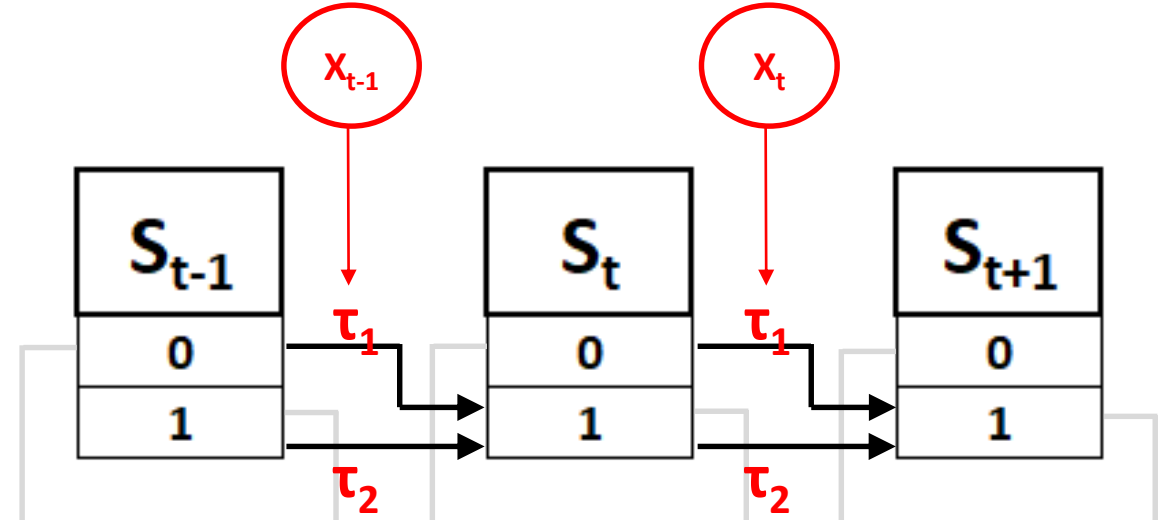
$$\tau_1 = p(T_t=1 \mid T_{t-1}=0)$$

$$\tau_2 = p(T_t=1 \mid T_{t-1}=1)$$

$$p(S_t) = (1 - S_{t-1})\tau_1 + S_{t-1}\tau_2$$

Incorporation of risk factor to the HMM

- Herd level probability of new infection (τ_{1t}) modelled as a function of one or several risk factors (X_t) using logistic regression



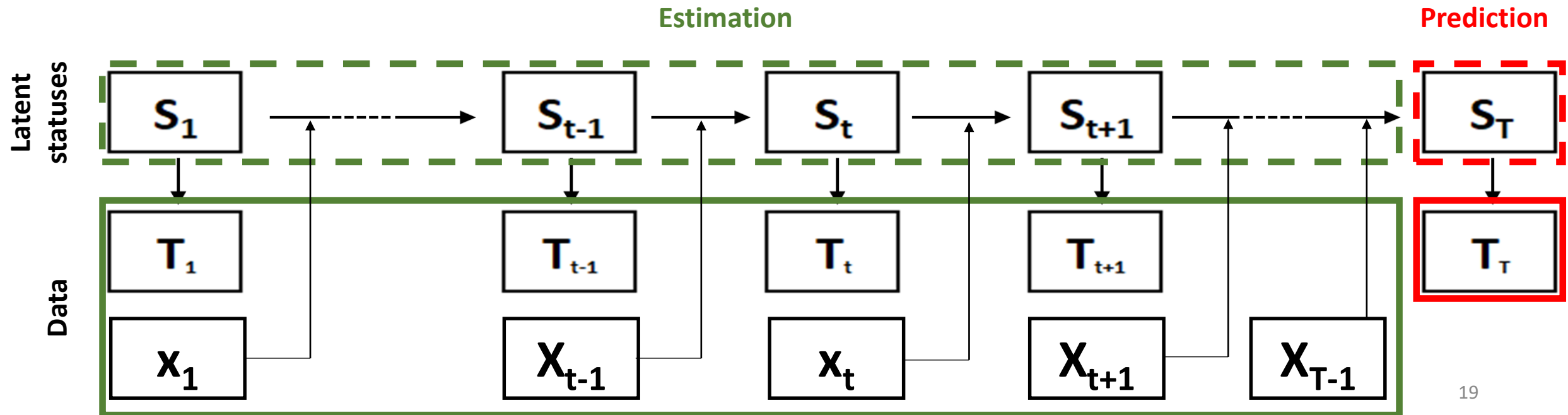
$$S_t \sim \text{Bernoulli}(p(S_t))$$

$$p(S_t) = (1 - S_{t-1})\tau_{1t} + S_{t-1}\tau_2$$

$$\ln\left(\frac{\tau_{1t}}{1-\tau_{1t}}\right) = X_t\theta$$

Predicted probability of status positive

- The model predicts a herd level probability of being status positive at time T
- All the data collected available up to T are used for parameter estimation



Latent status

Latent status

- The latent status is:
 - usually the presence of infection, but can be something else
 - in a herd
 - on a given month
- In the model, the latent status is defined by:
 - Hypotheses on test characteristics
 - Hypotheses on monthly status dynamics
 - Data on test results
 - Data on risk factors

Test characteristics as conditional probabilities

- Test sensitivity and specificity are probabilities of obtaining a given test result conditionally on the latent status of interest
 - $Se = p(T^+ | S^+)$
 - $Sp = p(T^- | S^-)$
- In the model, the hypotheses we make on these will determine what the latent status is

Infection by the BVDV as an example

- We consider infections infection by the bovine viral diarrhoea virus (BVDV)
- The latent status of interest is: *the presence of at least one persistently infected (PI) animal*
- In order to determine this status:
 - Different groups of animals can be tested
 - Different tests can be used
- Usually PI animals
 - are born in the herd
 - rarely reach adulthood

