Diagnostic Tests II

Solymosi Norbert

Quantitative veterinary epidemiology

Department of Microbiology and Infectious Diseases
University of Veterinary Medicine Budapest

Lecture 4

- To improve diagnostic accuracy, tests can be **repeated**, or **additional** tests may be involved.
- In fact, most diagnoses are based on multiple tests (e.g., medical history, physical examination, laboratory tests).
- Multiple tests can be applied simultaneously or consecutively, and the results can be interpreted in **parallel** or **serial**.
- Sensitivity and specificity values of test combinations **differ** from the sensitivity and specificity of individual tests.
- In interpreting results from a combination of tests, it is a fundamental assumption the tests must be **independent** of each other.
- If this independence is not satisfied, then the accuracy improvement will be **lower** than theoretically expected.

- Such **correlated** results are expected when the combined tests measure the same/similar characteristics of the sample, but less likely when different biological responses (e.g. histopathological and serological testing) are the target of the tests.
- In **parallel** testing, the sensitivity will be higher than the sensitivity of any individual test used.

$$Se_{par} = 1 - (1 - Se_1) \times (1 - Se_2)$$

$$Sp_{par} = Sp_1 \times Sp_2$$

- In **serial** testing, the specificity will be higher than the specificity of any individual test used.

$$Se_{ser} = Se_1 \times Se_2$$

$$Sp_{ser} = 1 - (1 - Sp_1) \times (1 - Sp_2)$$

- If two tests are used, one of the following four results is possible:
  - both are positive
  - both are negative
  - the first test is negative and the second one is positive
  - the first test is positive and the second one is negative
- In a parallel interpretation, an animal is considered positive if one test is positive - this increases the sensitivity of the combined tests, but reduces its specificity.
- This parallel testing strategy is useful when none of the tests has a particularly high sensitivity, but they can detect different types of the disease (e.g. early - late, fast-slowly progressing).
- Culturing may be more sensitive than serological tests in the early stages of infection. Still, serology may be more sensitive in a later stage of it when the amount of pathogens is lower.

- In serial testing, both consecutive tests must be positive to identify the animal as positive - this increases the specificity of the combined tests, but reduces its sensitivity.
- The first test can be high sensitivity and inexpensive; the result can be followed by a high-specificity test to determine false positives.
- It is a cost-effective approach if the first test negatives are not tested by the second test.
- This strategy allows the vet to use fewer tests to rule out the disease; however, it is more time-consuming.
- When both tests are positive, for the estimation of disease probability, the first test positive predictive value will be the pre-test probability of the second test.

- For example, for a test A, the positive predictive value was 67.9% applied in a herd with prevalence 20%.
- If an animal tested by A was positive and was retested with another test B (sensitivity of 45.9%, specificity of 96.9%) the positive predictive value of 67.9% is considered as the pre-test probability for B-test.
- Using the Bayes theorem, the positive predictive value after the application of the second test will be 96.9%, assuming that the results of A and B are not correlated.
- If the condition of uncorrelation of the tests can be maintained, then the two positive values obtained by using A and B together are a stronger indicator of infection than it was predicted by test A alone.

Serological and molecular detection of *Theileria equi* infection in horses in Hungary

R. Farkas[a,*], B. Tánczos[a], M. Gyurkovszky[a], G. Földvári[a], N. Solymosi[b], R. Edelhofer[c], S. Hornok[a]

[a] *Department of Parasitology and Zoology, Faculty of Veterinary Science, Szent István University, H-1078, Budapest, István u. 2, Hungary*
[b] *Department of Animal Hygiene, Herd-health and Veterinary Ethology, Faculty of Veterinary Science, Szent István University, H-1078, Budapest, István u. 2, Hungary*
[c] *Institute of Parasitology and Zoology, Department of Pathobiology, University of Veterinary Medicine, Veterinärplatz 1, 1210-Vienna, Austria*

ARTICLE INFO

Article history:
Received 22 July 2012
Received in revised form
25 September 2012
Accepted 27 September 2012

Keywords:
T. equi
Horse
cELISA
IFAT
PCR
Hungary

ABSTRACT

The prevalence of *Theileria equi* infection was studied in 324 healthy horses from 27 farms in Hungary with cELISA and IFAT and the blood samples of 101 horses selected randomly were also examined by PCR. The results indicate that there are many stud farms where one or more horses are infected with *T. equi*. Among 27 farms 17 (67.9%) were found to have seropositive horses. The seroprevalence of theileriosis among the tested stud farms ranged between 0 and 100%. No marked differences were found in seropositivity between geographical areas. The overall prevalence of positive samples was 32.0% with cELISA as well as with IFAT. The results obtained with cELISA and IFAT in this study had the strongest agreement, except for 9 samples in which the two serological tests gave different results. The prevalence of infection among 101 horses was 49% with PCR. All 14 sequenced samples were found by BLAST analysis to be closest to the *T. equi* 18S rRNA gene sequences in GenBank with a similarity of $\geq$99%.

No significant association was found between the seropositivity and the age of horses. Horses below 5 years of age had three times higher chance to be PCR-positive, than older ones. There was no significant association between the gender and the results of diagnostic tests (cELISA: $p = 0.40$; IFAT: $p = 0.25$; PCR: $p = 0.41$). Based on the findings, the prevalence of equine theileriosis is much higher than expected and it occurs in many regions of the country unlike equine babesiosis. To the authors' knowledge, this is the first report of the serological and molecular survey of *T. equi* infection in horses in Hungary.

Farkas et al. (2013)

Elanco Keto-Test

Study	Herds (n)	Sample size (n)	DIM	SE %	SP %	Prev
Belanger et al. (2003)	1	55	2-21	93	68	25.4
Carrier et al. (2004)	1	850	2-15	73	96	7.6
Geishauser et al. (2000)	21	469	1-7	80	76	12.0
Oetzel (2004)	17	221	?	87	83	17.2
Osborne et al. (2002)	1	248	1-15	95	69	16.5

IDEXX Milk Pregnancy Test

Pregnant: 923, open: 392

SE: 98.8% (95% CI: 97.7-99.3%); SP: 97.4% (95% CI: 95.2-98.6%)

IDEXX SNAP BVDV Ag Test (serum)

Positive: 211, negative: 215

SE: 95.9% (95% CI: 92.3-97.9%); SP: 100% (95% CI: 97.7-100%)

Prevalence represents the fraction of **existing cases** in a population:

- the ratio between the number of diseased animals and the total number of animals at risk
- the probability that a randomly-chosen animal is diseased
- pre-test probability, post-test probability

Point prevalence is the proportion of infected individuals in a defined population at a given time point.

Period prevalence is the proportion of infected individuals in a defined population **found over** a specified time period.

If the **case definition** is based on an imperfect test, the test bias should be taken into account in prevalence estimation.

