Chapitre 4

Évaluation des méthodes de diagnostic

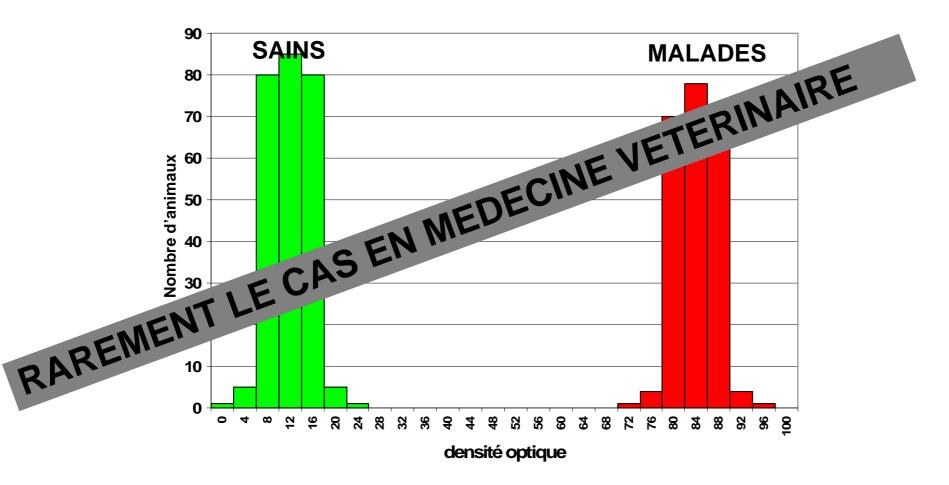
Attentes de l'utilisateur (vétérinaire) concernant les résultats d'analyse

- Estimation la plus proche possible de la réalité
- Cette estimation dépend de nombreux facteurs :
 - Qualité des prélèvements (type et méthode)
 - Qualité des commémoratifs (anamnèse)
 - Qualité de la technique d'analyse utilisée
 - Qualité de l'interprétation des résultats

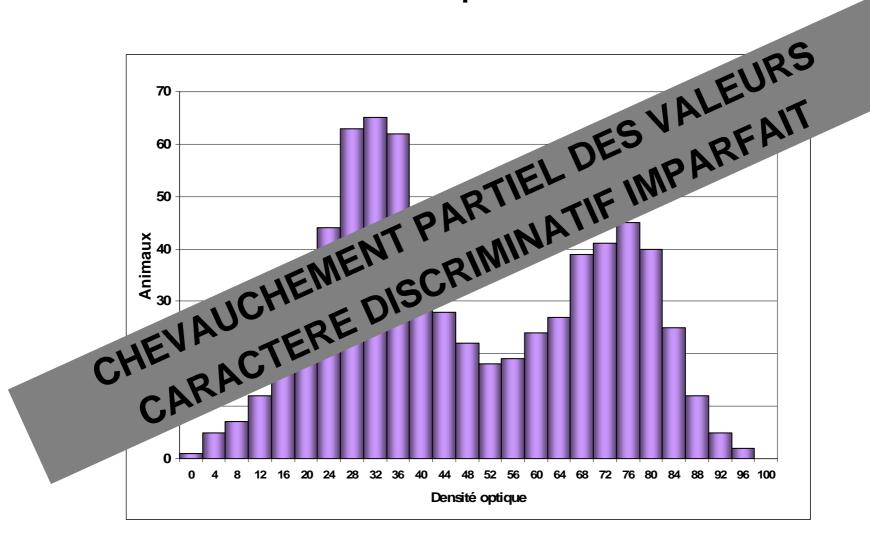
Interprétation d'un test de diagnostic en médecine vétérinaire

- Les données brutes peuvent être <u>continues</u> (densité optique d'un test ELISA)
- On se limite souvent à cataloguer le résultat sous la forme de données discontinues (jugements catégoriels)
 - positif négatif (exemple : test sérologique)
 - positif douteux négatif (exemple : intradermotuberculination)
 - diminué normal augmenté (exemple : signes cliniques)
- En réalité il n'est pas toujours facile de départager le « normal » de l'« anormal »
 - quelle est la limite inférieure de la normale ?
 - quelle est la limite supérieure de la normale ?
 - quelle est la limite de détection ?

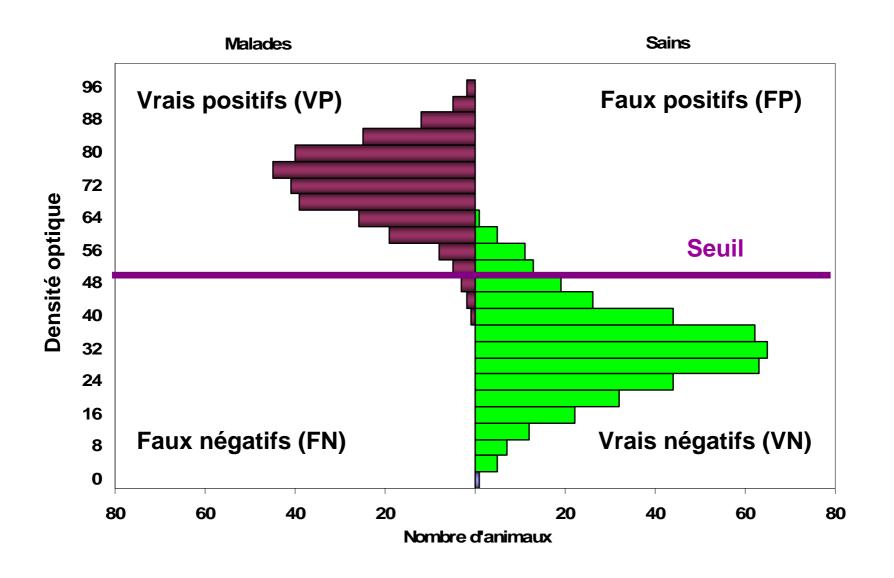
Limite de discrimination (cut-off) Test idéal = pouvoir discriminant optimal



