Diagnostic Tests

Solymosi Norbert

Quantitative veterinary epidemiology

Department of Microbiology and Infectious Diseases

University of Veterinary Medicine Budapest

Making a diagnosis

- assemble a list of differentials
- order the list as to which conditions are important, or more likely
- choose appropriate tests (and the order in which to do them)
- interpret the results

What is a test?

- any procedure that reduces uncertainty about the state of disease
- includes routine examination, questions posed during history taking, clinical signs, laboratory findings (haematology, serology, biochemistry, histopathology), post mortem findings

- Is a subject (individual, group, region, country)
  - diseased (abnormal, affected)
  - not diseased (normal, unaffected)?
- The testing process is imperfect
  - clarity of case definition (what is positive, what is negative)
  - quality of measurement tools
- Applications of testing
  - establish the presence of disease in an individual
  - establish the level of disease in a population
  - monitoring, screening and surveillance
- Test evaluation
  - The perfect diagnostic test allows us to differentiate between disease-positive and disease-negative individuals without error
  - Most tests that we use are imperfect
  - Imperfect tests should be compared with a 'gold standard' so we know how good they are

Stevenson 2012

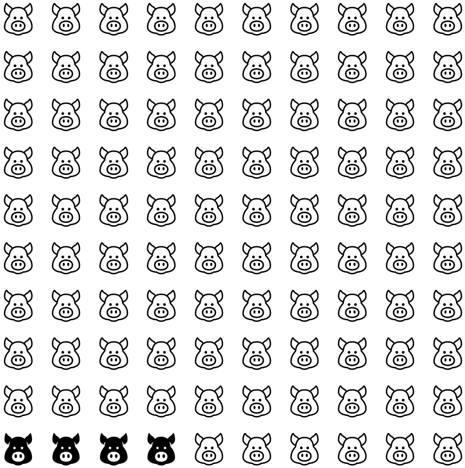
Contingency table

	Infection $+$	Infection $-$	Total
Test $+$	True positive a	False positive b	Test positive $a+b$
Test $-$	False negative c	True negative d	Test negative $c+d$
Total	$a+c$	$b+d$	$a+b+c+d$

infected pig

uninfected pig

let's say the prevalence of infection is 4% (4/100)

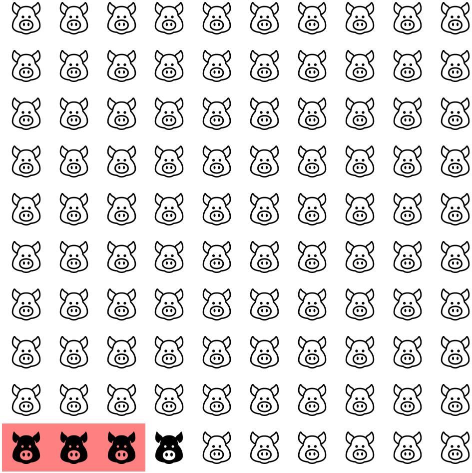


infected pig:

test positive

uninfected pig

with a diagnostic test 3 of 4 infected animals are positive



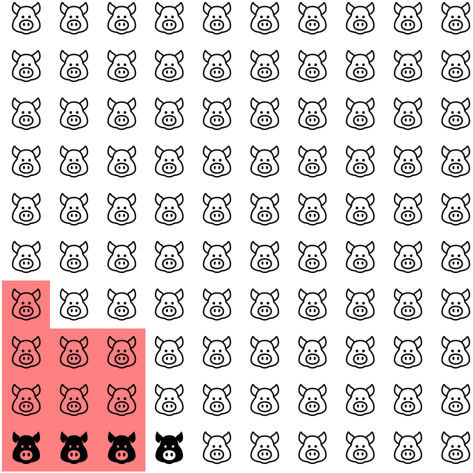
infected pig:

test positive

uninfected pig:

test positive

further 7 non infected animals were test positive



uninfected pig:

- test positive
- test negative

infected pig:

- test positive
- test negative

89 non infected and 1 infected animals were test negative

uninfected pig:

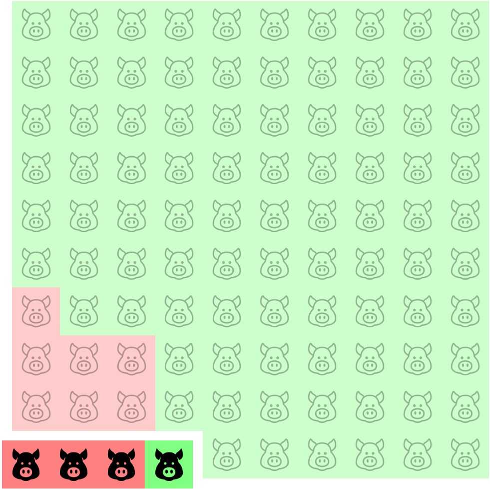
- test positive
- test negative

infected pig:

- test positive
- test negative

Sensitivity is the proportion of infected animals identified as positive by the test:

$$\frac{3}{4} = 0.75$$

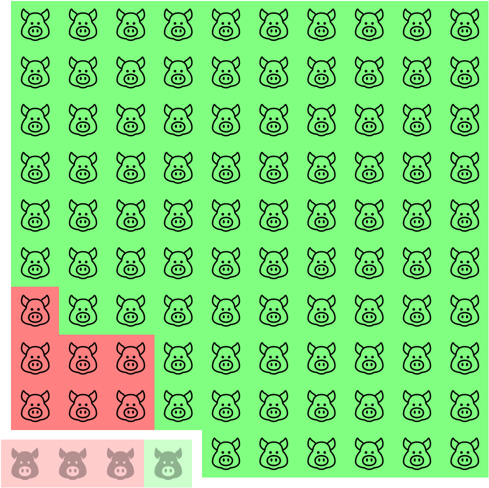


uninfected pig:

- test positive
- test negative

infected pig:

- test positive
- test negative



Specificity is the proportion of uninfected animals identified as negative by the test:

$$\frac{89}{96} = 0.93$$

Contingency table

	Infection $+$	Infection $-$	Total
Test $+$	True positive a	False positive b	Test positive $a+b$
Test $-$	False negative c	True negative d	Test negative $c+d$
Total	$a+c$	$b+d$	$a+b+c+d$

	Infection $+$	Infection $-$	$\sum$
Test $+$	3	7	10
Test $-$	1	89	90
$\sum$	4	96	100

Sensitivity is the proportion of infected animals identified as positive by the test.
Or the probability that a truly diseased animal will be classified as diseased.

	Infection $+$	Infection $-$	Total
Test $+$	True positive a	False positive b	Test positive $a+b$
Test $-$	False negative c	True negative d	Test negative $c+d$
Total	$a+c$	$b+d$	$a+b+c+d$

Sensitivity is the proportion of infected animals identified as positive by the test.
Or the probability that a truly diseased animal will be classified as diseased.

	Infection $+$	Infection $-$	Total
Test $+$	True positive a	False positive b	Test positive $a+b$
Test $-$	False negative c	True negative d	Test negative $c+d$
Total	$a+c$	$b+d$	$a+b+c+d$

$$Se = \frac{a}{a+c} = \frac{3}{3+1} = 0.75$$

Specificity is the proportion of uninfected animals identified as negative by the test.
Or the probability that a truly non-diseased animal will be classified as non-diseased.

	Infection +	Infection −	Total
Test +	True positive a	False positive b	Test positive $a+b$
Test −	False negative c	True negative d	Test negative $c+d$
Total	$a+c$	$b+d$	$a+b+c+d$

Specificity is the proportion of uninfected animals identified as negative by the test.
Or the probability that a truly non-diseased animal will be classified as non-diseased.

	Infection +	Infection −	Total
Test +	True positive a	False positive b	Test positive $a+b$
Test −	False negative c	True negative d	Test negative $c+d$
Total	$a+c$	$b+d$	$a+b+c+d$

$$Sp = \frac{d}{b+d} = \frac{89}{7+89} = 0.93$$

- We've just considered two tests where differentiation between 'positive' and 'negative' is clear
- Some tests are measured on a continuous scale
  - blood pressure
  - blood chemistry
  - serum enzymes
  - serology
- So we need to set a cutpoint to distinguish between positives and negatives

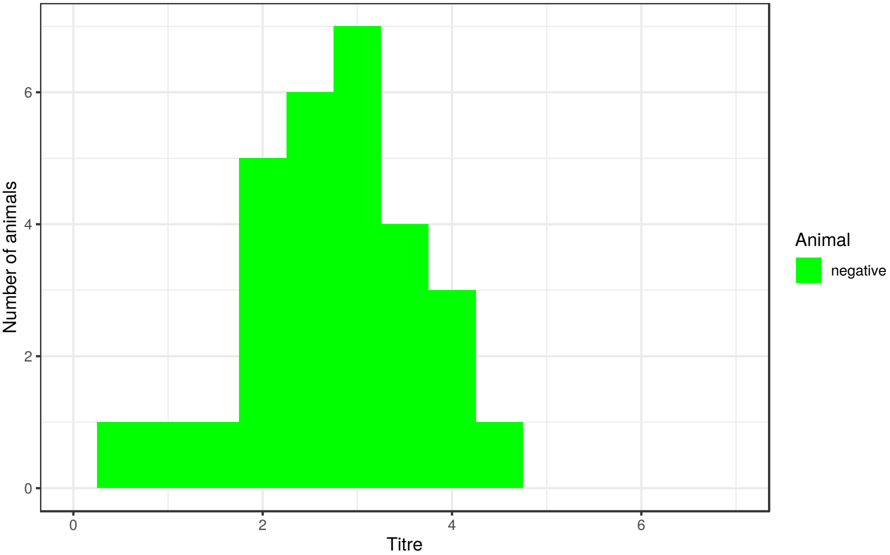
Stevenson 2012

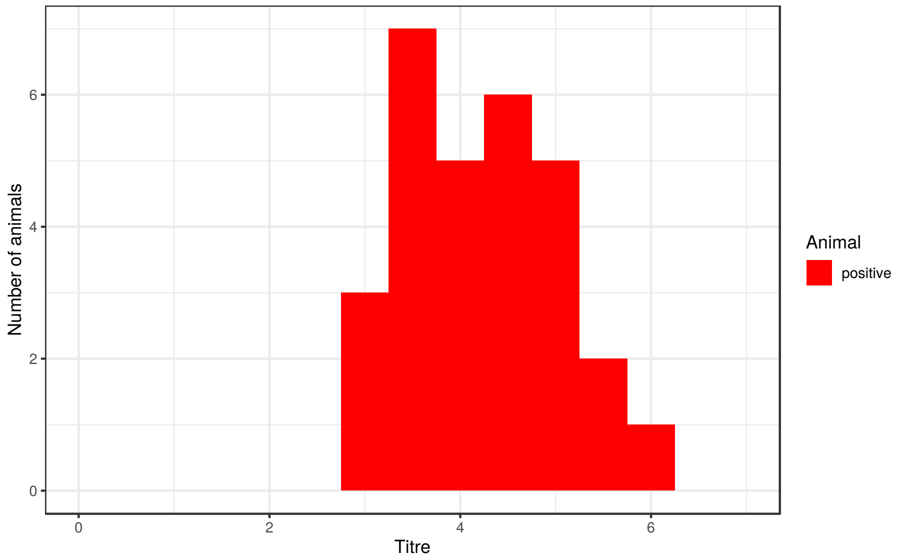
Titres of negative animals: 3, 2, 3.5, 3, 2, 2.5, 3, 1.5, 2.5, 3.5, 4, 2.5, 4, 2, 2, 3, 1, 2.5, 3.5, 0.5, 2.5, 4.5, 2.5, 3, 3, 3, 4, 2, 3.5

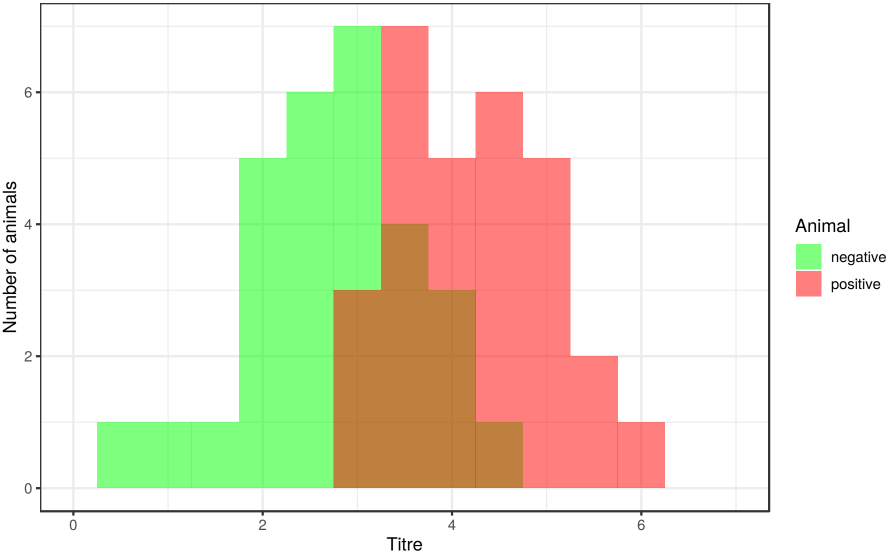
Titres of positive animals: 3.5, 3.5, 3.5, 4.5, 4, 5, 5.5, 4, 3.5, 4.5, 5, 3.5, 6, 4, 4.5, 3.5, 4, 3, 5, 5.5, 5, 3, 4.5, 4.5, 5, 3, 3.5, 4.5, 4

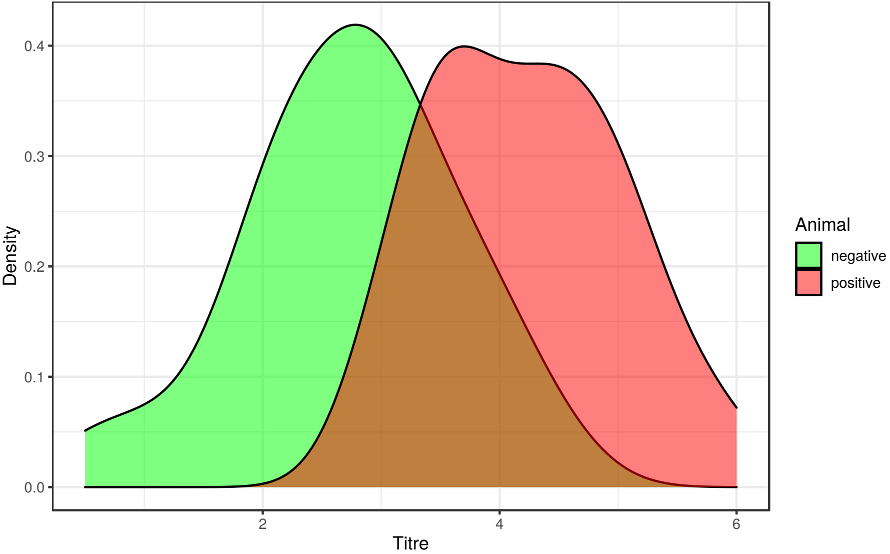
Titre	Number of animals negative	Number of animals positive
0.5	1	
1.0	1	
1.5	1	
2.0	5	
2.5	6	
3.0	7	3
3.5	4	7
4.0	3	5
4.5	1	6
5.0		5
5.5		2
6.0		1

