



Modelling the Herd-specific Force Of Infection for BoHV-1 The Influence of Clustering.

Christel Faes¹, Niel Hens¹, Marc Aerts¹, Ziv Shkedy¹, Helena Geys², Koen Mintiens³, Hans Laevens³ and Frank Boelaert⁴.

contact: niel.hens@luc.ac.be; christel.faes@luc.ac.be; marc.aerts@luc.ac.be

¹Center For Statistics, Limburgs Universitair Centrum, Universitaire Campus, 3590 Diepenbeek, Belgium

²Johnson and Johnson, PRD Biometrics and Reporting, Belgium; ³Veterinary and Agrochemical Research Centre, Belgium; ⁴European Food Safety Authority, Italy

Abstract

Modelling infectious diseases often has to deal with clustering in the data. When animals are grouped within herds, one cannot ignore the possibility of animals within herds being more alike than between herds. Since independence among sample observations is a key assumption underlying standard statistical procedures, the presence of clustering in the data may raise important statistical issues, which should be addressed in the analysis. This poster shows the effect of clustering in the data on the force of infection of the bovine herpesvirus-1 in Belgian Cattle. Both herd-specific as well as population-averaged force of infections are derived.

Bovine HerpesVirus-1 (BoHV-1)

- ▷ transmissible disease
- ▷ serological survey 1997-1998 in Belgium
- ▷ province-stratified random sample
- ▷ 11284 cows in 309 herds
- ▷ ELISA-test for BoHV-1 glycoprotein B
- ▷ restriction to 10363 cases for which age > 6 months
- ▷ interest in age-specific sero-prevalence

$$\pi(a) = P(gB = 1|a)$$

- ▷ interest in age-specific force of infection (FOI), which is the rate at which susceptible animals get infected

$$\ell(a) = \frac{\pi'(a)}{1 - \pi(a)}$$

Ignore Clustering

- Logistic regression
- Assumption: independency
- Act as if animals are independent

$$Y_{ij} \sim \text{Bernoulli}(\pi_{ij})$$

$$\eta_{ij} = \log\left(\frac{\pi_{ij}}{1 - \pi_{ij}}\right) = f(a)$$

- Use fractional polynomials (Royson and Altman 1994) together with constraint of monotonicity:

$$f(a) = \beta_0 + \beta_1 a^{p_1} + \beta_2 a^{p_2}, \quad f'(a) > 0$$

- Force of infection:

$$\ell(a) = \pi(a)f'(a)$$

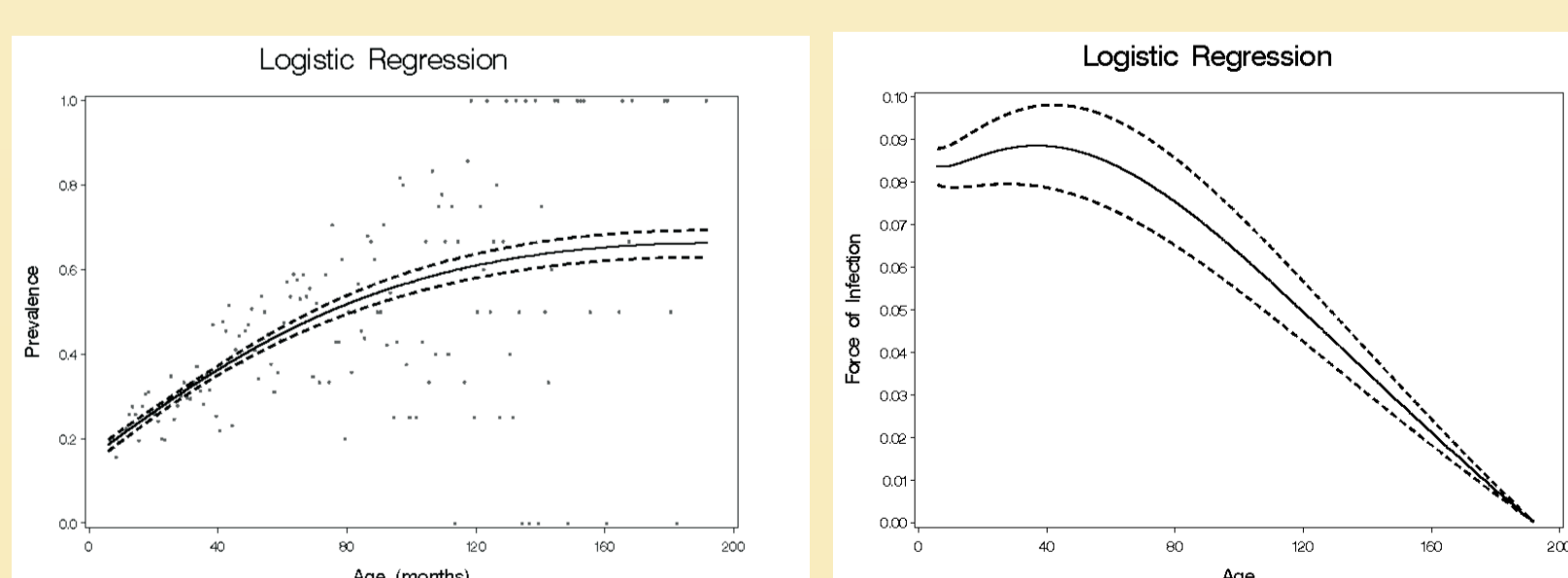
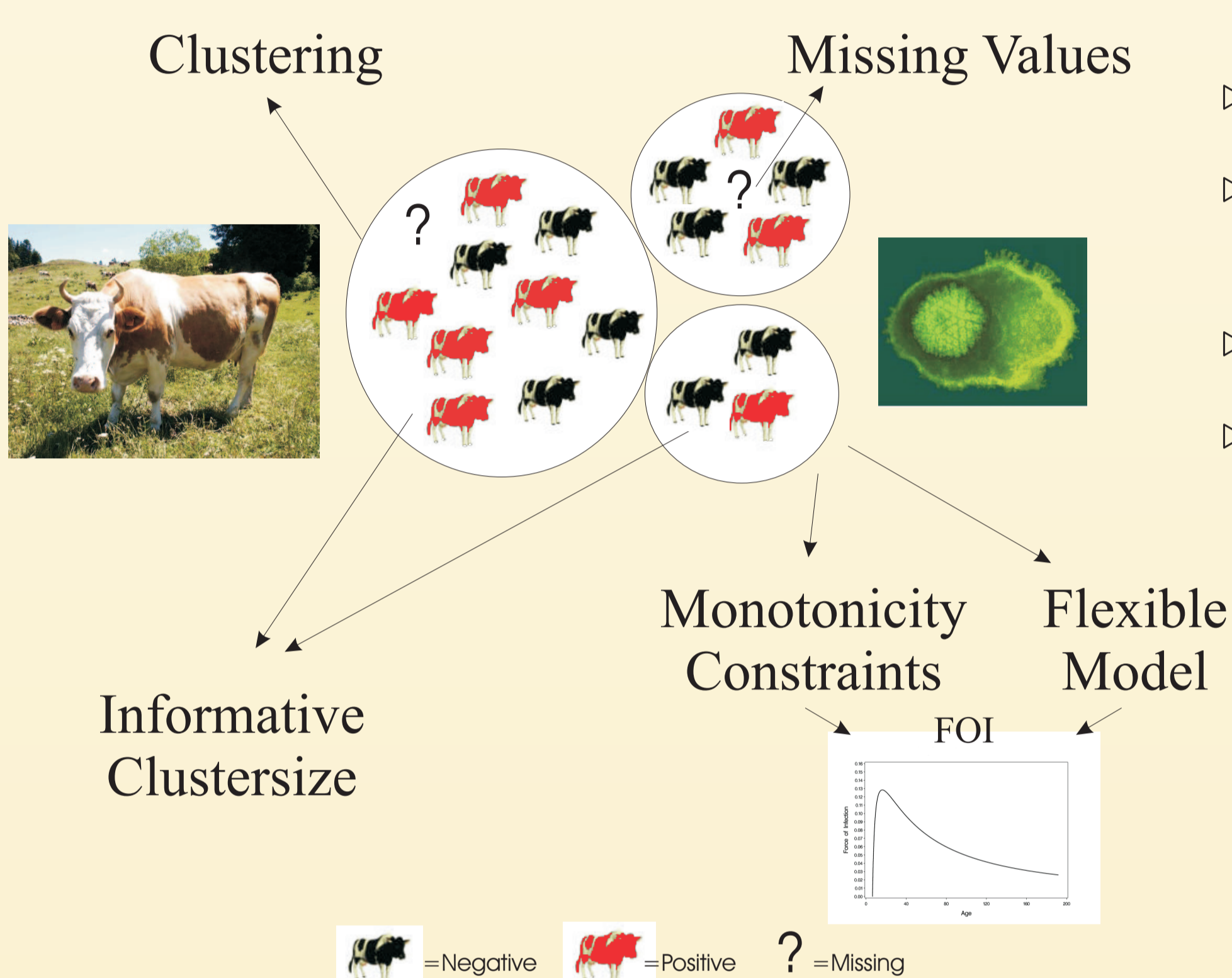