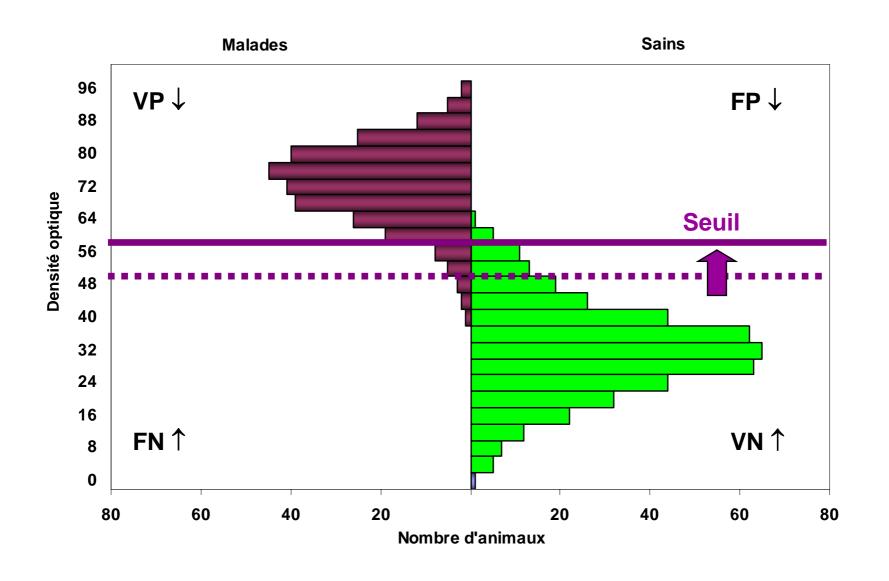
Limite de discrimination (cut-off) Situation usuelle : que constatez-vous?



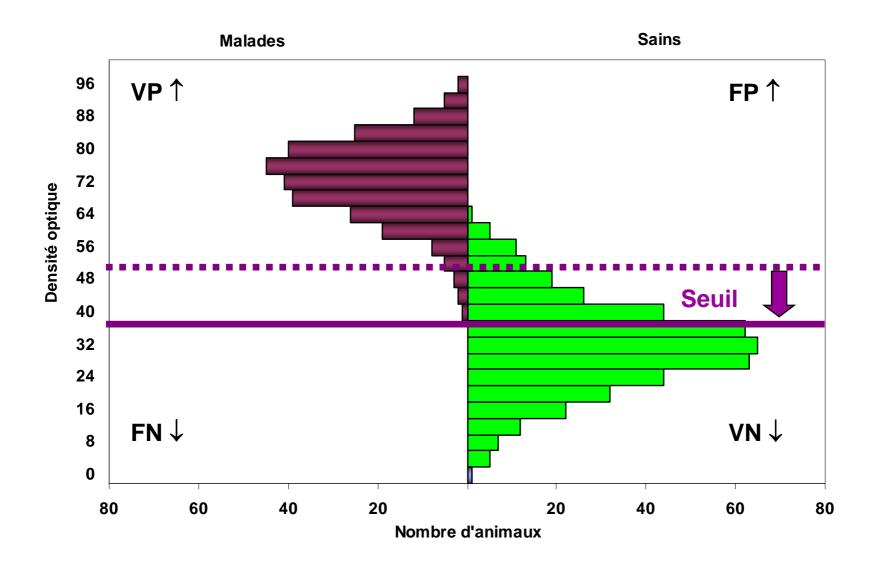
Que se passe-t-il quand je fixe un seuil?



et si je déplace le seuil vers le haut ?



et si je déplace le seuil vers le bas ?



Quelles sont les réponses exactes et erronées ?

		Statut de l'animal		
		Infecté	Sain	
Réponse au test	Positive	Vrai positif	Faux positif	
	Négative	Faux négatif	Vrai négatif	

Sensibilité d'un test

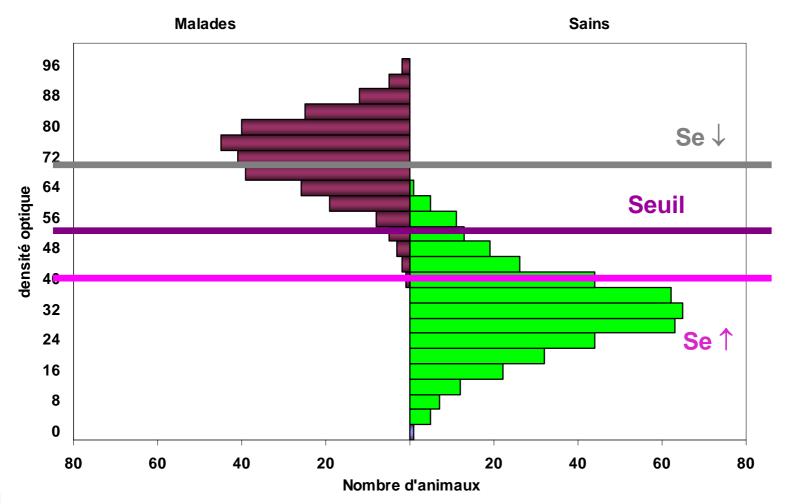
Sensibilité = aptitude d'un test à fournir une réponse positive chez un individu malade ou infecté

		Statut infe		
		Infectés	Indemnes	Totaux
Réponse du test	positive	VP		
	négative	FN		
	Totaux	VP + FN		

Sensibilité =
$$\frac{VP}{VP + FN}$$

= probabilité conditionnelle d'obtenir
 un résultat positif par un test de diagnostic
 chez un animal réellement infecté = P(T+/M+)

Variation de la sensibilité en fonction du cut-off?



•

La sensibilité d'un test

- ☐ nécessite un test de référence (statut réel infectieux)
- ☐ nécessite un échantillon représentatif issu de la population cible des individus "infectés"
- □ la sensibilité ne varie pas en fonction de la fréquence de la maladie (valeur dite "intrinsèque")
- ☐ la sensibilité peut varier au cours du temps (cinétique de la réponse immunitaire spécifique)

Spécificité d'un test

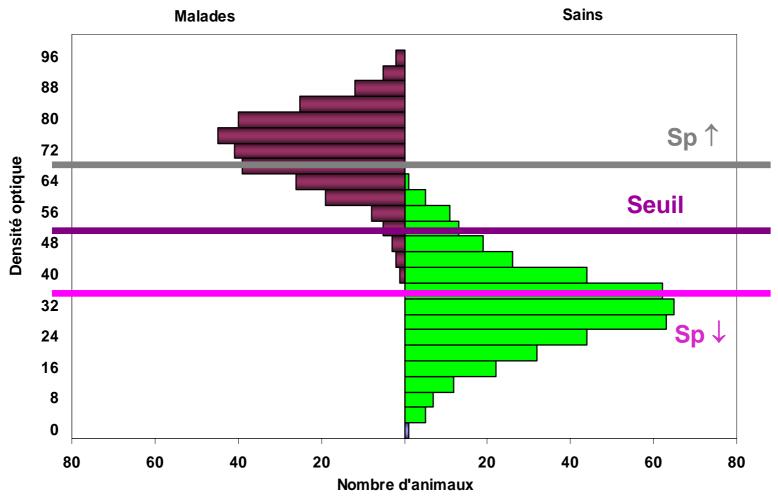
Spécificité = aptitude d'un test à fournir une réponse négative chez un individu sain

		Situation		
		Infectés	Indemnes	Totaux
Réponse du test	positive		FP	
	négative		VN	
	Totaux		VN + FP	

Spécificité =
$$\frac{VN}{VN + FP}$$

= probabilité conditionnelle d'obtenirun résultat négatif par un test de diagnosticchez un animal sain = P(T-/M-)

Variation de la spécificité en fonction du cut-off?



