Locksley Messam
Conor McAloon
Eleftherios Meletis

Diagnostic test evaluation with Bayesian latent class models

UCD DUBLIN

cost
EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

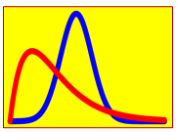
COST is supported by the EU Framework Programme Horizon 2020

HARMONY
Novel tools for test evaluation and disease prevalence estimation

1

UCD DUBLIN

Diagnostic Tests



cost
EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

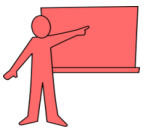
COST is supported by the EU Framework Programme Horizon 2020

HARMONY
Novel tools for test evaluation and disease prevalence estimation

	D+	D-
T+	TP	FP
T-	FN	FN

Locksley L. McV. Messam Dr. Med. Vet., PhD.
Lecturer in Veterinary Epidemiology
University College Dublin, School of Veterinary Medicine,
Section: Herd Health and Animal Husbandry

2



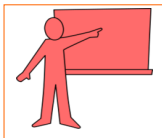
Epictetus (AD 55 – 135)

"**Appearances to the mind are of four kinds:**

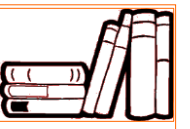
- Things either are what they appear to be; **TP**
- or they neither are, nor appear to be; **TN**
- or they are, and do not appear to be; **FN**
- or they are not, and yet appear to be. **FP**

Rightly to aim in all these cases is the wise man's task."

3



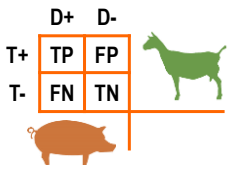
Topics



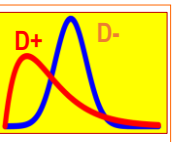
- Diagnostic tests (Definition, Types, Uses)
- Diagnostic Test Characteristics
- Applications of Diagnostic Tests

	D+	D-
T+	TP	FP
T-	FN	FN

4

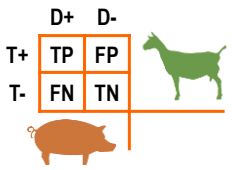


Diagnostic Tests

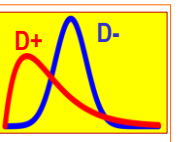


"Any device or process designed to detect a sign, substance, tissue change or response."

5



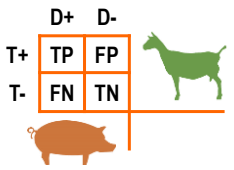
Diagnostic Tests



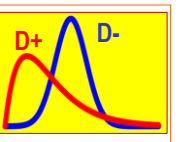
"…sign, substance, tissue change or response."

- Not necessarily disease
- Outcome of interest: Pregnancy checks via palpations, X-rays, Ultrasounds,….., Confounders too!!!

6



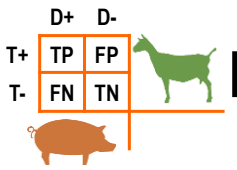
Diagnostic Tests



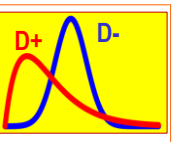
- **Dichotomous:** Fever, bacteria present?, Broken limb – Yes or No.
- **Ordinal:** Serologic Titres (1:50, 1:250, 1:500), lameness severity (1 - 5), Body condition score, Pain level (1-10)
- **Continuous scale:** WBC count, Temperature

Ordinal & continuous ⇒ Dichotomous *(i.e., +/- at given cut off).*

7

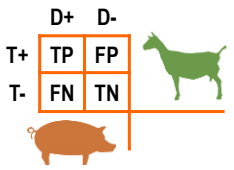


Diagnostic Tests: Uses

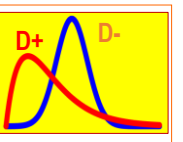


- **Detect (screen) or confirm** disease, infection, infectiousnes (individual or group)
  - Isolation and eradication (test and removal of reactors – Vet. Med.)
- **Institute preventive measures**: Human (Covid-19) & Vet. Med.
- **Monitor vaccination status:** Rabies, Tetanus

8



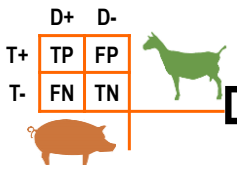
Diagnostic Tests: Gold Standard



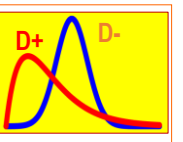
"A diagnostic method (or combination of methods) which determines without error, whether or not the organism has the condition (outcome).."