Figure 2: Sero-prevalence plot (left) and force of infection (right) based on logistic regression.

- Underestimates variance
- Spuriously 'significant' risk factors



Clustering in BoHV-1

- ▷ Animals are grouped within herds
- ▷ Animals in same herd share same environment, type of farm, ...
- ▷ Animals within herd more alike than between herds
- ▷ Animals within herd not independent

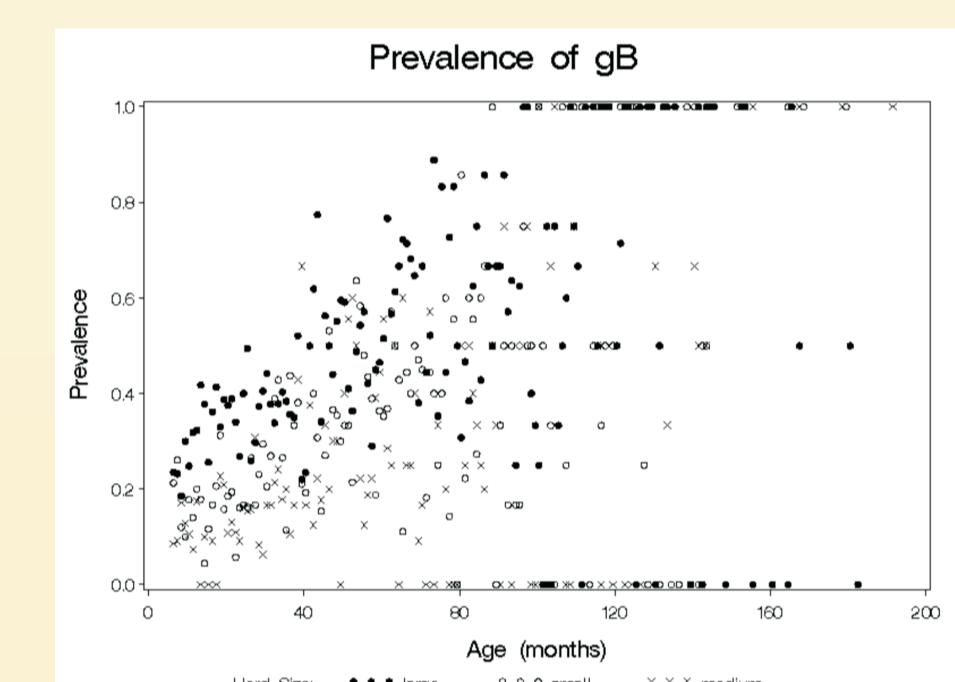


Figure 1: Sero-prevalence plot as a function of age based on the complete cases for small (circles), medium (stars) and large herds (triangles).

