RESEARCH REPORT
SWINE HEALTH

Title: Design and analysis of PRRSv surveillance: temporal and spatial sampling, mapping, monitoring and automated rapid detection of outbreak – **NPB #11-165**

Investigator: Chong Wang

Institution: Iowa State University

Date Submitted: 03/06/2014

Industry Summary:

Innovation of cost effective methods for eliminating the PRRSv from individual herds has stimulated hope that the industry may someday eliminate the virus from the U.S. However, exploiting this innovation has been hampered by the industries lack of progress on preventing the frequent transmission of the virus from one herd to another. Recently, several regional elimination projects have been initiated by producers and veterinarians to overcome this hurdle. These projects are efforts to reduce the frequency of transmission from one herd to another by better understanding what the PRRS virus is doing in the region, improving biosecurity and reducing the prevalence of the virus in the region. Surveillance is a critical element of all of these projects for better understanding what the virus is doing in the region and for measuring progress.

Although sequential testing is widely needed and used in practice for PRRSv surveillance, there is little guidance on how frequently to sample. This study provides cost effective methods for PRRSv surveillance including guidance on the frequency of sampling with a firm theoretical basis. The methodologies we develop can be expected to provide a standard framework for design and analysis of PRRSv surveillance studies. We study the disease progression and transmission of PRRS through a proposed statistical model and provide rigorous, detailed, data-based statistical framework for design and analysis of PRRS surveillance. The proposed adaptive design will help detect PRRS outbreak earlier. The methodologies developed are essential for effective control and elimination of PRRS virus on individual farms and for regional elimination projects.

We develop a web-based interface called SSF (Sample Size and Frequency), built upon the Shiny web-application framework. SSF provides easy-to-use and instantly displayed calculation of sample size and frequency based on a custom-defined scheme of their own choosing.

Keywords: Sample size, sample frequency, surveillance, temporal, spatial

These research results were submitted in fulfillment of checkoff-funded research projects. This report is published directly as submitted by the project's principal investigator. This report has not been peer-reviewed.

For more information contact:
National Pork Board • PO Box 9114 • Des Moines, IA 50306 USA • 800-456-7675 • Fax: 515-223-2646 • pork.org

Scientific Abstract:

In animal disease testing at the population level, traditional calculation methods relate the power of detection to sample size using probability models, e.g., binomial distribution.1 These probability models are based on the assumption that all individuals have the same chance of acquiring disease. In contemporary animal production systems, however, a hierarchical structure exists. For example, a farm may consist of multiple buildings and each building usually contains multiple pens of animals. Farm sample size may be calculated using classic formulas, but there is a lack of guidance about how to subsequently allocate sampling across strata. Despite the fact that strata are not necessarily homogenous in terms of disease status, it is common for samples to be collected from selected strata, leaving other strata unsampled: (1) for convenience, (2) for lack of formal guidance regarding sample allocation across strata, and (3) under the assumption that certain strata are most representative of the farm's disease status. For optimal disease detection at the farm level, a mathematically more intuitive way to allocate samples would be even distribution across the different strata. But does sampling from one stratum vs. multiple strata differ in the power of disease detection? If multiple strata sampling is preferable, is even distribution of samples across the strata the optimal strategy? If not, what are the formulas for sample size calculation and allocation for populations with a two-level structure? The objective of this project is to address these questions from a mathematical perspective.

A pivotal part of disease surveillance is the repeated disease diagnostic testing to monitor the status of the disease. Despite their importance, little work has been done to address the common questions of sampling frequency as well as sample size for repeated testing. We develop mathematical relationships between the sample size, sample frequency and the other parameters in disease detection, ie, the prevalence, the desired detection time and the desired power of detection. We also develop a web application called SSF (Sample Size and Frequency), built upon the Shiny web-application framework. SSF provides easy-to-use and instantly displayed calculation of sample size and frequency based on a custom-defined scheme of their own choosing.

Introduction:

Porcine reproductive and respiratory syndrome (PRRS), caused by the PRRS virus, is a major health, production and financial problem for swine producers in nearly every country. PRRS costs the United States swine industry around $560 million annually (Neumann et al., 2005). PRRS outbreaks in China caused pork prices to increase by 85 percent in 2006 (Li et al., 2007). Surveillance of PRRS has been of great interest.

