## What drives sample size for detecting a pathogen? Part 2.

Key Points:

- Risk of lateral infection over the sampling period and correlation of results between consecutive weeks have a major influence on our confidence that a herd that is testing negative is actually free of infection.
- Detection of clinical signs can increase confidence, but is obviously ineffective for strains of low virulence.
- Models of this type can be used as an aid for establishing optimal sample strategy.

Last week we described a sampling model that can be used for pathogen detection that is being developed by Dr. Ana Alba Casals, a post doc in our Swine Group. Ana used the model to compare sampling approaches for two scenarios; (1) a multiplier that intends to become PRRSv negative (status 4), and (2) a commercial herd that intends to be stable (status 2vx or 2fvi). At 12 weeks after infection, we take our first sample of 10 sera and 2 pools were negative by PCR. We then increase the sample size as follows:

- Multiplier with very low incidence (once every 10 years)
  - 10 →30, 30, and then 60 / month thereafter
  - 10 → and then 30 / month thereafter
- Commercial herd – with high incidence of lateral infection over the sampling period (once every 3 years)
  - 10 / month
  - 10 → and then 30 / month thereafter

Other than sample size, all other factors such as population size, test sensitivity and specificity, initial probability of being infected or design prevalence, are assumed to be equal for each scenario. We can compare our overall confidence that the herd has <5% prevalence by computing the "Area Under the Curve (AUC)". First, we compare AUC for each scenario under the assumption that consecutive weeks are perfectly correlated. That is more likely to be the case late in the closure period when virus may have been eliminated. And then we compare AUC when consecutive weeks are moderately correlated (r=0.71). This is the case in the middle weeks of closure emphasizing the need for consecutive samples.

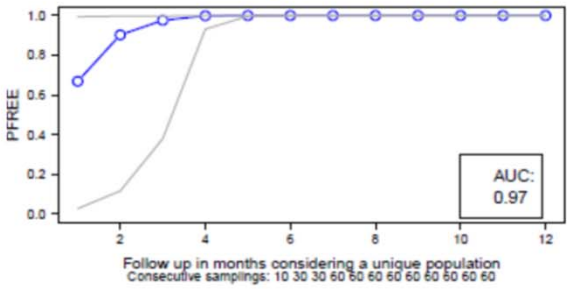
| | SCENARIOS | | | |
|---|---|---|---|---|
| **INPUTS** | **A** | **B** | **C** | **D** |
| Number of pigs | 3000 | 3000 | 3000 | 3000 |
| Detection prevalence | 0.05 | 0.05 | 0.05 | 0.05 |
| Total # samples | 610 | 340 | 120 | 340 |
| Test sensitivity | 1 | 1 | 1 | 1 |
| Test specificity | 1 | 1 | 1 | 1 |
| Initial probability of infection | 0.49 | 0.47 | 0.47 | 0.47 |
| Probability of incursion between consecutive samplings | 0.0083 | 0.0083 | 0.0278 | 0.0278 |
| Median correlation between groups | 0.71 | 0.71 | 0.71 | 0.71 |
| **OUTPUTS** | | | | |
| Cost of samplings | $3,050 | $1,700 | $600 | $1,700 |
| Area under curve (with correlation of 1between consecutive weeks ) | 97% | 97% | 89% | 96% |
| Area under curve (our overall confidence, correlation <1 between consecutive weeks ) | 93% | 85% | 43% | 84% |

**References**

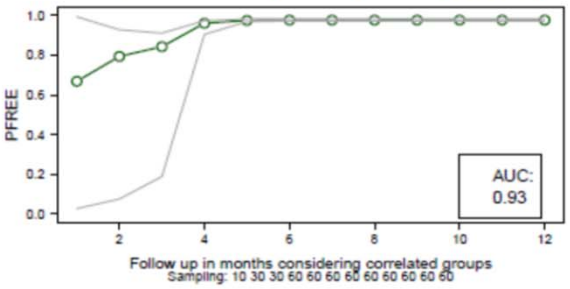
1. Cameron, A.R., Baldock, F.C., 1998. A new probability formula for surveys to substantiate freedom from disease. Prev.Vet.Med. 34, 1-17.
2. Martin, P.A.J., Cameron, A.R., Greiner, M., 2007. Demonstrating freedom from disease using multiple complex data sources. A new methodology based on scenario trees. Prev.Vet. Med. 79, 71-97.
3. P.A.J. Martin. 2008. Current value of historical and ongoing surveillance for disease freedom: Surveillance for bovine Johne's disease in Western Australia. Prev. Vet. Med. 84, 291-309