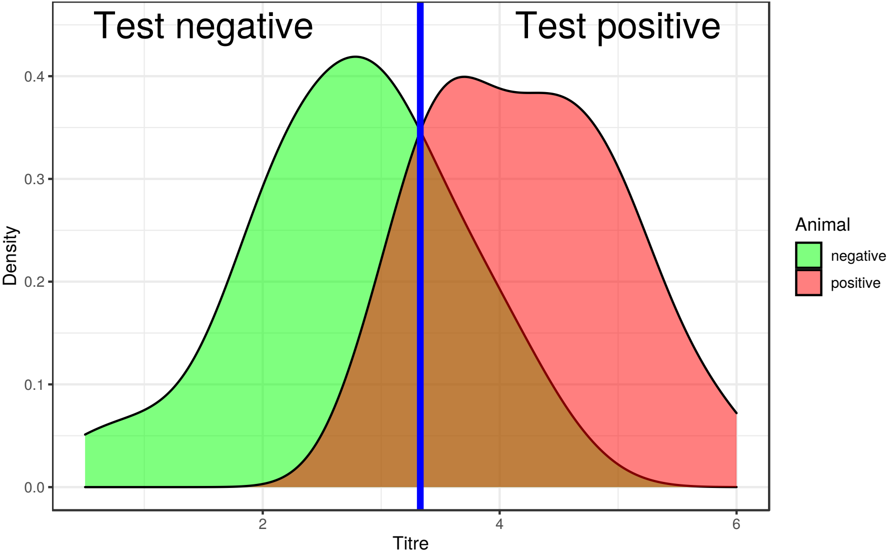
Noordhuizen et al. (2001)

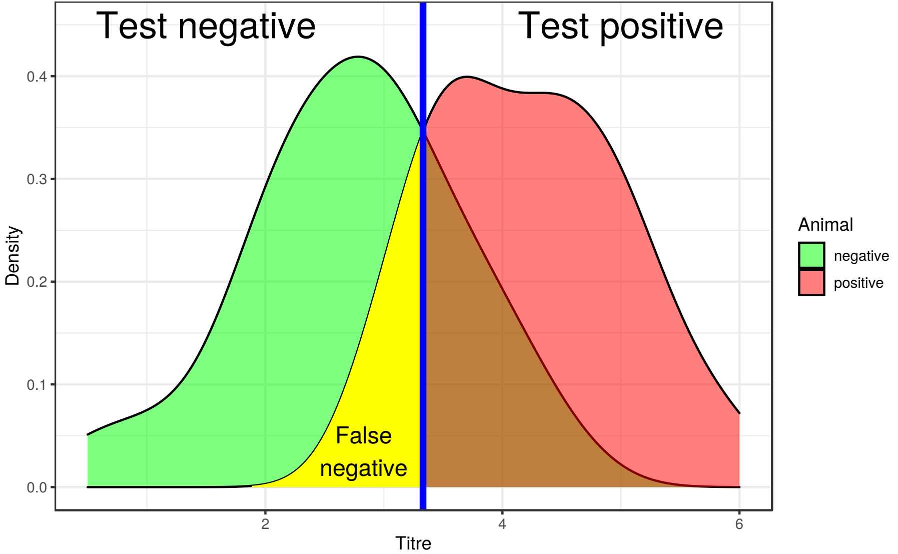

Number of animals
Titre
Animal
negative
0
2
4
6
0
2
4
6

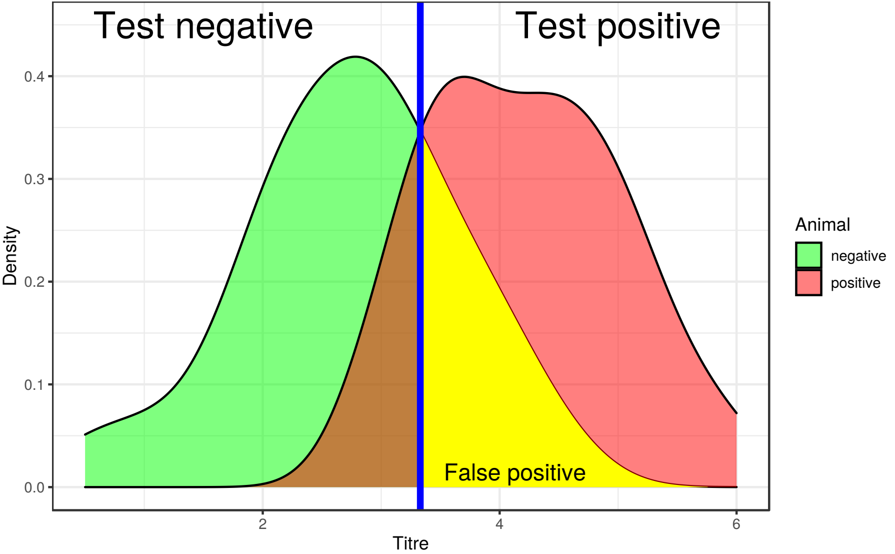

Number of animals
0
2
4
6
Titre
Animal
positive


Number of animals
6
4
2
0
0
2
4
6
Titre
Animal
negative
positive

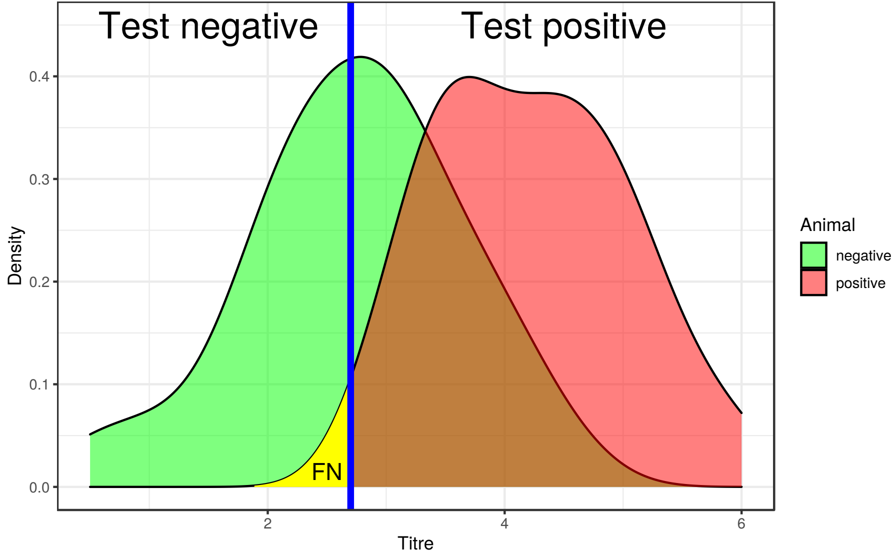

0.4
0.3
0.2
0.1
0.0
Density
2
4
6
Titre
Animal
negative
positive


Test negative
Test positive
0.4
0.3
0.2
0.1
0.0
Density
2
4
6
Titre
Animal
negative
positive

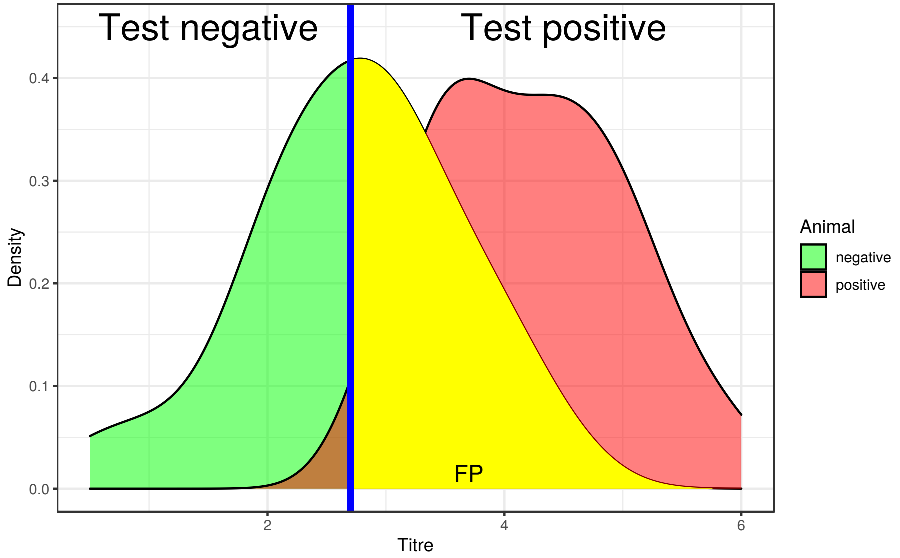

Test negative
Test positive
False
negative
0.4
0.3
0.2
0.1
0.0
Density
2
4
6
Titre
Animal
negative
positive


Test negative
Test positive
False positive
Density
0.0
0.1
0.2
0.3
0.4
Titre
2
4
6
Animal
negative
positive

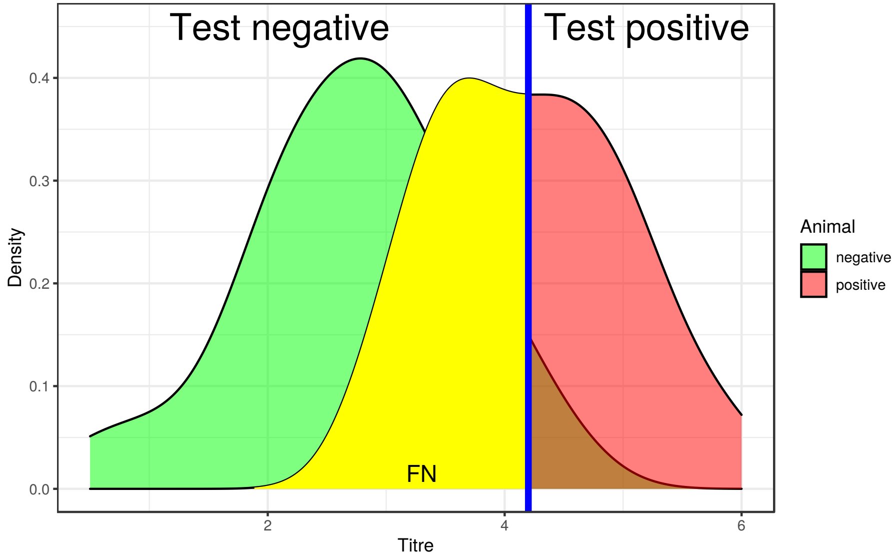
Se↑, Sp↓ → FN↓ (TN↓)


Test negative
Test positive
FN
Density
0.0
0.1
0.2
0.3
0.4
2
4
6
Titre
Animal
negative
positive

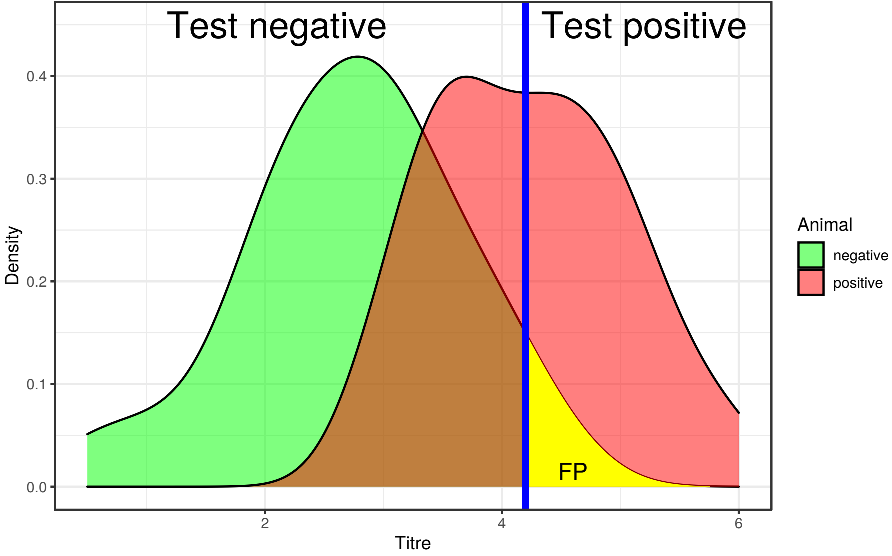
Se↑, Sp↓ → FP↑ (TP↑)


Test negative
Test positive
FP
Density
0.0
0.1
0.2
0.3
0.4
Titre
2
4
6
Animal
negative
positive

Se↓, Sp↑ → FN↑ (TN↑)


Test negative
Test positive
FN
Density
0.4
0.3
0.2
0.1
0.0
2
4
6
Titre
Animal
negative
positive

Se↓, Sp↑ → FP↓ (TP↓)


Test negative
Test positive
FP
Density
0.4
0.3
0.2
0.1
0.0
2
4
6
Titre
Animal
negative
positive

What threshold should be used?

- It depends on the purpose of testing.

Paolo Veronese (1582) Youth between Virtue and Vice

If the threshold is changed to increase the sensitivity, the specificity is decreased at the same time.

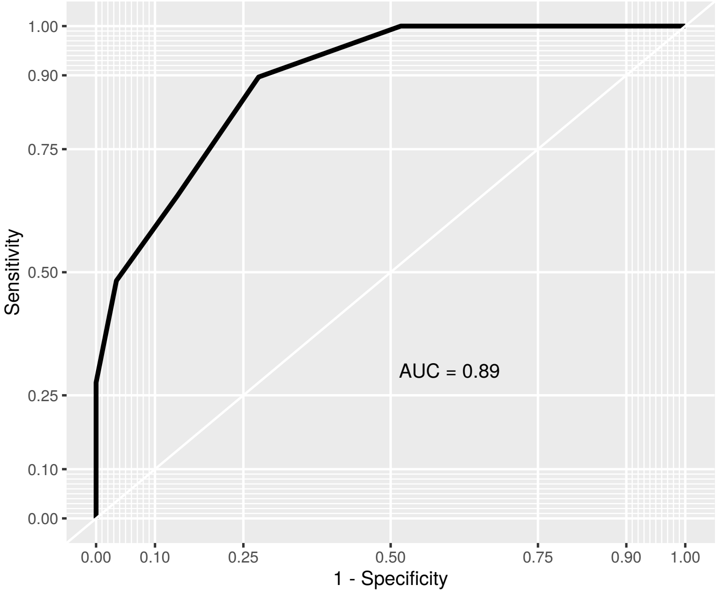
Se↑

- If we want to find all infected animals, we need to increase sensitivity.
- More infected animals will test positive and more uninfected animals will test positive.
- It increases the probability that test-negative animals are indeed not infected.
- When screening breeding pigs for purchase into a herd, a false-positive result would be much less harmful to a client than a false negative, which might allow infected pigs to enter a noninfected herd.