- Apparent (P_A): the probability that a randomly-choosen unit of observation will test positive

$$P_A = \frac{x}{n}$$

- Apparent (P_A): the probability that a randomly-choosen unit of observation will test positive

$$P_A = \frac{x}{n}$$

- $x = 5$, $n = 20$

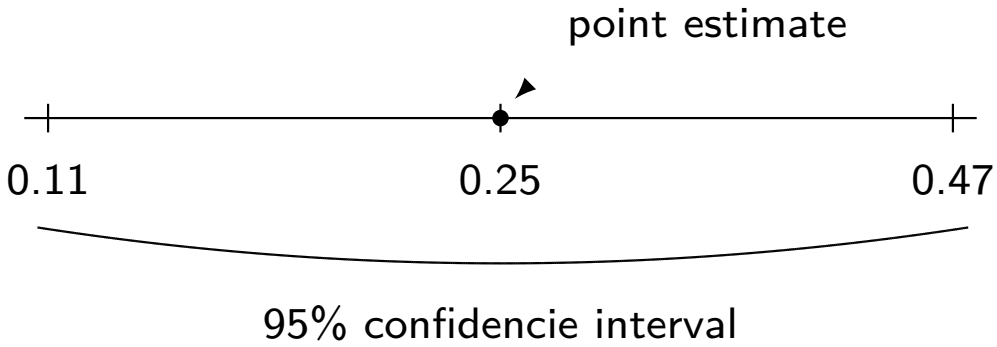
$$P_A = \frac{5}{20} = 0.25$$

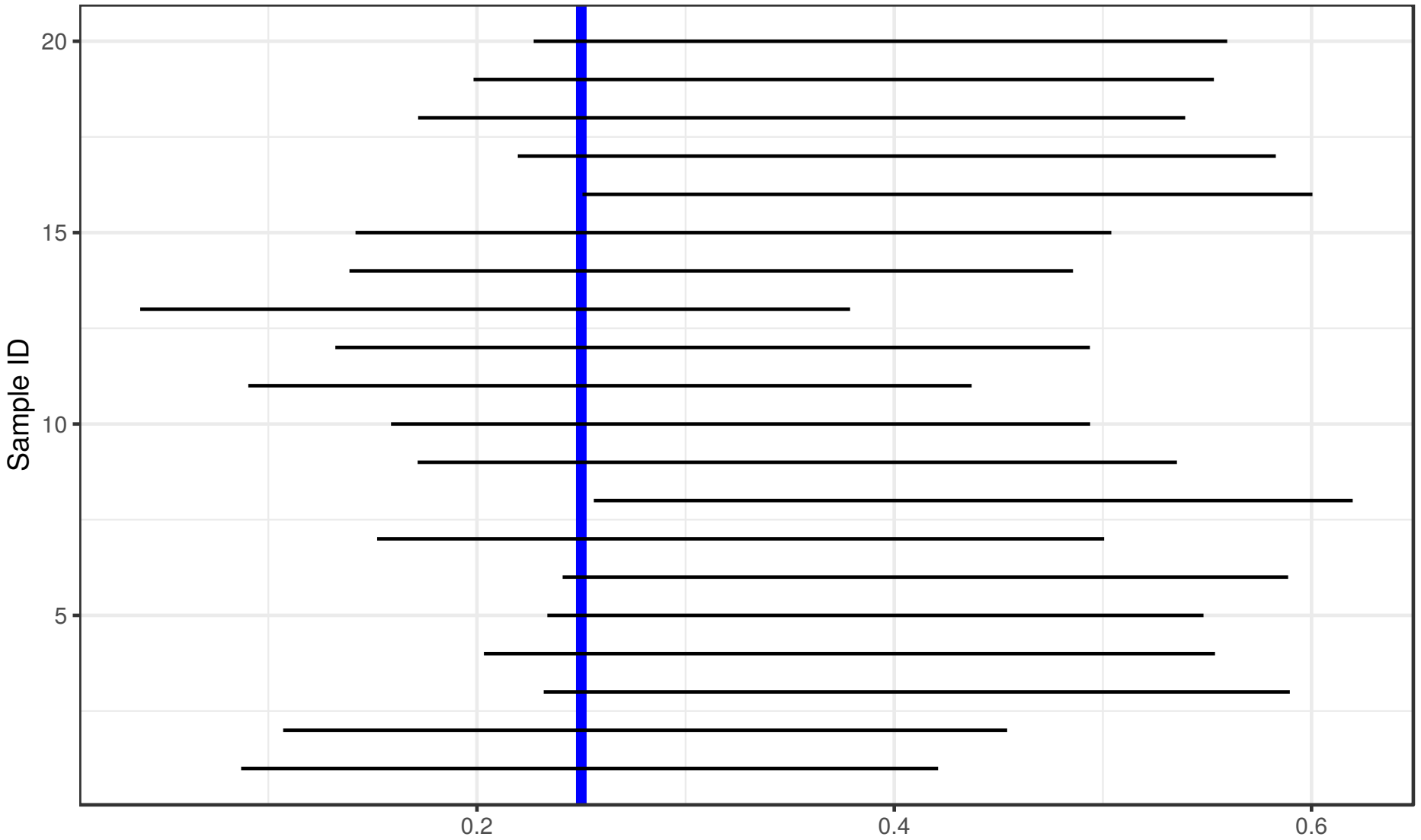
- Apparent (P_A): the probability that a randomly-choosen unit of observation will test positive