A pivotal part of disease surveillance is the sequential disease diagnostic testing, which establishes patterns of disease prevalence progression for the study population. In design of sequential diagnostic testing, two questions have to be answered: 1. How many animals need to be tested? and 2. How often does the testing need to be done? Although there exist classic approaches to address the first, sample size question for testing at a single time point, little has been done to address the second, frequency question or answer the two questions together. Sample size calculation for disease freedom testing at a single time point has been extensively used in designing infectious disease studies. For disease testing with imperfect diagnostic tests, Cameron and Baldock (1998) developed mathematical formulas to calculate samples sizes, through probabilistic modeling of the relationship among population size, disease prevalence, diagnostic test sensitivity and specificity, hypothesis testing confidence and power, and sample size. Cannon (2001) derived fast approximation formulas for the above calculation. Such works provide a rigorous theoretical basis for design and analysis of cross-sectional animal disease testing studies.

In disease surveillance, series of testing needs to be carried out to establish trend/history of disease status. In a surveillance study, the confidence in disease status is not only related to sample size at each testing time, but also the frequency of repeating the diagnostic testing. For example, in the recent PRRS-free certification pilot project protocol of the Canadian Swine Health Board, testing frequency has a direct positive relationship with the farm’s score to gain

PRRS-free certification. Unfortunately very little work has been done to address the sampling frequency question in design of animal disease surveillance. The only work that the investigators are aware of is the simulation based method by Rovira et al. (2007). Such simulation-based methods, however, are limited by the simulation size, computing speed, lack of exactness and lack of rigorous theoretical basis.

The theory of sequential testing in surveillance can be traced back to quality control studies in engineering, where the question of interest was fault detection (Lai 1995). Yet such quality control studies do not involve transmission of disease pathogens, thus are not appropriate for surveillance of infectious disease. In human medicine, statistical algorithms have been proposed to analyze infectious disease surveillance data (Farrington et al. 1996), but not much work has been done on the design aspect. This is because human diagnostic testing is on a voluntary basis, as patients intentionally seek disease diagnoses and treatments. In animal disease surveillance, testing is initiated and costs are paid by producers thus the design needs to be cost-efficient.

Objectives:

The project will focus on development of sequential sampling schemes for PRRSv surveillance, the analysis of sequentially sampled PRRSv surveillance data and adaptive designs to rapidly detect change in disease prevalence based on analysis of present data. These algorithms will be made available to producers and swine veterinary practitioners in a user-friendly web-based application.

Materials & Methods:

1. Spatial Sample Allocation

Two strategies for sample allocation with perfect tests:

In a population with k strata, e.g., a site with k buildings, denote the disease prevalence in each stratum by p_i, i =1,…,k. Let n_i be the sample size for the i'th stratum, e.g., pen within a building, in a certain sampling strategy. The number of diseased animals m_i sampled from the i'th stratum is commonly modeled using a binomial distribution, $m_i \sim Bin(n_i, p_i)$. The population is classified positive for disease if at least one sample is test positive, i.e., $m_1 + ... + m_k \geq 1$. In this section, we compare the following two strategies in terms of their power of detection, $P(m_1 + ... + m_k \geq 1)$.

Method 1. Sample n animals from each stratum, i.e., $n_1 = n_2 = ... = n_k = n$.

Method 2. Sample k×n animals from one randomly selected stratum, i.e., $n_j = kn$ for some j and $n_i = 0$ for i≠j.

The power of detection for Method 1 can be derived as

$$P(m_1 + ... + m_k \geq 1) = 1 - P(m_1 = ... = m_k = 0) = 1 - P(m_1 = 0)... P(m_k = 0)$$
$$= 1 - (1 - p_1)^n ...(1 - p_k)^n \quad (1)$$

If Method 2 is used, then each stratum has 1/k chance to be selected. If the i'th pen is chosen, the power of detection in this stratum is then $P(m_i \geq 1) = 1 - P(m_i = 0) = 1 - (1 - p_i)^{kn}$. Overall, the chance of detection is

$$P(m_1 + ... + m_k \geq 1) = \frac{1}{k}(1 - (1 - p_1)^{kn}) + ... + \frac{1}{k}(1 - (1 - p_k)^{kn})$$
$$= 1 - \frac{1}{k}\sum_{i=1}^{k} (1 - p_i)^{kn} \quad (2)$$