Account for Clustering

- Generalized Estimating Equations (Liang and Zeger 1986)
- Population-averaged approach
- Correct for clustering: working correlation structure

$$Y_{ij} \sim \text{Bernoulli}(\pi_{ij})$$

$$\eta_{ij} = \log\left(\frac{\pi_{ij}}{1 - \pi_{ij}}\right) = f(a),$$

- Use fractional polynomials (Royson and Altman 1994) together with constraint of monotonicity:

$$f(a) = \beta_0 + \beta_1 a^{p_1} + \beta_2 a^{p_2}, \quad f'(a) > 0$$

- Force of infection in population:

$$\ell(a) = \pi(a)f'(a)$$

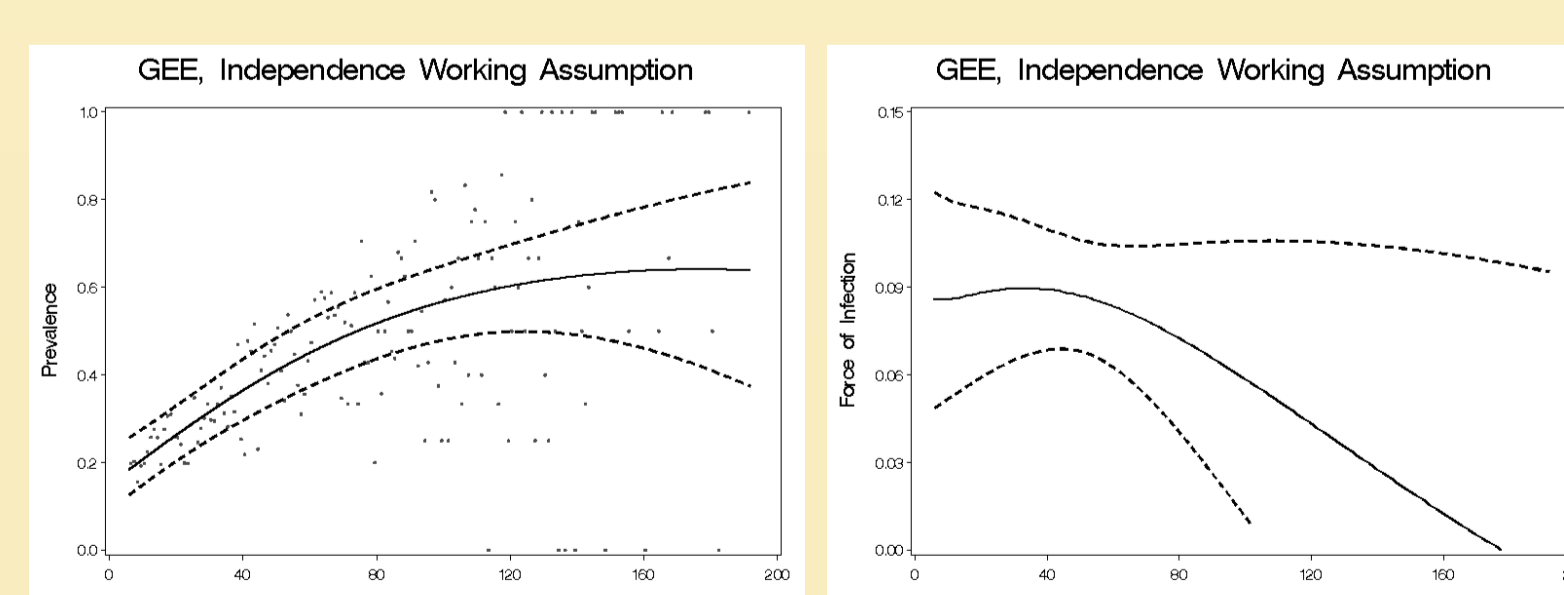


Figure 3: Sero-prevalence plot (left) and force of infection (right) based on generalized estimating equations.