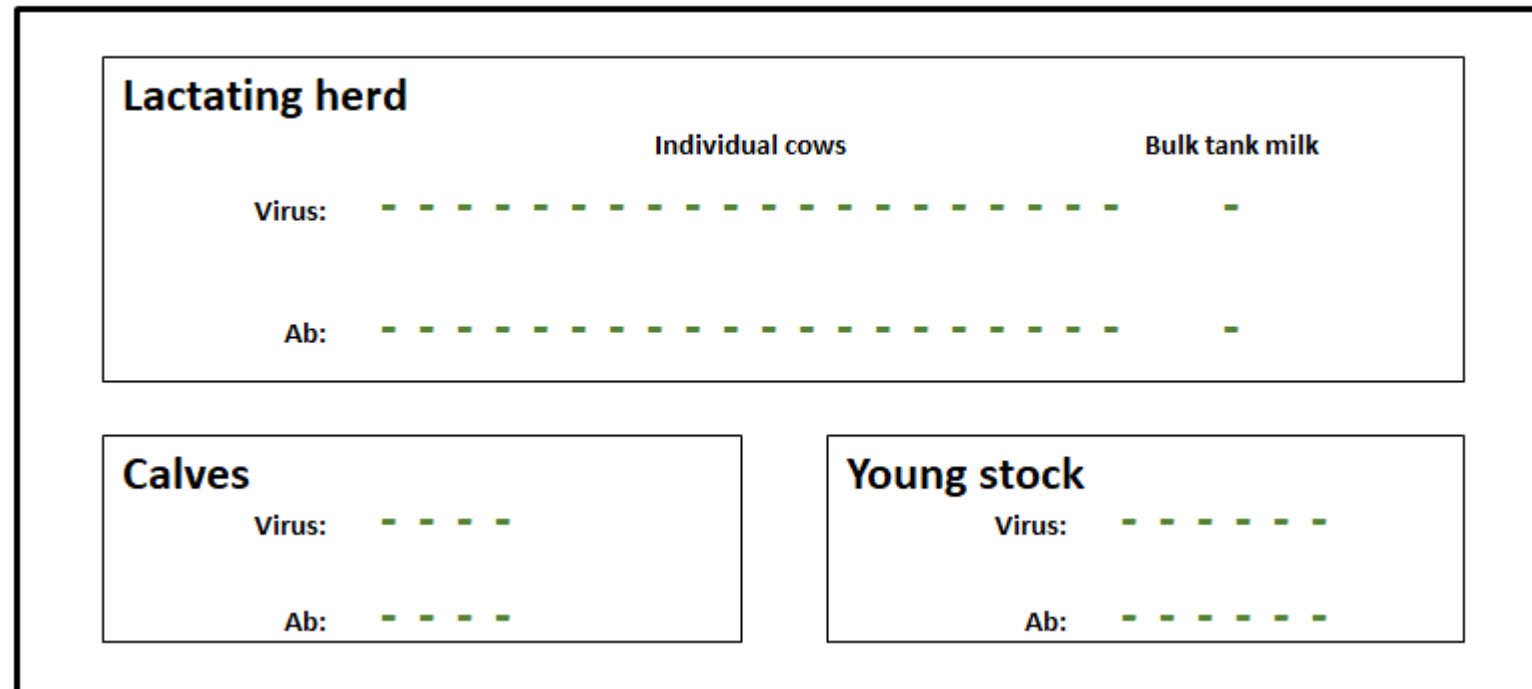
Infection by the BVDV as an example

- In order to explore the relationship between the latent status of interest and the results of different testing regimes, we describe a very simple scenario
 - a herd made of 3 age groups: lactating herd, calves and young stock
 - an initial situation where all animals are both virus and antibody negative
 - infection is introduced in the lactating herd and results in the birth of PI calves 6 months later

Infection by the BVDV as an example

- Month 1: all animals are both virus and antibody negative

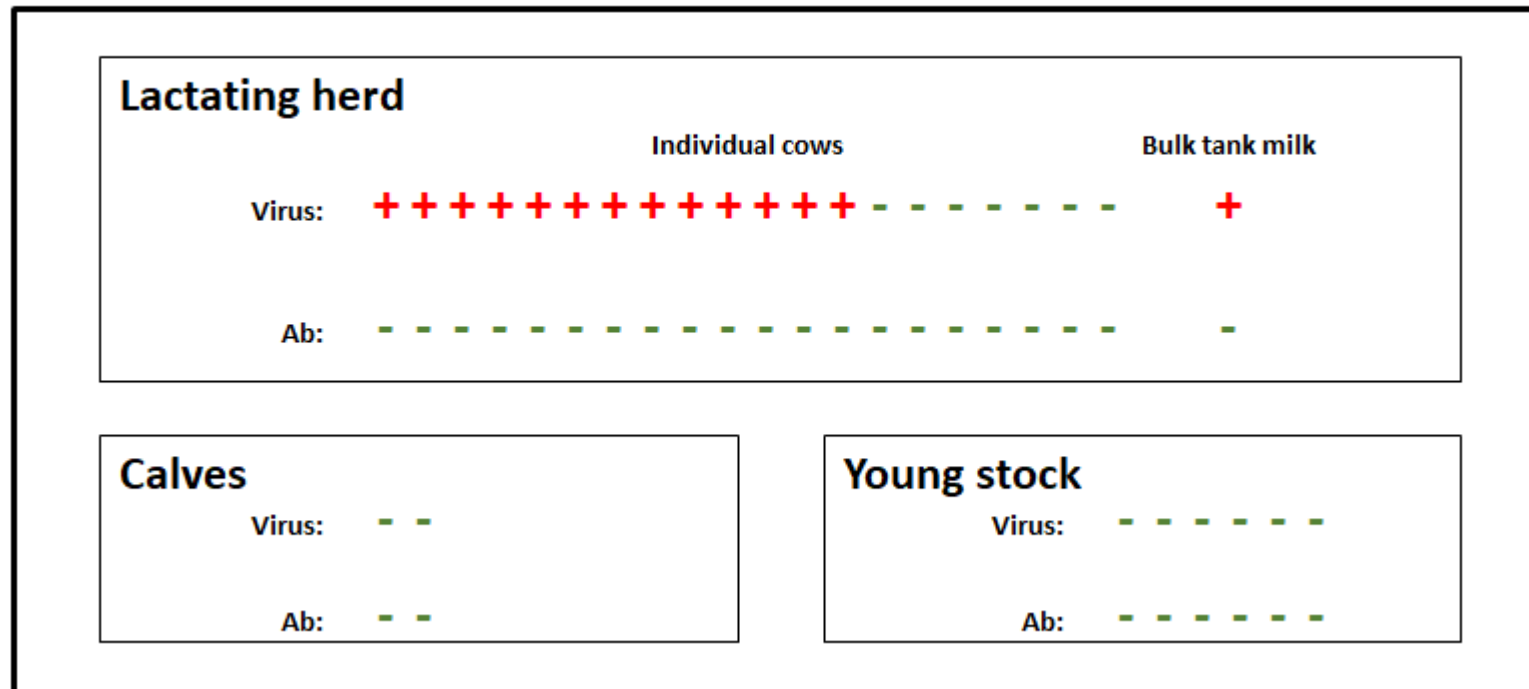
$t = 1$



Infection by the BVDV as an example

- Month 2: infection is introduced into the herd
 - Most cows are virus positive
 - The virus is found in milk