14

La spécificité d'un test

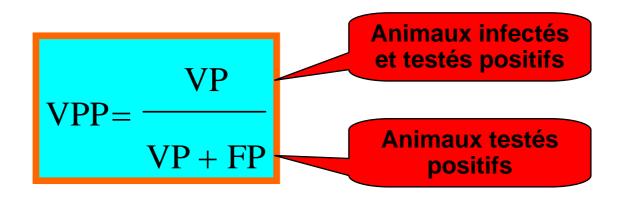
- ☐ nécessite un test de référence
- ☐ nécessite un échantillon représentatif issu de la population des individus "indemnes"
- ☐ la spécificité ne varie pas en fonction de la fréquence de la maladie (valeur dite "intrinsèque")
- ☐ la spécificité peut varier dans l'espace et dans le temps (microbisme ambiant)

Relativité des valeurs de Se et Sp

- Les valeurs de sensibilité et spécificité ne sont utilisables que dans une population ayant strictement les mêmes caractéristiques générales (espèce, race, âge) que la population de référence ayant permis de les définir.
- Il faut considérer la sensibilité et la spécificité avec assez de clairvoyance et de prudence pour éviter des conclusions excessives non justifiées

Valeur prédictive d'un résultat positif (VPP)

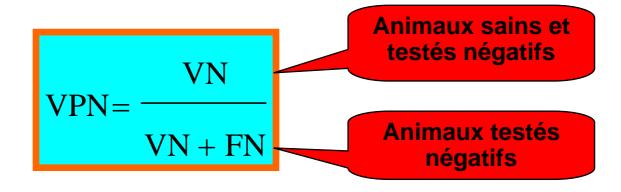
 Quelle est la chance que l'animal soit malade quand le résultat du test est positif ? = P(M+/T+)



• Parmi les animaux qui ont fourni une réponse positive au test de dépistage, on peut distinguer 2 catégories : les vrais positifs (VP) et les faux positifs (FP)

Valeur prédictive d'un résultat négatif (VPN)

 Quelle est la chance que l'animal soit sain quand le résultat du test est négatif ? = P(M-/T-)



 Parmi les animaux qui ont fourni une réponse négative au test de dépistage, on peut distinguer 2 catégories : les vrais négatifs (VN) et les faux négatifs (FN)

Comment évoluent la VPP et la VPN en fonction de la prévalence réelle ?

TABLEAU II.9

Calcul de VPP et VPN dans deux zones A et B où 1 000 animaux sont soumis à un même test de dépistage de sensibilité égale à 95 p. cent et de spécificité égale à 90 p. cent (taux d'infection des animaux : zone A : 20 p. cent ; zone B : 2 p. cent).

	Situati	Totaux	
	Infectés	Indemnes	
+	190	80	270
<u> </u>	10	720	730
Totaux	200	800	1 000
	+ - Totaux	+ 190 - 10	+ 190 80 - 10 720

Zone B		Situation réelle			Totaux		
Prévalence 2 p. c	ent		In	fectés	ī	Indemnes	n,
Réponse du test	3	+		19		98	117
de dépistage		14		1		882	883
		Totaux		<i>20</i>		980	1 000

Zone A:

VPP=190/270=70 %

Zone B:

VPP=19/117=16 %

VPP ↓ quand Pr ↓

Zone A:

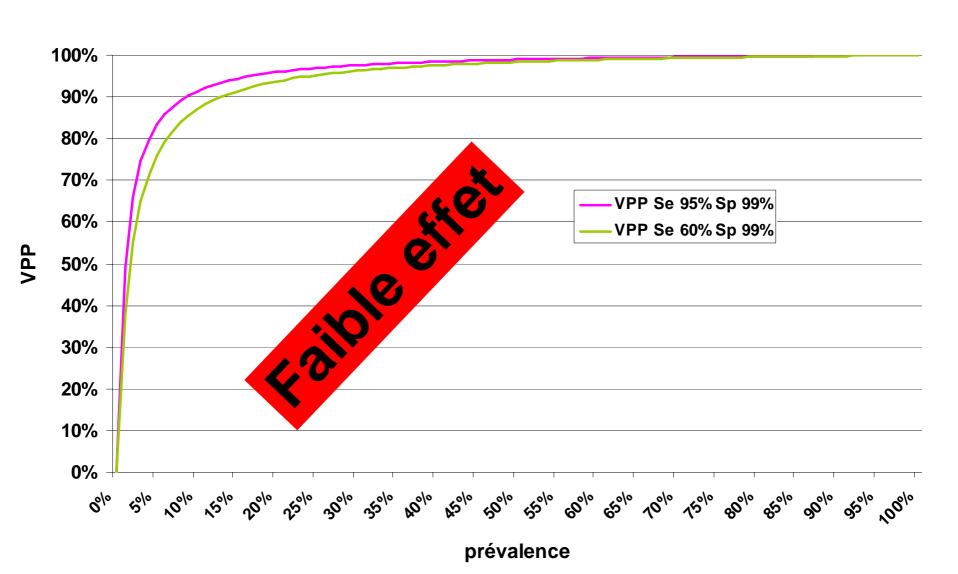
VPN=720/730=98,6 %

Zone B:

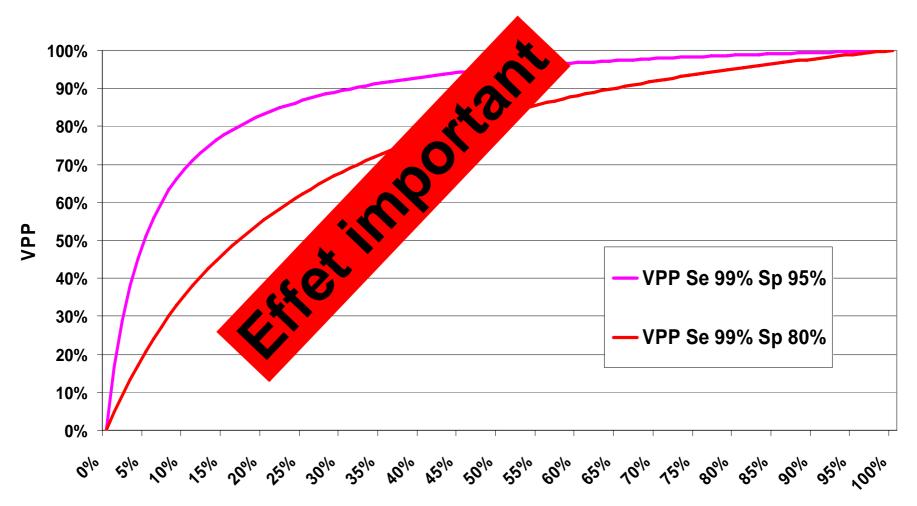
VPN=882/883=99,8 %

VPN ¬ quand Pr ↓

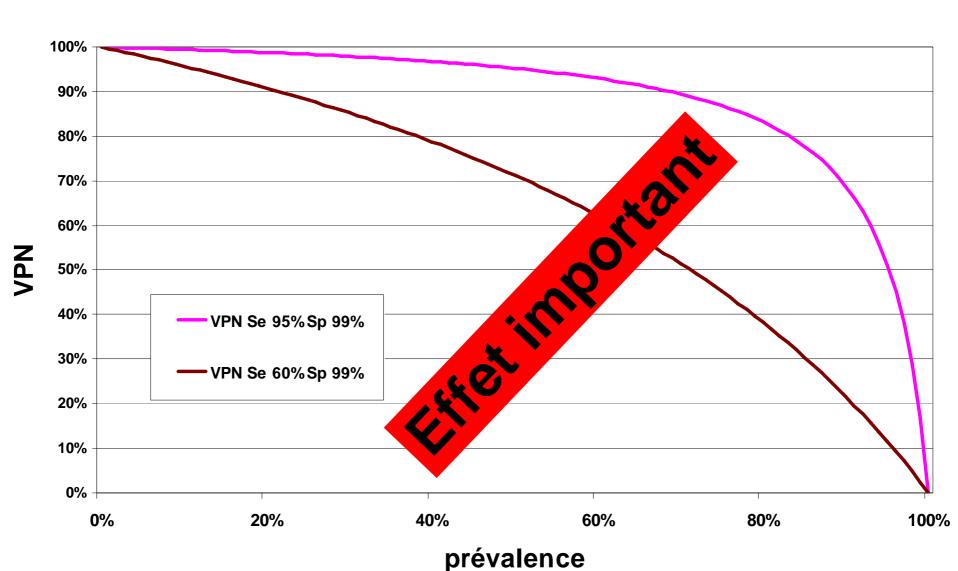
Effet de la sensibilité d'un test sur la VPP ?



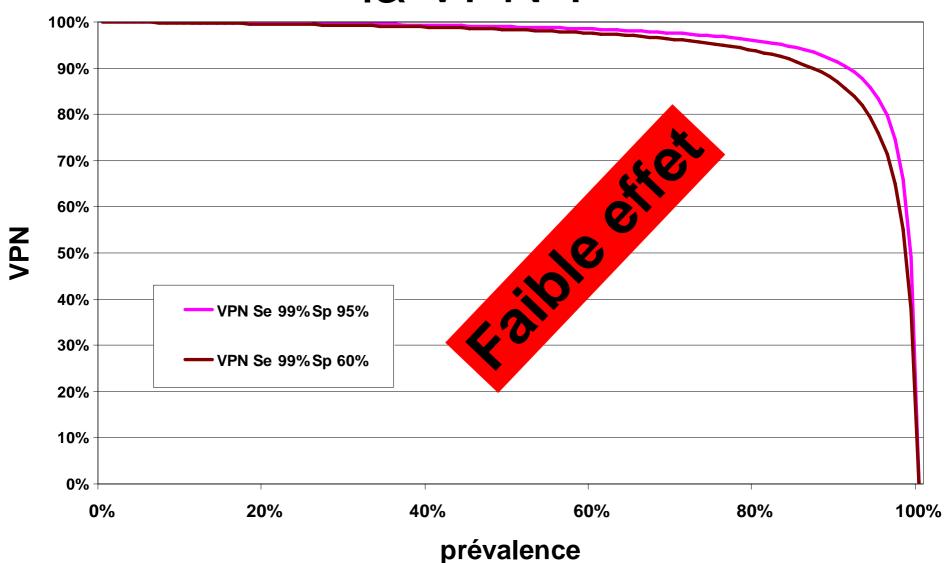
Effet de la spécificité d'un test sur la VPP ?



Effet de la sensibilité d'un test sur la VPN ?



Effet de la spécificité d'un test sur la VPN ?



Résumé concernant VPP et VPN

Confiance attribuée à un résultat positif au test

La VPP

- □ varie de 0 à 1
- ☐ dépend de Sp et de PR
- ☐ plus la maladie est rare, plus la VPP est faible

Confiance attribuée à un résultat négatif au test

La VPN

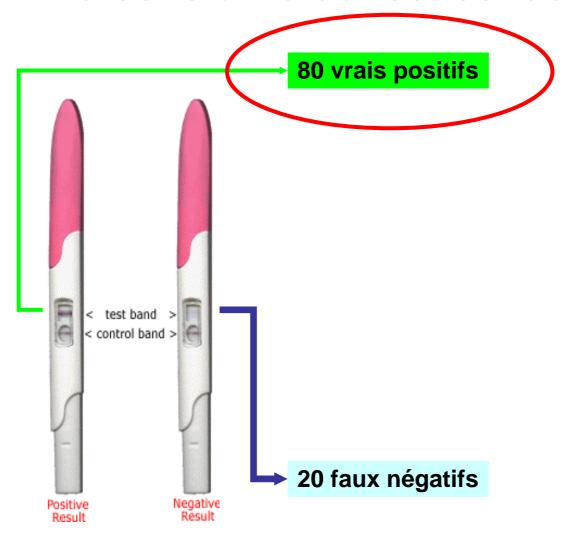
- □ varie de 0 à 1
- ☐ dépend de Se et de PR
- ☐ plus la maladie est rare, plus la VPN est élevée