- Larger confidence bounds are the correct ones
- Population-averaged trend

Clustering is of Interest

- Generalized random-effects model
- Herd-specific approach
- Interest in clustering

$$Y_{ij}|u_i \sim \text{Bernoulli}(\pi_{ij})$$

$$\eta_{ij} = \log\left(\frac{\pi_{ij}}{1 - \pi_{ij}}\right) = f(a) + u_i,$$

$$u_i \sim N(0, D)$$

- Use fractional polynomials (Royson and Altman 1994) together with constraint of monotonicity:

$$f(a) = \beta_0 + \beta_1 a^{p_1} + \beta_2 a^{p_2}, \quad f'(a) > 0$$

- Force of infection in herd:

$$\ell(a|u_i) = \pi(a|u_i)(f'(a) + u_i)$$

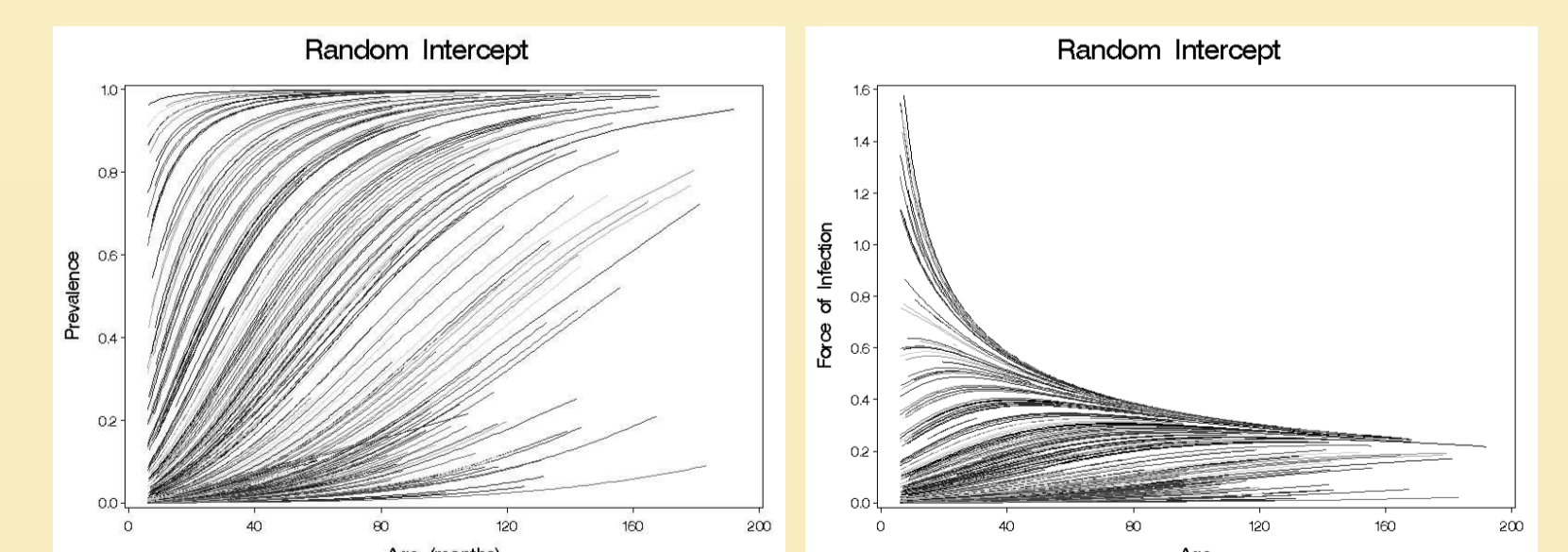


Figure 4: Sero-prevalence plot (left) and force of infection (right) based on random effects model.

- Large differences between herds exist
- Herd-specific characteristics have major impact on the disease dynamics

Conclusion

The BoHV-1 data were analyzed to determine the FOI. It is clear from the presented results that the dataset suffers from several complications. Correct modelling of the quantities of interest is crucial. Ignoring the correlations in the data underestimates the variances. Both population-averaged and herd-specific prevalence and FOI give information about the pattern of the disease, and should be considered.