t = 2



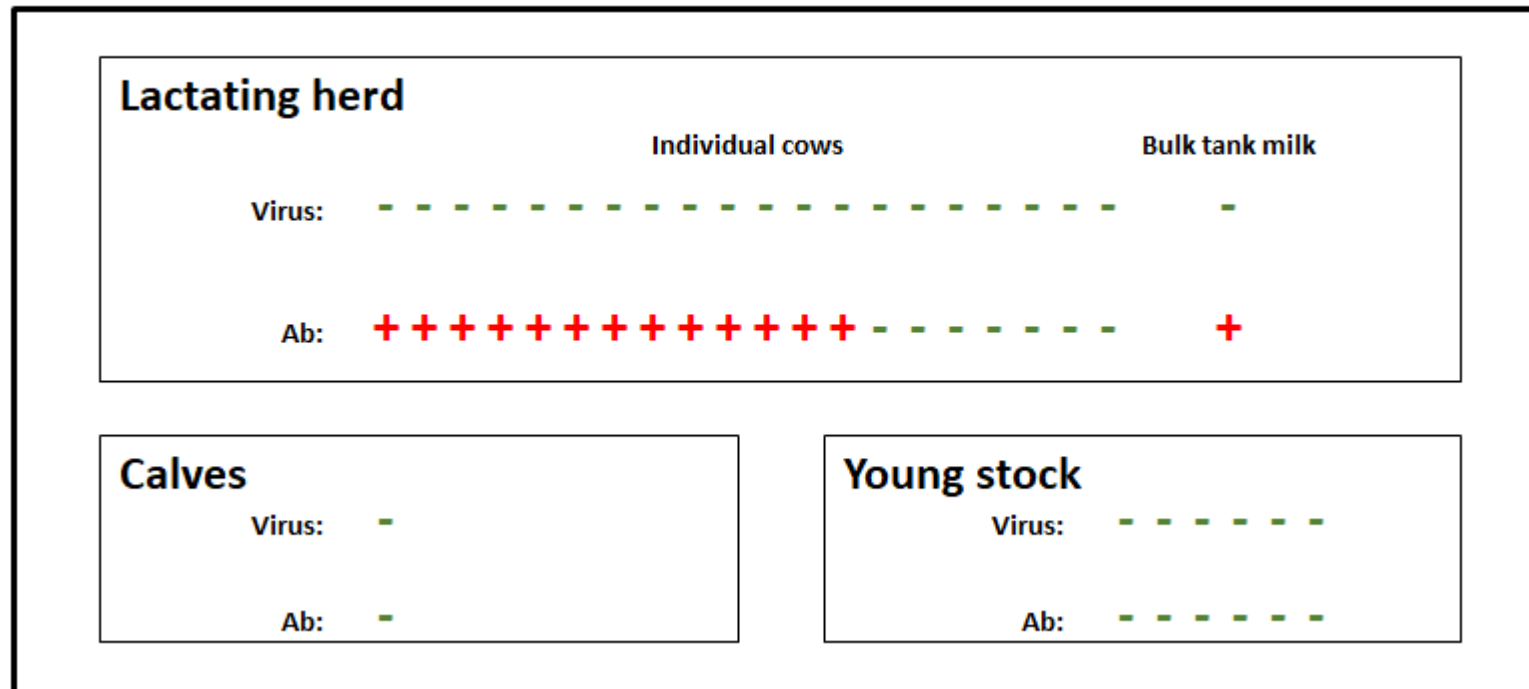
There are no PI animals: According to our definition, the herd is status negative!

Considering that the herd is status positive because infected cows could give birth PI calves in a few months would involve lowering the sensitivity of the other tests that are currently negative

Infection by the BVDV as an example

- Month 3: cows infected in the previous month become antibody positive

t = 3



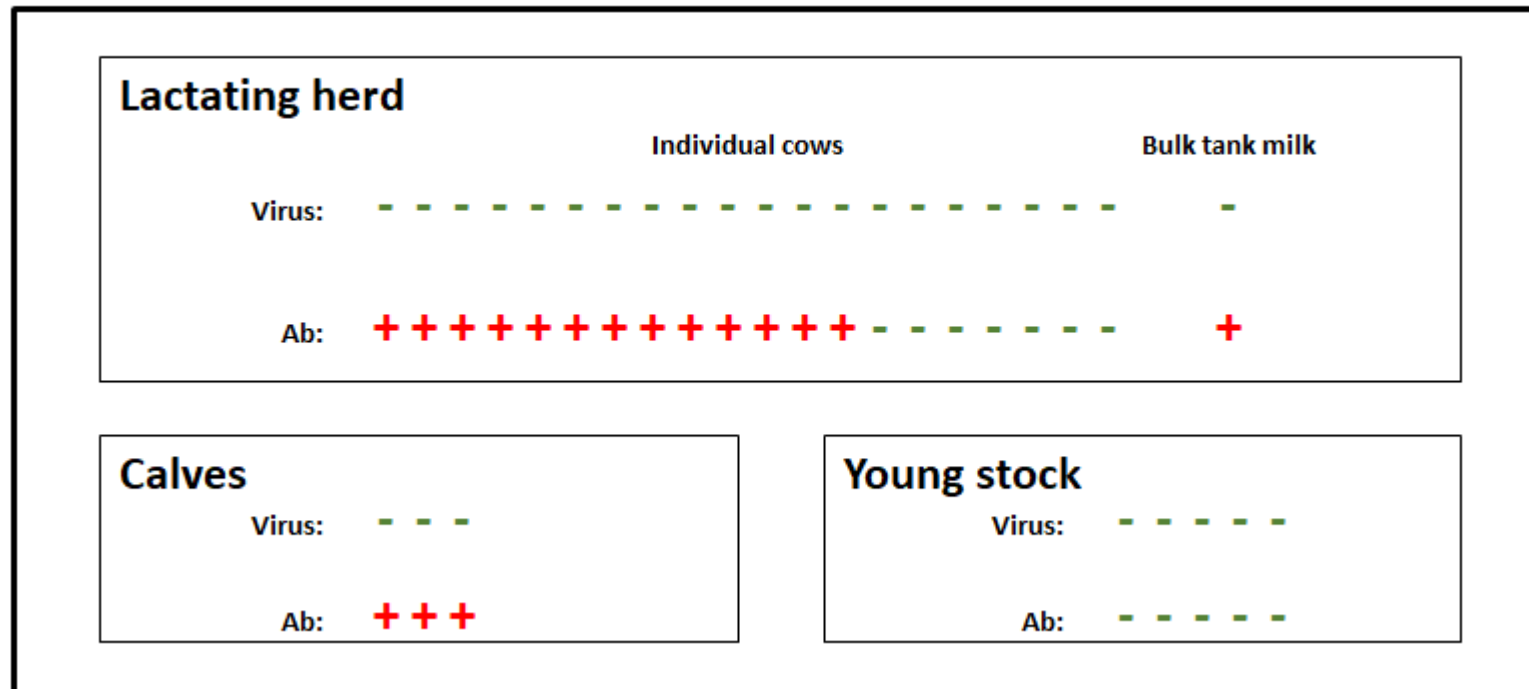
There are no PI animals: According to our definition, the herd is status negative!