Exemple: test de grossesse

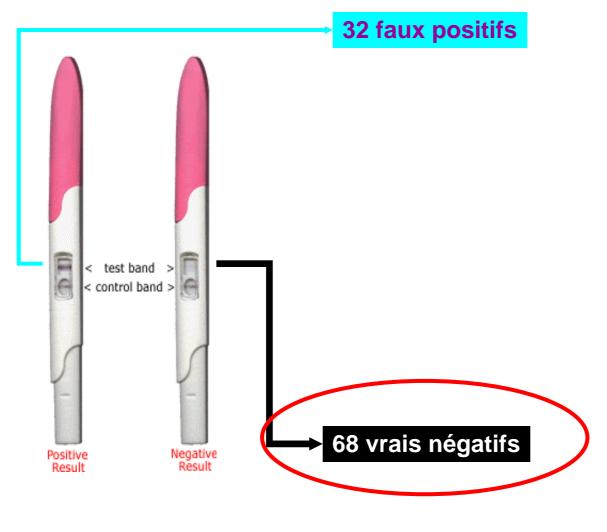


16/01/2006

Soit 100 femmes qui sont enceintes et la sensibilité du test est de 80 % ?

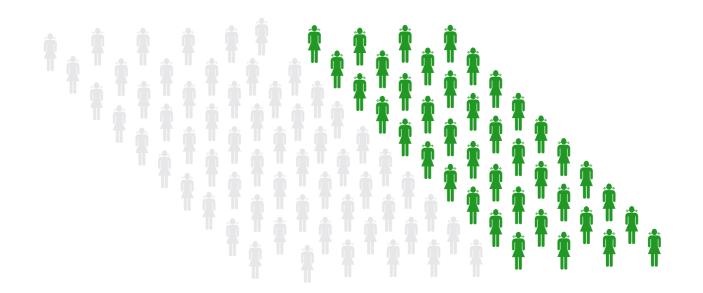


Soit 100 femmes qui ne sont pas enceintes et la spécificité du test est de 68 % ?



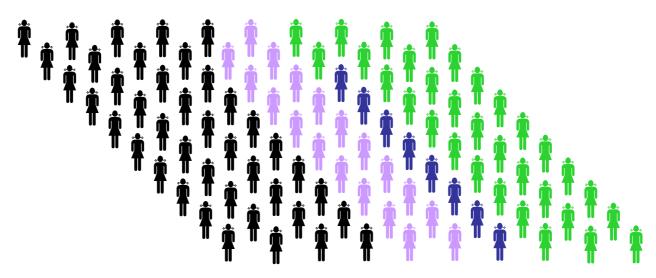
Prévalence réelle (inconnue) → supposons = 40 % :

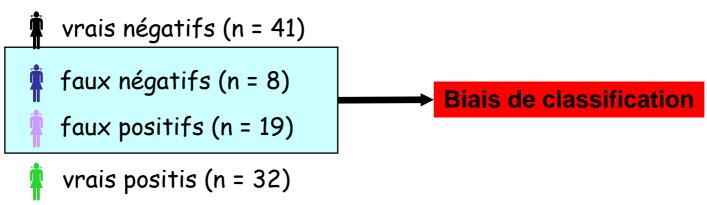
- 40 femmes qui sont enceintes
- 60 femmes qui ne sont pas enceintes



- femmes enceintes (n = 40)
 - femmes pas enceintes (n = 60)

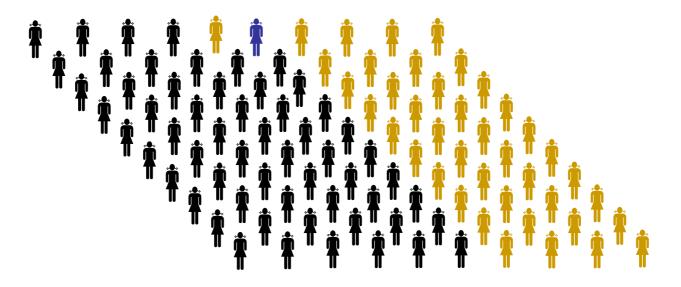
En réalité : des individus sont mal classés!





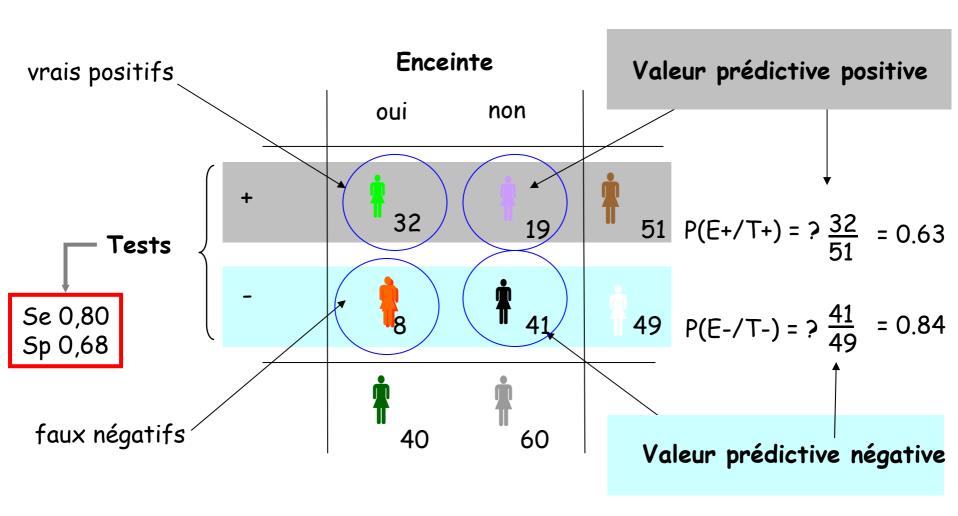
16/01/2006

Si on tient compte seulement des résultats du test (prévalence apparente)