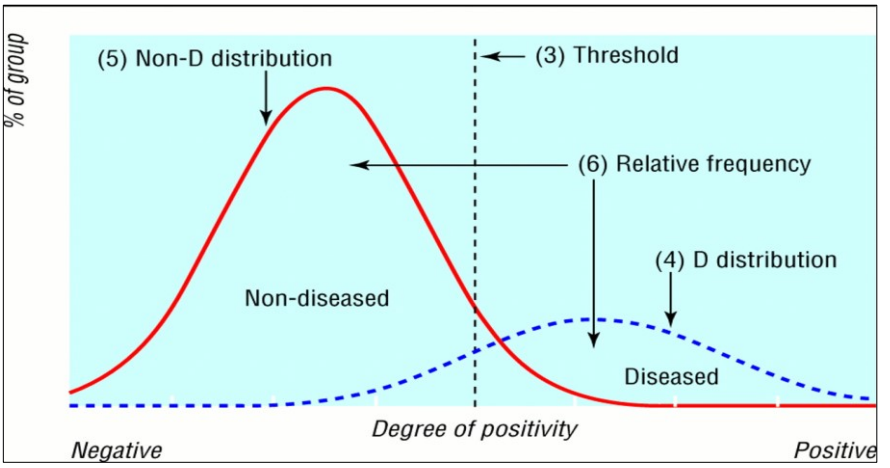
- Many diseases = no gold standard or not practical (costly, invasive)
- Vet. Med. - Few gold standards (if any)

9

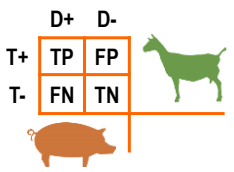


Diagnostic Tests: Diseased and Non-Diseased

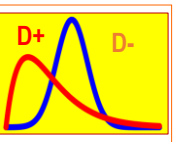




10

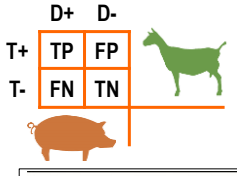


Diagnostic Tests

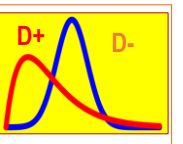


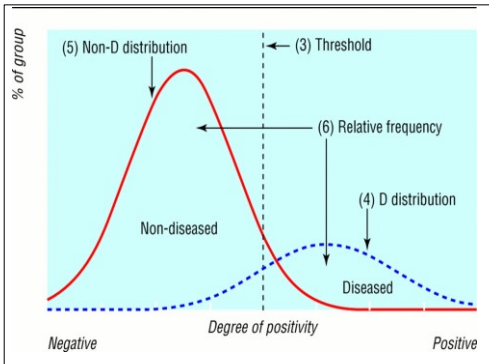
	D+	D-	
T +	Diseased and test Positive **(TP)**	Not diseased but test positive **(FP)**	TP + FP
T -	Diseased but test negative **(FN)**	Not diseased and test negative **(TN)**	TN +FN
			TP + FP + TN +FN

11



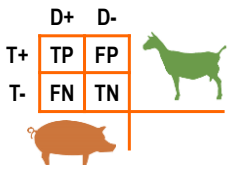
Diagnostic Tests



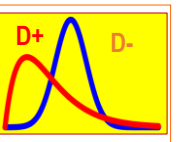


	D+	D-	
T +	TP	FP	TP + FP
T -	FN	TN	TN +FN
	TP+FN	FP+TN	TP+FP+TN+FN

12

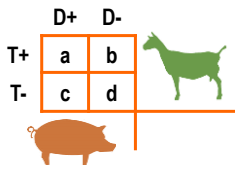


Diagnostic Tests: Gold Standard

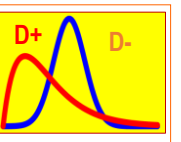


	D+	D-	
T +	Diseased and test Positive **(TP)**	**FP = 0**	**TP + FP**
T -	**FN = 0**	Not diseased and test negative **(TN)**	**TN +FN**
	TP + FN	**FP + TN**	**TP + FP + TN +FN**