Considering that the herd is status positive because infected cows could give birth PI calves in a few months would involve lowering the sensitivity of the other tests that are currently negative

Infection by the BVDV as an example

- Month 4: calves are antibody positive through passive transfer from colostrum

t = 4



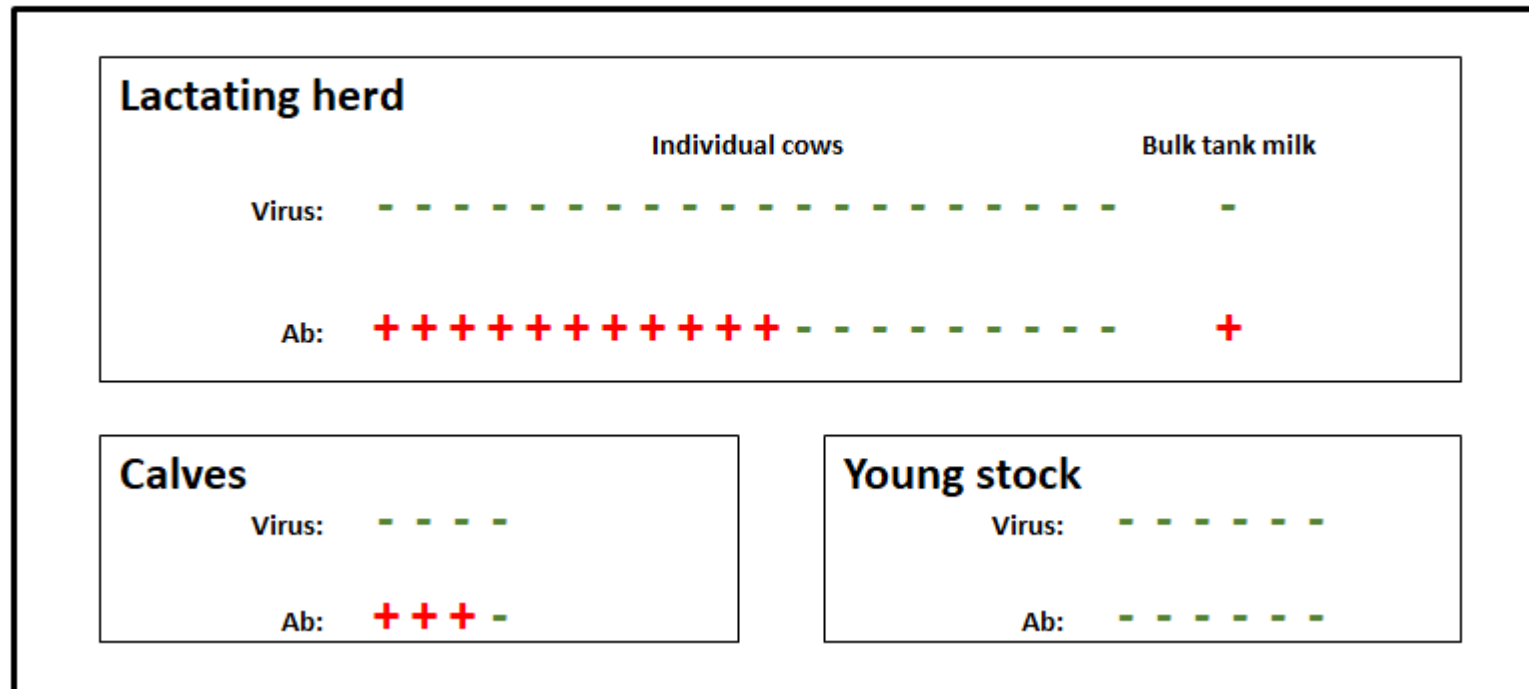
There are no PI animals: According to our definition, the herd is status negative!

Considering that the herd is status positive because infected cows could give birth PI calves in a few months would involve lowering the sensitivity of the other tests that are currently negative

Infection by the BVDV as an example

- Month 5: calves are antibody positive through passive transfer from colostrum

t = 5



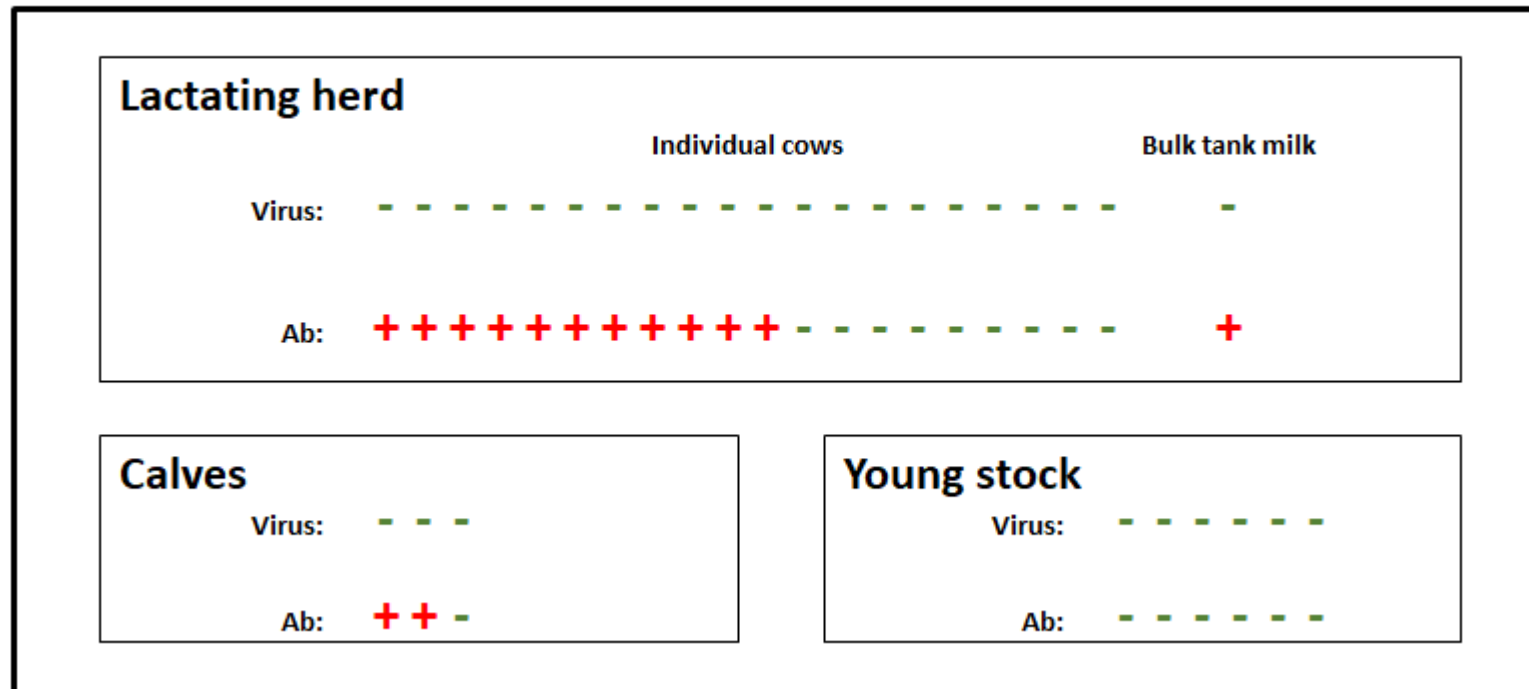
There are no PI animals: According to our definition, the herd is status negative!

