

# Breeding Value Predictions for Survival in Laying Hens Showing Cannibalism

T. Brinker, E. D. Ellen, R. F. Veerkamp, P. Bijma



WAGENINGENUR  
*For quality of life*



A Hendrix Genetics Company

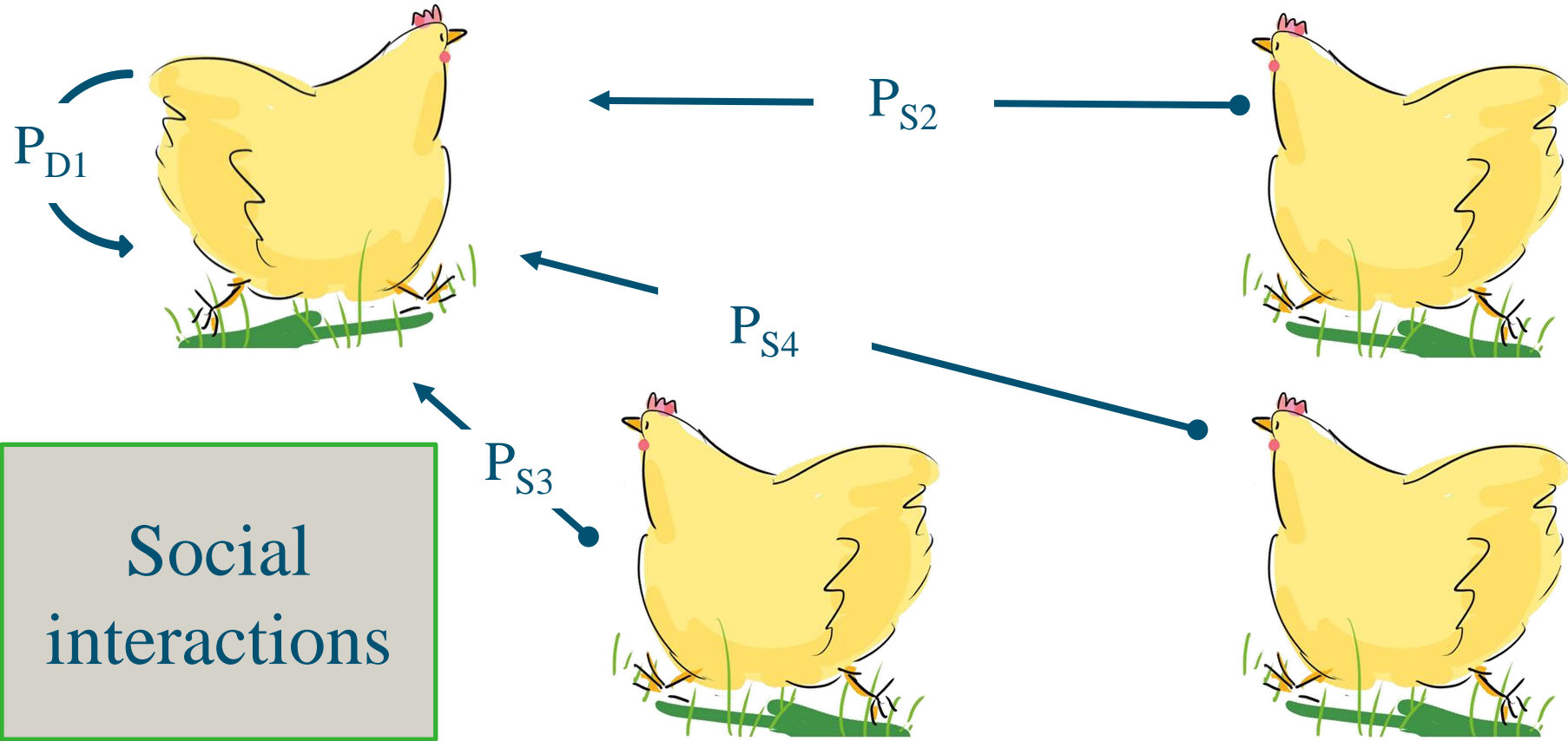