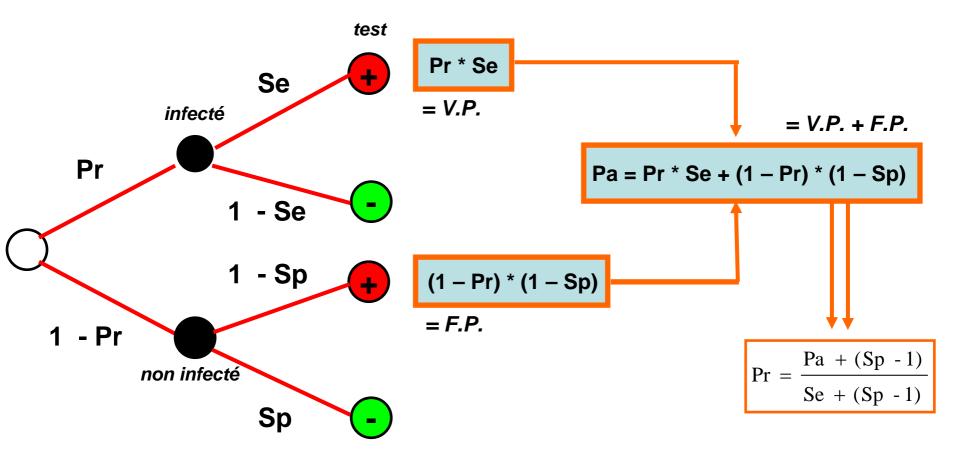


- 🛊 tests négatifs (n = 49)
- tests positifs (n = 51)

Soit une table de contingence 2 X 2



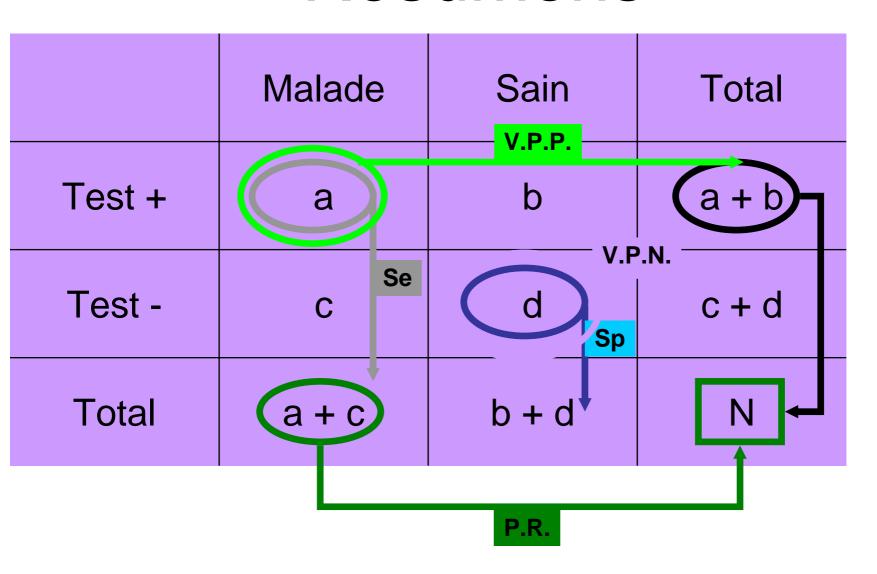
Relation entre la prévalence réelle et la prévalence apparente



Relations entre Se, Sp et P

	Malades	Sains	Total	
Test +	Se.P	(1-Sp) ₋ (1-P)	Se.P + (1-Sp).(1-P)	
10011		(· OP) · (· ·)		
Test -	(1-Se).P	Sp.(1-P)	(1-Se).P + Sp.(1-P)	
Total	Р	(1-P)	1	

Résumons



P.A.

16/01/2006

Claude Saegerman

Choix du cut-off d'un test

- Le cut-off = seuil de positivité
- Choix dans un test
- Optimaliser Se et Sp
- WinEpiscope apporte une aide pour cela :
 - indice de Youden
 - courbes ROC

Indice de Youden (J)

- Mesure de la qualité du test
- Basé sur la Se et la Sp
- Indice de Youden : J = Se + Sp 1
 - J maximal = 1
 - J varie de -1 à +1
 - lorsque la sensibilité et la spécificité d'un test sont nulles, alors la valeur de cet indice est -1
 - lorsque le test est parfait (Se et Sp = 1), cet indice est égal à 1

Exercice WinEpiscope (menu « tests », rubrique « evaluation »)

- Soit respectivement 29 porcs infectés et 171 porcs indemnes de Yersinia enterocolitica 0:9
- 23 porcs infectés sont positifs au test ELISA YOP
- 168 porcs indemnes sont testés négatifs au même test
- Quelle est la valeur de sensibilité du test ?
- Quelle est la valeur de spécificité du test ?
- Quelle est la valeur de l'indice de Youden ?

Courbe ROC (Receiver Operating Characteristic)

- Graphique
- Abscisse :
 - fraction des faux positifs (b/b+d)
 - (1-Sp)
- Ordonnée :
 - fraction des vrais positifs (a/a+c)
 - Se
- Aide graphique pour optimaliser le choix du cut-off en fonction de Se et de Sp

Exercice WinEpiscope (dans « test », rubrique « cut-off »)

Titre	Contaminé(s)	Non contaminé(s)
0.0	0	41
0.5	0	25
1.0	1	17
1.5	3	7
2.0	7	4
2.5	10	2
3.0	26	2
3.5	34	1
4.0	16	1
4.5	3	0

Exercice (suite)

- Comment importer des données (commandes « input of data » et « paste data ») ?
- Que se passe-t-il lorsque je diminue le cut-off (commande « result ») ?
- Que se passe-t-il lorsque j'augmente le cut-off?
- Comment choisir la valeur de cut-off optimale ?
- A quoi doit tendre un bon test de diagnostic (commande « area »)?

Sensibilité troupeau (« HSe »)

Définition :

- probabilité d'obtenir au moins un résultat positif dans un cheptel infecté
- HSe = 1 (1 fxSe)ⁿ
 - Hse = sensibilité troupeau (herd sensitivity)
 - n = nombre d'animaux infectés
 - Se = Se individuelle
 - f = fraction de sondage (n/N)
 - (1 fxSe)ⁿ : risque de ne pas identifier l'infection (FN)
- Validité: test individuel appliqué à tous les animaux du cheptel

Troupeau

- avec un seul animal infecté dans un troupeau : HSe = Se individuelle
- avec plusieurs animaux infectés dans un troupeau : HSe > Se ind.
- NB: cette formule suppose l'absence de faux positifs (FP): (1-Spⁿ)

Sensibilité d'un test de dépistage pour la détection de l'infection d'un troupeau, par rapport à la sensibilité individuelle du test

Le tableau fournit la sensibilité à l'échelle du troupeau, en fonction de la proportion d'animaux soumis au test de dépistage (100 p. cent et du nombre d'animaux infectés présents dans le troupeau (de 1 à 7). Les chiffres ont été calculés en considérant que la spécificité du test est de 100 p. cent.