Considering that the herd is status positive because infected cows could give birth PI calves in a few months would involve lowering the sensitivity of the other tests that are currently negative

Infection by the BVDV as an example

- Month 6: calves are antibody positive through passive transfer from colostrum

t = 6



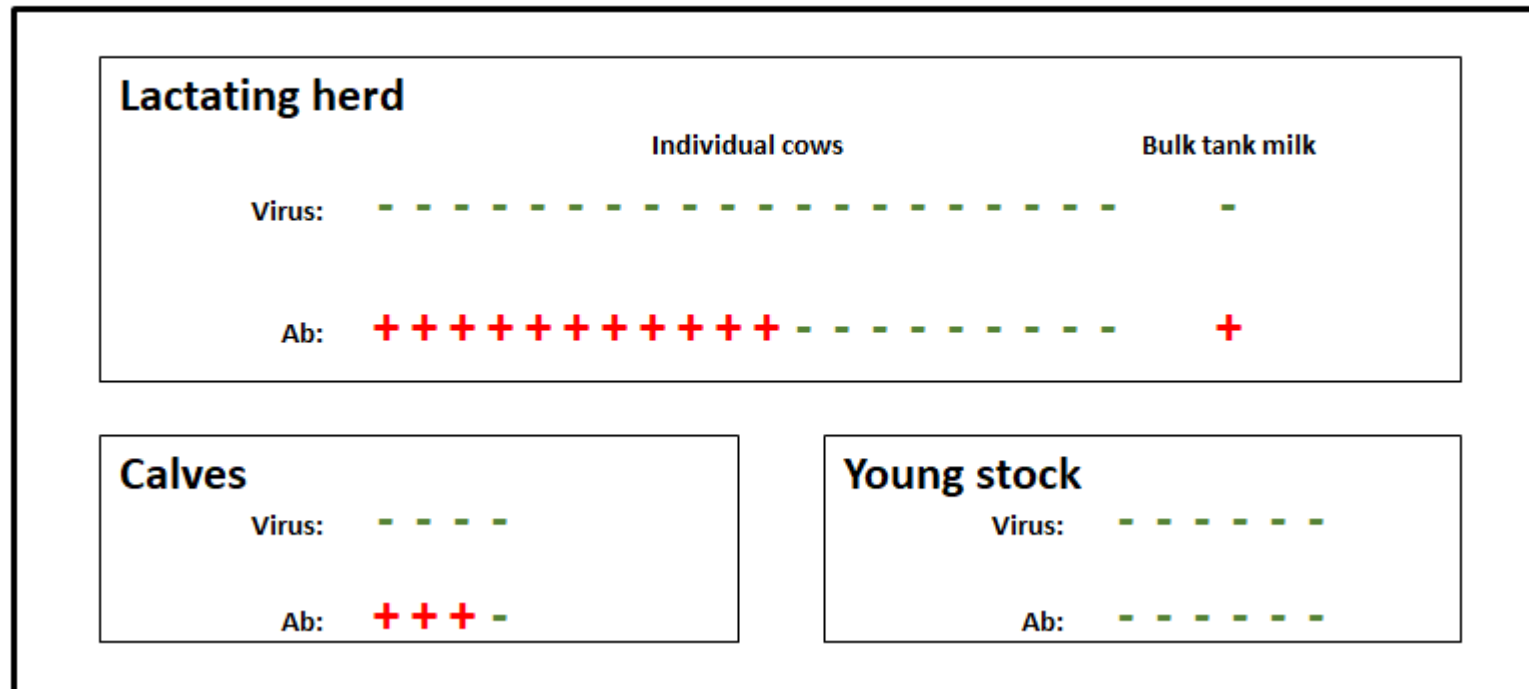
There are no PI animals: According to our definition, the herd is status negative!

Considering that the herd is status positive because infected cows could give birth PI calves in a few months would involve lowering the sensitivity of the other tests that are currently negative