# Feather pecking and cannibalism

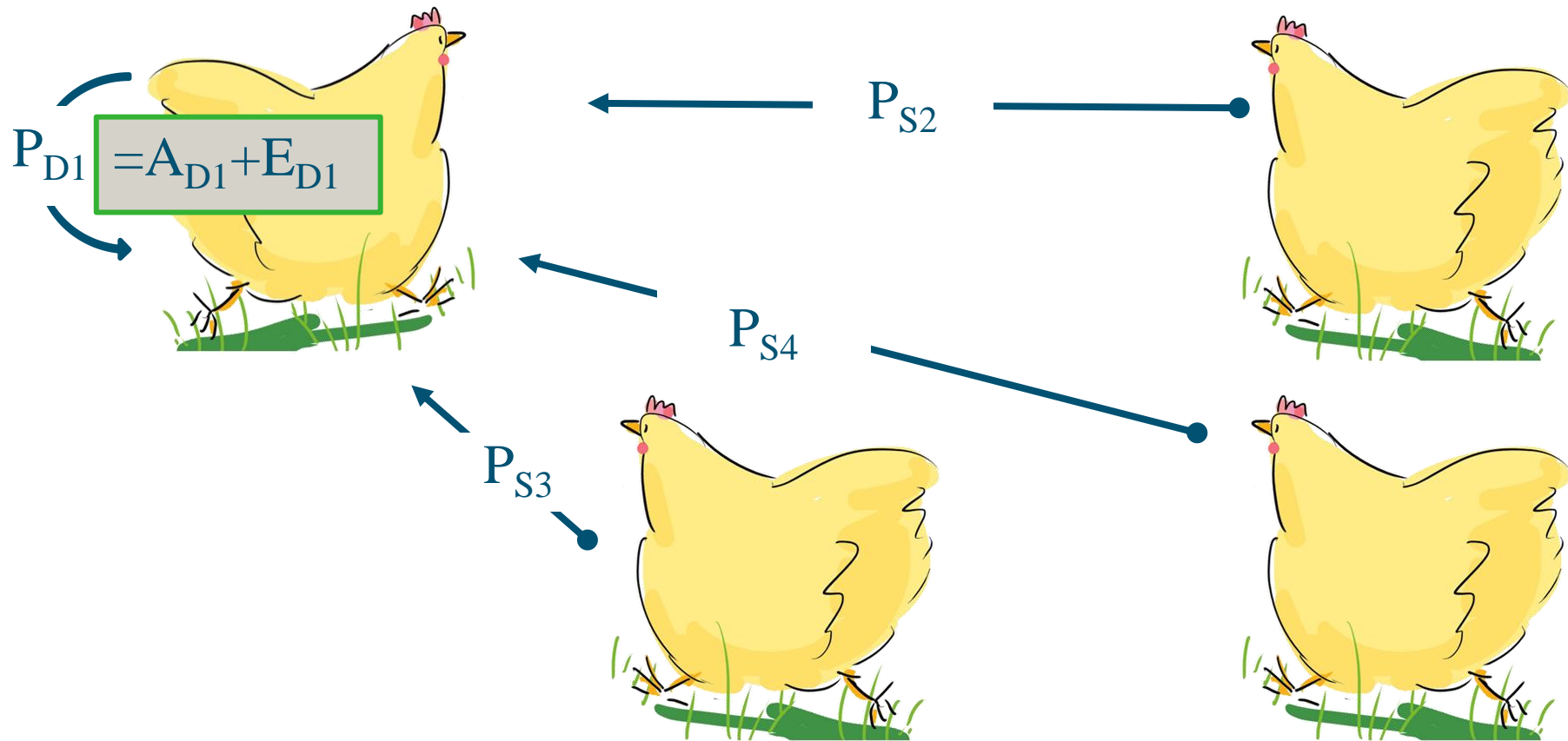


# Genetic selection for increased survival

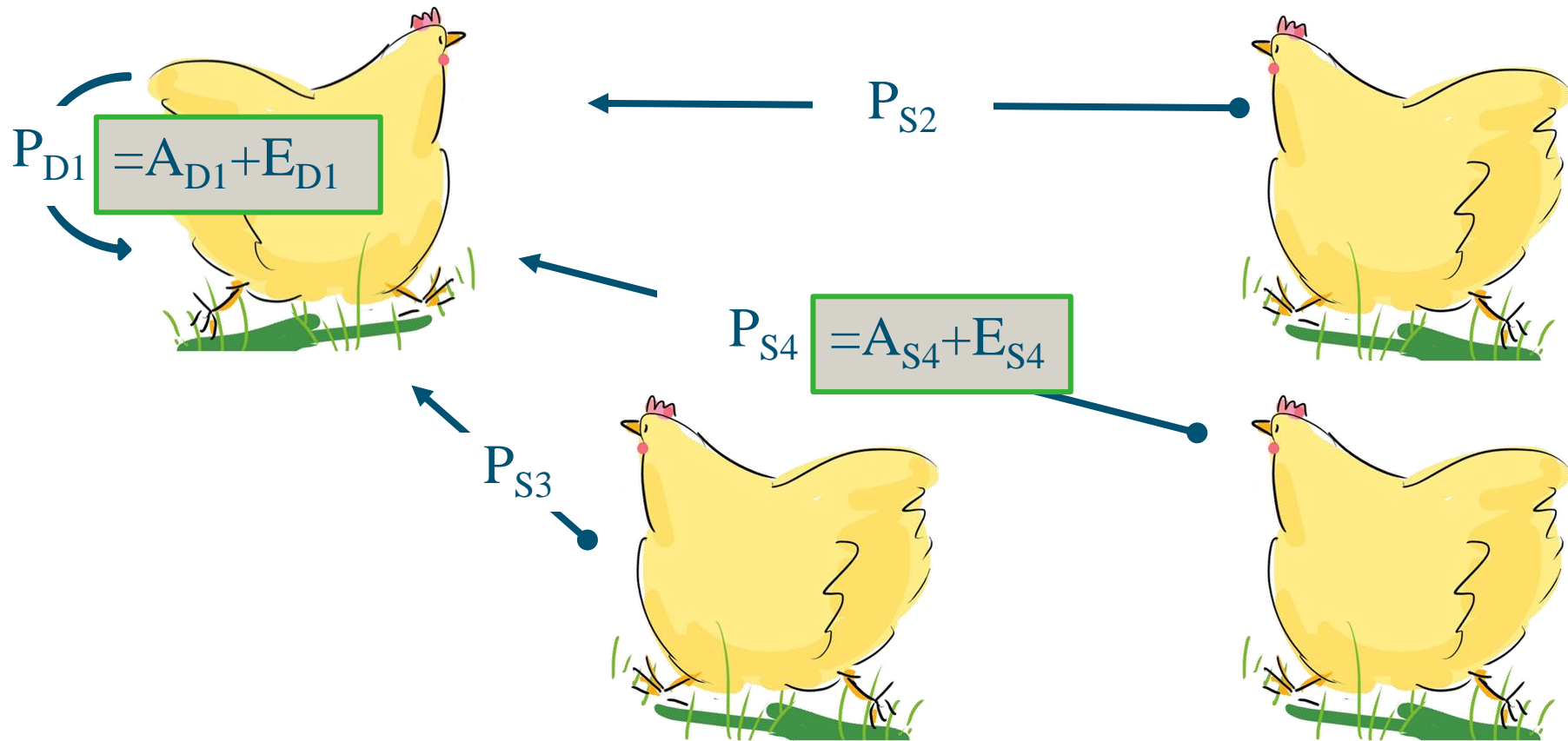
- Low  $h^2 \approx 2-10\%$
- High censoring
- Social interactions (IGE)