Two strategies for sample allocation with imperfect tests

Most diagnostic tests used in disease detection are imperfect. Denote the sensitivity of a diagnostic test by θ, where θ is not necessarily 1. Then, not all diseased animals sampled will be test positive. Denote the number of diagnostic test positive animals by x_i, i =1,…,k. Then $x_i \mid m_i \sim Bin(m_i, \theta)$. Together with the model $m_i \sim Bin(n_i, p_i)$, it can be derived that the distribution of x_i unconditional on m_i is $x_i \sim Bin(n_i, p_i\theta)$. Similar to the derivation in the case of a perfect diagnostic test in section 2.1, the power formulas for the two sampling strategies can be derived as follows:

Method1:

$$P(x_1 + ... + x_k \geq 1) = 1 - P(x_1 = ... = x_k = 0) = 1 - (1 - p_1\theta)^n ...(1 - p_k\theta)^n \quad (4)$$

Method2:

$$P(x \geq 1) = 1 - P(x = 0) = \frac{1}{k}(1 - (1 - p_1\theta)^{kn}) + ... + \frac{1}{k}(1 - (1 - p_k\theta)^{kn})$$

$$= 1 - \frac{1}{k}\sum_{i=1}^{k}(1 - p_i\theta)^{kn} \quad (5)$$

Optimal strategy for sample allocation

For one-level population disease detection, the common goal can be stated as to test against a certain non-zero disease prevalence p>0. Analogically, the goal of a two-level disease detection can be formed as testing again a distribution of the disease prevalence, p~f(p), with the mean prevalence E(p)>0 and certain heterogeneity Var(p)>0. Under the assumption that the prevalences follow such distribution of heterogeneity, $p_1,..., p_k \sim f(p)$, we can derive the power of detection as a function of the sample allocation n_i, i=1,2,…,k:

$$P(X_1 + \ldots + X_k \geq 1) = 1 - \int (1 - p_1\theta)^{n_1} ...(1 - p_k\theta)^{n_k} f(p_1)... f(p_k) dp_1 ... dp_k$$

$$= 1 - \prod_{i=1}^{k} \int_0^1 (1 - p_i\theta)^{n_i} f(p_i) dp_i$$

$$= 1 - E(1 - p\theta)^{n_1} ... E(1 - p\theta)^{n_k} \quad (6)$$

2. Sample Size and Frequency for Repeated Sampling

Assume that there are *N* sampling units in the population among which *M* are diseased. Thus the population disease prevalence is *M/N*. Suppose the sampling scheme is to sample *n* units at each time, and sample every *d* days. Here we derive the relationship between the time of detection, *T*, and the sampling parameters *n* and *d*. Let *m* denote the random number of positive pens being sampled, then *m* has a hypergeometric distribution with probability mass function $P(m) = \binom{M}{m}\binom{N-M}{n-m} \Big/ \binom{N}{n}$. Considering detection with an imperfect diagnostic test with sensitivity θ (chance of a diseased unit to be test positive). We prove the chance of detection at a single time of sampling is

$$p = P(\text{Detection at a single time}) = \sum_{x=1}^{\min(m,n)} \binom{M}{x}\binom{N-M}{n-x} \Big/ \binom{N}{n} \cdot \left[1 - (1 - \theta)^x\right]. \quad (7)$$

The chance of detection at a single time might be low for detection of disease of low prevalence with small sample size *n*. With repeated testing every *d* days, the chance of detection will increase. The time of detection *T*, can be shown to follow a geometric distribution with probability $P(T = t) = (1 - p)^{t/d} p$, where *t = 0, d, 2d, 3d, …* and $p = P(\text{Detection at a single time})$ in

formula (1). Based on this, we can derive the cumulative chance of detection by certain time t to be:

$$P(\text{Detection by time } t) = 1 - (1 - p)^{\lfloor t/d \rfloor + 1}, \quad (8)$$

where $\lfloor t/d \rfloor$ indicates the largest integer not greater than t/d. The expected time of detection is derived to be:

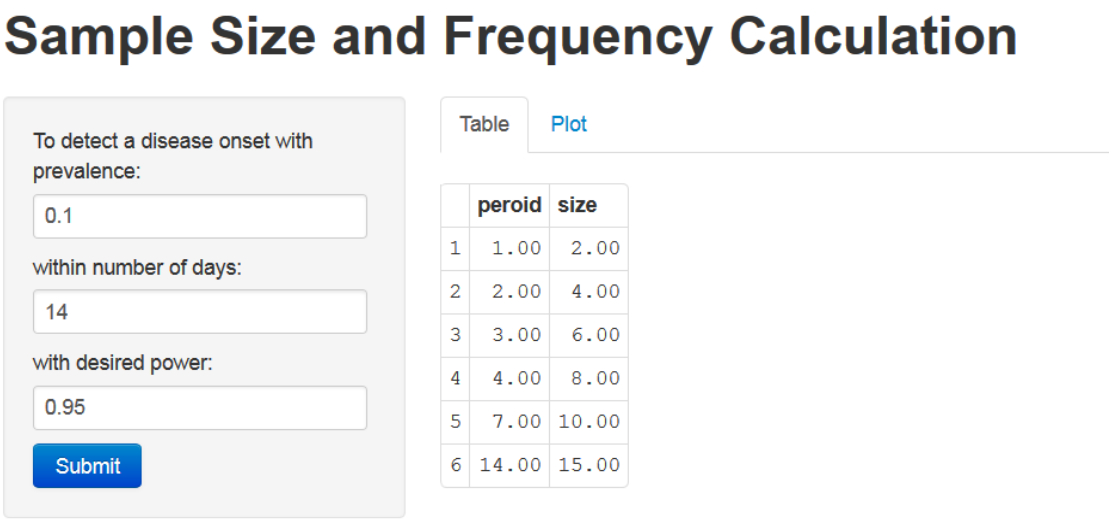
$$E(T) = d(1 - p)/p. \quad (9)$$

Formula (8) shows that chance of detection by certain time is an increasing function of sample size n and sample frequency $1/d$. In practice, if we want to detect the disease with certain chance by a desired time, we can use formula (8) to calculate the required combination of sampling time and frequency.

We have also developed a web application called SSF (Sample Size and Frequency), built upon the Shiny web-application framework. SSF is an application of Shiny, a framework for writing web-applications in the R language. SSF uses this framework to provide quick and convenient calculations of sample size and frequency for design of repeated diagnostic testing. For instance, when the user of SSF changes a value of parameter in setting, such as the prevalence, the result tables and plots will instantly update to reflect the new setting. SSF is currently hosted on the server provided by Iowa State University.

Similar to other Shiny applications, SSF consists of three primary UI components, the Configuration Panel, the Tab Panel, and the Results Panel, as shown in Figure 1. The Configuration Panel is located along the left-hand column. This panel allows for various parameters to be adjusted. The top of the right panel is the Tab Panel, which contains two tabs corresponding to table and plot, respectively. The bottom of the right panel is the Results panel, which will contain the results of the analysis depending on the tab and configuration options selected.

Figure 1. Screenshot of SSF, with tabular results.



Results:

1. Spatial Sample Allocation

Comparison of two strategies for sample allocation with perfect tests

Comparison between formulas (1) and (2) reveals a direct relationship between the power of Method 1 and Method 2:

Theorem 1. Method 1 is superior in detection power to Method 2.

Proof. By mathematical theory, the arithmetic mean is no less than the geometric mean. Thus $\frac{1}{k}(1-p_1)^{kn} + \ldots + \frac{1}{k}(1-p_k)^{kn} \geq \left((1-p_1)^{kn} \ldots (1-p_k)^{kn}\right)^{\frac{1}{k}}$. This directly leads to Theorem 1.

The equality between arithmetic mean and geometric mean holds if and only if $p_1 = \ldots = p_k$, i.e., all the strata have exactly the same prevalence. This is not a practical assumption. Thus generally, Method 1 should be preferred to Method 2 for the detection of disease in a two-level structure.

Comparison of the two strategies for detection with imperfect tests

Theorem 2. Method 1 is superior in detection power to method 2 when an imperfect diagnostic test is used.

Proof. By mathematical relationship between the arithmetic mean and the geometric mean, $\frac{1}{k}(1-p_1\theta)^{kn} + \ldots + \frac{1}{k}(1-p_k\theta)^{kn} \geq \left((1-p_1\theta)^{kn} \ldots (1-p_k\theta)^{kn}\right)^{\frac{1}{k}}$. Theorem 2 is proved.

Again, the equality holds if and only if all strata have exactly the same prevalence, which is not a practical assumption. Therefore, Method 1 is always equal to, or better than, Method 2.