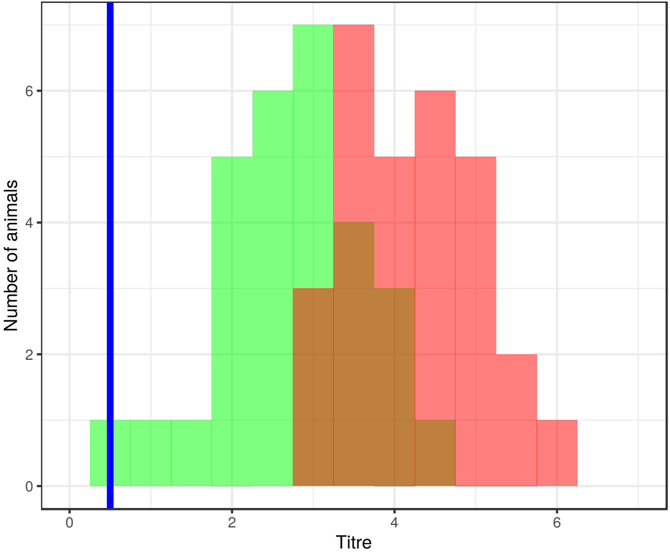
If the threshold is changed to increase the sensitivity, the specificity is decreased at the same time.

Sp$\uparrow$

- If the goal is to make sure the test positives are truly infected, we need to increase the specificity.
- Fewer infected animals will be test positive, and parallel fewer uninfected animals will be test positive too.
- A veterinarian relying on the results of a test to decide to cull a sow probably wants to minimize the chance of a false-positive diagnosis by using a highly specific test, especially if the sow is asymptomatic and pregnant and there are no other reasons for culling.



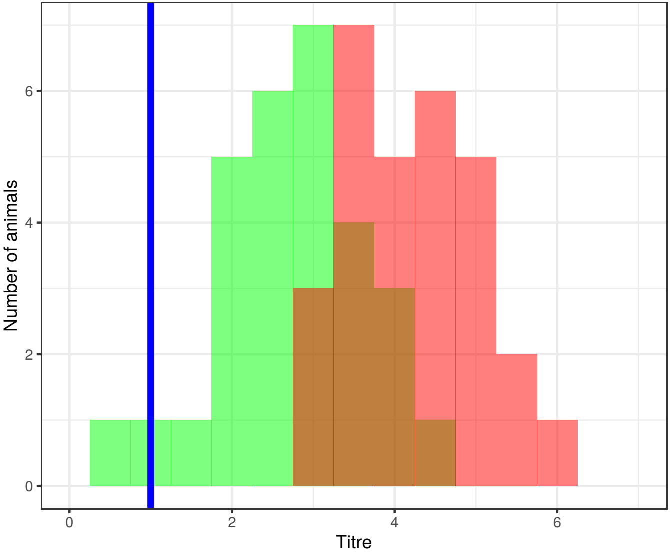
Receiver operating characteristic (ROC) curve