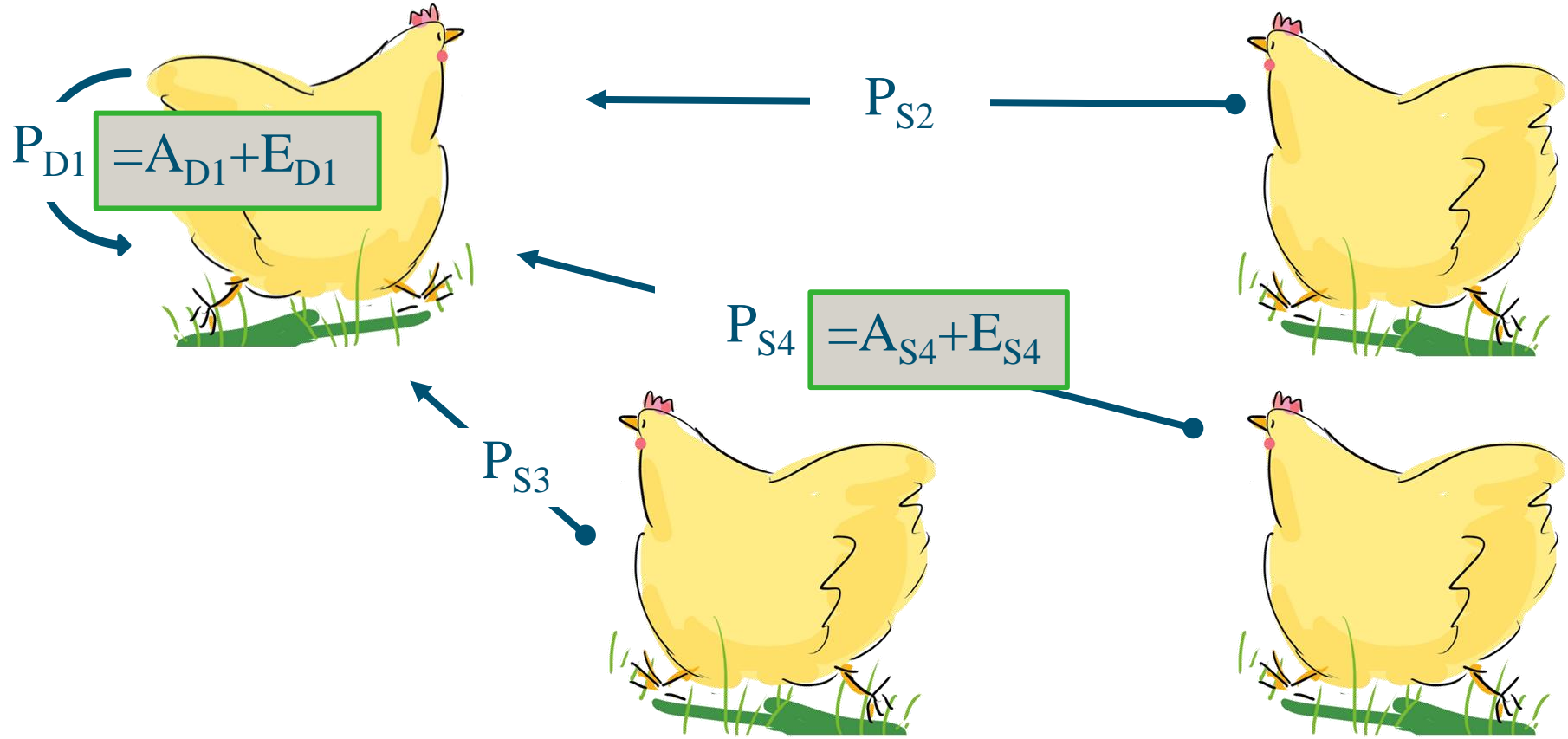
$$P_1 = P_{D1} + P_{S2} + P_{S3} + P_{S4}$$