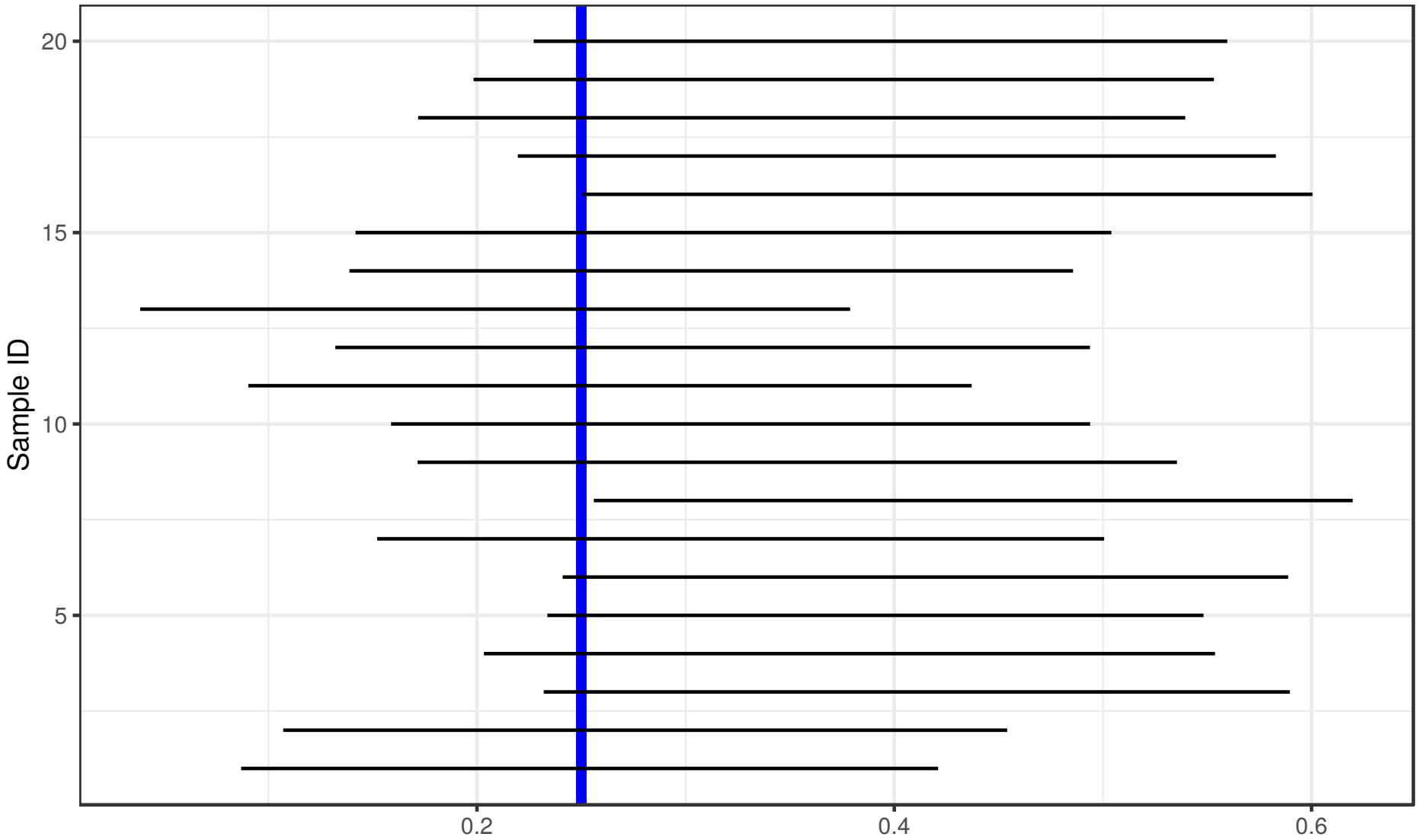
$$P_A = \frac{x}{n}$$

- $x = 5$, $n = 20$

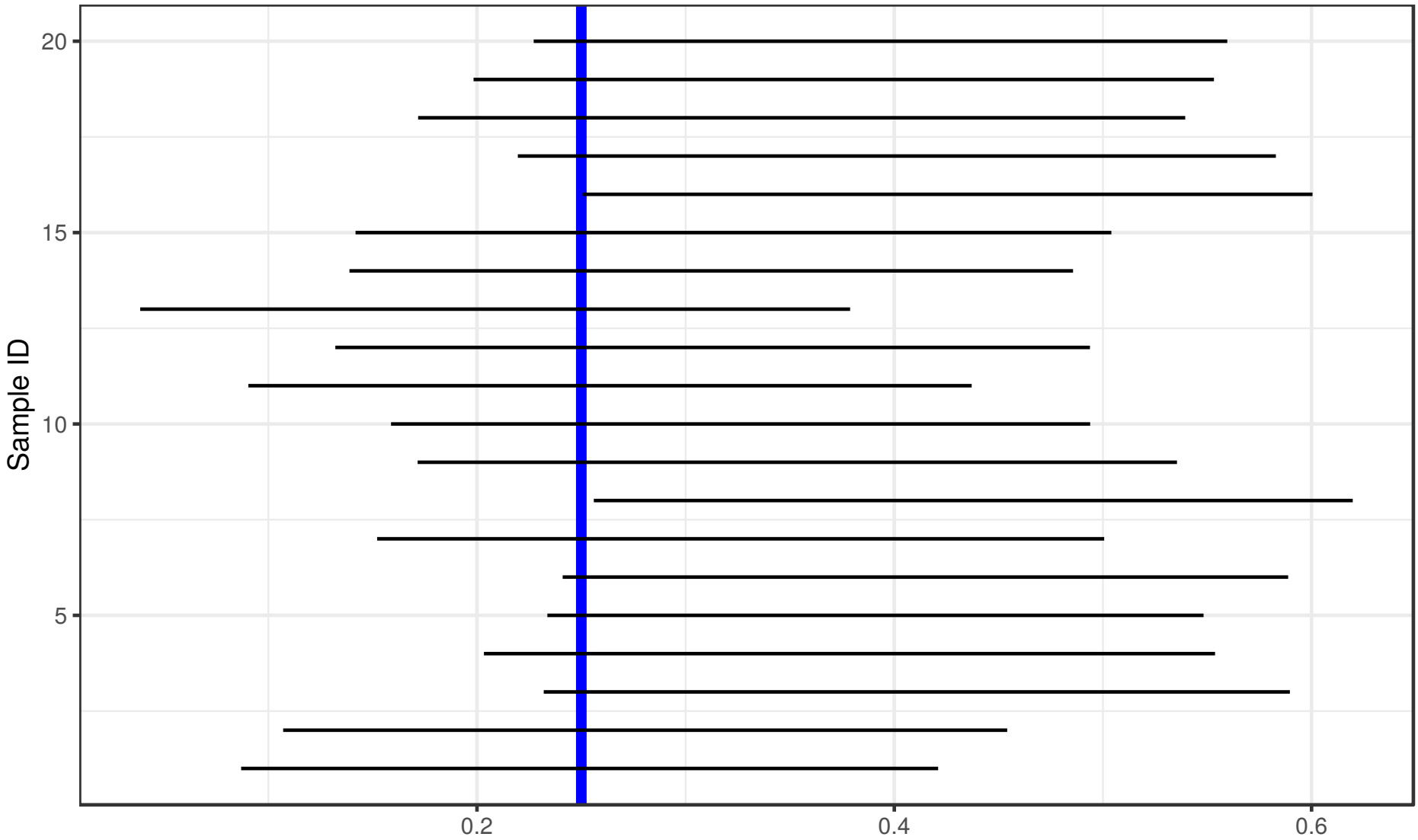
$$P_A = \frac{5}{20} = 0.25$$



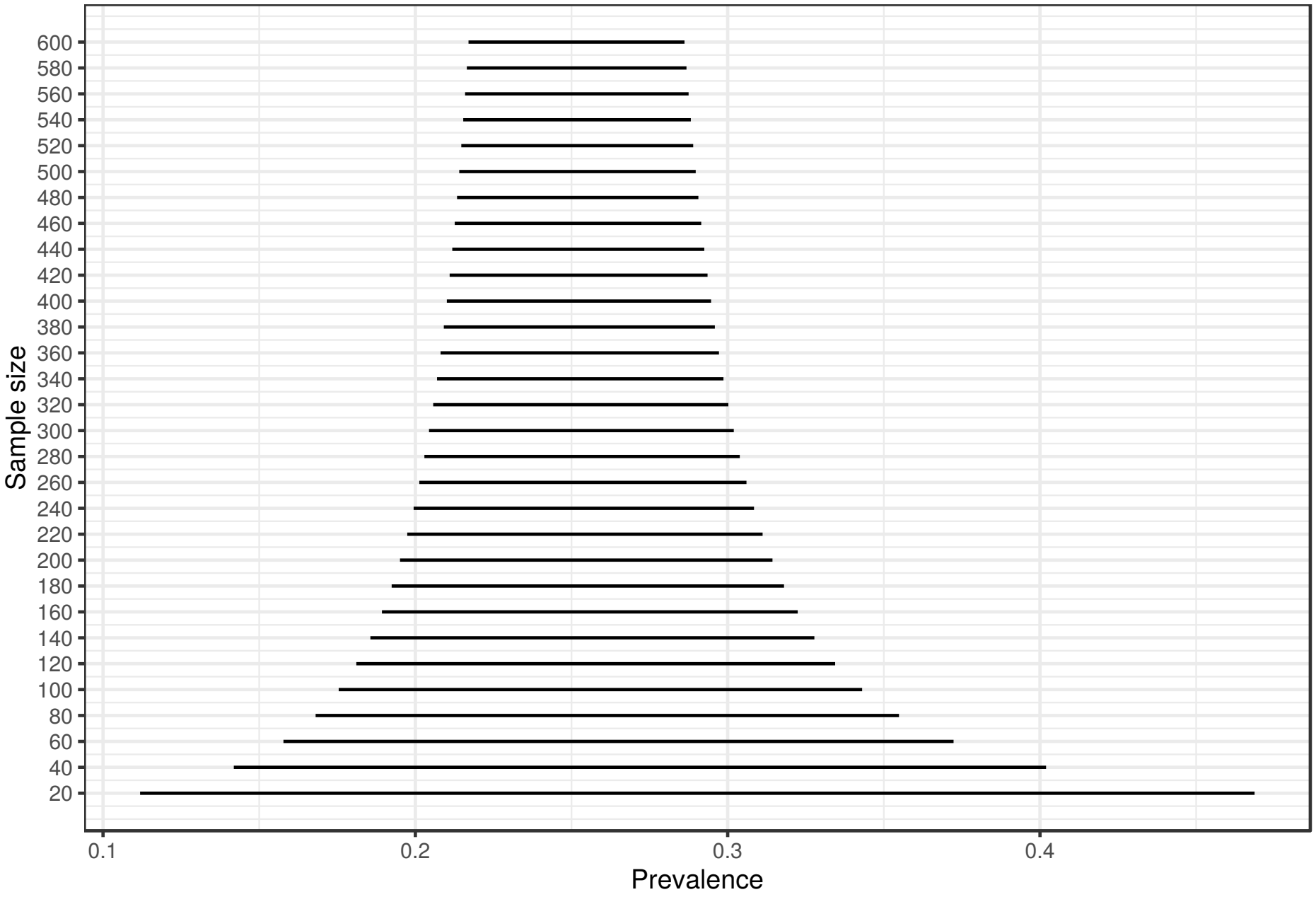

Sample ID
20
15
10
5
0.2
0.4
0.6



95% CI: 95 of 100 repeated samples will contain the prevalence 0.25



95% CI: 19 of 20 repeated samples will contain the prevalence 0.25


Sample size
600
580
560
540
520
500
480
460
440
420
400
380
360
340
320
300
280
260
240
220
200
180
160
140
120
100
80
60
40
20
0.1
0.2
0.3
0.4
Prevalence

- Apparent (P_A): the probability that a randomly-choosen unit of observation will test positive

$$P_A = \frac{x}{n}$$

- Apparent (P_A): the probability that a randomly-choosen unit of observation will test positive

$$P_A = \frac{x}{n}$$

- Bayesian estimation:
  - $n/N \leq 0.1$: $x \sim binomial(n, P_A)$
  - $n/N > 0.1$: $x \sim hypergeometric(N, n, P_A)$