13



Diagnostic Tests (Characteristics)

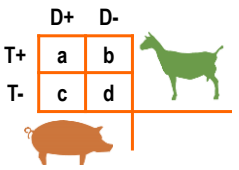


Sensitivity:

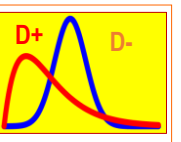
- Probability that randomly chosen diseased individual will test positive = P (T+|D+)
- Proportion of diseased that test positive
- Range 0 - 1 (0 – 100%)
- Gold Standard: Se = 100

$$Se = \frac{\text{Number of diseased individuals testing positive}}{\text{Total number of diseased individuals}} = \frac{a}{a + c} = \frac{TP}{TP + FN}$$

14



Diagnostic Tests (Characteristics)

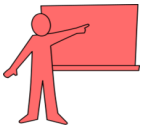


Specificity:

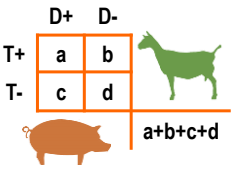
- Probability that randomly chosen non-diseased individual will test negative = P(T-|D-)
- Proportion of non-diseased that test negative
- Range: 0 - 1 (0 - 100%)
- Gold Standard: Sp = 100%

$$Sp = \frac{\text{Number of non-diseased individuals testing negative}}{\text{Total number of non-diseased individuals}} = \frac{d}{b+d} = \frac{TN}{TN+FP}$$

15



Diagnostic Tests: (Factors affecting Se and Sp)

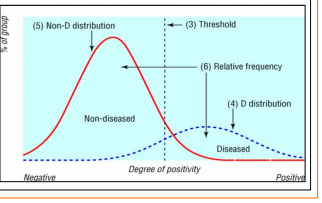


$$Se = \frac{a}{a+c}$$

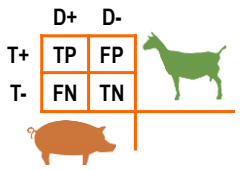
- **Source population (disease stage)** – e.g. TB lower sensitivity - *early, late stages, after calving*
- **Choice of cut-off (continuous tests)** – Se decreases with increase in cut-off (vice-versa)
- **Prevalence** – complex - distribution of biological factors related to infection also related to prevalence.

$$Sp = \frac{d}{b+d}$$

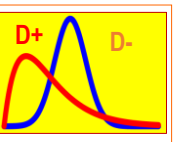
- **Cross-reacting agents** (Serologic tests) – Similar Antibodies (*E. coli*, *Pasteurella, Hemophilus spp.*) - False positives
- **Choice of cut-off (continuous tests)** – Sp increases with increase in cut-off (vice-versa)



16



Diagnostic Tests:
(Sensitivity and Specificity)



Answers 4 questions for the clinician

Chance that a **diseased** individual will test positive?

Chance that a **diseased** individual will test negative?

Chance that a **non-diseased** individual will test positive?

Chance that a **non-diseased** individual will test negative?

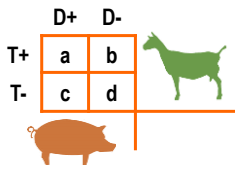
NOT HELPFUL!

Clinician asks:

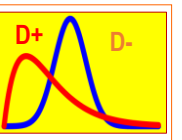
If an individual **tests positive,** chance that it is diseased (or non-diseased)?

If an individual **tests negative,** chance that it is non-diseased (or diseased)?