Exemple: Si l'on soumet tous les animaux d'un troupeau à un test de dépistage de sensibilité individuelle de 90 p. cent, la probabilité de détection de l'infection du troupeau sera de 90 p. cent si un seul animal est infecté dans le troupeau, 99 p. cent si deux animaux sont infectés, 99,9 p. cent si trois animaux le sont, etc.

Proportion d'a	Nombre d'animaux infectés							
prélevés : 100 p. cent		1	2	3	4	5	6	7
:	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
	99,50	99,50	100,00	100,00	100,00	100,00	100,00	100,00
	99,00	99.00	99,99	100,00	100,00	100,00	100,00	100,00
Sensibilité	90,00	90,00	99,00	99,90	99,99	100,00	100,00	100,00
individuelle	80,00	80,00	96,00	99,20	99,84	99,97	99,99	100,00
(p. cent)	70,00	70,00	91,00	97,30	99,19	99,76	99,93	99,98
	60,00	60,00	84,00	93,60	97,44	98,98	99,59	99,84
	50,00	50,00	75,00	87,50	93,75	96,88	98,44	99,22

Toma et al., 2001

Spécificité troupeau (« HSp »)

définition :

- probabilité de n'obtenir que des résultats négatifs dans un cheptel indemne
- $HSp = 1 (1 Sp^n)$, soit $HSp = Sp^n$
- (1- Spⁿ): risque d'obtenir au moins une réponse positive et de considérer le troupeau comme infecté (FP)
- validité: test individuel appliqué à tous les animaux du cheptel (n) ou à n animaux de l'échantillon

troupeau

- avec un seul animal testé : HSp = Sp individuelle
- avec plusieurs animaux testés : HSp < Sp individuelle

Spécificité d'un test de dépistage à l'échelle du troupeau par rapport à la spécificité individuelle

Le tableau donne la spécificité à l'échelle du troupeau en fonction de sa taille (nombre d'animaux soumis au test) et de la spécificité individuelle.

Ce tableau n'est valable que si les réactions non spécifiques ne sont pas groupées particulièrement dans certains troupeaux (à cause de communautés antigéniques).

Exemple: Pour un test de spécificité individuelle de 90 p. cent, la spécificité à l'échelle d'un troupeau de 10 animaux soumis à ce test est de 34,9 p. cent; seuls 34,9 p. cent de ces troupeaux ne fourniront que des réponses négatives; 65,1 p. cent de ces troupeaux fourniront donc au moins une réponse positive.

Taille	du	Spécificité du test individuel								
troupeau	eau _	99,1	99	98	97	95	93	91	90	80
1	0 9	91,36	90,44	81,71	73,74	59,87	48,40	38,94	34,87	10,74
2	0 8	6,46	81,79	66,76	54,38	35,85	23,42	15,16	12,16	1,15
3	0 7	76,24	73,97	54,55	40,10	21,46	11,34	5,91	4,24	0,12
4	0 6	59,65	66,90	44,57	29,57	12,85	5,49	2,30	1,48	0,01
5	0 6	53,63	60,50	36,42	21,81	7,69	2,66	0,90	0,52	0,00
80	0 4	18,52	44,75	19,86	8,74	1,65	0,30	0,00	0,02	0,00
10	0 4	10,49	36,60	13,26	4,76	0,59	0,07	0,00	0,00	0,00
150	0 2	25,77	22,15	4,83	1,04	0,05	0,00	0,00	0,00	0,00
ma et al								-,00	0,00	0,00

Les tests multiples

- Tests
 - en série
 - en parallèle
- Modification
 - de la sensibilité
 - de la spécificité

Tests en parallèle

- 2 ou plusieurs tests sont appliqués au même moment
- effet : prouver que l'animal est indemne
 - Se élevée
 - VPN élevée
- un animal négatif pour tous les tests est déclaré indemne

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

 $Sp_{par} = Sp_1 \times Sp_2$

Tests en série

- 2 ou plusieurs tests sont appliqués consécutivement
- effet : prouver que l'animal est malade
 - Sp élevée
 - VPP élevée
- un animal est déclaré malade s'il est positif pour tous les tests

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

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

Interprétation en série ou en parallèle : partiquement...

	Caté	gories	Total			
Test 1	+	+	-	-		
Test 2	+	-	+	-		
Malades	134	4	29	9	176	∑ 750
Sains	0	28	12	534	574	
Interprétation en série	+	-	-	-		
Interprétation en parallèle	+	+	+	-		↓

P.R. = 23,47 %

Se test
$$1 = 138/176 = 0,784$$

Se test2 =
$$(134+29)/176 = 0,926$$

Se interp.série =
$$(134)/176 = 0,761$$

Se interp.paral. =
$$(134+4+29)/176 = 0,949$$

$$Sp test1 = (12+534)/574 = 0,951$$

$$Sp test2 = (28+534)/574 = 0,979$$

Sp paral. =
$$534/574 = 0.930$$

Exercice WinEpiscope (menu « multiple tests »)

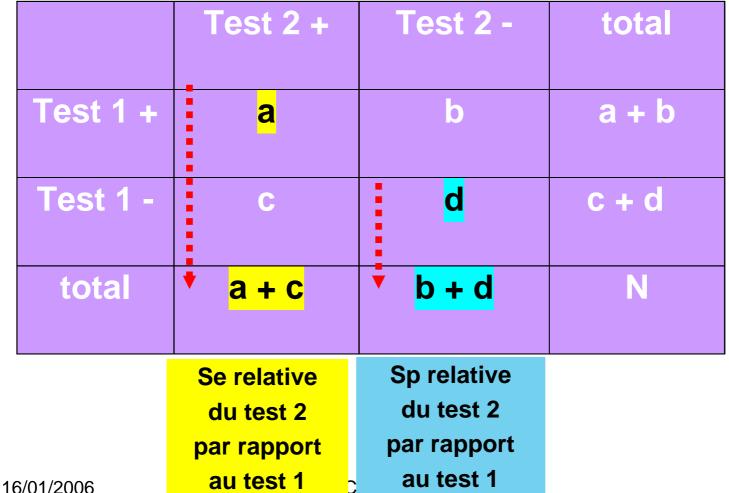
- Au départ de l'exemple précédent
- Nombre d'animaux = 750
- Prévalence réelle = 23,47 %
- Se1 = 78,4 % et Se2 = 92,6 %
- Sp1 = 95,1 % et Sp2 = 97,9 %
- Que vaut Se et Sp si l'interprétation se fait en série et en parallèle ?
- Comparer ces valeures estimées aux valeurs réelles de l'exemple précédent ?
- A quoi ces différences seraient-elles dues ?
- Vous référant à vos connaissances d'immunologie, donnez un exemple où deux tests ont le plus de chance d'être indépendants (pensez aux différents effecteurs immunologiques)?