$$P_1 = P_{D1} + P_{S2} + P_{S3} + P_{S4}$$



$$P_1 = P_{D1} + P_{S2} + P_{S3} + P_{S4}$$

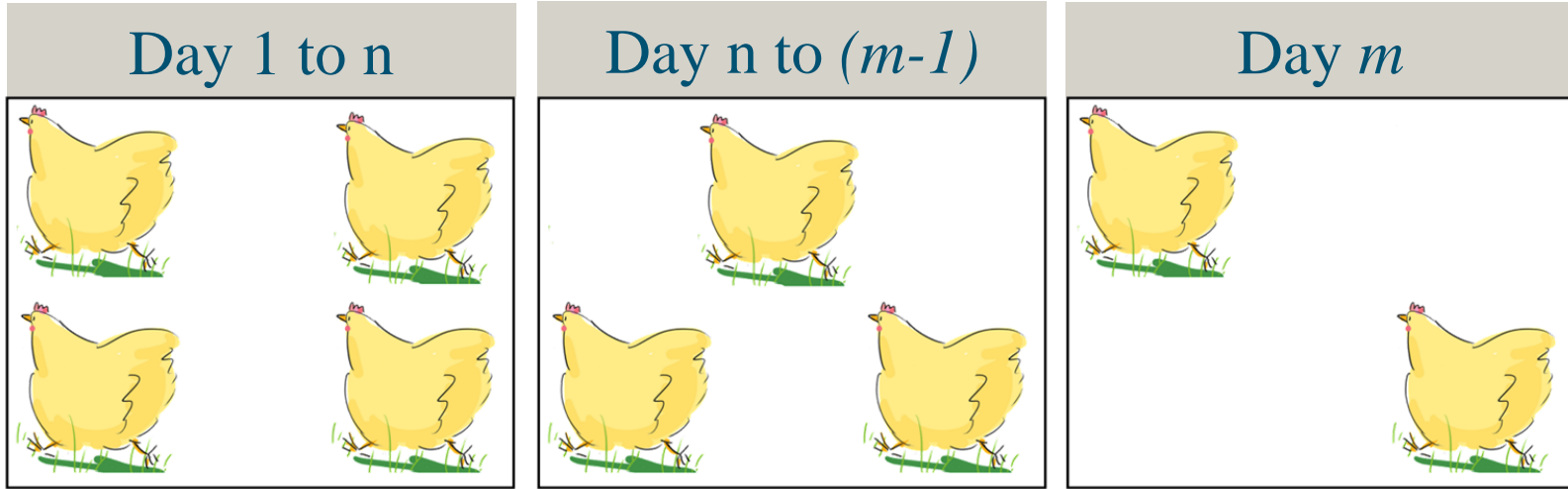


$$P_1 = A_{D1} + E_{D1} + A_{S2} + E_{S2} + A_{S3} + E_{S3} + A_{S4} + E_{S4}$$

# DGE-IGE Models

Focus on survival time

Censoring



Changing cage composition



# Objective

Improve breeding value predictions for survival time in laying hens showing cannibalism



# DATA



# Genetic stock

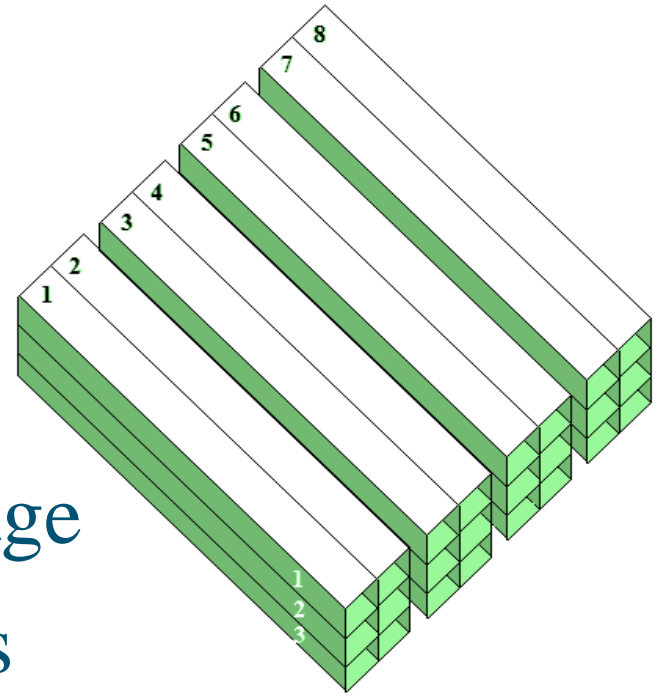
2 purebred White Leghorn layer lines

- 6 276 W1
- 6 916 WB
- Intact beaks

# Housing

## 2 laying houses

- 4 bird battery cages
  - Hens same line and age
  - Unrelated individuals



Example laying house

# Data collection

- Survival Time
- Survival each month (max=13)  
alive (1) or dead (0)

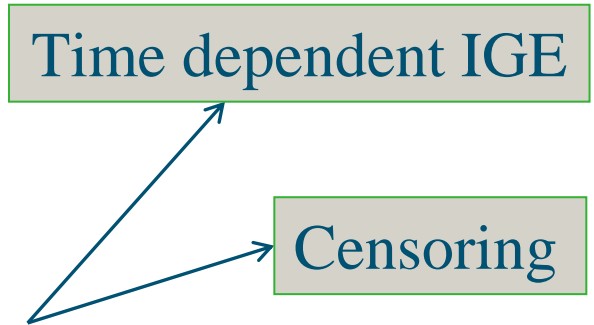
# METHODS



# Four models

## Compare breeding values

- DGE & IGE for survival time
  - Linear mixed model
- DGE & IGE for survival (0/1)
  - 3 repeated measures models



<b>Abbr</b>	<b>Trait</b>	<b>Model</b>	<b>Obs/Ind</b>
STM	Survival time	Simple linear regression	1
RMM.t	Survival (0/1)	Linear DGE and IGE regressions on time	13
RMM.p	Survival (0/1)	Linear DGE and IGE regressions on a function of mean survival	13
GLMM	Survival (0/1)	Generalized linear mixed model	13