17



Diagnostic Tests
Characteristics

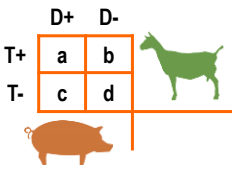


Positive Predictive Value (PPV):

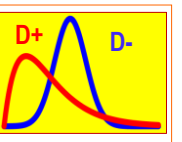
- Probability that an individual that tests positive is truly diseased = P(D+|T+)
- Proportion of T+ that are truly D+
- Range 0 - 1 (0 – 100%)
- Gold Standard: PPV = 1

$$PPV = \frac{\text{Number of T+ individuals that are D+}}{\text{Total number of D+ individuals}} = \frac{a}{a+b} = \frac{TP}{TP+FP}$$

18



Diagnostic Tests
(Characteristics)

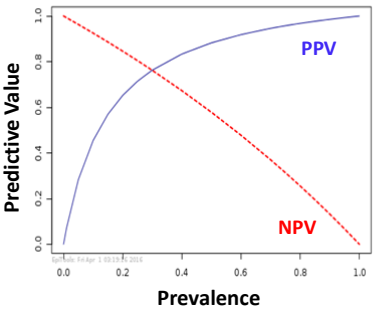


PPV: Dependent on Se, Sp and Prevalence (P).

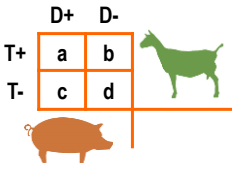
$$PPV = \frac{P \times Se}{(P \times Se) + [(1 - P) \times (1 - Sp)]}$$

Disadvantage: For fixed Se, Sp very dependent on prior probability of outcome (i.e., prevalence)

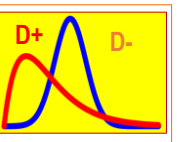
Advantage: Preserves diagnostic logic order



19



Diagnostic Tests
(Characteristics: Predictive Values)

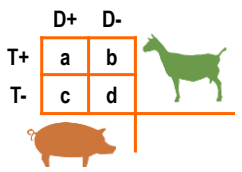


Negative Predictive Value (NPV):

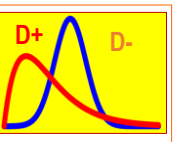
- Probability that an individual that tests negative is truly non-diseased = Prob (D-|T-)
- Proportion of T- that are truly D-
- Range 0 -1 (0 – 100%)
- Gold Standard: NPV = 1

$$NPV = \frac{\text{Number of T- individuals that are D-}}{\text{Total number of T- individuals}} = \frac{d}{c + d} = \frac{TN}{TN + FN}$$

20



Diagnostic Tests
(Characteristics: Predictive Values)

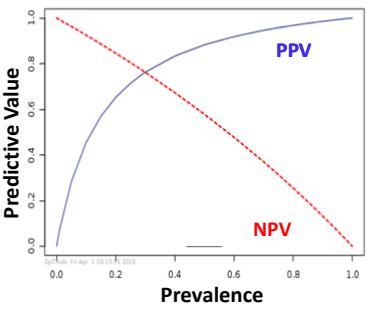


NPV: Dependent on Se, Sp and Prevalence (P)

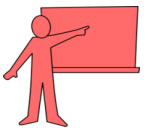
$$NPV = \frac{(1 - P) \times Sp}{[(1 - P) \times Sp] + [P \times (1 - Se)]}$$

Disadvantage: For fixed Se, Sp, very dependent on prior probability of outcome (i.e., prevalence)

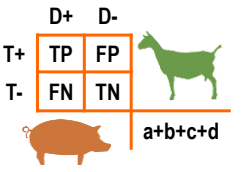
Advantage: Preserves diagnostic logic order



21



Diagnostic Tests:
(Likelihood Ratios)



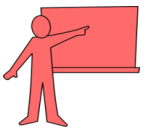
$$LR+ = \frac{Se}{1 - Sp}$$

- Chance of diseased animal tests positive divided by chance that a non-diseased animal tests positive
- How many **more times** will a diseased animal test positive than a non-diseased animal
- Range: 1 to ∞
- Unique for each test
- Association between D+ and T+ ↑ → LR+ ↑