Titre threshold:
T+: > 0.5
T−: ≤ 0.5

	I+	I−
T+	29	28
T−	0	1

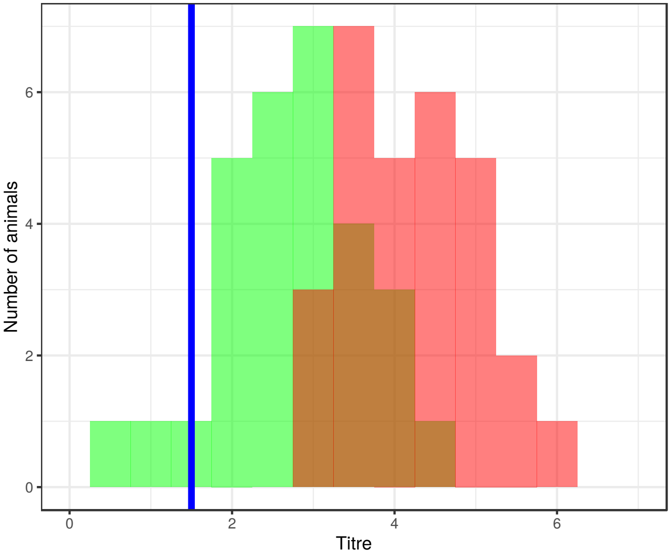
Se=1.00
Sp=0.03



Titre threshold:
T+: > 1.0
T−: ≤ 1.0

	I+	I−
T+	29	27
T−	0	2

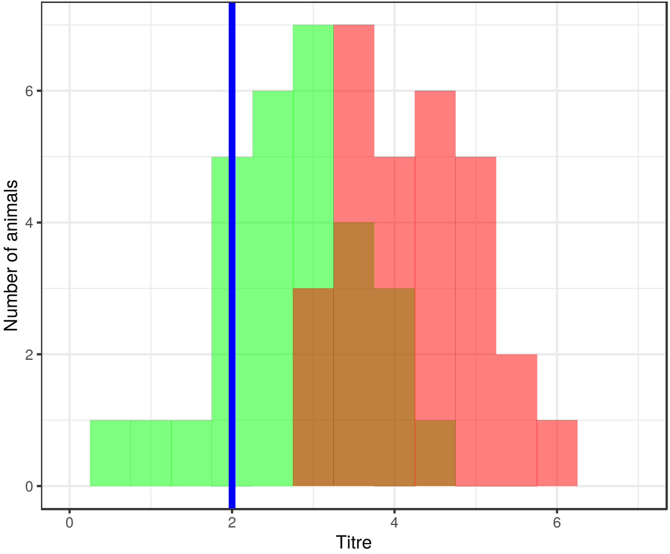
Se=1.00
Sp=0.07



Titre threshold:
T+: > 1.5
T−: ≤ 1.5

	I+	I−
T+	29	26
T−	0	3

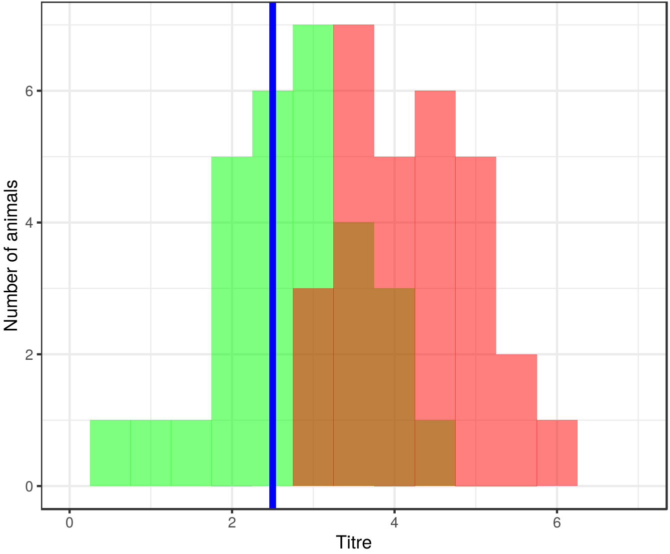
Se=1.00
Sp=0.10



Titre threshold:
T+: > 2.0
T−: ≤ 2.0

	I+	I−
T+	29	21
T−	0	8

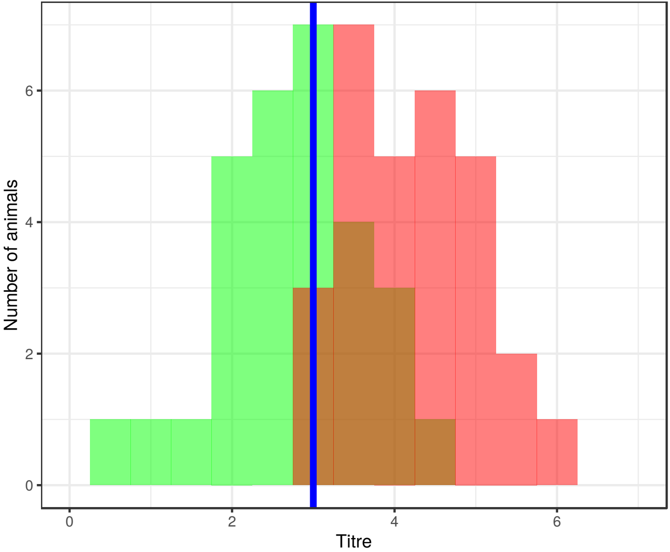
Se=1.00
Sp=0.28



Titre threshold:
T+: > 2.5
T−: ≤ 2.5

	I+	I−
T+	29	15
T−	0	14

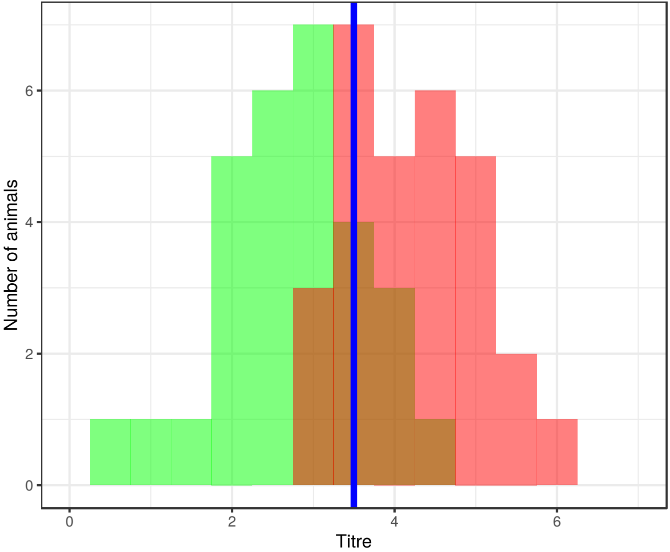
Se=1.00
Sp=0.48



Titre threshold:
T+: > 3.0
T−: ≤ 3.0

	I+	I−
T+	26	8
T−	3	21

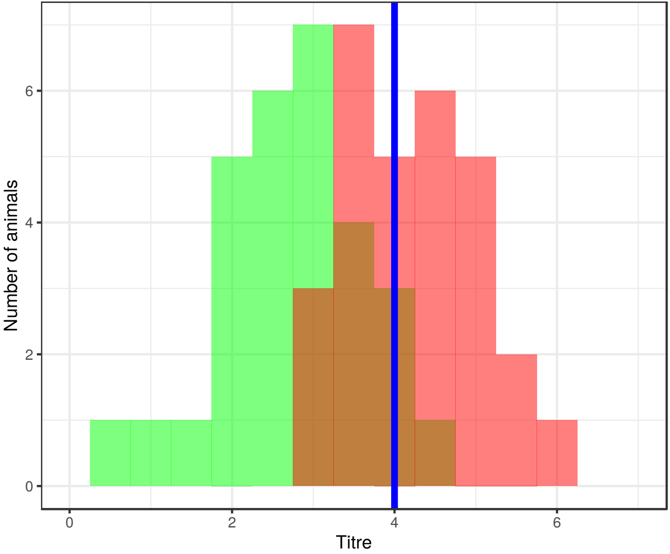
Se=0.90
Sp=0.72



Titre threshold:
T+: > 3.5
T−: ≤ 3.5

	I+	I−
T+	19	4
T−	10	25

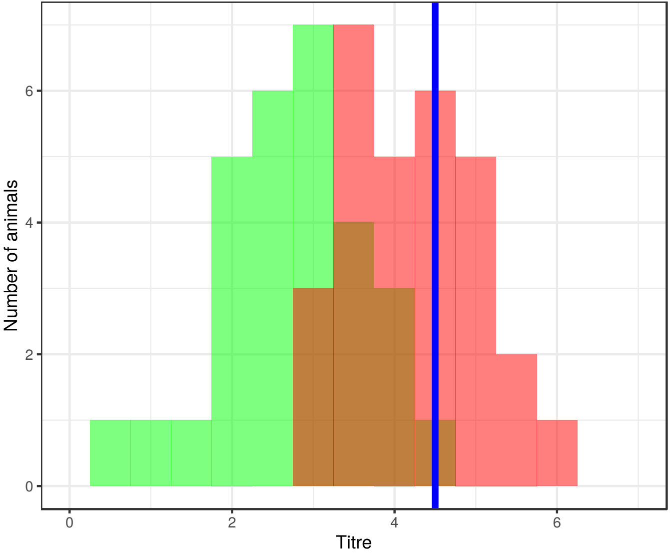
Se=0.66
Sp=0.86



Titre threshold:
T+: > 4.0
T−: ≤ 4.0

	I+	I−
T+	14	1
T−	15	28

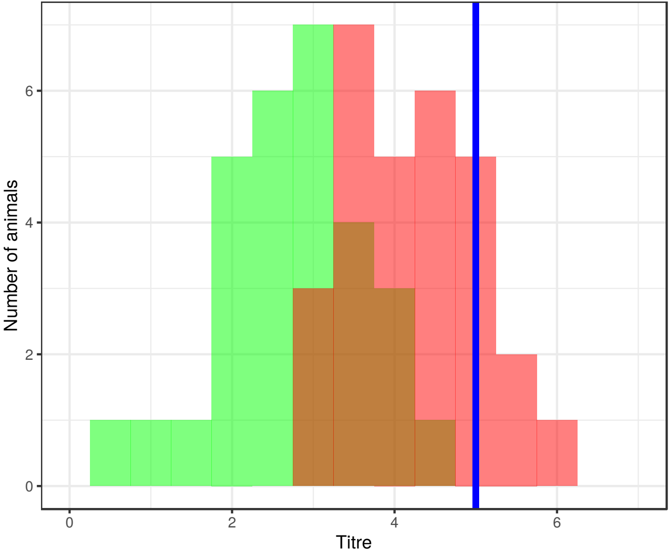
Se=0.48
Sp=0.97



Titre threshold:
T+: > 4.5
T−: ≤ 4.5

	I+	I−
T+	8	0
T−	21	29

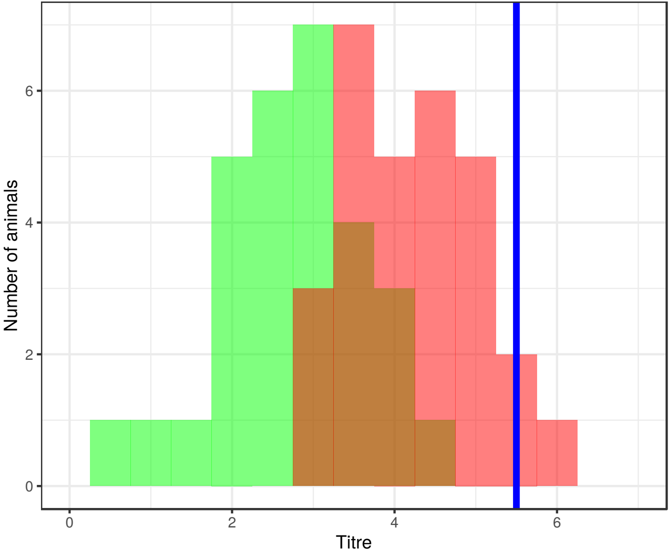
Se=0.28
Sp=1.00



Titre threshold:
T+: > 5.0
T−: ≤ 5.0

	I+	I−
T+	3	0
T−	26	29

Se=0.10
Sp=1.00



Titre threshold:
T+: > 5.5
T−: ≤ 5.5

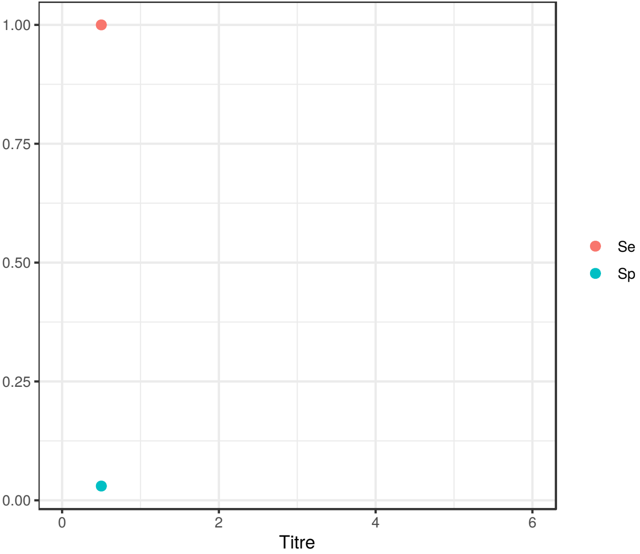
	I+	I−
T+	1	0
T−	28	29

Se=0.03
Sp=1.00

Titre threshold:
T+: > 0.5
T−: ≤ 0.5

	I+	I−
T+	29	28
T−	0	1

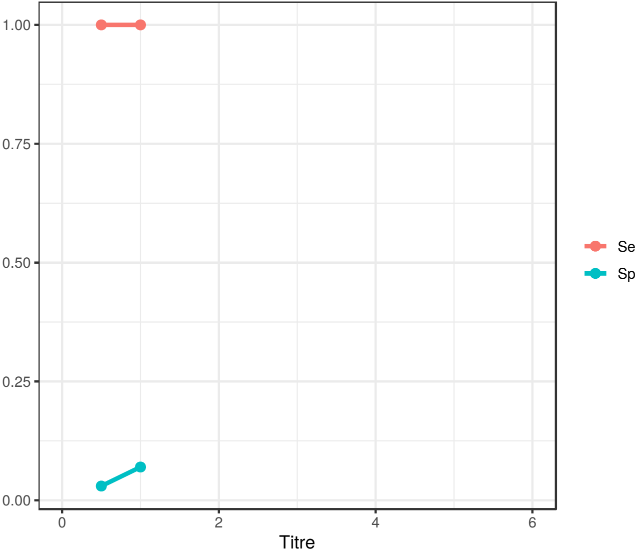
Se=1.00
Sp=0.03



Titre threshold:
T+: > 1.0
T−: ≤ 1.0

	I+	I−
T+	29	27
T−	0	2

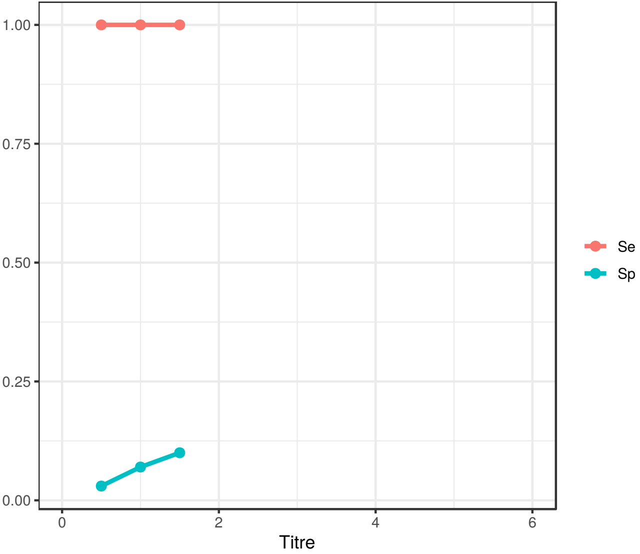
Se=1.00
Sp=0.07



Titre threshold:
T+: > 1.5
T−: ≤ 1.5

	I+	I−
T+	29	26
T−	0	3

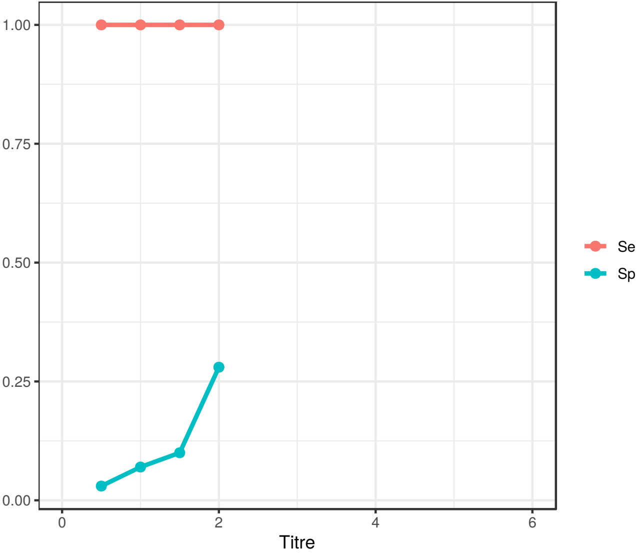
Se=1.00
Sp=0.10



Titre threshold:
T+: > 2.0
T−: ≤ 2.0

	I+	I−
T+	29	21
T−	0	8

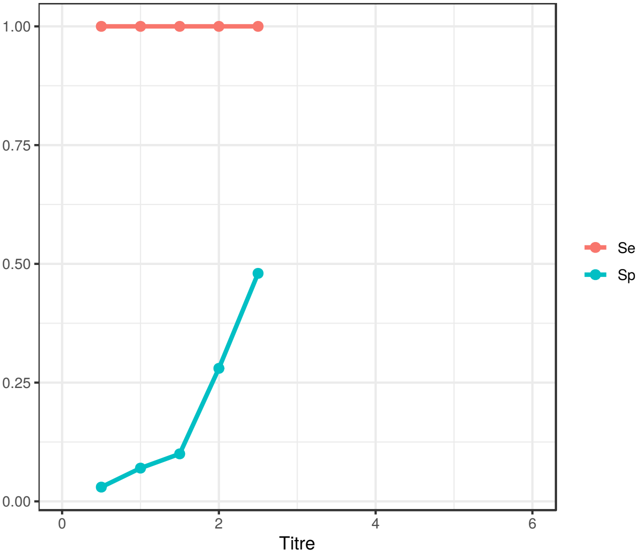
Se=1.00
Sp=0.28



Titre threshold:
T+: > 2.5