# RMM.t vs RMM.p

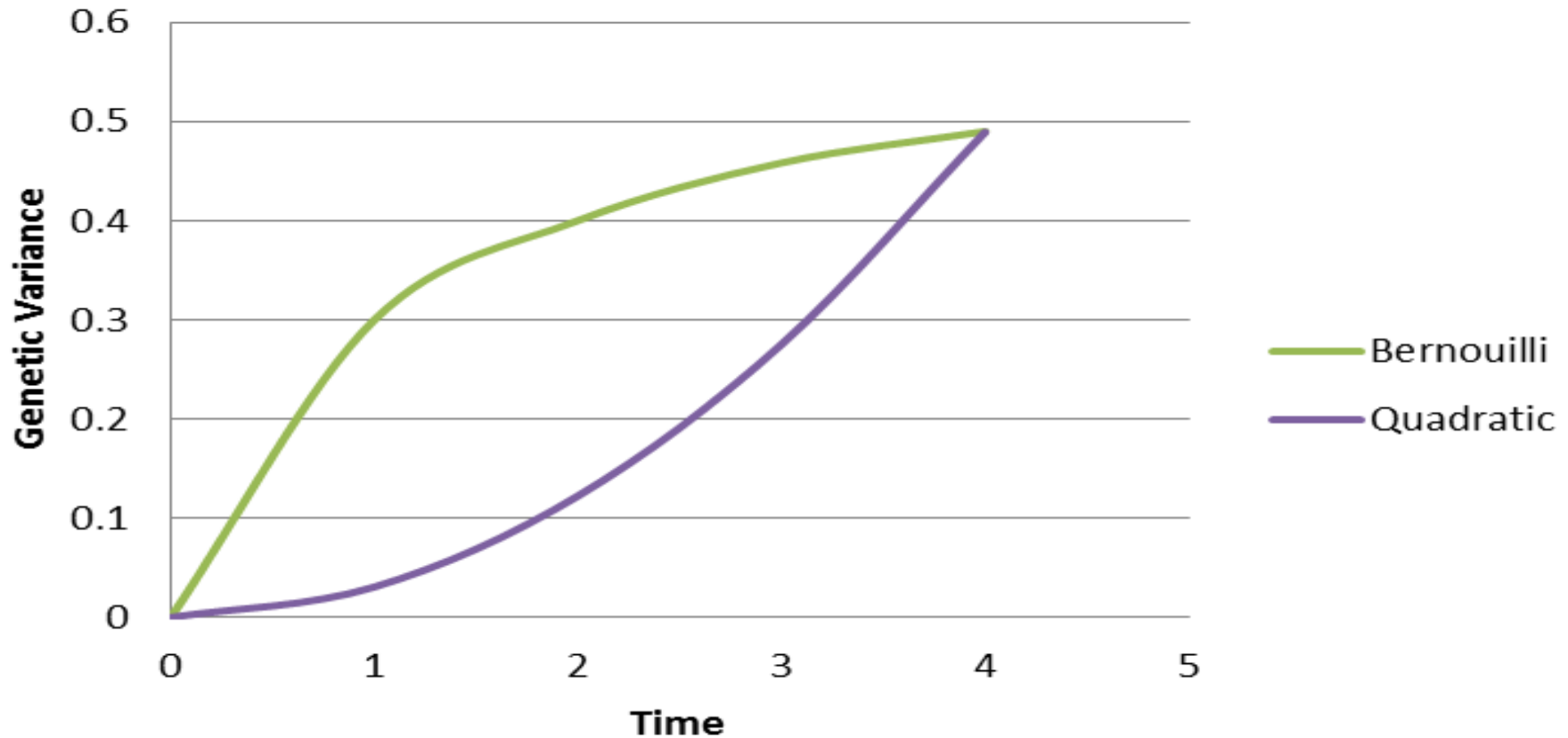
*RMM.t*  $\rightarrow$  time (months); *RMM.p*  $\rightarrow \sqrt{p(1-p)}$

Variances estimated in *RMM.t* are quadratic:

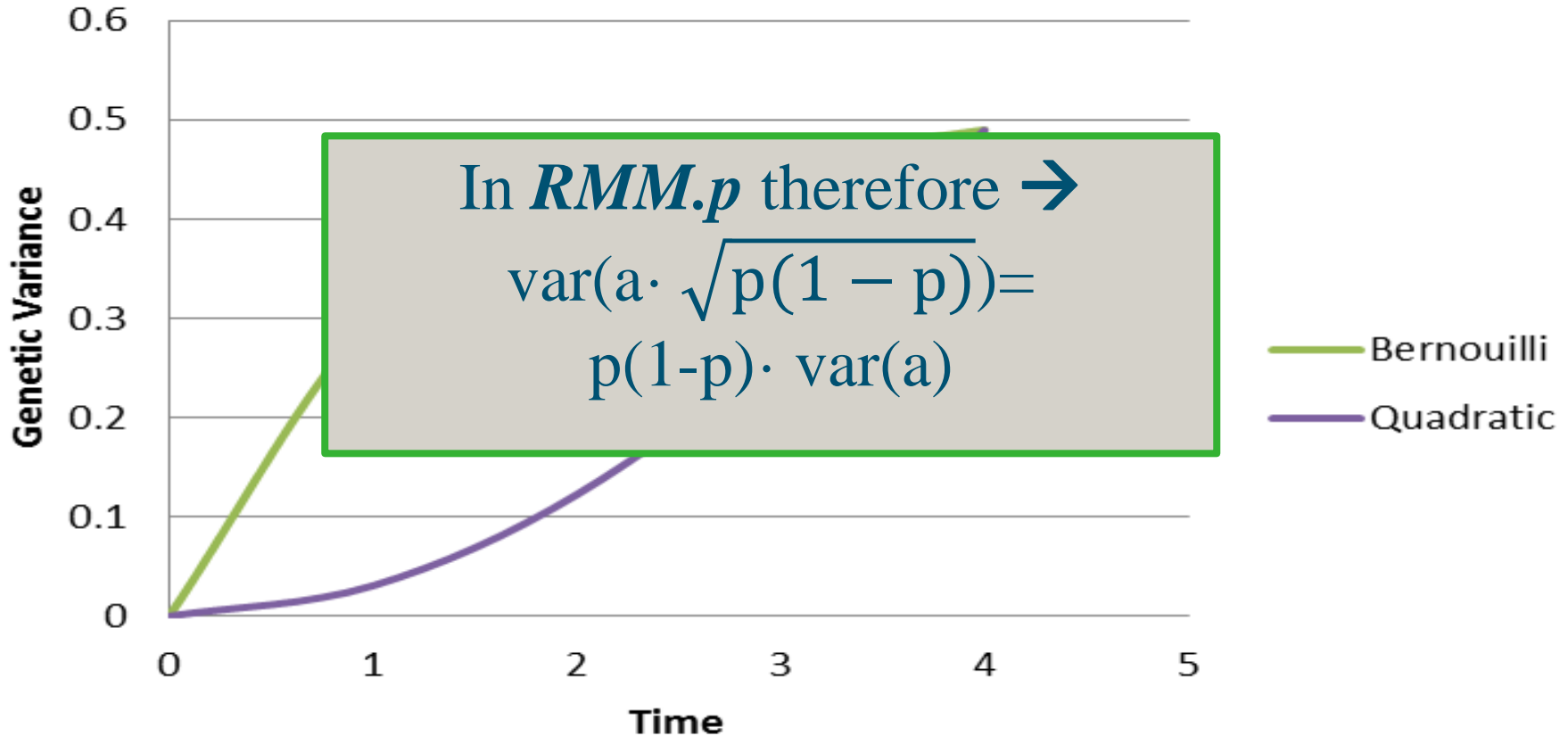
$$\text{var}(a \cdot t) = t^2 \text{var}(a)$$