- Apparent (P_A): the probability that a randomly-choosen unit of observation will test positive

$$P_A = \frac{x}{n}$$

- Bayesian estimation:
  - $n/N \leq 0.1$: $x \sim binomial(n, P_A)$
  - $n/N > 0.1$: $x \sim hypergeometric(N, n, P_A)$
- Diagnostic misclassification:
  - Sensitivity: $p(+|Infected) \neq 100\%$
  - Specificity: $p(-|Not\ infected) \neq 100\%$
  - Rogan-Gladen estimator:

$$P_T = \frac{P_A + Sp - 1}{Sp + Se - 1}$$

  - Bayesian binomial:

$$x|P_A, Se, Sp \sim binomial(n, P_T Se + (1 - P_T)(1 - Sp))$$

- Rogan-Gladen estimator:
- $x = 5,\ n = 20,\ Se = 0.3,\ Sp = 0.96,\ N = 675$

$$P_A = 5/20 = 0.25$$

$$P_T = \frac{P_A + Sp - 1}{Sp + Se - 1} = \frac{0.25 + 0.96 - 1}{0.3 + 0.96 - 1} = 0.807$$

- Rogan-Gladen estimator:
- $x = 5,\ n = 20,\ Se = 0.3,\ Sp = 0.96,\ N = 675$

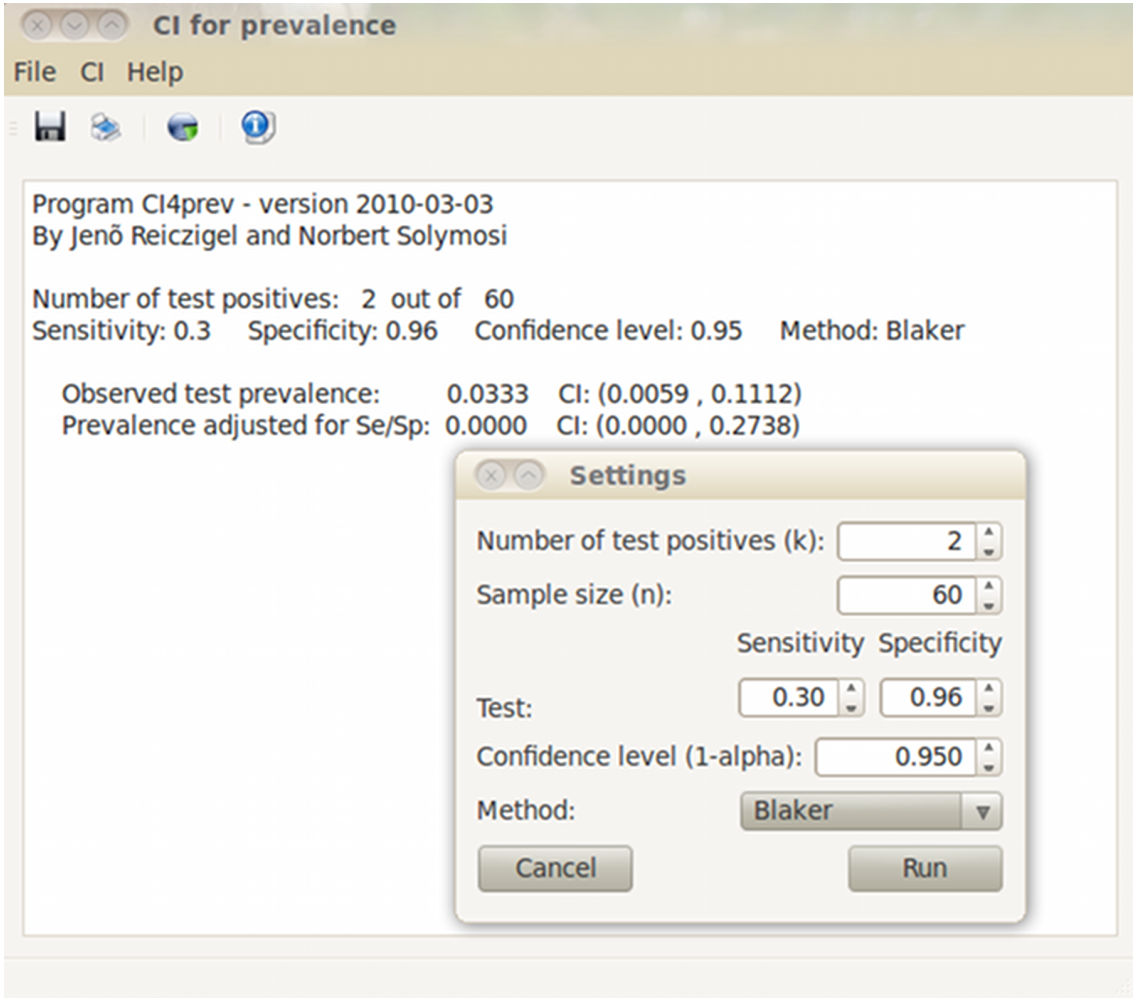
  $P_A = 5/20 = 0.25$

$$P_T = \frac{P_A + Sp - 1}{Sp + Se - 1} = \frac{0.25 + 0.96 - 1}{0.3 + 0.96 - 1} = 0.807$$