Optimal strategy for sample allocation

Sections 2.1 and 2.2 compared two specific strategies and proved that even allocation (Method 1) is better. Is Method 1 the best among all sample allocation strategies? Intuitively, if we knew the exact prevalence values, $p_1,\ldots, p_k$, then we could improve the power of Method 1 by taking all samples from the most prevalent stratum. However, if we knew the exact prevalences, it would be unnecessary to test in the first place. In practice, it can only be assumed that prevalence varies among strata.

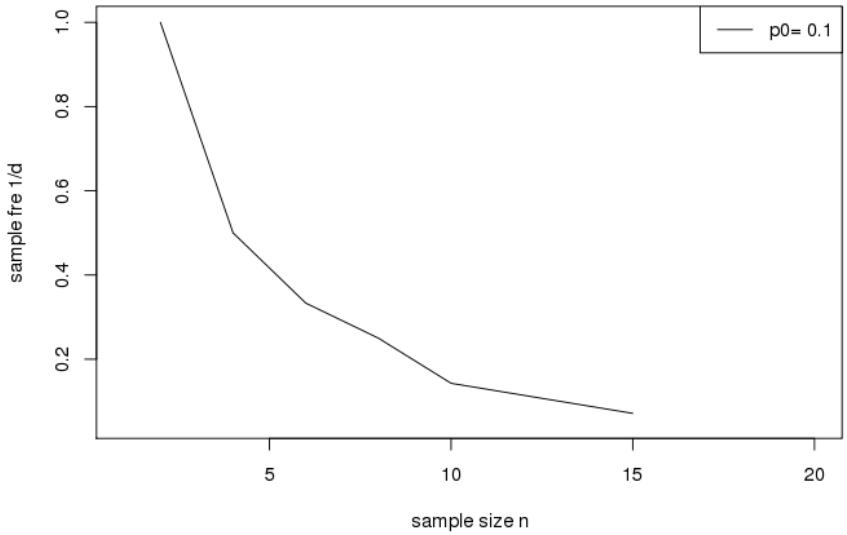
Theorem 3. In case $p_1,\ldots, p_k \sim f(p)$, the power of detection $P(X_1 + \ldots + X_k \geq 1)$ is optimized when $n_1 = n_2 = \ldots = n_k$.

Proof. Based on the power expression in equation (6), we need to prove $1-(E(1-p\theta)^n)^k \geq 1 - E(1-p\theta)^{n_1} \ldots E(1-p\theta)^{n_k}$, for any allocation strategy n_i, i=1,2,…,k with the same overall sample size. This is equivalent to $E(1-p\theta)^{n_1} \ldots E(1-p\theta)^{n_k} \geq (E(1-p\theta)^n)^k$. The proof can be achieved through mathematical induction and Hölder's Inequality.

2. Sample Size and Frequency for Repeated Sampling

SSF provides easy-to-use and instantly displayed calculation of sample size and frequency based on a custom-defined scheme of their own choosing. The Configuration Panel allows users to input parameter values in calculation of sample size and frequency. The parameters include the prevalence, the desired detection time and the desired power of detection. For example, if the user specifies the parameter values for detecting a disease onset of prevalence 0.1 within 14 days with desired power 0.95, combinations of sample size and frequency are instantly calculated and displayed in table (Figure 1) or plot (Figure 2).

Figure 2. Sample size vs frequency plot in result section.



Discussion:

We have mathematically proved that even allocation of sample size to strata (eg building, pen) has the optimal power, among all possible sample allocation strategies. This implies that in practice, disease surveillance should not be done by focusing on one or a few second-level strata, even if these few strata are randomly chosen. For any given total number of samples, sampling as many strata as possible provides the best power for disease detection.

We have also developed mathematical relationships between the sample size, sample frequency and the other parameters in disease detection, ie, the prevalence, the desired detection time and the desired power of detection. We develop a web application called SSF (Sample Size and Frequency), built upon the Shiny web-application framework. By utilizing the R language and the web application framework Shiny, SSF allows for a quick and convinient calculation of sample size and frequency in repeated diagnostic testing. It is flexible enough to allow investigation of a wide variety of different scenarios with varying prevalence, desired detection time and detection power. The application is web-based and easy to use, making the algorithm available to practitioners outside the field of statistics.