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

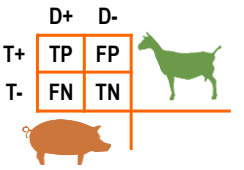
- Chance of diseased animal tests negative divided by chance that a non-diseased animal tests negative
- How many **more times** will a diseased animal test negative than a non-diseased animal
- Range: 0 to 1
- Unique for each test
- Association between D- and T- ↑ → LR+ ↓

22



Fagan's Nomogram:

Fagan (NEJM, 1975)



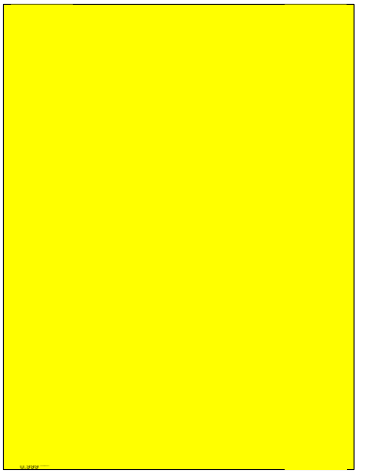
Fagan's Nomogram

- Two – dimensional, graphical analogue computation device
- **Three scales:**

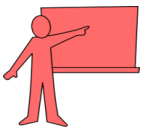
  i. **Pre-test probability** = Opinion before testing

  i. **Likelihood ratio** = How much more likely is a D+ to have a given test result than a D-

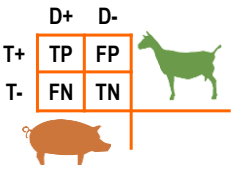
  iii. **Post-test probability** = Opinion after testing



23



Fagan's Nomogram:
(Evidence-Based Medicine)



Fagan's Nomogram:

Two important elements:

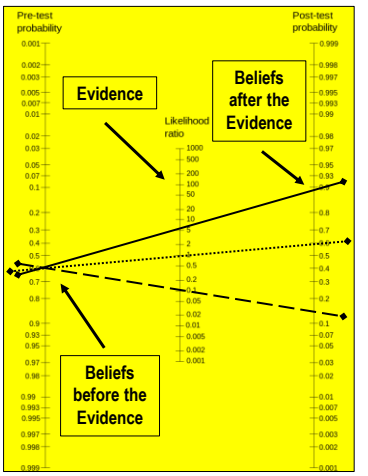
i. Clinical expertise – Prior beliefs, opinion

+

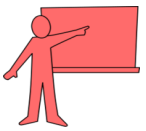
ii. External sources – Evidence: results of the diagnostic test

⬇

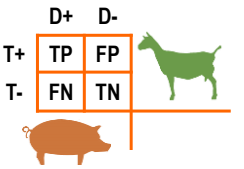
Diagnosis = Final beliefs, opinion



24



Fagan's Nomogram:
(Evidence-Based Medicine)

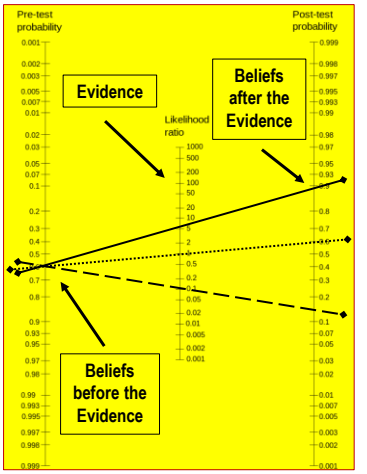


Fagan's Nomogram

- Links pre-test probabilities (prevalences) to post-test probabilities

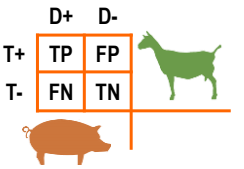
Pre-Test Odds × LR = Post-Test Odds

$$\frac{P_{Pre\text{-}Test}}{1 - P_{Pre\text{-}Test}} \times LR = \frac{P_{Post\text{-}Test}}{1 - P_{Post\text{-}Test}}$$