- $x = 2,\ n = 60,\ Se = 0.3,\ Sp = 0.96,\ N = 675$

  $P_A = 2/60 = 0.033$

$$P_T = \frac{P_A + Sp - 1}{Sp + Se - 1} = \frac{0.033 + 0.96 - 1}{0.3 + 0.96 - 1} = -0.0269$$



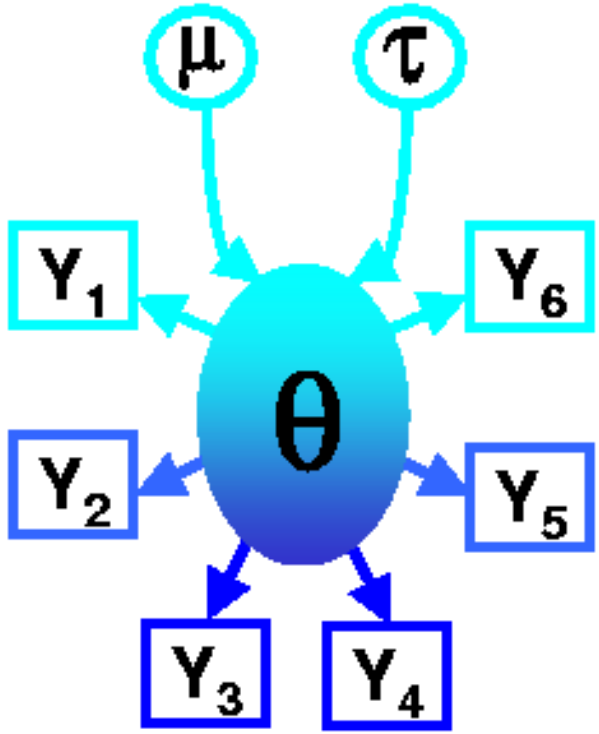
http://solymosin.github.io/CI4prev/

https://epitools.ausvet.com.au/trueprevalence

- $x = 2$, $n = 60$, $Se = 0.3$, $Sp = 0.96$, $N = 675$
- Bayesian approach accounts uncertainity of P_T, Se, Sp:
  - 95% certain that $Se < 0.5$ and $Sp > 0.94$
  - Beta prior distributions
  - Posterior distributions
    - $P_T = 0.02$, 95% credible interval $0 - 0.456$
    - $Se = 0.29$, 95% credible interval $0.11 - 0.52$
    - $Se = 0.96$, 95% credible interval $0.94 - 0.98$
    - 97.5% certain $P_T < 0.456$
    - 58% certain population is infected

Thomas Bayes (1702 – 1761)

$$p(\theta|x) = \frac{p(x|\theta)}{p(x)} \times p(\theta)$$



BUGS

https://www.mrc-bsu.cam.ac.uk/software/bugs

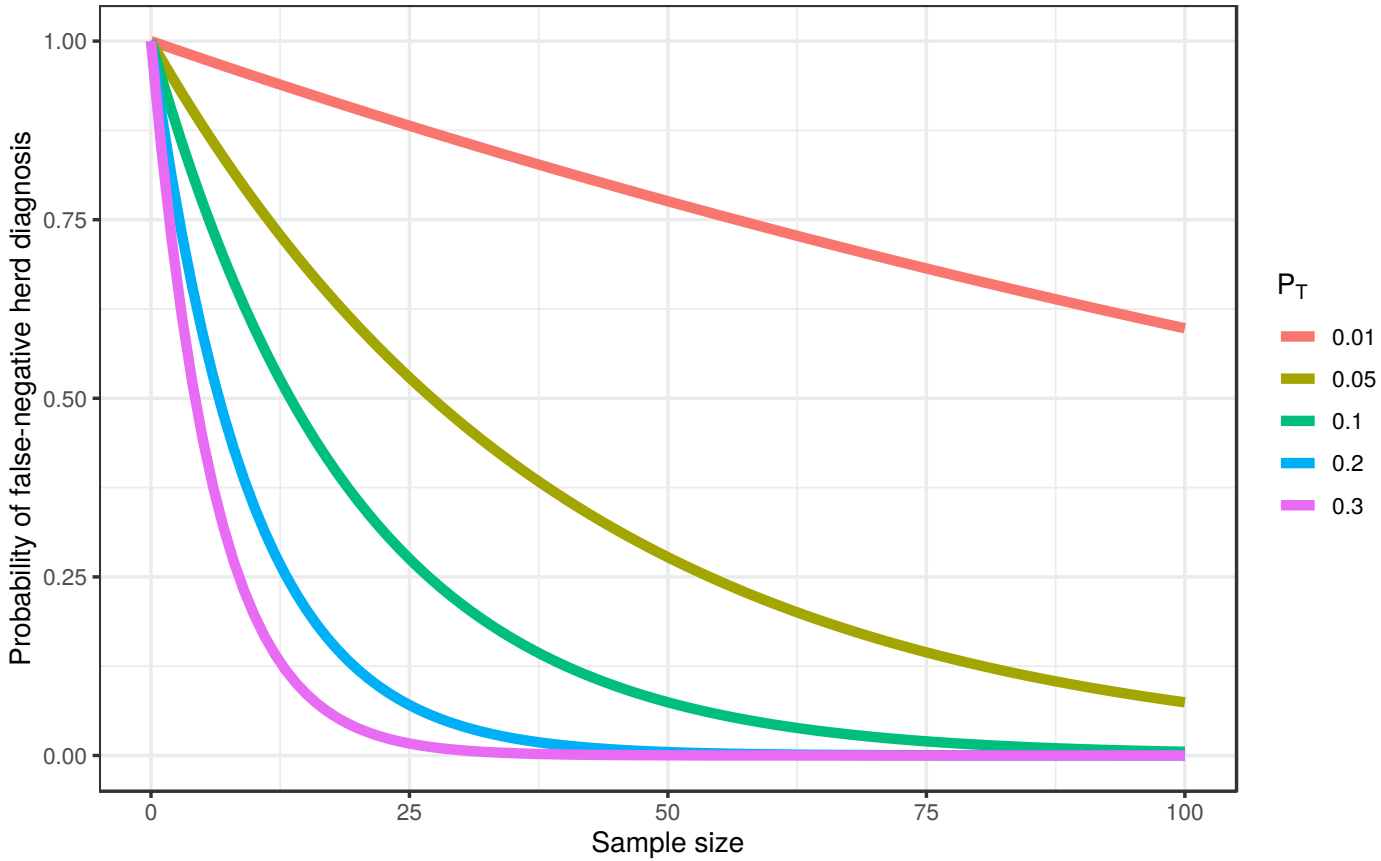
https://cadms.vetmed.ucdavis.edu/diagnostic-tests

- In many cases, the health state of a population unit (farm, barn, litter or other groups) is the subject of study, not the state of the individuals.
- It is not commonly known the heard-level tests should be interpreted differently than individual tests.
- Herd-level interpretation of tests is often more complex, especially if the tests used have no perfect specificity.
- As in the case of individual test results, for the correct decision on the state of a herd, the herd-level sensitivity and specificity of applied tests must be known.
- Usually, the most likely performance of herd tests is estimated based on the individual level sensitivity and specificity of applied tests.