CONTINUED ON NEXT PAGE

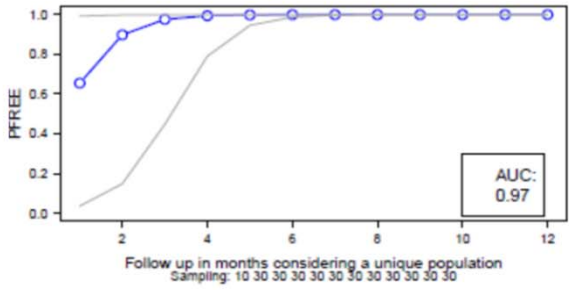
A: Multiplier farm with very low incidence taking 610 samples



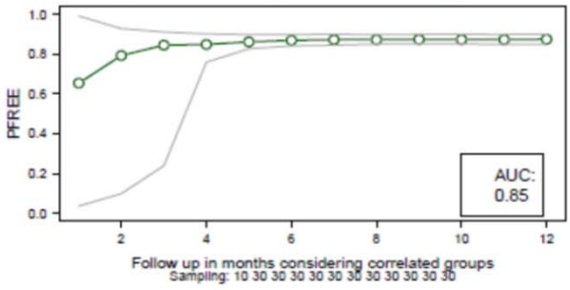

A: Multiplier farm with very low incidence taking 610 samples



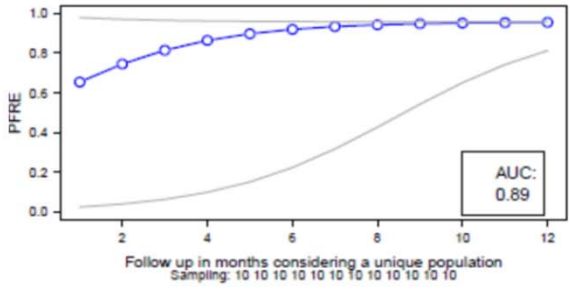

B: Multiplier farm with very low incidence taking 340 samples



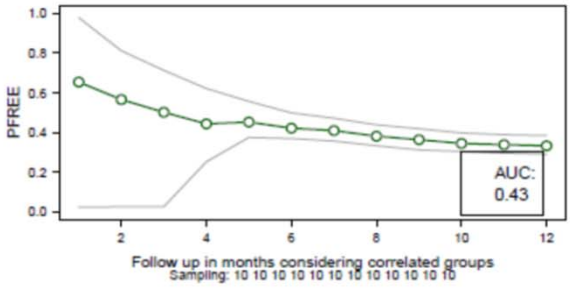

B: Multiplier farm with very low incidence taking 340 samples



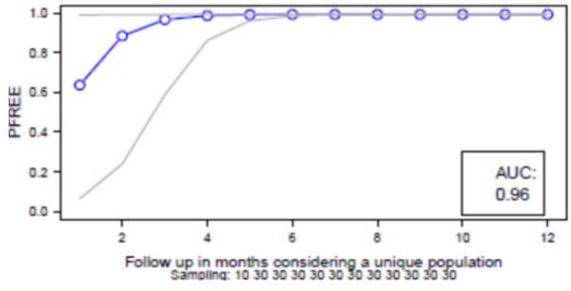

C: Commercial herd with medium incidence taking 120 samples




C: Commercial herd with medium incidence taking 120 samples




D: Commercial herd with medium incidence taking 340 samples




D: Commercial herd with medium incidence taking 340 samples