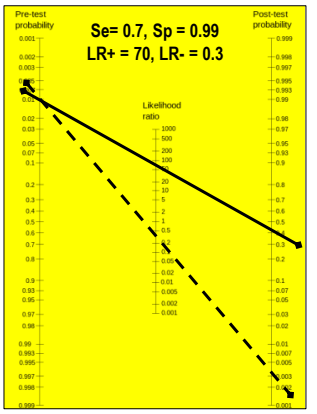
25

Covid -19 (Hypothetical Example)

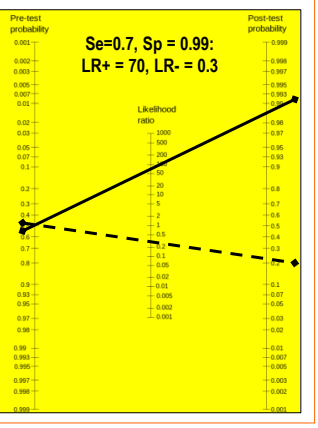


Interpretation depends on pre-test probability

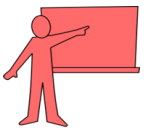
Low Exposure Risk



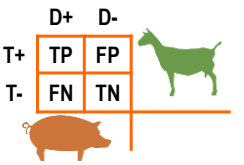
High Exposure Risk



26



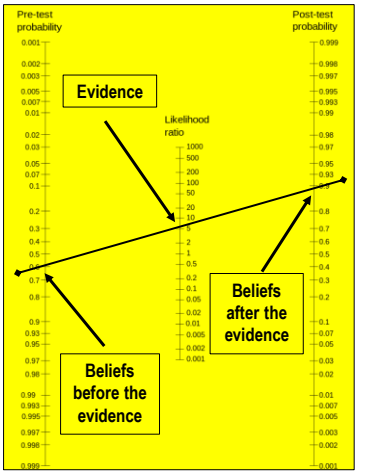
Diagnostic Tests
(Likelihood ratios, Nomogram)



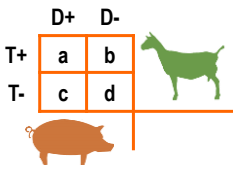
Experience, Skill
- Prior beliefs regarding disease

Diagnostic tests: Likelihood Ratios
- Weight of (external) evidence

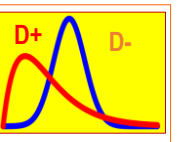
Clinical Nomogram
- **Initial Beliefs + Weight of Evidence = Final beliefs**
- Helps to concretize assumptions
- Formalizes cognitive processes
- Helps in the recognition of illogical thought processes
- Easy.
- No need for knowledge of math behind it (Bayes theorem)
- Easy communication with client



27



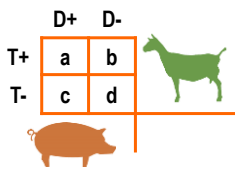
Diagnostic Test: Applications
(True and Apparent Prevalence)



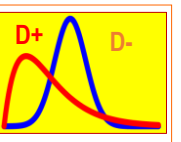
True and Apparent (or Test) Prevalence

- Test imperfect ⇒ Disease misclassification
- Some D+ classified as D- (FN)
- Some D- classified as D+ (FP)

28



Diagnostic Test Applications
(True and Apparent Prevalence)



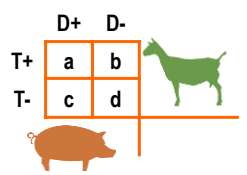
True and Apparent (or Test) Prevalence.

$$\text{Apparent Prevalence (AP)} = \frac{\text{Number T+ @ time "t"}}{\text{Total population @ risk @ time "t"}}$$

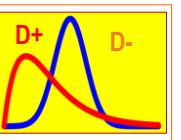
$$AP = \frac{\text{Number of T+ @ time "t"}}{\text{Total No. Tested @ time "t"}}$$

$$AP = \frac{a + b}{a + b + c + d}$$

29



Diagnostic Test Applications
(True and Apparent Prevalence)