T−: ≤ 2.5

	I+	I−
T+	29	15
T−	0	14

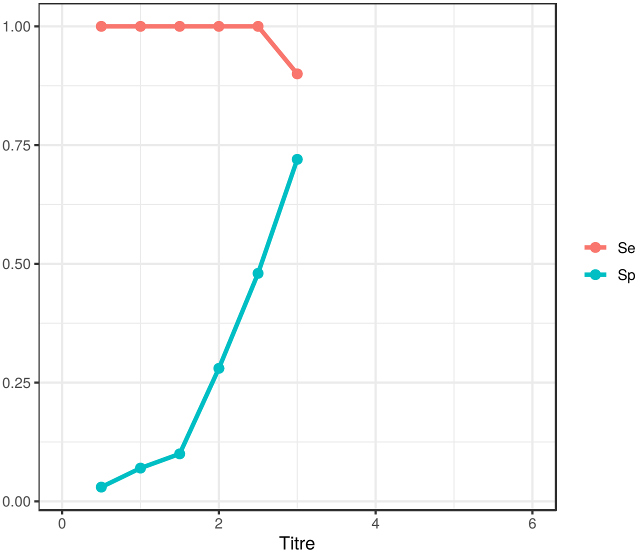
Se=1.00
Sp=0.48



Titre threshold:
T+: > 3.0
T−: ≤ 3.0

	I+	I−
T+	26	8
T−	3	21

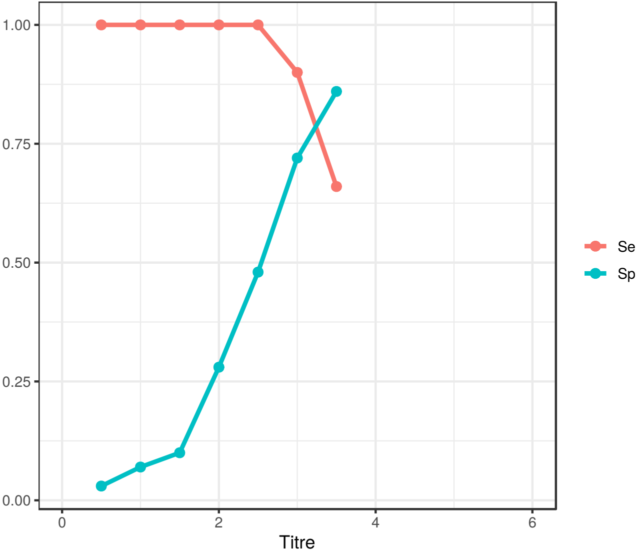
Se=0.90
Sp=0.72



Titre threshold:
T+: > 3.5
T−: ≤ 3.5

	I+	I−
T+	19	4
T−	10	25

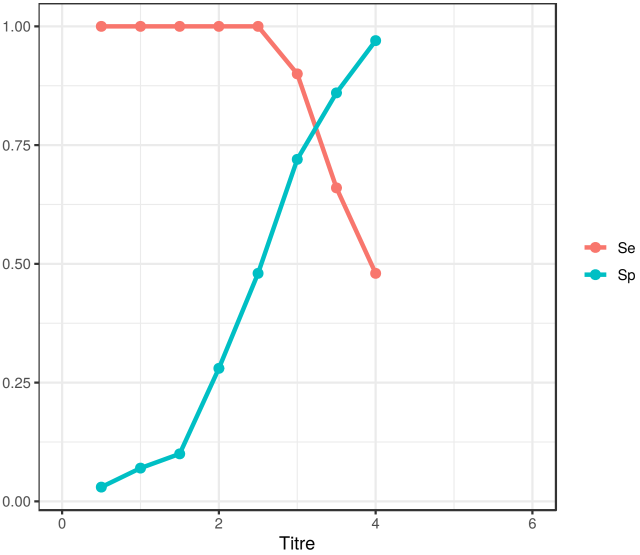
Se=0.66
Sp=0.86



Titre threshold:
T$+$: > 4.0
T$-$: ≤ 4.0

	I$+$	I$-$
T$+$	14	1
T$-$	15	28

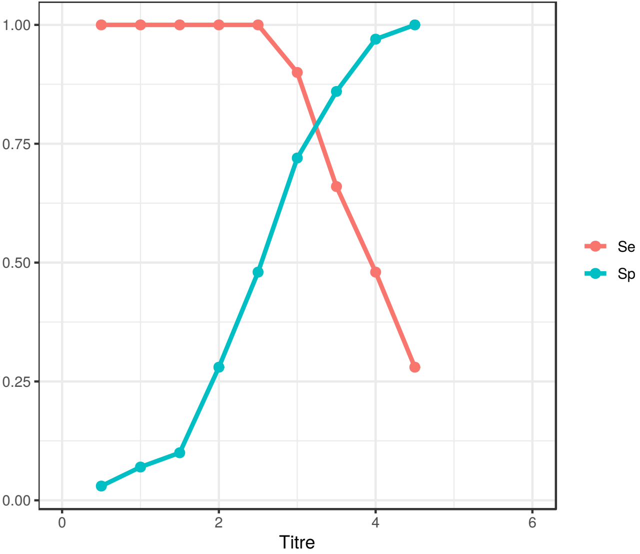
Se=0.48
Sp=0.97



Titre threshold:
T+: > 4.5
T−: ≤ 4.5

	I+	I−
T+	8	0
T−	21	29

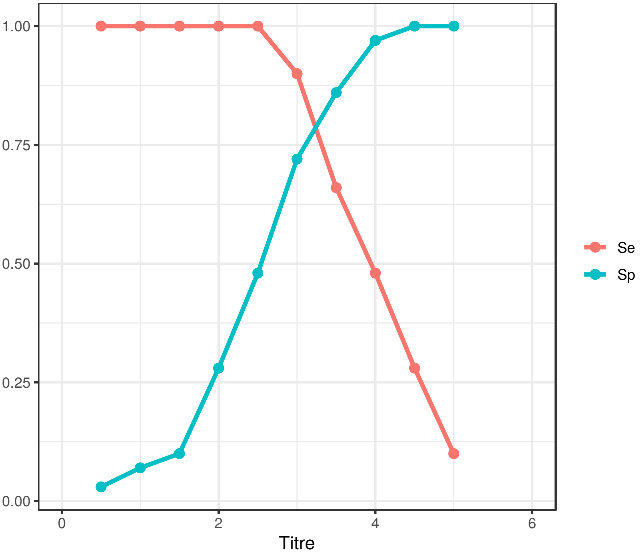
Se=0.28
Sp=1.00



Titre threshold:
T+: > 5.0
T−: ≤ 5.0

	I+	I−
T+	3	0
T−	26	29

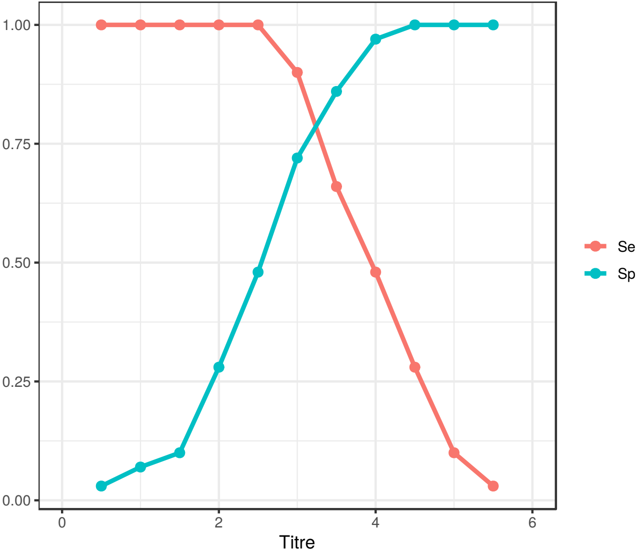
Se=0.10
Sp=1.00



Titre threshold:
T+: > 5.5
T−: ≤ 5.5

	I+	I−
T+	1	0
T−	28	29

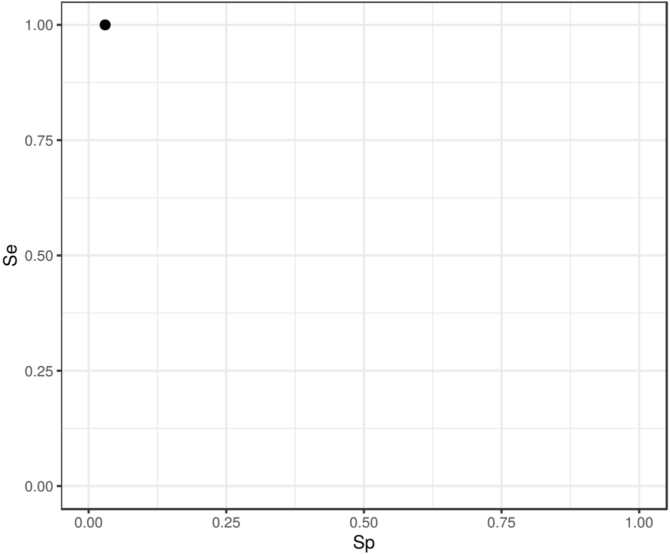
Se=0.03
Sp=1.00



Titre threshold:
T+: > 0.5
T−: ≤ 0.5

	I+	I−
T+	29	28
T−	0	1

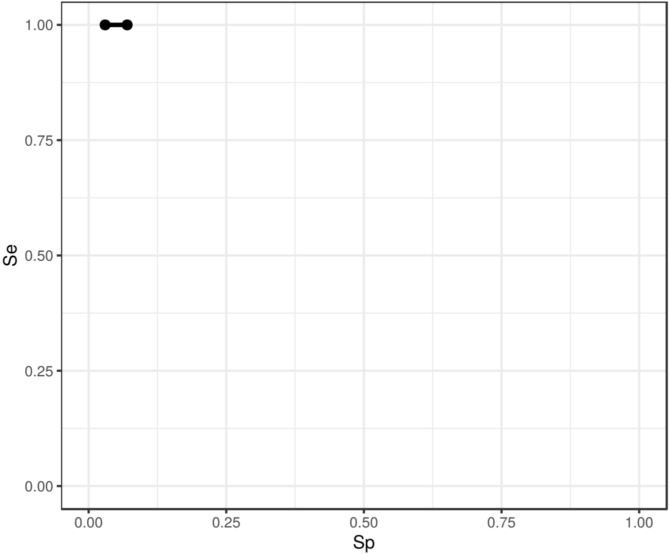
Se=1.00
Sp=0.03



Titre threshold:
T+: > 1.0
T−: ≤ 1.0

	I+	I−
T+	29	27
T−	0	2

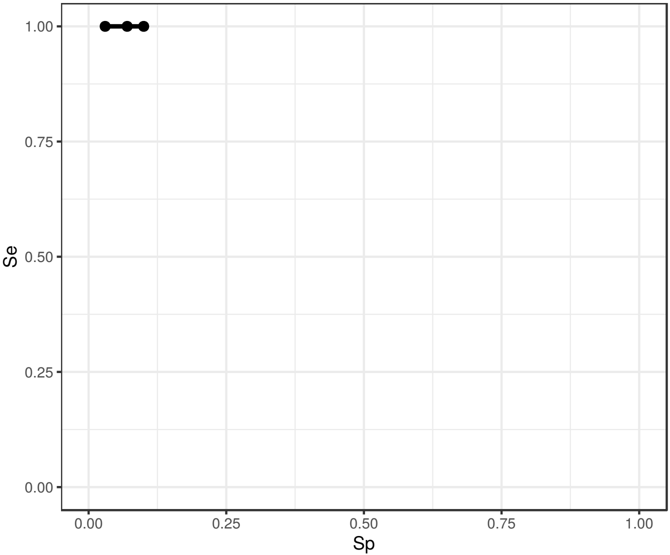
Se=1.00
Sp=0.07



Titre threshold:
T+: > 1.5
T−: ≤ 1.5

	I+	I−
T+	29	26
T−	0	3

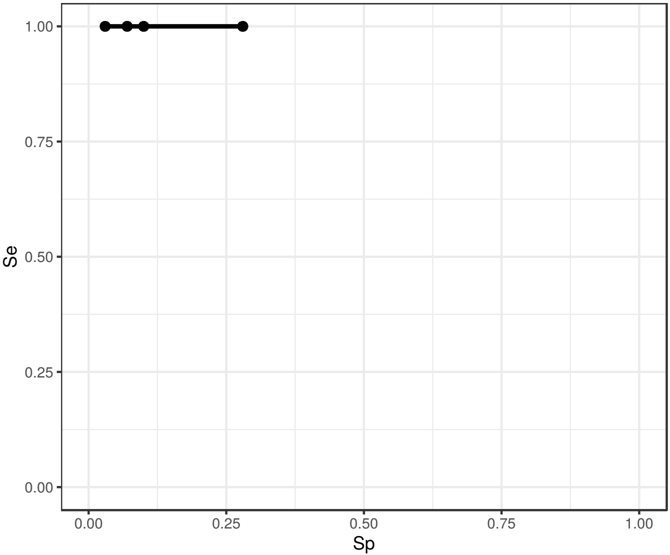
Se=1.00
Sp=0.10



Titre threshold:
T+: > 2.0
T−: ≤ 2.0

	I+	I−
T+	29	21
T−	0	8

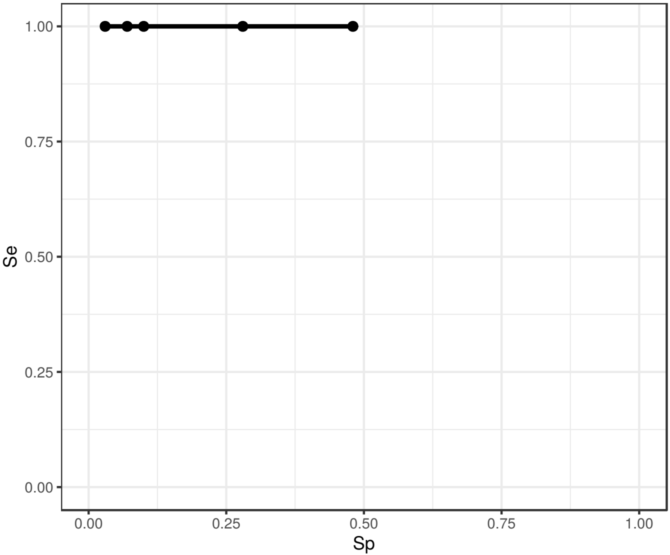
Se=1.00
Sp=0.28



Titre threshold:
T+: > 2.5
T−: ≤ 2.5

	I+	I−
T+	29	15
T−	0	14

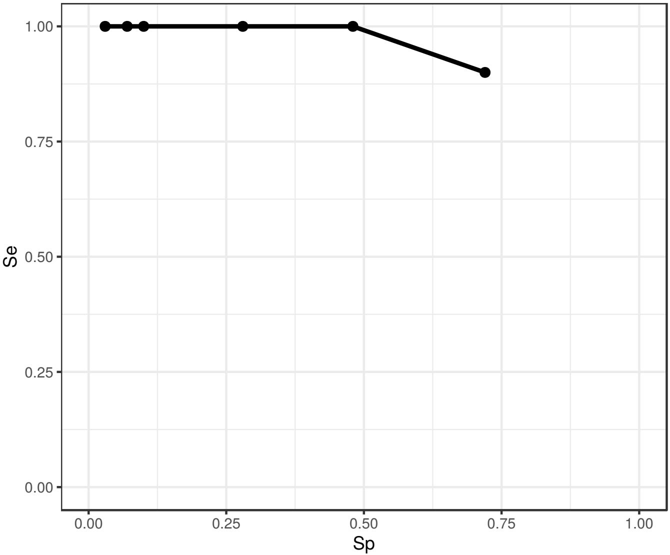
Se=1.00
Sp=0.48



Titre threshold:
T+: > 3.0
T−: ≤ 3.0

	I+	I−
T+	26	8
T−	3	21

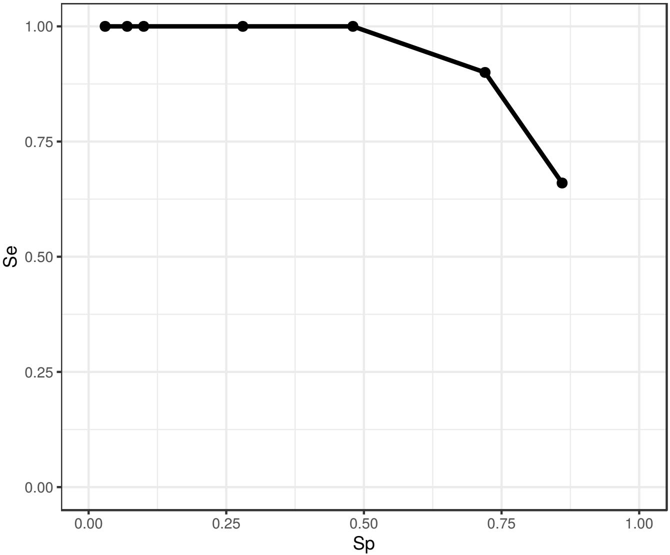
Se=0.90
Sp=0.72



Titre threshold:
T+: > 3.5
T−: ≤ 3.5

	I+	I−
T+	19	4
T−	10	25

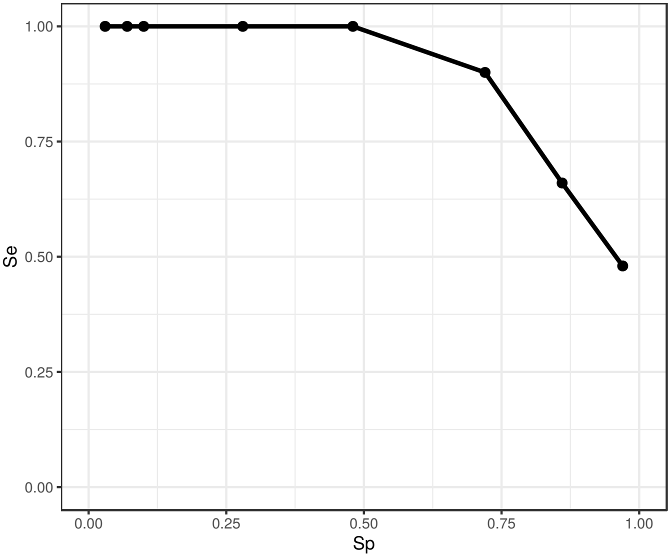
Se=0.66
Sp=0.86



Titre threshold:
T+: > 4.0
T−: ≤ 4.0

	I+	I−
T+	14	1
T−	15	28

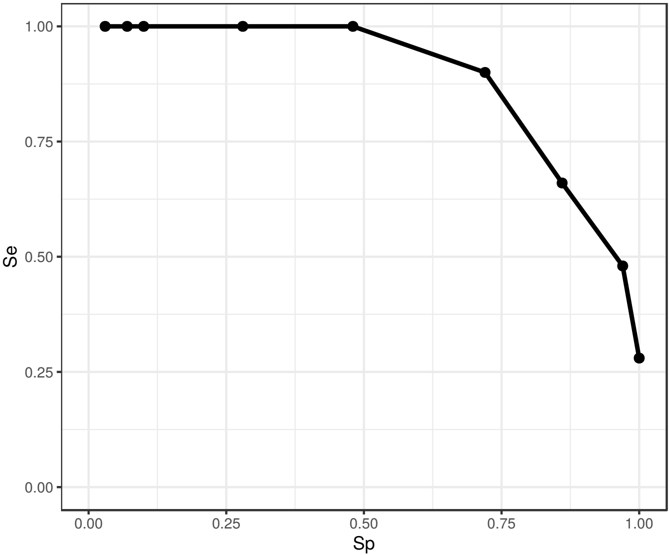
Se=0.48