- Herd-level sensitivity (HSE) is the probability that an infected herd will be detected as positive by the test.
- Herd-level specificity (HSP) is the probability that an uninfected herd will be detected as negative by the test.
- Herd-level sensitivity and specificity besides the individual level sensitivity and specificity of the applied test depend on further factors:
  - sample size (n)
  - prevalence within the farm
  - critical number:
    - how many animals (1,2,3, etc.) should be positive in the population to consider the herd as positive?
    - as the critical number increases, HSP increases and HSE decreases

number of tested animals ↑:

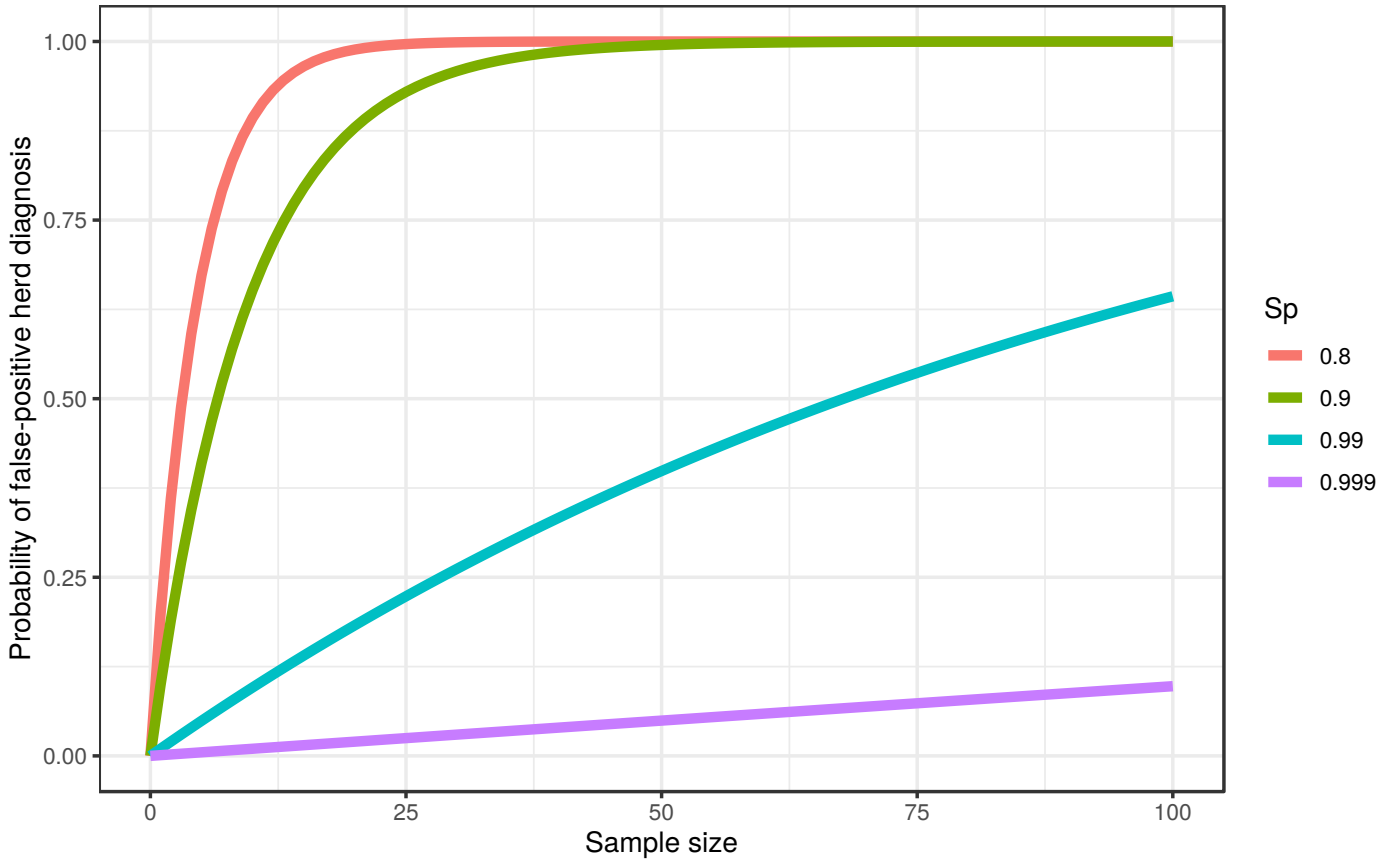
- HSE↑
- probability of the false negative herd-detection (1-HSE) ↓
- the probability of the false negative herd-detection decreases faster when the within-herd prevalence is higher



it is easier to distinguish between infected and uninfected herd by the same sample size when the within-herd prevalence is higher

number of tested animals ↑:

- the probability of founding at least one false positive animal is increasing, that is HSP↓
- usually, from larger populations, more samples are taken so that we can get false positivity more often in large farms



If the critical number is 1

$$HSE = 1 - (1 - P_A)^n$$

$$P_A = P_T Se + (1 - P_T)(1 - Sp)$$

- $1 - P_A$ is the probability of testing one animal as negative
- $(1 - P_A)^n$ is the probability of testing all n animals as negative
- $1 - (1 - P_A)^n$ is the probability to test at least 1 animal out of n as positive

$$HSP = Sp^n$$

- if $Sp = 0.95$ and $n = 1$, then $HSP = 0.95^1 = 0.95$
- if $Sp = 0.95$ and $n = 5$, then $HSP = 0.95^5 = 0.774$

If the critical number is larger than 1

$$HSP = \sum_{i=0}^{c-1} \frac{n!}{i! \times (n-i)!} \times p_f^i \times q_f^{n-i}$$

$$HSE = 1 - \sum_{i=0}^{c-1} \frac{n!}{i! \times (n-i)!} \times p_d^i \times q_d^{n-i}$$

- n sample size
- c critical number
- p_f AP if the herd is truly free of the disease
- q_f $1 - p_f$
- p_d AP if the herd is truly disease
- q_d $1 - p_d$

Census: if all animals in a population are investigated.

If a survey is designed well, then a reasonably accurate and acceptable estimate of a variable can be made by examining some of the animals in the relevant population; that is, a **sample**.

The **target population** is the total population about which information is required.

The **study population** is the population from which a sample is drawn.

The study population consists of **elementary units**, which cannot be divided further.

A collection of elementary units, grouped according to a common characteristic, is a **stratum**.

Veterinarians often need to determine whether an infection is or has ever been present in the herd or a subpopulation of the herd.

For tests of 100% specificity, a **single positive** is usually considered sufficient to class the herd as positive, although for serological tests of imperfect specificity, more than one positive might be necessary.