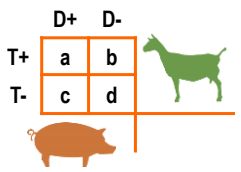


True Prevalence (TP):

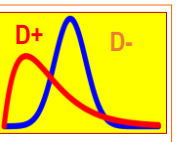
- Latent (Unobserved) Variable
- TP ≠ AP (except for Gold Standard)
- Can be > AP or < AP

$$TP = \frac{AP + Sp - 1}{Se + Sp - 1}$$

30



Diagnostic Test Applications
(True and Apparent Prevalence)



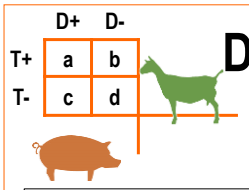
True and Apparent (or Test) Prevalence.

$$TP = \frac{\text{Number having outcome @ time "t"}}{\text{Total population @ risk @ time "t"}}$$

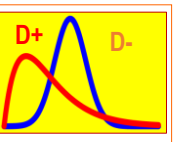
$$TP = \frac{\text{No. of D+ @ time "t"}}{\text{Total No. Tested @ time "t"}}$$

$$TP = \frac{a + c}{a + b + c + d}$$

31



Diagnostic Test Applications
(True and Apparent Prevalence)

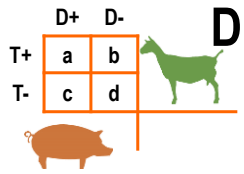


Estimates of true prevalence of F. hepatica infection corresponding to combinations of sensitivity (0.61-0.75), specificity (0.85-0.91) and AP values of 0.21 and 0.21

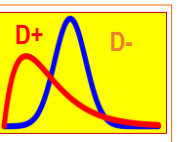
	Specificity									
	Apparent Prevalence = 0.21					Apparent Prevalence = 0.26				
Sensitivity	**0.85**	**0.87**	**0.88**	**0.89**	**0.91**	**0.85**	**0.87**	**0.88**	**0.89**	**0.91**
0.61	0.13	0.17	0.18	0.20	0.23	0.24	0.27	0.29	0.31	0.33
0.63	0.13	0.16	0.18	0.19	0.22	0.23	0.26	0.27	0.30	0.31
0.65	0.12	0.15	0.17	0.19	0.21	0.22	0.25	0.26	0.28	0.30
0.67	0.12	0.15	0.16	0.18	0.21	0.21	0.24	0.25	0.26	0.29
0.68	0.11	0.15	0.16	0.18	0.20	0.20	0.23	0.25	0.26	0.28
0.69	0.11	0.14	0.16	0.17	0.20	0.20	0.23	0.25	0.26	0.28

In shaded areas TP > AP (Carrol et. al (2017) PVM. 144: 57-65)

32



Diagnostic Test Applications
(Screening and Confirmatory Tests)

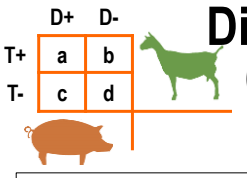


Screening Tests:

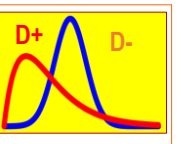
- Detect asymptomatic individuals
- High false negative cost: (Covid-19, bTB, Breast Cancer, HIV-Aids, Strangles, MAP)
- Se high (as possible)
- Rule out test negatives
- FN~0
- Cheaper (hopefully!)

	D+	D-
T+	TP	FP
T-	**FN~0**	FN

33



Diagnostic Test Applications
(Screening and Confirmatory Tests)



Diagnostic Tests:

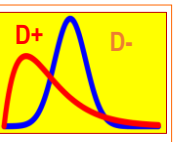
- Final diagnosis (previous T+'s)
- High FP cost: Avoid unnecessary culling, isolation, costly or (painful) treatment (Cancer, HIV-AIDs)
- Sp high (as possible)
- Rule in Test positives
- FP~0

	D+	D-
T+	TP	**FP~0**
T-	FN	TN

34

	D+	D-
T+	a	b
T-	c	d

Diagnostic Test Applications
(Screening and Confirmatory Tests)