Sp=0.97



Titre threshold:
T+: > 4.5
T−: ≤ 4.5

	I+	I−
T+	8	0
T−	21	29

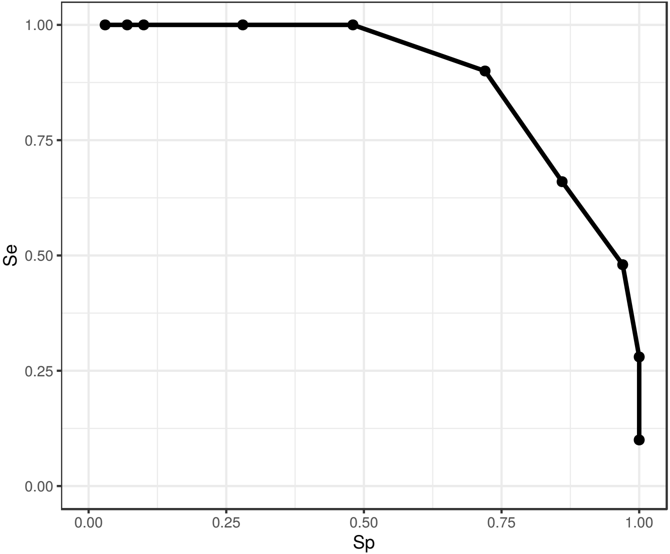
Se=0.28
Sp=1.00



Titre threshold:
T+: > 5.0
T−: ≤ 5.0

	I+	I−
T+	3	0
T−	26	29

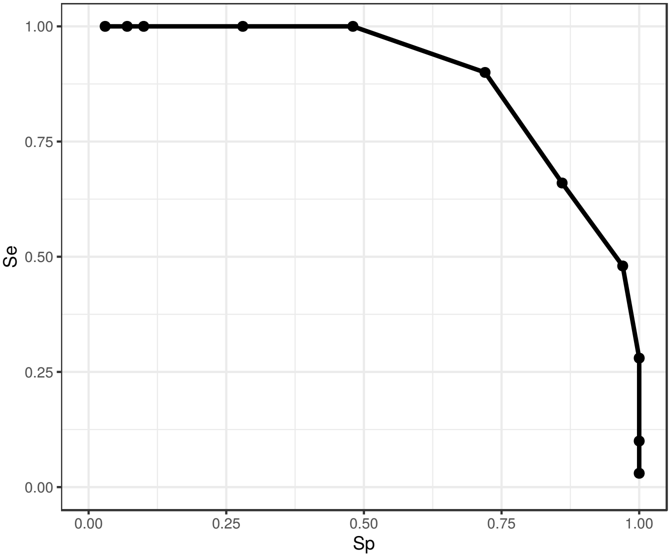
Se=0.10
Sp=1.00



Titre threshold:
T+: > 5.5
T−: ≤ 5.5

	I+	I−
T+	1	0
T−	28	29

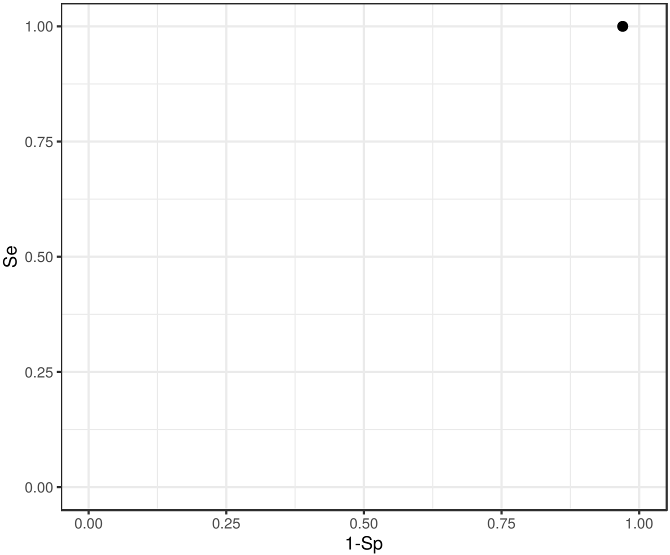
Se=0.03
Sp=1.00



Titre threshold:
T+: > 0.5
T−: ≤ 0.5

	I+	I−
T+	29	28
T−	0	1

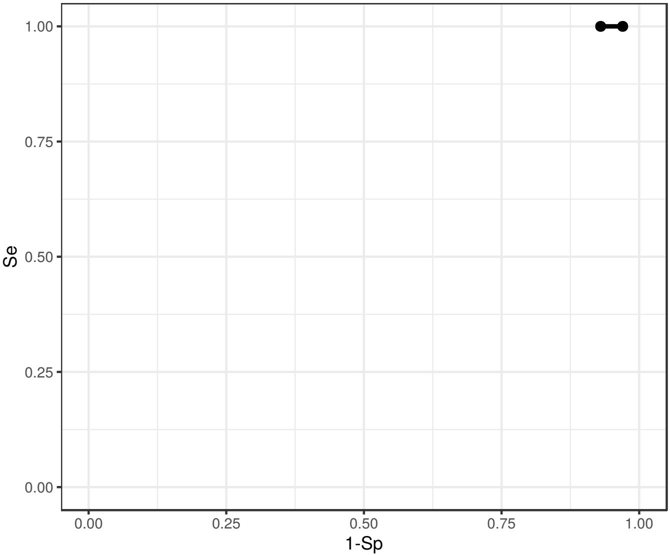
Se=1.00
Sp=0.03



Titre threshold:
T+: > 1.0
T−: ≤ 1.0

	I+	I−
T+	29	27
T−	0	2

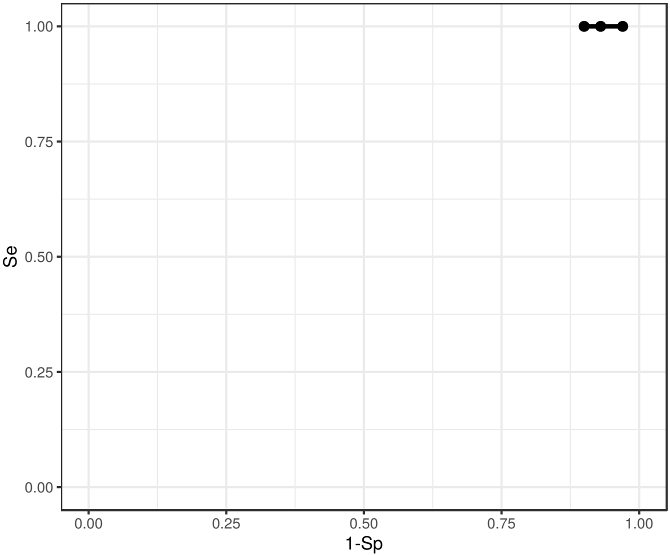
Se=1.00
Sp=0.07



Titre threshold:
T+: > 1.5
T−: ≤ 1.5

	I+	I−
T+	29	26
T−	0	3

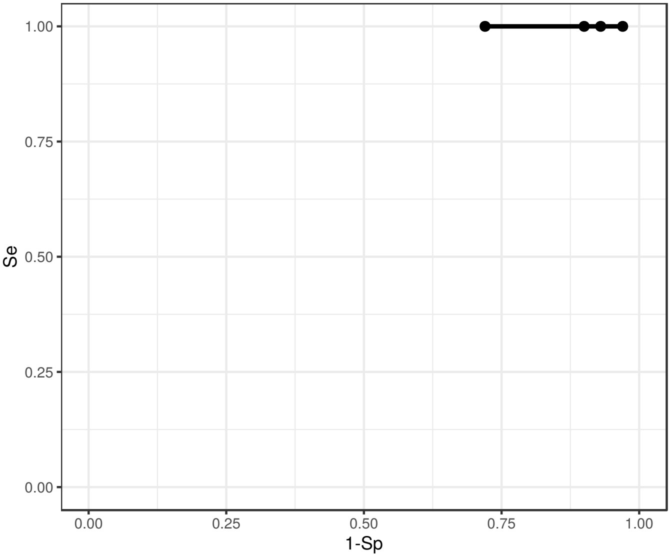
Se=1.00
Sp=0.10



Titre threshold:
T+: > 2.0
T−: ≤ 2.0

	I+	I−
T+	29	21
T−	0	8

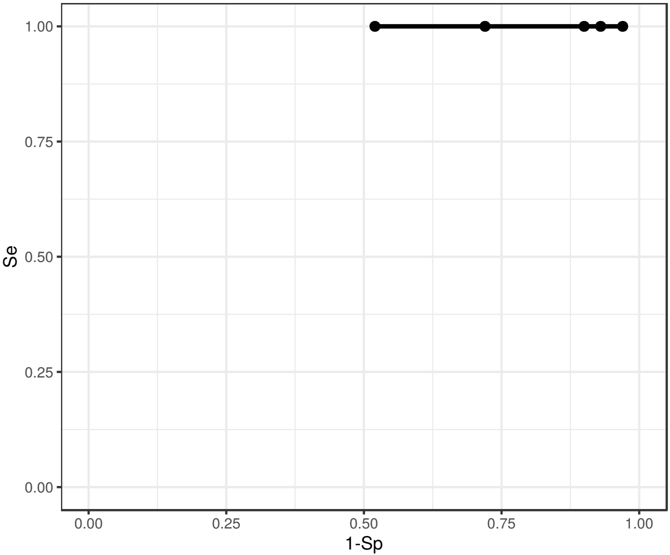
Se=1.00
Sp=0.28



Titre threshold:
T+: > 2.5
T−: ≤ 2.5

	I+	I−
T+	29	15
T−	0	14

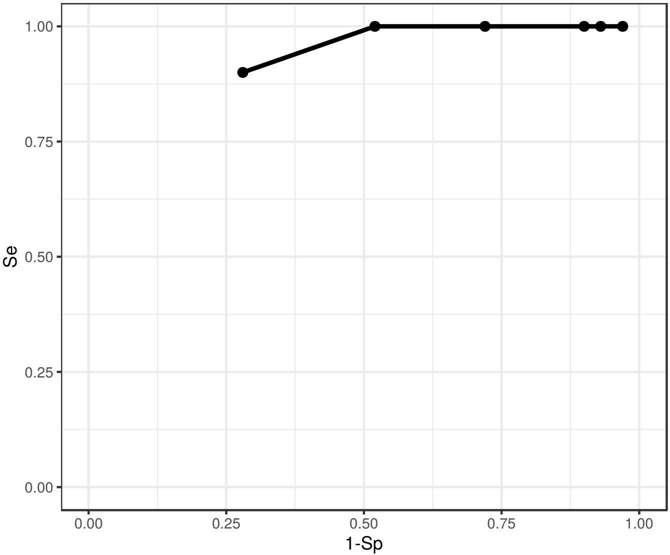
Se=1.00
Sp=0.48



Titre threshold:
T+: > 3.0
T−: ≤ 3.0

	I+	I−
T+	26	8
T−	3	21

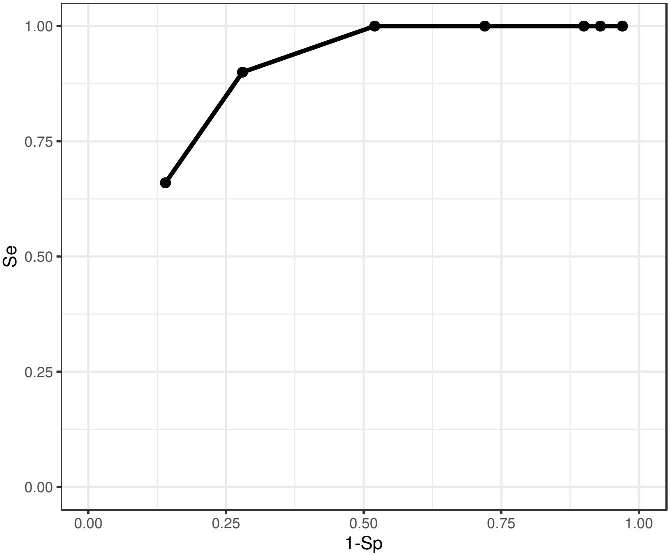
Se=0.90
Sp=0.72



Titre threshold:
T+: > 3.5
T−: ≤ 3.5

	I+	I−
T+	19	4
T−	10	25

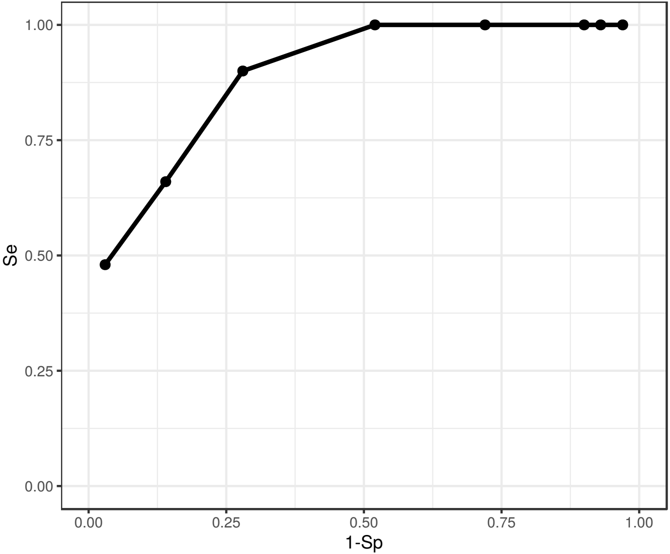
Se=0.66
Sp=0.86



Titre threshold:
T+: > 4.0
T−: ≤ 4.0

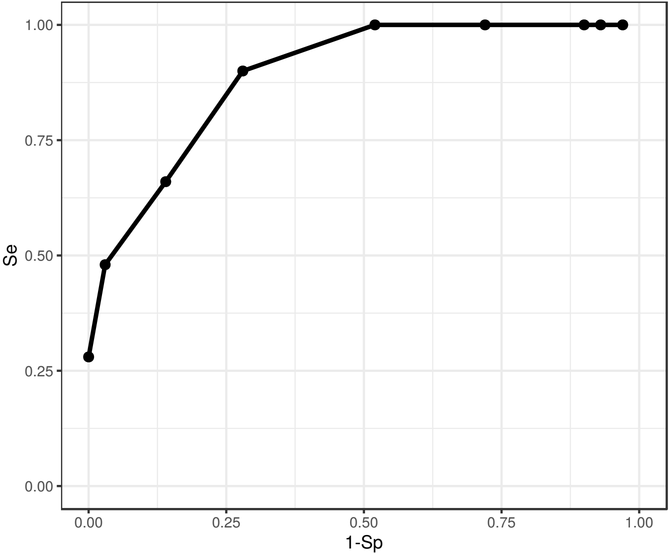
	I+	I−
T+	14	1
T−	15	28

Se=0.48
Sp=0.97



Titre threshold:
T+: > 4.5
T−: ≤ 4.5

	I+	I−
T+	8	0
T−	21	29

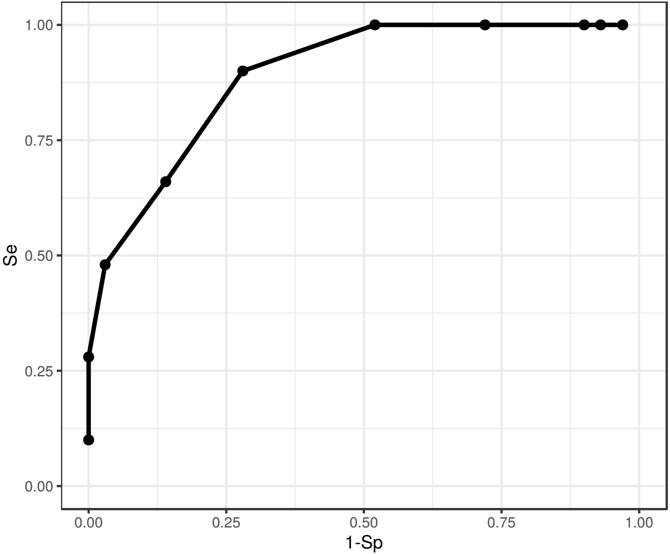


Se=0.28
Sp=1.00

Titre threshold:
T+: > 5.0
T−: ≤ 5.0

	I+	I−
T+	3	0
T−	26	29

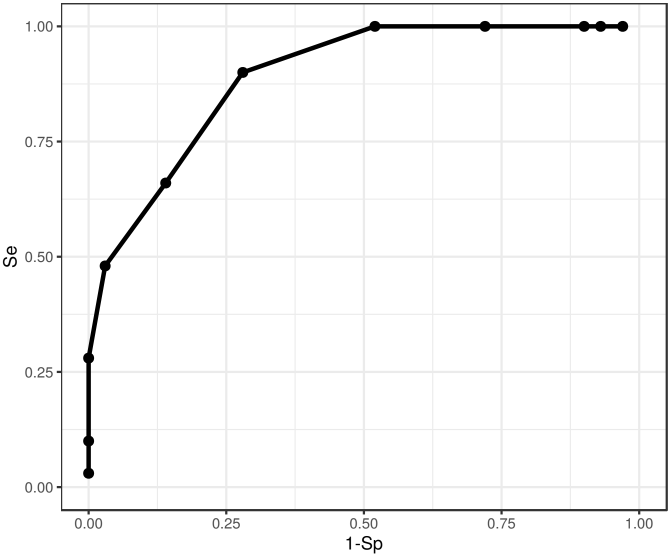
Se=0.10
Sp=1.00



Titre threshold:
T+: > 5.5
T−: ≤ 5.5

	I+	I−
T+	1	0
T−	28	29

Se=0.03
Sp=1.00



Practical question: If the test result is positive or negative, what is the probability that the subject is affected or not, respectively?