Infection by the BVDV as an example

- Month 7: calves are antibody positive through passive transfer from colostrum

t = 7



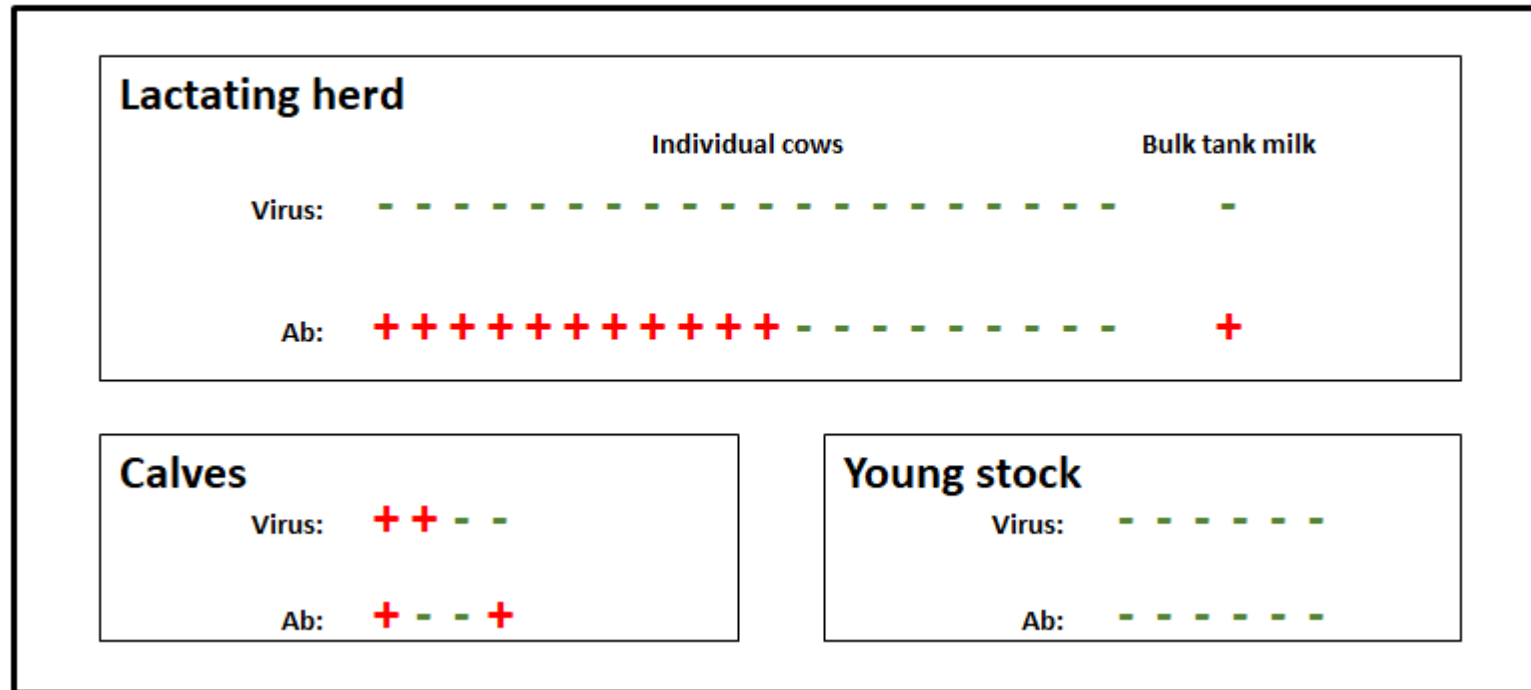
There are no PI animals: According to our definition, the herd is status negative!

Considering that the herd is status positive because infected cows could give birth PI calves in a few months would involve lowering the sensitivity of the other tests that are currently negative

Infection by the BVDV as an example

- Month 8: two PI calves are born
 - The herd becomes status positive

t = 8

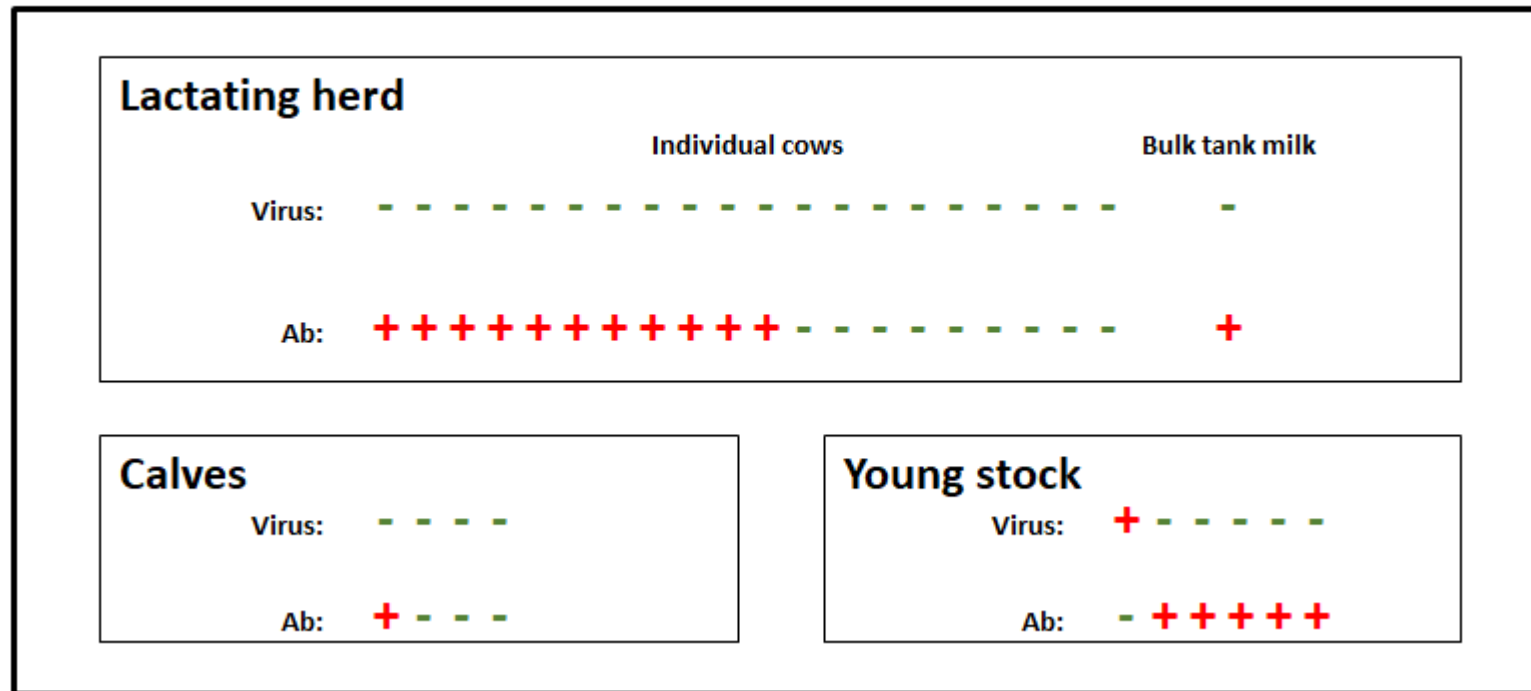


The herd becomes status positive

Infection by the BVDV as an example

- Month 9: one of the PI calves moves to the young stock group
 - The other calves from the group become antibody positive