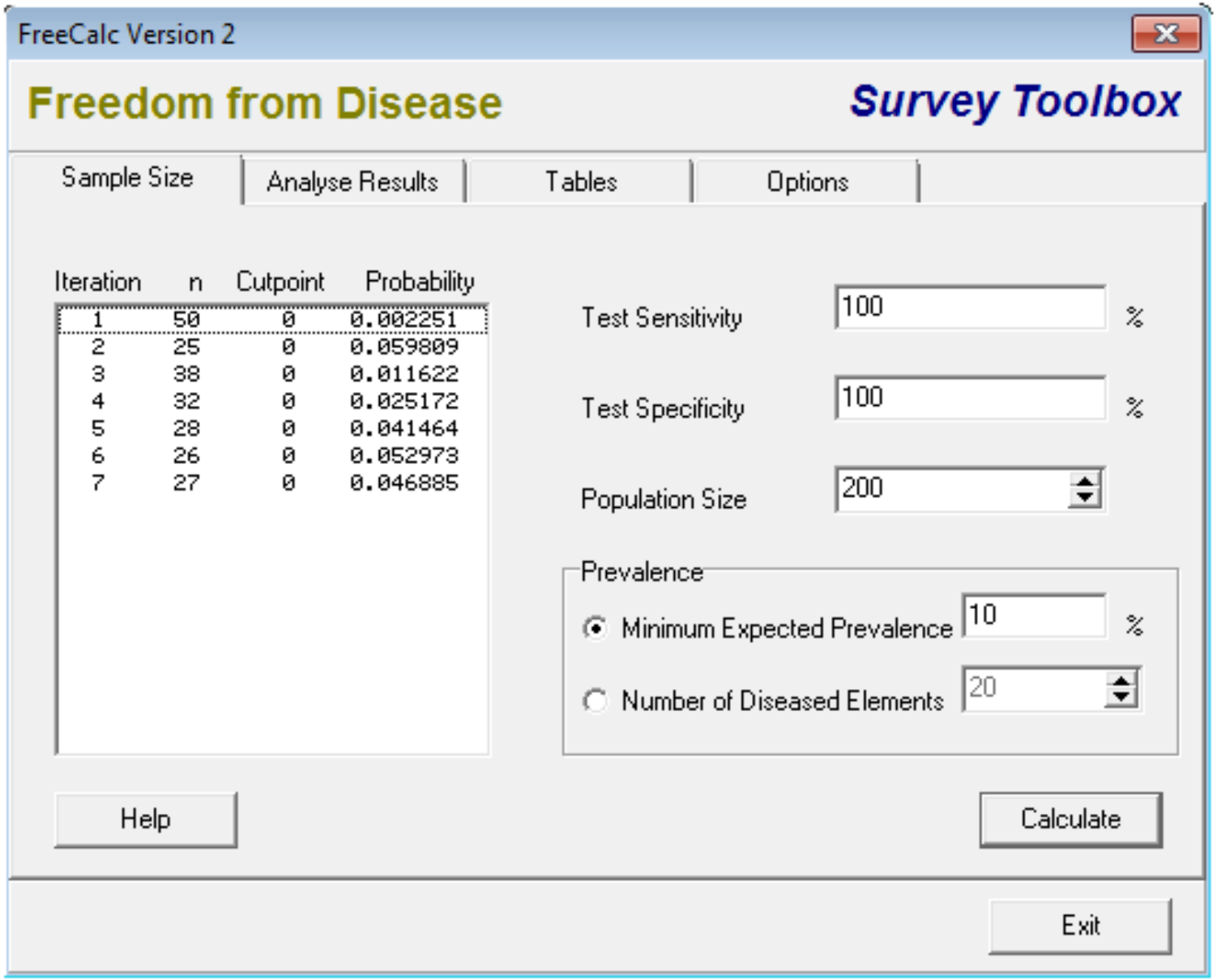
To estimate required numbers to detect infection, the following values are necessary: the **required level of confidence**, usually 95%, the likely **prevalence** of infection in the herd or in the specific group of pigs being evaluated, the population size.

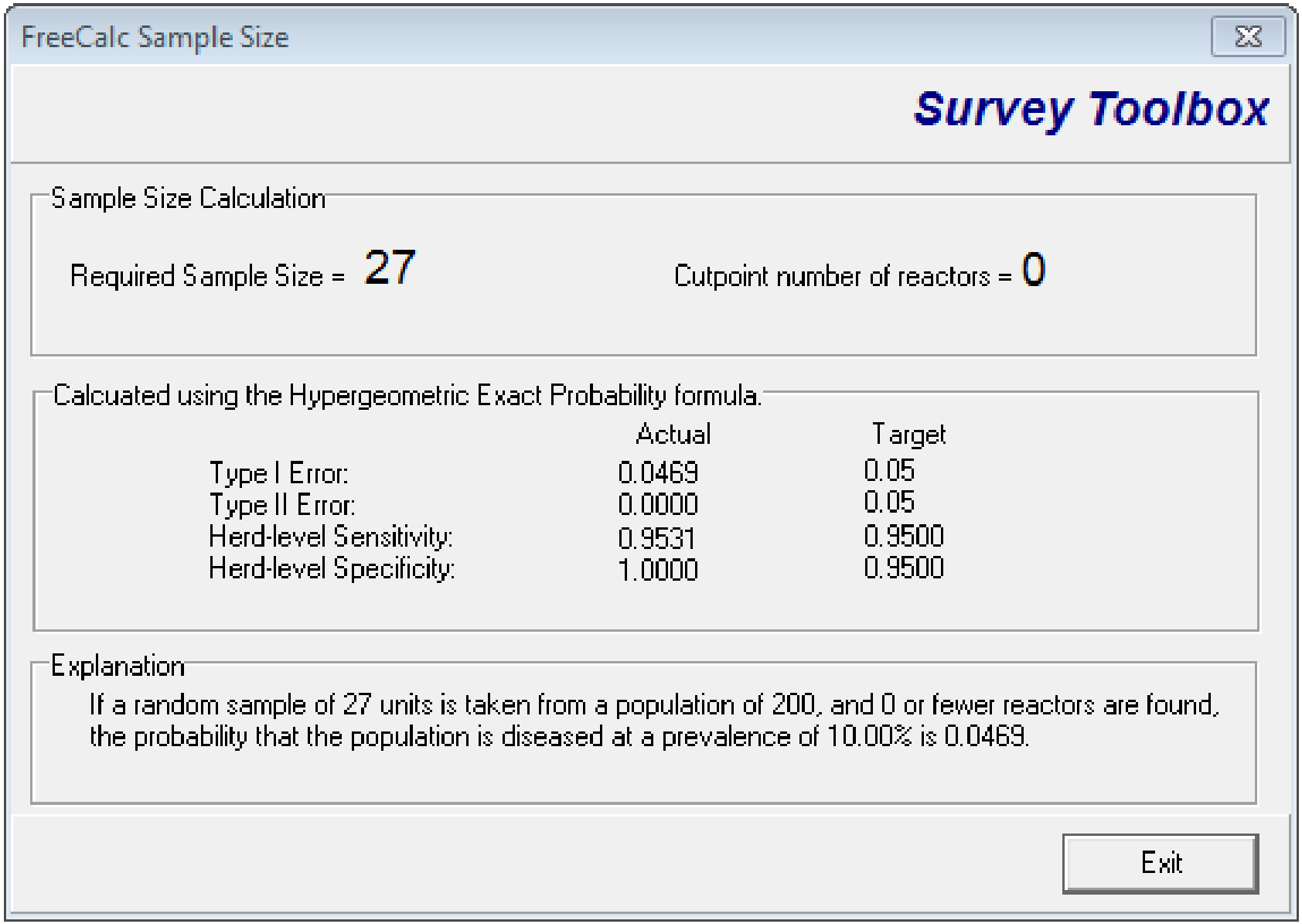
The selected prevalence value should be realistic, but if there is doubt, erring toward a **lower** prevalence is preferable to ensure that adequate numbers of pigs are sampled.

If a veterinarian's only goal is to detect infection, sampling does not need to be random but can be directed to **higher-prevalence** groups, for example, different age groups when there is an age-related risk of infection or clinically affected versus otherwise healthy pigs.