Andrew got a tattoo. Two months later he was refused as a blood donor. The phlebotomist explained that he had to wait a year to make sure he didn't get hepatitis B from the tattoo. That got him worried, so he ordered a home test kit for hepatitis B virus (HBV) from a website. The website said that the sensitivity of the test was 0.99 and the specificity was 0.995.

Hepatitis B is rare among those who are not intravenous drug users – about 2 cases per 100,000 people. Studies suggest that getting a tattoo from an operator who follows accepted hygenie standards does not greatly increase the risk. Let's assume that Andrew believed that his risk was about 3 in 100,000.

If Andrew expect 10 million people as population at risk, then about 300 would have HBV, and the rest would not. As we know HBV test has 99% sensitivity, which means that it will catch 99% of the HBV cases (297 of the 300 cases) and miss the rest.

The test has 99.5% specificity, which means that 99.5% of the noninfected people will test negative, but 0.5% of them will be false positives.

	HBV +	HBV −	$\sum$
Test +	297	49,998	50,295
Test −	3	9,949,702	9,949,705
$\sum$	300	9,999,700	10,000,000

Suppose Andrew tests negative. There are 9,949,705 people like him – negative. Of these only 3 have HBV, so there are 3 chances in 9,949,705 (about 1 in 3.3 million) that a person who tests negative actually is infected.

On the other hand, suppose Andrew tests positive. There are 50,295 people like him – positive. Out of this group, only 297 really do have HBV (about 1 of 170). That means that even if Andrew tests positive, there is still only about 0.6% chance that he is actually infected.

Another example: `http://yudkowsky.net/rational/bayes`

- uninfected pig:
  - test positive
  - test negative
- infected pig:
  - test positive
  - test negative

uninfected pig:

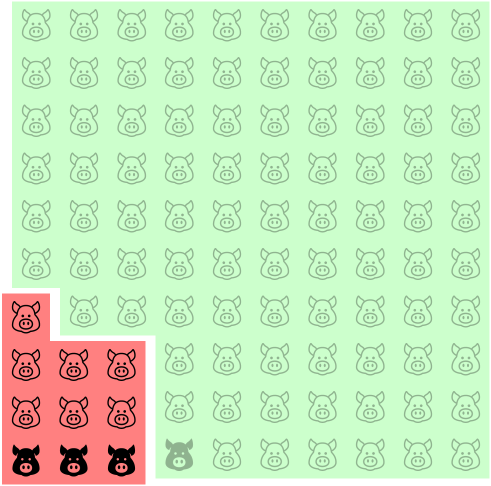
- test positive
- test negative

infected pig:

- test positive
- test negative

Positive predictive value is the proportion of test positives who actually infected:

$$3/10 = 0.30$$



uninfected pig:

- test positive
- test negative

infected pig:

- test positive
- test negative

Negative predictive value is the proportion of test negatives who are noninfected:

$$89/90 = 0.99$$

Positive predictive value is the proportion of animals tested positive while they are truly infected.
Or the probability that animals with a positive test result truly infected.

	Infection +	Infection −	Total
Test +	True positive a	False positive b	Test positive $a+b$
Test −	False negative c	True negative d	Test negative $c+d$
Total	$a+c$	$b+d$	$a+b+c+d$

Positive predictive value is the proportion of animals tested positive while they are truly infected.
Or the probability that animals with a positive test result truly infected.

	Infection +	Infection −	Total
Test +	True positive a	False positive b	Test positive $a+b$
Test −	False negative c	True negative d	Test negative $c+d$
Total	$a+c$	$b+d$	$a+b+c+d$

$$PPV = \frac{a}{a+b} = \frac{3}{3+7} = 0.30$$

Negative predictive value is the proportion of animals tested negative while they are truly uninfected.
Or the probability that animals with a negative test result truly uninfected.

	Infection $+$	Infection $-$	Total
Test $+$	True positive a	False positive b	Test positive $a+b$
Test $-$	False negative c	True negative d	Test negative $c+d$
Total	$a+c$	$b+d$	$a+b+c+d$

Negative predictive value is the proportion of animals tested negative while they are truly uninfected.
Or the probability that animals with a negative test result truly uninfected.

	Infection +	Infection −	Total
Test +	True positive a	False positive b	Test positive $a+b$
Test −	False negative c	True negative d	Test negative $c+d$
Total	$a+c$	$b+d$	$a+b+c+d$

$$NPV = \frac{d}{c+d} = \frac{89}{1+89} = 0.99$$

Predictive values

- evaluate a test's ability to correctly identify the condition of interest: that is, is the test useful?
- depend on how much disease is in the population you're testing

Example

- bovine tuberculosis
- the caudal skin fold test has a sensitivity of 0.75 to 0.85 and a specificity of 0.99
- what is the positive predictive value in a herd with TB prevalence of 10%?

Stevenson 2012

Bovine tuberculosis
Farm size: 10000

	Infection +	Infection −	Total
Test +	a	b	$a+b$
Test −	c	d	$c+d$
Total	$a+c$	$b+d$	10000 $a+b+c+d$

Bovine tuberculosis
Farm size: 10000, Prevalence: 0.10

	Infection +	Infection −	Total
Test +	a	b	$a+b$
Test −	c	d	$c+d$
Total	1000 $a+c$	9000 $b+d$	10000 $a+b+c+d$

Bovine tuberculosis
Farm size: 10000, Prevalence: 0.10, Se: 0.80

	Infection +	Infection −	Total
Test +	800 a	b	$a+b$
Test −	200 c	d	$c+d$
Total	1000 $a+c$	9000 $b+d$	10000 $a+b+c+d$

Bovine tuberculosis
Farm size: 10000, Prevalence: 0.10, Se: 0.80, Sp: 0.99

	Infection +	Infection −	Total
Test +	800	90	
	a	b	$a+b$
Test −	200	8910	
	c	d	$c+d$
Total	1000	9000	10000
	$a+c$	$b+d$	$a+b+c+d$

Bovine tuberculosis
Farm size: 10000, Prevalence: 0.10, Se: 0.80, Sp: 0.99

	Infection $+$	Infection $-$	Total
Test $+$	800 a	90 b	890 $a+b$
Test $-$	200 c	8910 d	9110 $c+d$
Total	1000 $a+c$	9000 $b+d$	10000 $a+b+c+d$

Bovine tuberculosis

Farm size: 10000, Prevalence: 0.10, Se: 0.80, Sp: 0.99

	Infection +	Infection −	Total
Test +	800	90	890
	a	b	$a+b$
Test −	200	8910	9110
	c	d	$c+d$
Total	1000	9000	10000
	$a+c$	$b+d$	$a+b+c+d$

$$NPV = \frac{d}{c+d} = \frac{8910}{200+8910} = \frac{8910}{9110} = 97.8\%$$

$$PPV = \frac{a}{a+b} = \frac{800}{800+90} = \frac{800}{890} = 89.9\%$$

Bovine tuberculosis

Farm size: 10000, Prevalence: 0.01, Se: 0.80, Sp: 0.99

	Infection +	Infection −	Total
Test +	80 a	99 b	179 $a+b$
Test −	20 c	9801 d	9821 $c+d$
Total	100 $a+c$	9900 $b+d$	10000 $a+b+c+d$

$$NPV = \frac{d}{c+d} = \frac{9801}{20+9801} = \frac{9801}{9821} = 99.8\%$$

$$PPV = \frac{a}{a+b} = \frac{80}{80+99} = \frac{80}{179} = 44.7\%$$

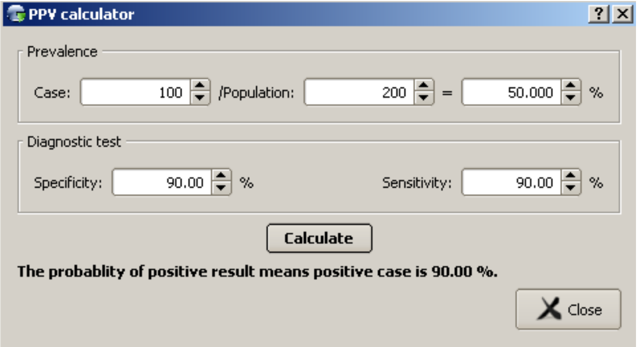
Bovine tuberculosis
Prevalence: 0.10, Se: 0.80, Sp: 0.99

$$
\begin{aligned}
NPV = \frac{d}{c+d} &= \frac{(1-P)\times Sp}{P\times(1-Se)+(1-P)\times Sp} \\
&= \frac{(1-0.10)\times 0.99}{0.10\times(1-0.80)+(1-0.10)\times 0.99} = 97.8
\end{aligned}
$$

$$
\begin{aligned}
PPV = \frac{a}{a+b} &= \frac{P\times Se}{P\times Se+(1-P)\times(1-Sp)} \\
&= \frac{0.10\times 0.80}{0.10\times 0.80+(1-0.10)\times(1-0.99)} = 89.9
\end{aligned}
$$

- as prevalence falls
    - PPV decreases regardless of sensitivity and specificity of the test
    - even very good tests become very poor at predicting the presence of disease

"According to our calculation the positive predictive value of the best Lyme antibody tests if applied in this way is 9.1%."



http://kullancs.hu/doc/PPVcalc_setup.exe

Lakos et al. (2010)

How much the **odds of the disease increase** when a test is positive?

- likelihood ratio for positive test result

$$LR^+ = \frac{Se}{1 - Sp}$$

How much the **odds of the disease decrease** when a test is negative?

- likelihood ratio for negative test result

$$LR^- = \frac{1 - Se}{Sp}$$

Bayes' Theorem:
Post-test probability = pre-test probability × likelihood

$$\text{Odds of event} = \frac{\text{Probability of event}}{\text{1-Probability of event}}$$

$$\text{Probability of event} = \frac{\text{Odds of event}}{\text{1+Odds of event}}$$

- California Mastitis Test Se: 68.8%, Sp: 71.5%
- $LR^+ = \dfrac{0.688}{1-0.715} = 2.414$
- $LR^- = \dfrac{1-0.688}{0.715} = 0.436$
- The pre-test probability of mastitis: $\dfrac{50}{1000} = 0.05$
- The pre-test odds of mastitis: $\dfrac{0.05}{1-0.05} = 0.053$
- The post-test odds of mastitis given a positive test result: pre-test odds $\times\ LR^+ = 0.053 \times 2.414 = 0.1279$
- The post-test probability of mastitis given a positive test result: $\dfrac{0.1279}{1+0.1279} = 0.11$

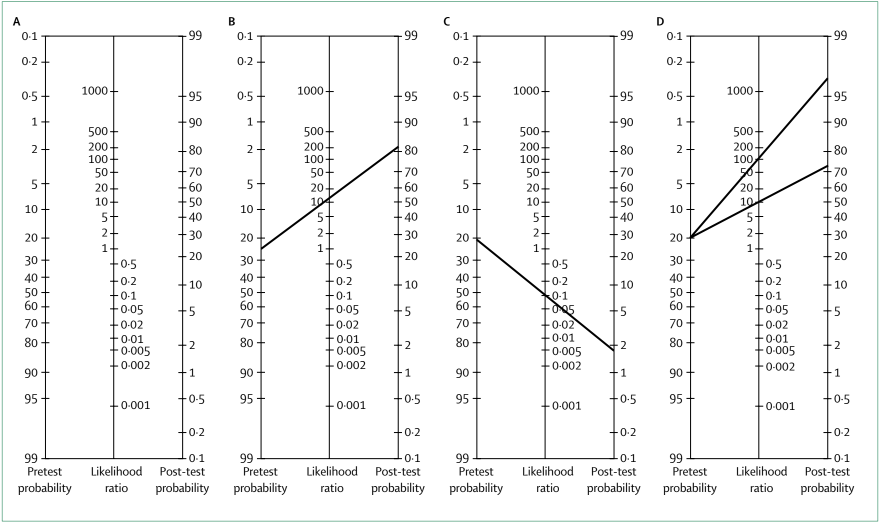


***Figure* 2: Nomograms for probabilities and likelihood ratios**[13]

(A) Nomogram reprinted from reference 13 with permission of the Massachusetts Medical Association. (B) Straight edge applied for pretest probability of 0·25 and likelihood ratio of 13. (C) Straight edge applied for pretest probability of 0·20 and likelihood ratio of 0·1. (D) Effect of likelihood ratios of 10 and 100 on pretest probability of 0·2.

Grimes and Schulz (2005)

- 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.