True variances of binomial data:  $p(1-p)$

## The distributions do not fit!



## The distributions do not fit!



# GLMM

*GLMM* → Logit Link  
function

$$\text{logit}(E(y)) = \log\left(\frac{p}{1-p}\right)$$

So that the underlying scale  
goes from  $-\infty$  to  $+\infty$

In *GLMM* therefore

$$y=0,1$$

but with a linear predictor

$$-\infty < y < \infty$$

# Cross validation

5 mutually exclusive subsets

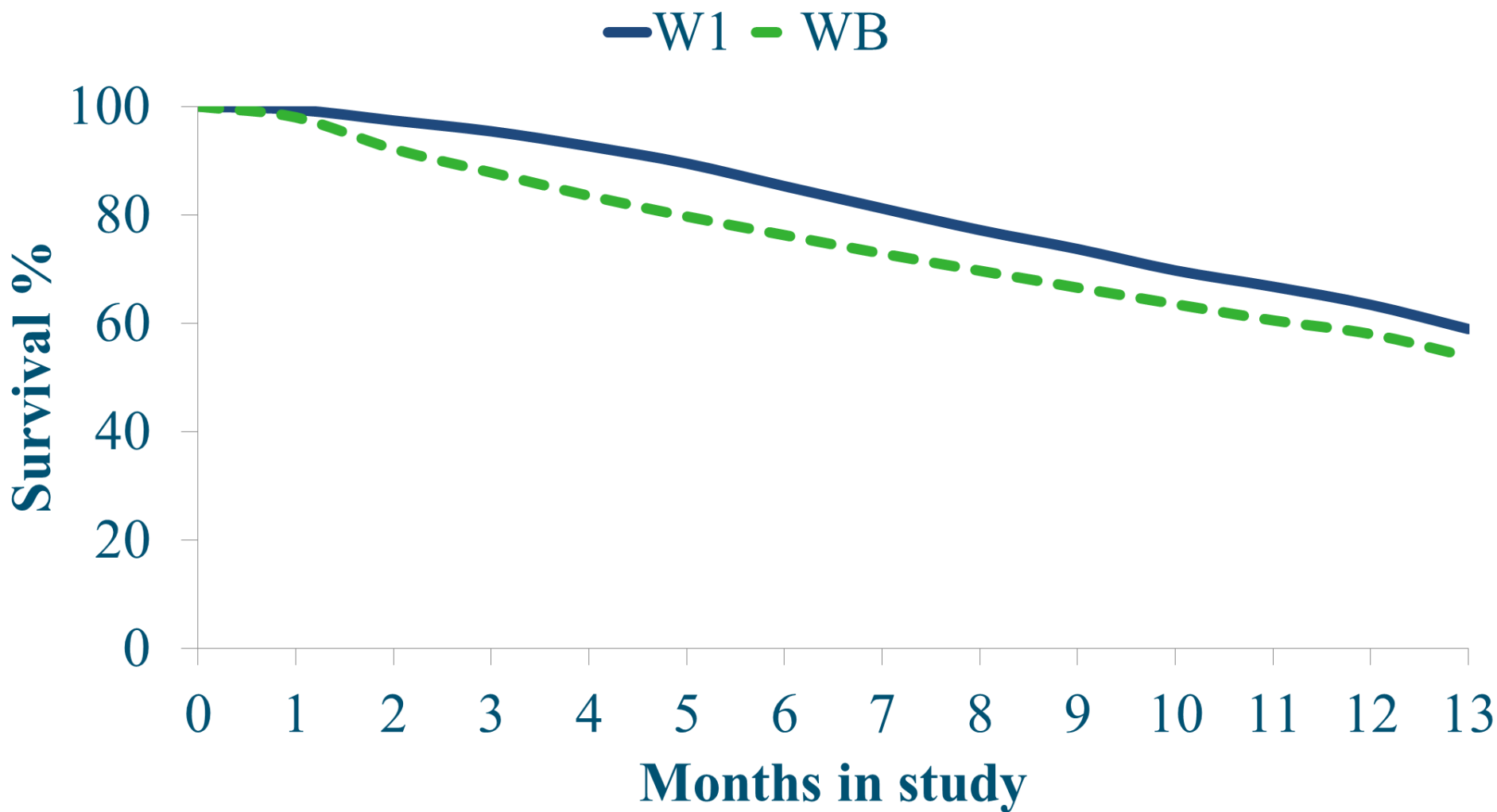
- 20% known phenotypes set missing
- Predict missing phenotypes