t = 9

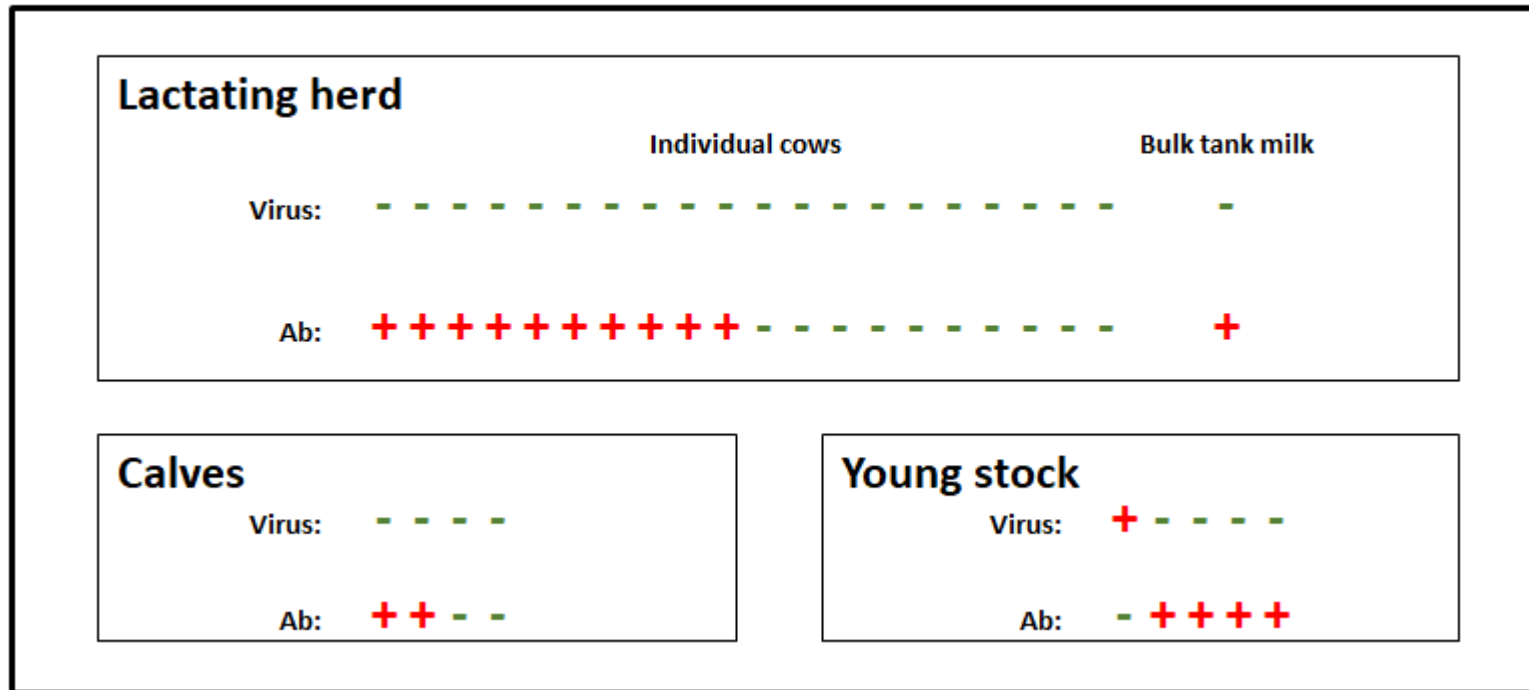


➔ The herd is still status positive

Infection by the BVDV as an example

- Month 10: one of the PI calves is still present

t = 10

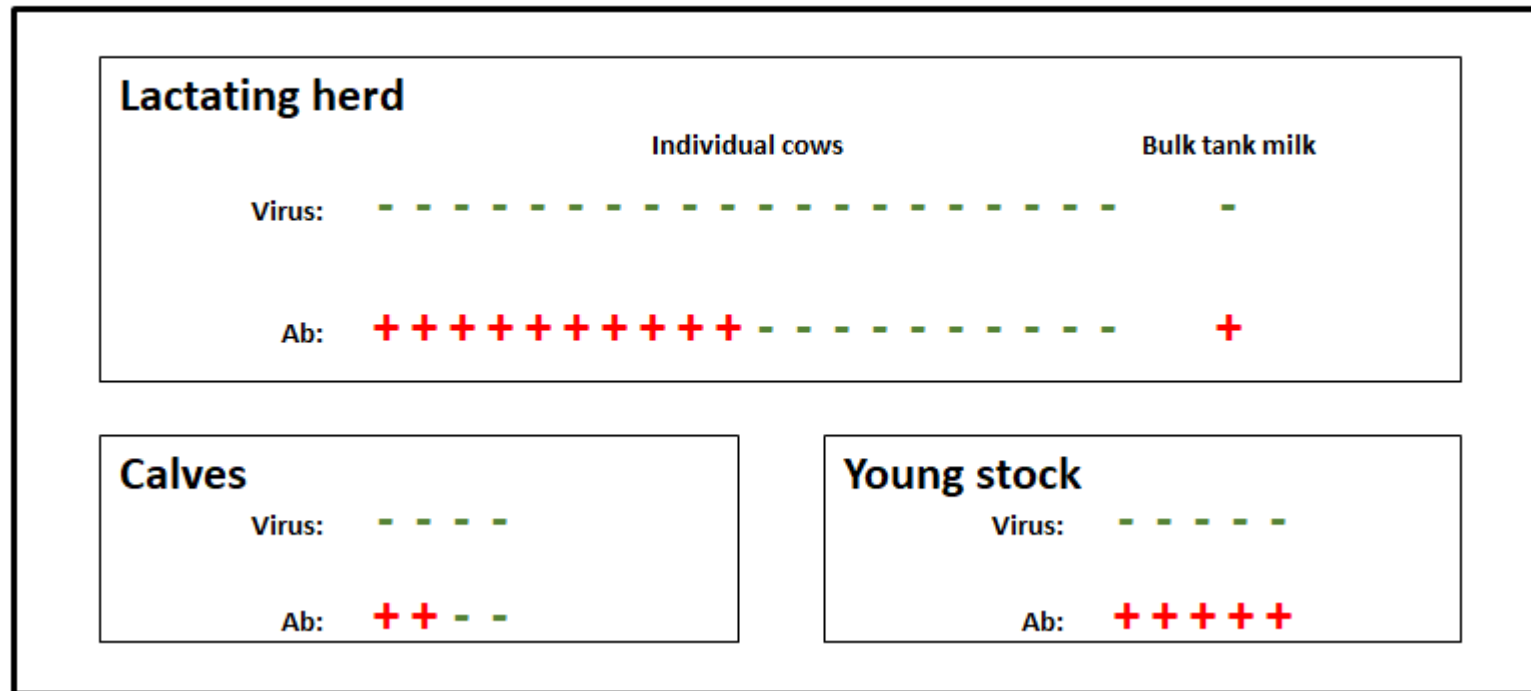


→ The herd is still status positive

Infection by the BVDV as an example

- Month 11: there is no remaining PI calf
 - The herd becomes status negative
 - There remain antibody positive animals in all groups => false positives