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

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

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

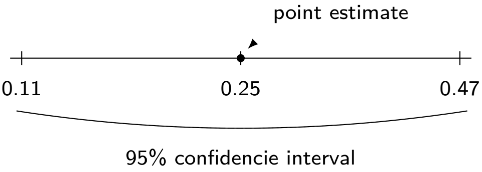
$$P_A = \frac{5}{20} = 0.25$$

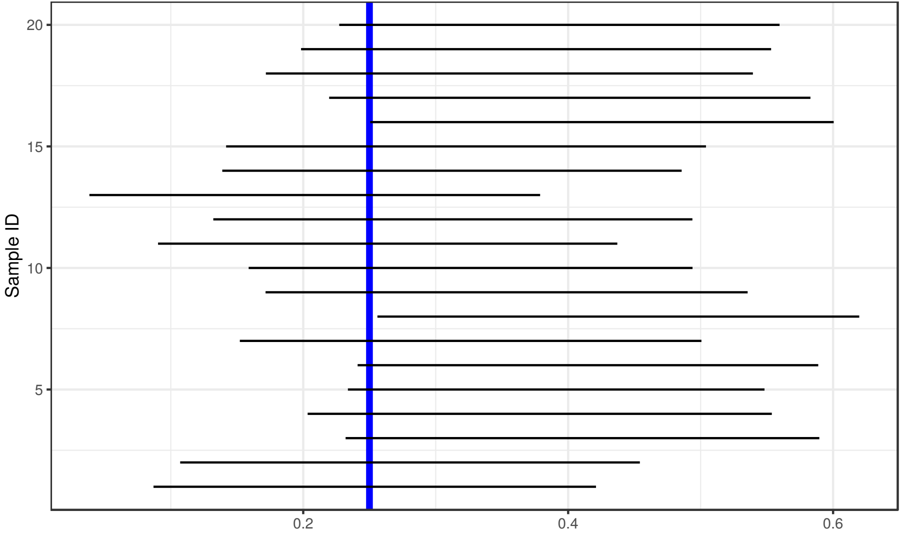
- Apparent (P_A): the probability that a randomly-choosen unit of observation will test positive

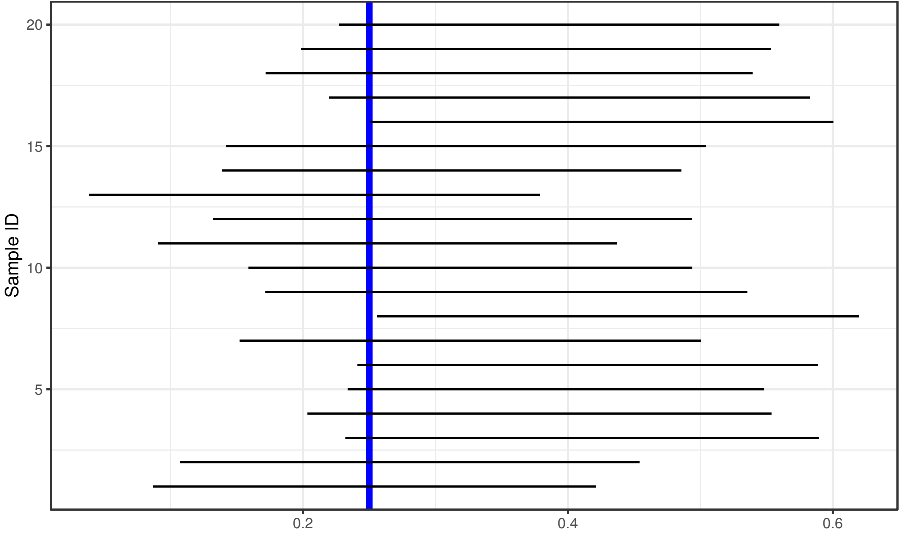
$$P_A = \frac{x}{n}$$

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

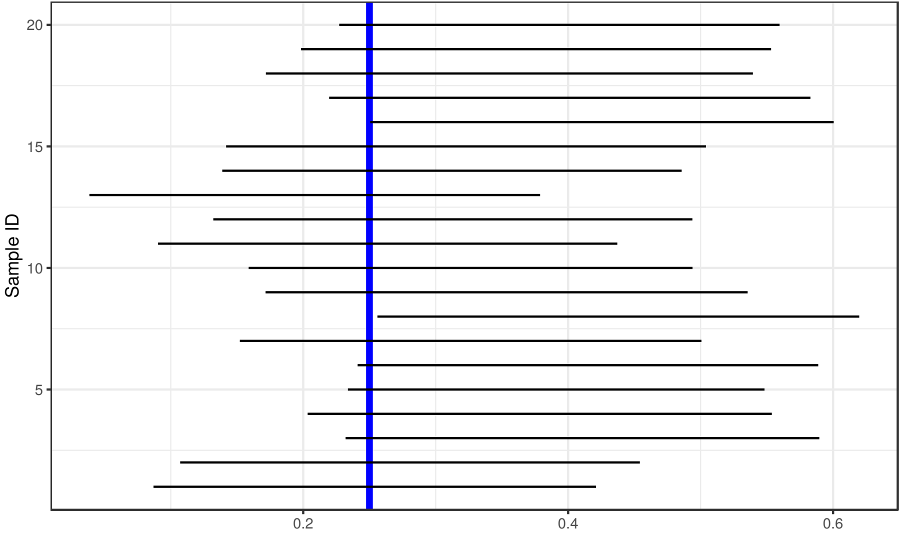
$$P_A = \frac{5}{20} = 0.25$$



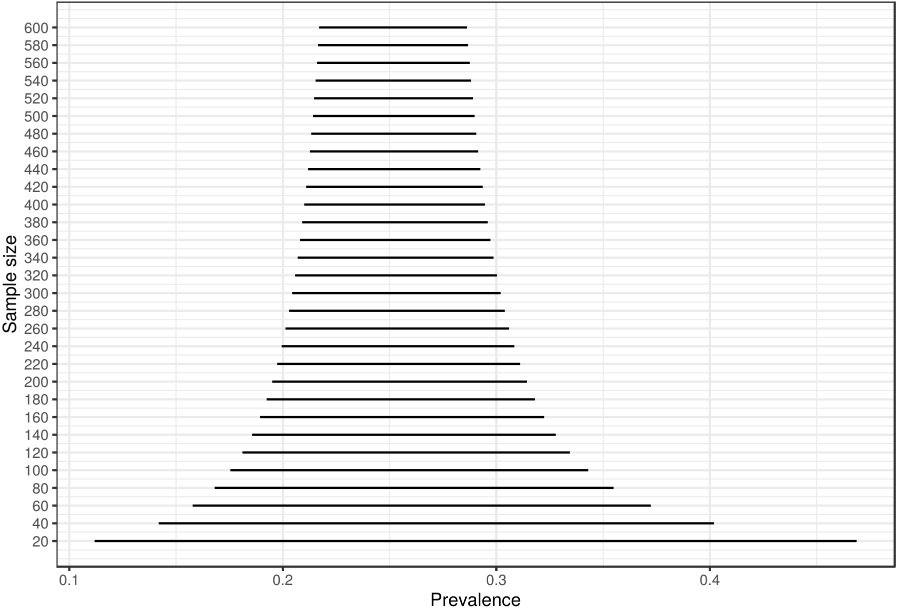

Sample ID
20
15
10
5
0.2
0.4
0.6



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

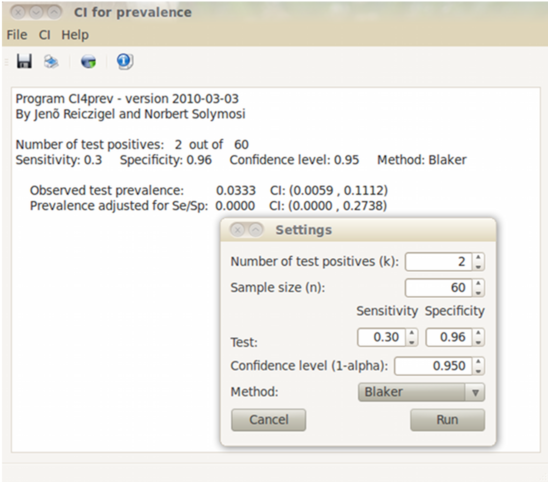
  $P_A = 5/20 = 0.25$

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

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

  $P_A = 2/60 = 0.033$

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



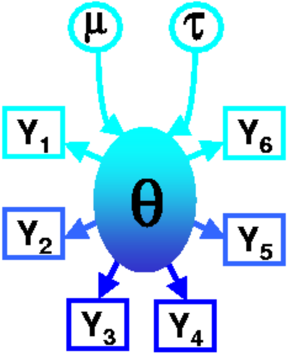
http://solymosin.github.io/CI4prev/

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

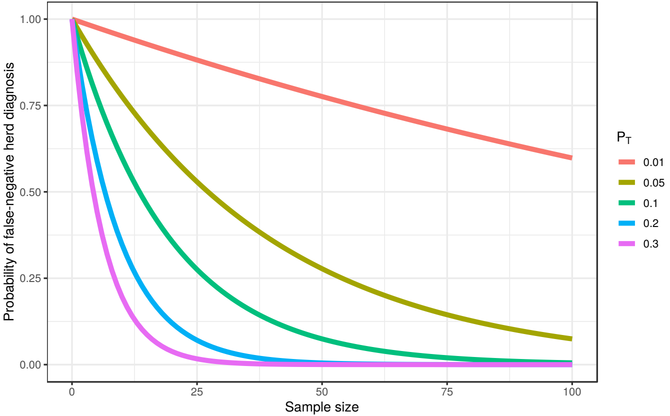
https://www.mrc-bsu.cam.ac.uk/software/bugs
https://cadms.vetmed.ucdavis.edu/diagnostic-tests

- 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

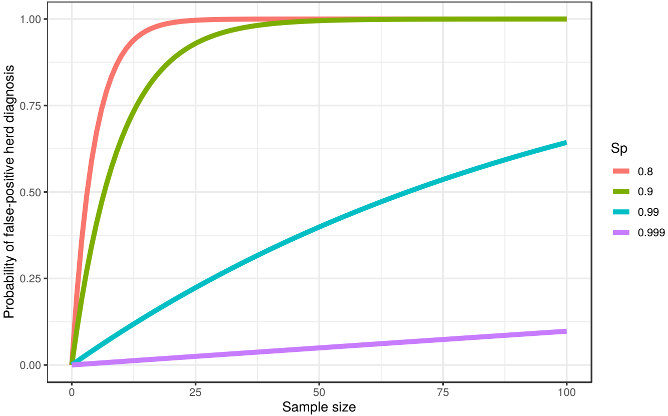
number of tested animals ↑:

- 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.

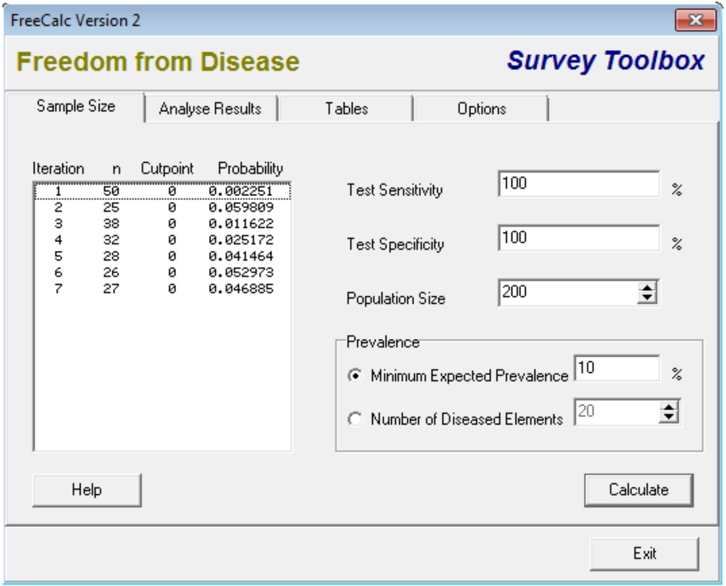
To estimate required numbers to detect infection, the following values are necessary: the **required level of confidence**, usually 95%, the likely **prevalence** of infection in the herd or in the specific group of pigs being evaluated, the population size.

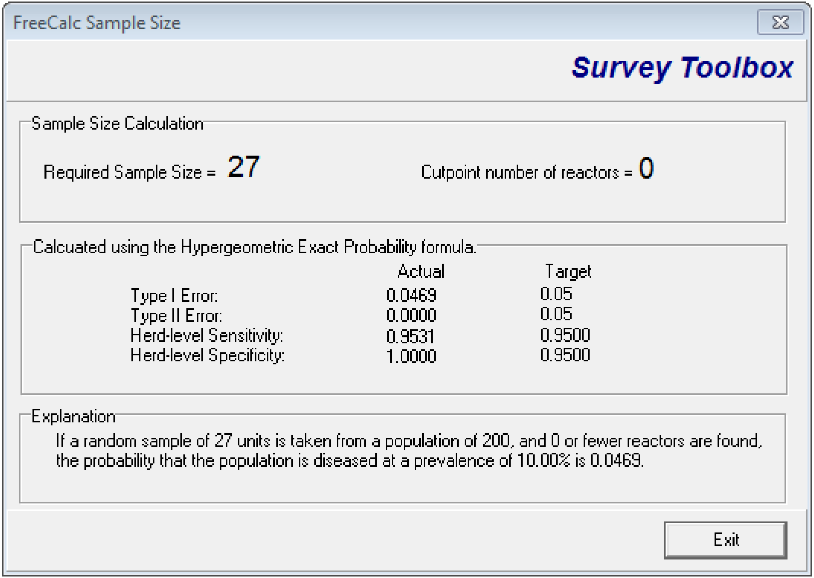
The selected prevalence value should be realistic, but if there is doubt, erring toward a **lower** prevalence is preferable to ensure that adequate numbers of pigs are sampled.

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.

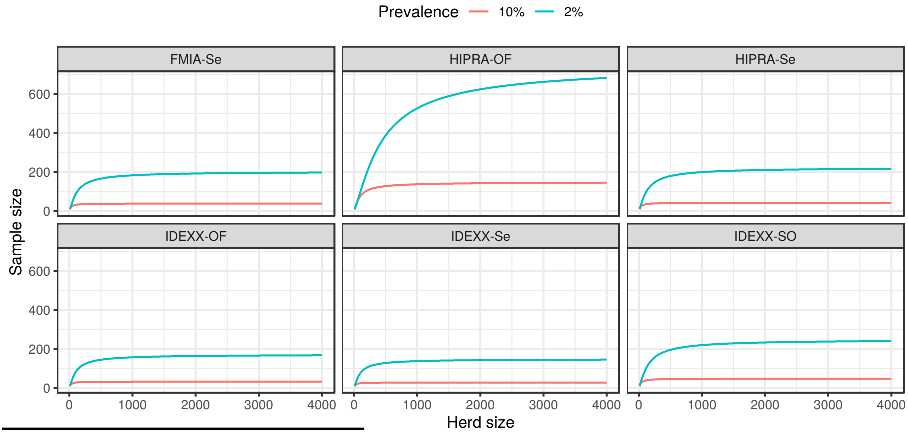
To detect *T. gondii* in a herd, sows are a better population to sample because prevalence is likely to be higher than in grower-finisher pigs.

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.

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.

Grimes, D. A. and K. F. Schulz (2005). Refining clinical diagnosis with likelihood ratios. *The Lancet 365*(9469), 1500–1505.

Lakos, A., J. Reiczigel, and N. Solymosi (2010). The positive predictive value of borrelia burgdorferi serology in the light of symptoms of patients sent to an outpatient service for tick-borne diseases. *Inflammation research 59*(11), 959–964.

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.

Stevenson, M. (2012). An introduction to veterinary epidemiology. EpiCentre, IVABS, Massey University, Palmerston North, New Zealand.

Woodworth, G. G. (2004). *Biostatistics: A Bayesian Introduction*. Probability and Statistics Series. Hoboken, New Jersey, USA: Wiley & Sons. ISBN 0471468428.