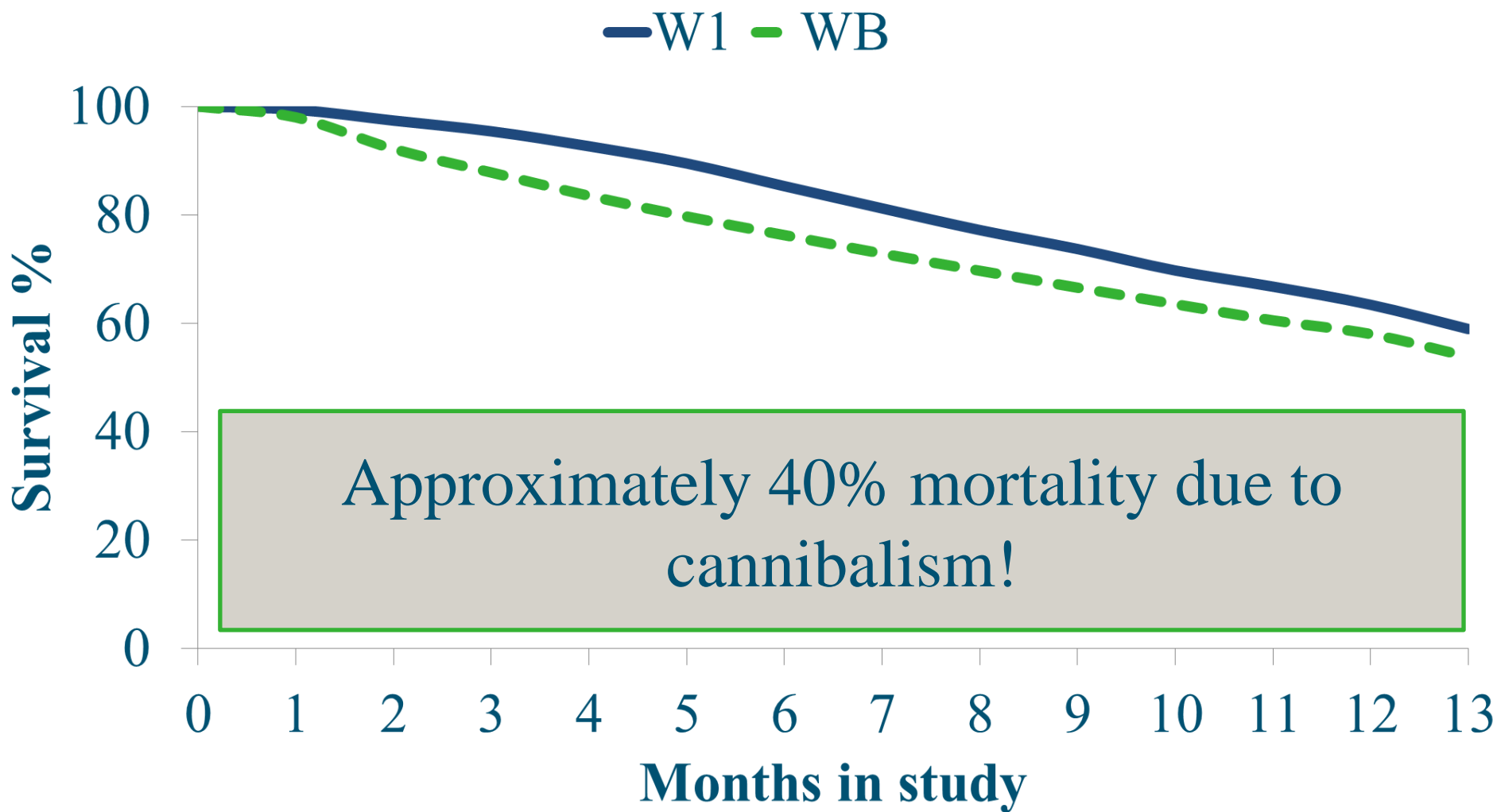
Compare predicted to observed phenotypes

- Spearman rank correlation

# RESULTS









<b>LINE W1</b>	<b>TD</b>	<b>Rank correlation</b>	<b>% Improved</b>	<b>Approximate Accuracy</b>
STM	-	0.135±0.012	-	0.44
SM.t	No	0.148±0.012	+10	0.48
SM.p	No	0.162±0.012	+20	0.53
GLMM	No	0.150±0.012	+11	0.49
SM.t	Yes	0.063±0.013	-53	0.20
SM.p	Yes	0.049±0.013	-64	0.16
GLMM	Yes	0.081±0.013	-41	0.26

<b>LINE W1</b>	<b>TD</b>	<b>Rank correlation</b>	<b>% Improved</b>	<b>Approximate Accuracy</b>
STM	-	0.135±0.012	-	0.44
SM.t	No	0.148±0.012	+10	0.48
SM.p	No	0.162±0.012	+20	0.53
GLMM	No	0.150±0.012	+11	0.49
SM.t	Yes	0.063±0.013	-53	0.20
SM.p	Yes	0.049±0.013	-64	0.16
GLMM	Yes	0.081±0.013	-41	0.26

<b>LINE WB</b>	<b>TD</b>	<b>Rank correlation</b>	<b>% Improved</b>	<b>Approximate Accuracy</b>
STM	-	0.170±0.012	-	0.46
SM.t	No	0.185±0.012	+9	0.51
SM.p	No	0.174±0.012	+2	0.47
GLMM	No	0.190±0.012	+12	0.52
SM.t	Yes	0.134±0.012	-21	0.37
SM.p	Yes	0.124±0.012	-27	0.34
GLMM	Yes	0.149±0.012	-12	0.41

<b>LINE WB</b>	<b>TD</b>	<b>Rank correlation</b>	<b>% Improved</b>	<b>Approximate Accuracy</b>
STM	-	0.170±0.012	-	0.46
SM.t	No	0.185±0.012	+9	0.51
SM.p	No	0.174±0.012	+2	0.47
GLMM	No	0.190±0.012	+12	0.52
SM.t	Yes	0.134±0.012	-21	0.37
SM.p	Yes	0.124±0.012	-27	0.34
GLMM	Yes	0.149±0.012	-12	0.41

# Discussion

Max rank correlation?