t = 11

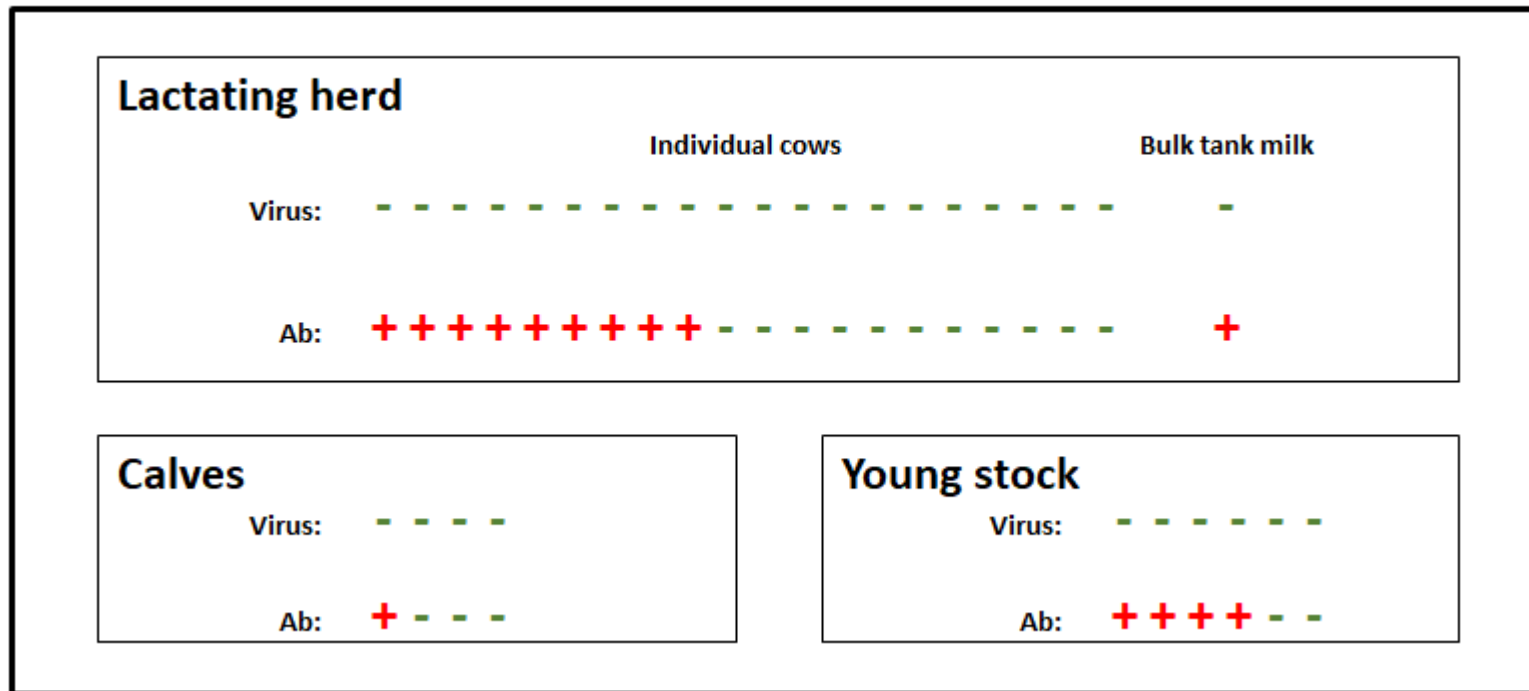


The herd is status negative

Infection by the BVDV as an example

- Month 12: there is no remaining PI calf
 - The herd remains status negative
 - There remain antibody positive animals in all groups => false positives

t = 12



The herd is status negative

Infection by the BVDV as an example

- In this very simple scenario:
 - Latent status: > 1 PI animal
 - Testing calves was the best test
 - Virus testing in the lactating herd had $Se = 0$ and $Sp = 0$
 - Testing for antibodies had a low specificity

	Introduction of infection into the lactating herd							Birth of PI calves			No remaining PI animal	
Month	1	2	3	4	5	6	7	8	9	10	11	12
Latent status	0	0	0	0	0	0	0	1	1	1	0	0
BTM virus	0	1	0	0	0	0	0	0	0	0	0	0
BTM Ab	0	0	1	1	1	1	1	1	1	1	1	1
Calves virus	0	0	0	0	0	0	0	1	0	0	0	0
Calves Ab	0	0	0	1	1	1	1	1	1	1	1	1
Young stock virus	0	0	0	0	0	0	0	0	1	1	0	0
Young stock Ab	0	0	0	0	0	0	0	0	1	1	1	1

Infection by the BVDV as an example

- In this very simple scenario:
 - Latent status: from introduction of infection to removal last PI animal
 - No single test performs well

		Introduction of infection into the lactating herd ↓						Birth of PI calves ↓			No remaining PI animal ↓	
Month	1	2	3	4	5	6	7	8	9	10	11	12
Latent status	0	1	1	1	1	1	1	1	1	1	0	0
BTM virus	0	1	0	0	0	0	0	0	0	0	0	0
BTM Ab	0	0	1	1	1	1	1	1	1	1	1	1
Calves virus	0	0	0	0	0	0	0	1	0	0	0	0
Calves Ab	0	0	0	1	1	1	1	1	1	1	1	1
Young stock virus	0	0	0	0	0	0	0	0	1	1	0	0
Young stock Ab	0	0	0	0	0	0	0	0	1	1	1	1

Infection by the BVDV as an example

- With this example, we see that it is possible to give contradictory information to the model
 - If we say that both finding the virus in the lactating herd and finding the virus in calves have a high sensitivity and a high specificity, the model will have to deal with contradictory information which may hamper convergence
 - This may explain that the model works better in the Irish context where the testing protocol is simpler than in the Dutch case and where the group tested is the most likely to contain PI animals

Estimation and prediction

Estimation and prediction

- Performed in a Bayesian framework, using JAGS
- Prior distributions required for all model parameters
 - Prior for τ_1 only required when no risk factor
 - Beta priors for Se, Sp, τ_1 and τ_2
 - Normal priors on the logit scale for the coefficients of the logistic model for new infection
 - The exponentials of these coefficients are odds ratios of new infection

Markov chain Monte Carlo

- Burn-in for convergence
- Number of iterations, thinning...

Model outputs

- For each herd
 - model output = distribution for the predicted probability of infection
- Needs to be converted into infected / free from infection