Example:

200 Individuals, TP = 40/200 = 20%

- Test A: Se = 98%, Sp = 90%, €15
- Test B: Se = 85%, Sp = 99%, €40

Phase I (Screening)

- Only 55 test positive ⇒ Phase II

Phase II (Confirmation)

- 33/39 D+ test positive ⇒ isolated, treated, culled

- Assuming cure, TP ⇒ 7/200 = 3.5%
- Assuming culling, TP ⇒ 7/167 = 4.2%

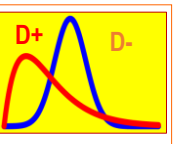
I	D+	D-	
T+	39	16	55
T-	1	144	145
	40	160	**200**

II	D+	D-	
T+	33	0	33
T-	6	16	22
	39	16	55

35

	D+	D-
T+	a	b
T-	c	d

Diagnostic Test Applications
(Series and Parallel)



Series

	Test I	Test II
Positive	+	+
Negative	+	-
	-	+
	-	-

Parallel

	Test I	Test II
Positive	+	+
	+	-
	-	+
Negative	-	-

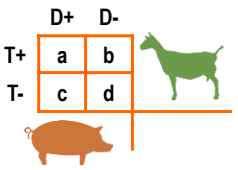
Series

- Joint Se lower
- Joint Sp higher
- Independence ↑ ⇒ Net difference ↑

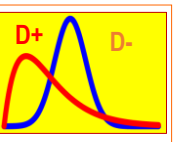
Parallel

- Joint Se higher
- Joint Sp lower
- Independence ↑ ⇒ Net difference ↑

36



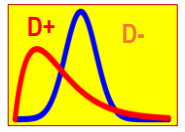
Diagnostic Tests
(Applications: Parallel Testing)



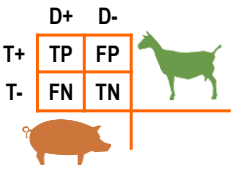
		Strangles Status S. equi+	S. equi-	
Antigen A	+	66	1	Se = 74.2%,
	-	23	138	Sp = 99.3%
	Total	89	139	
Antigen C	+	53	0	Se = 59.6%,
	-	36	139	Sp = 100.0%
Total		89	139	
Antigen A & C	+	83	1	Se = 93.3%,
	-	6	138	Sp = 99.3%
Total		89	139	

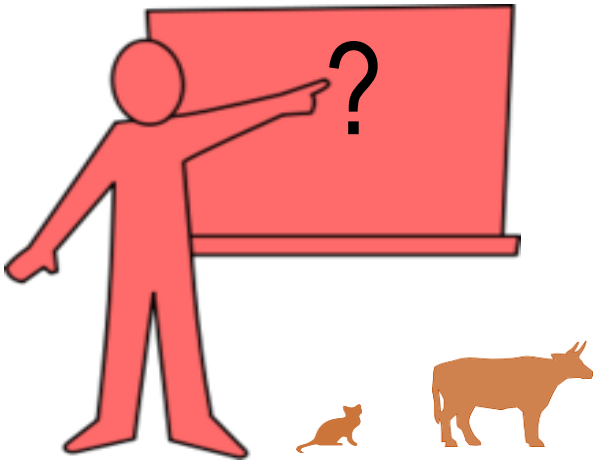
Taken from Robinson et. al (2013), TVJ 197:188-191

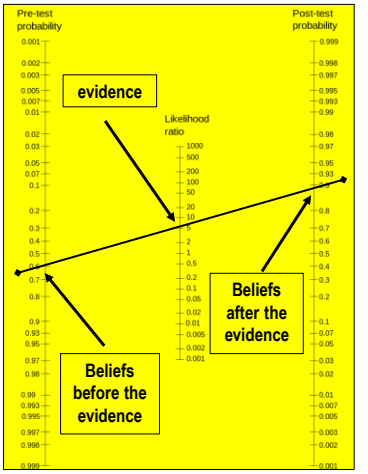
37



Diagnostic Tests







38