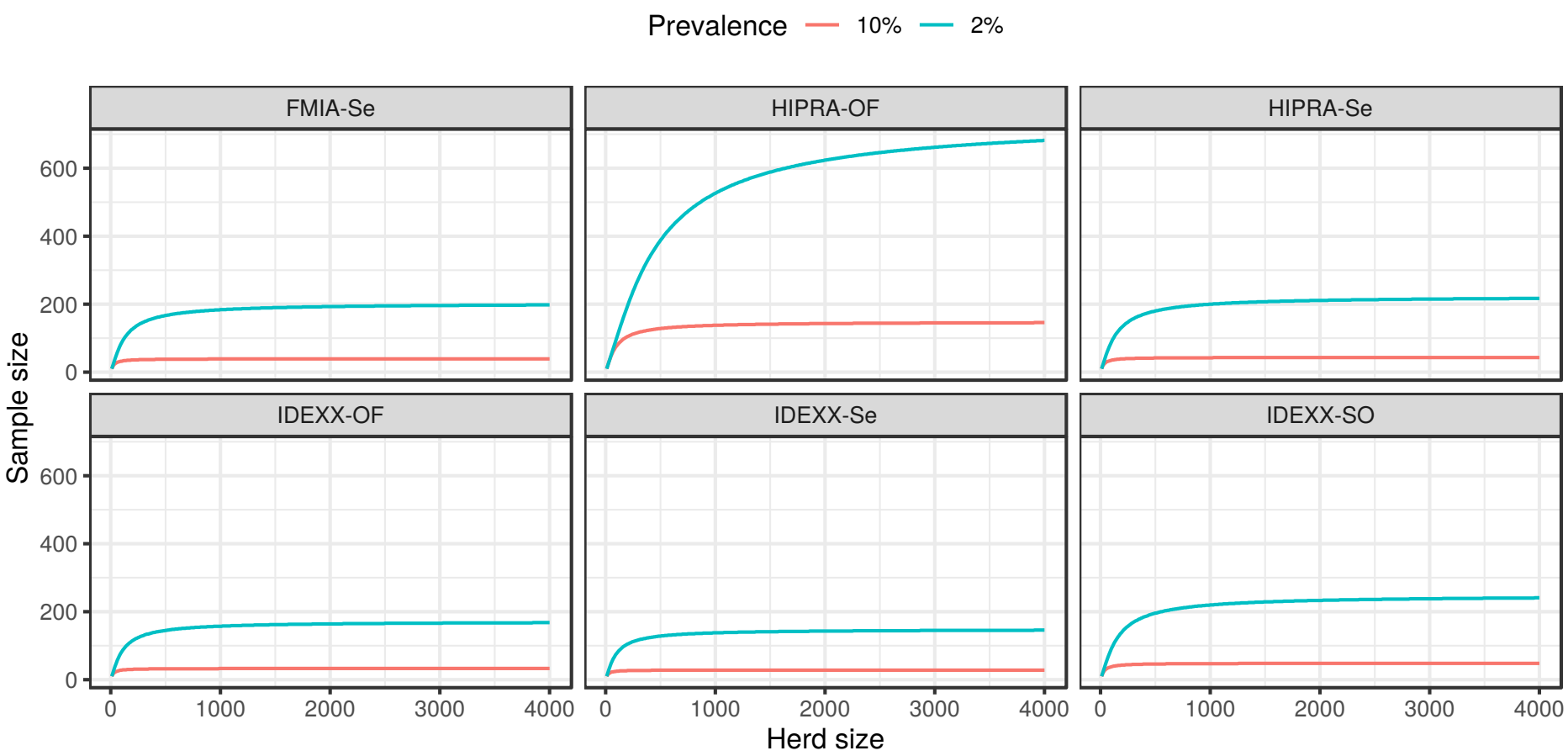
To detect *T. gondii* in a herd, sows are a better population to sample because prevalence is likely to be higher than in grower-finisher pigs.

To detect enteric pathogens by fecal culture or antigen detection methods, preference should be given to sampling pigs with diarrhea rather than pigs with normal feces.


FreeCalc Version 2
Freedom from Disease
Survey Toolbox
Sample Size
Analyse Results
Tables
Options
Iteration n Cutpoint Probability
1 50 0 0.002251
2 25 0 0.059809
3 38 0 0.011622
4 32 0 0.025172
5 28 0 0.041464
6 26 0 0.052973
7 27 0 0.046885
Test Sensitivity
100
%
Test Specificity
100
%
Population Size
200
Prevalence
Minimum Expected Prevalence
10
%
Number of Diseased Elements
20
Help
Calculate
Exit



https://epitools.ausvet.com.au/freecalctwo?page=FreeCalc2



PRRS Test	Sample	Se	Sp
FMIA-Se	serum	73.3%	73.3%
HIPRA-OF	saliva	20.0%	100.0%
HIPRA-Se	serum	66.7%	93.3%
IDEXX-OF	saliva	86.7%	100.0%
IDEXX-Se	serum	100.0%	100.0%
IDEXX-SO	saliva	60.0%	93.3%

Belanger, A., L. Descoteaux, Y. Couture, J. Baril, and R. Bagg (2003). Evaluation of a milk strip test for detection of subclinical ketosis at cow level. In *36th Annu. Am. Assoc. Bovine Pract.(AABP) Conf. Auburn, AL: AABP*, Volume 175.

Carrier, J., S. Stewart, S. Godden, J. Fetrow, , and P. Rapnicki (2004). Evaluation and use of three cowside tests for detection of subclinical ketosis in early postpartum cows. *J Dairy Sci 87*, 3725–3735.

Dinya, E., L. D. Molnár, J. Mészáros, and N. Solymosi (2019). *Biometria a klinikumban. Feladatok bayesi megoldása*. Budapest: Medicina.

Farkas, R., B. Tánczos, M. Gyurkovszky, G. Földvári, N. Solymosi, R. Edelhofer, and S. Hornok (2013). Serological and molecular detection of Theileria equi infection in horses in Hungary. *Veterinary parasitology 192*(1-3), 143–148.

Gardner, I. A. (2012). Analysis and use of diagnostic data. In *Diseases of Swine*, pp. 94–105. John Wiley & Sons, Ltd.

Geishauser, T., K. Leslie, J. Tenhag, and A. Bashiri (2000). Evaluation of eight cow-side ketone tests in milk for detection of subclinical ketosis in dairy cows. *Journal of dairy science 83*(2), 296–299.

Gerber, P. F., L. G. Giménez-Lirola, P. G. Halbur, L. Zhou, X.-J. Meng, and T. Opriessnig (2014). Comparison of commercial enzyme-linked immunosorbent assays and fluorescent microbead immunoassays for detection of antibodies against porcine reproductive and respiratory syndrome virus in boars. *Journal of virological methods 197*, 63–66.

Messam, L. L. M., A. J. Branscum, M. T. Collins, and I. A. Gardner (2008). Frequentist and Bayesian approaches to prevalence estimation using examples from Johne's disease. *Animal Health Research Reviews 9*(1), 1–23.

Noordhuizen, J. P. T. M., K. Frankena, M. Thrusfield, and E. A. M. Graat (2001). *Application of Quantitative Methods in Veterinary Epidemiology*. Wageningen, The Netherland: Wageningen Pers.

Oetzel, G. R. (2004). Monitoring and testing dairy herds for metabolic disease. *Veterinary Clinics: Food Animal Practice 20*(3), 651–674.

Osborne, T., K. Leslie, T. Duffield, C. Petersson, J. Ten Hag, and Y. Okada (2002). Evaluation of keto-test in urine and milk for the detection of subclinical ketosis in periparturient holstein dairy cattle. In *Proceedings of the 35th conference of the American Association of Bovine Practitioners, Rome, GA, USA*, pp. 188–9.