*“the square root of the proportion of phenotypic variance explained by breeding values” (Ellen et al., 2010)*

# Discussion

Max rank correlation?

*“the square root of the proportion of phenotypic variance explained by breeding values” (Ellen et al., 2010)*

E.g. max rank correlation for STM line W1: 0.309

The estimated was: 0.135

# Discussion

- Improvement of models because of censoring issue
- Time dependent IGE are detrimental

# Discussion

- Improvement of models because of censoring issue
- Time dependent IGE are detrimental

*Lipschutz-Powell et al. (2012) adjusted the IGE model; infected individuals only express IGE on susceptible group mates.*



# Conclusion

Using repeated measurement models, accuracies of EBVs were improved

- 10%-20% in W1
- 9%-12% in WB

Implication: response to selection can be improved accordingly

# Conclusion

# Thank you!

Using repeated measurement models, accuracies of EBVs were improved

- 10%-20% in W1
- 9%-12% in WB

Implication: response to selection can be improved accordingly

# EXTRA SLIDES



# What are the genetic parameters?



	<b>W1</b>	<b>WB</b>
$\sigma_{A_D}$	$28 \pm 3$	$41 \pm 4$
$\sigma_{A_I}$	$10 \pm 2$	$16 \pm 3$
$\sigma_{A_{DI}}$	$57 \pm 67$	$-158 \pm 120$
$\sigma_{TBV}$	$45 \pm 8$	$55 \pm 9$
$\sigma_P$	$107 \pm 1$	$135 \pm 1$
$T^2$	$0.18 \pm 0.06$	$0.16 \pm 0.05$
$r_A$	$0.20 \pm 0.22$	$-0.24 \pm 0.18$

# How to model time dependent IGE?



# Data – No time-dependent IGE

Self	Mate1	Mate2	Mate3	Time	Event
1	2	3	4	6	1
1	2	3	4	7	1
1	2	3	4	8	1
1	2	3	4	9	1
2	1	3	4	6	1
2	1	3	4	7	0
2	1	3	4	8	0
2	1	3	4	9	0

# Data – Time dependent IGE

Self	Mate1	Mate2	Mate3	Time	Event
1	2	3	4	6	1
1	2	3	4	7	1
1	0	3	4	8	1
1	0	3	4	9	1
2	1	3	4	6	1
2	1	3	4	7	0
2	1	3	4	8	0
2	1	3	4	9	0



# How do the models look like?



# 1. Survival Time Model → STM

$$y_{ijkl} = \text{fixed} + DGE_i + \sum_{j \neq i} IGE_j + \text{cage}_{k2} + e_{ijkl}$$

# 2/3. Repeated measurement model (RMM)

$$y_{ijkm} = \text{fixed} + DGE_i \cdot x_m + \sum_{j \neq i} IGE_j \cdot x_m + \text{cage}_{k1} + \text{cage}_{k2} \cdot x_m + PE_i \cdot x_m + e_{ijklm}$$

**RMM.t** →  $x$  : time (months) = function of mean survival

**RMM.p** →  $x$ :  $\sqrt{p(1-p)}$

# 4. Generalized linear mixed model → GLMM

$$\eta(E(y_{ijk})) = \text{fixed} + DGE_i + \sum_{j \neq i} IGE_j + \text{cage}_{k2} + PE_i$$

How do you calculate observed phenotypes?



**Model****Predicted phenotype**

STM

$$\hat{P}_i = DGE_i + \sum_{j \neq i} IGE_j$$

RMM.t/RMM.  
p

$$\hat{P}_i = \int_{start}^{end} DGE_i \cdot x \, dx + \sum_{j \neq i} \int_{start}^{end} IGE_j \cdot x \, dx$$

GLMM

$$\hat{P}_i = \sum_{start}^{end} \frac{e^{x_m + DGE_i + \sum_{j \neq i} IGE_j}}{1 + e^{x_m + DGE_i + \sum_{j \neq i} IGE_j}} (x)$$

**Back transformation LOGIT → Linear Scale**

$$\hat{p}_{i,m} = \frac{e^{x_m + DGE_i + \sum_{j \neq i} IGE_j}}{1 + e^{x_m + DGE_i + \sum_{j \neq i} IGE_j}}$$

How do you calculate the max correlation?



## Max rank correlation *(Ellen et al., 2010)*

$$\sqrt{r^2} = \sqrt{[\sigma_{A_D}^2 + (n - 1)\sigma_{A_S}^2] / \sigma_P^2}$$

$$\sqrt{\hat{r}_{IH}} = \text{corr}(\text{rank}(P_i - \bar{P})_i, \text{rank}(\hat{P}_i)) / \sqrt{r^2}$$

■ E.g. STM:  $\sigma_{A_D}^2 = 784$ ,  $\sigma_{A_S}^2 = 104$ ,  $\sigma_P^2 = 11449$

Max rank:  $\text{SQRT}((784 + 3 * 104) / 11449) = 0.309$

The estimated was: 0.135