Comparer deux tests de diagnostic entre eux

Se et Sp relatives

- coefficient Kappa
 - comparer deux tests
 - comparer le diagnostic de deux cliniciens

Sensibilité et spécificité relatives



La concordance entre jugements catégoriels appariés

dans WinEpidscope:
voir « tests », rubrique « agreement »
(coéfficient Kappa)

Le coefficient kappa

- Principe : mesurer l'accord en excluant le hasard
- Meilleur que le calcul de la corrélation
- Utilisation
 - comparer deux tests dont l'un est nouveau
 - comparer les capacités de diagnostic de deux cliniciens
- attention
 - les deux tests peuvent se tromper

Objectifs

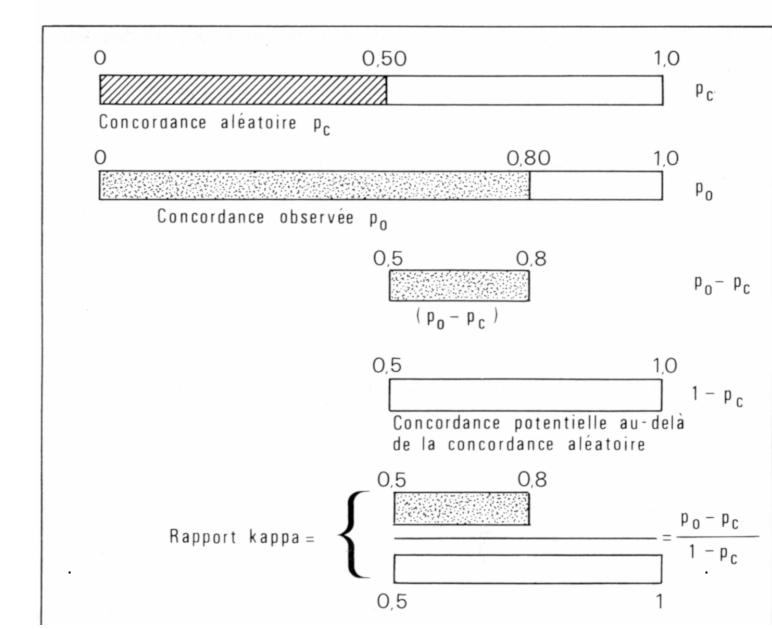
- Définir ce qu'est la concordance entre jugements catégoriels appariés
- Décrire la démarche pour calculer le coefficient kappa
- Appliquer la démarche à un exemple concret
- Mise en situation et exercices

Définir = analyse structurelle

- Concordance = conformité, similitude, de 2 ou de plusieurs informations se rapportant au même sujet
- Jugements catégoriels = jugements qualitatifs
 - **+**, +/-, -
 - présence, absence
 - conforme, non conforme
- Apparier = assortir les objets par paires
 - ☐ Ex. 1 : 2 tests de dépistage d'une même maladie
 - ☐ Ex. 2: intradermotuberculination lue

CONCORDANCE ENTRE DEUX JUGEMENTS CATÉGORIELS





Le coéfficient Kappa

	Test 2	Test 2	total
	+	-	
Test 1	a	b	a + b
+			
Test 1	C	d	c + d
-			
total	a + c	b + d	N

```
Kappa =

[(a+d)/N] - (Ea + Ed)/N

1 - (Ea + Ed)/N

soit

a + d - Ea - Ed

N - Ea - Ed
```

- □ (a+d)/N : accord incluant le hasard et le jugement correct
- □ Ea = [(a+b).(a+c)]/N
- $\Box \quad \mathsf{Ed} = [(\mathsf{c+d}).(\mathsf{b+d})]/\mathsf{N}$
- ☐ (Ea + Ed)/N : accord dû au hasard
- □ 1 (Ea + Ed)/N : accord maximum en dehors du hasard
- □ [(a+d)/N] (Ea + Ed)/N : accord observé en dehors du hasard

(WinEpiscope : voir « test » puis rubrique « agreement »)

Grille d'interprétation

 La concordance est d'autant plus élevée que le coef. Kappa est proche de 1

 La concordance est estimée en fonction de la grille ci-jointe

Kappa	Concordanc e
1 à 0,81	Excellente
0,8 à 0,61	Bonne
0,6 à 0,41	Moyenne
0,4 à 0,21	Faible
0,2 à 0,01	Négligeable
< 0	mauvaise

Décrire la démarche (1)

a) Soit une table de contingence 2 x 2

	Tes	st A	Total
	+	-	
Test B +	а	b	f ₁
-	С	d	f_2
Total	n_1	n_2	N

f₁, f₂, n₁, n₂ sont les fréquences marginales

Kappa =
$$(p_o - p_c) / (1 - p_c)$$

- P_o: taux de concordance observé
 (a + d) / N
- P_c: taux de concordance aléatoire
 (a' + d') / N
- Avec : a' = $(n_1 * f_1) / N$ et d' = $(n_2 * f_2) / N$

Coefficient kappa est le rapport :

 (p_o – p_c) : <u>différence</u> entre la concordance observée et la concordance aléatoire

divisée par

(1 – p_c): la concordance <u>disponible au-</u>
 <u>delà</u> de la concordance aléatoire

Décrire la démarche (2)

a) Soit une table de contingence 3 x 3

		L	Total		
		+	+/-	-	
	+	а	b	С	f_1
Lecteur B	+/-	d	е	f	f_2
	-	g	h	i	f_3
Total		n_1	n_2	n ₃	N

f₁, f₂, f₃, n₁, n₂ n₃ sont les fréquences marginales

Kappa = $(p_o - p_c) / (1 - p_c)$

- P_o: taux de concordance observé
 (a + e + i) / N
- P_c: taux de concordance aléatoire

$$(a' + e' + i') / N$$

• Avec : a' = $(n_1 * f_1) / N$ et e' = $(n_2 * f_2) / N$ et i' = $(n_3 * f_3) / N$

Appliquer la démarche (exemple)

- Un candidat vétérinaire et son maître de stage
- Le maître de stage (référence) montre au candidat comment tuberculiner dans les règles de l'art
- 209 bovins > 6 mois sont tuberculinés (IDS)
- La lecture des IDS est réalisée par le VT (dans un sens) et le candidat (en sans inverse)
- Chaque opérateur <u>note les observations</u> recueillies sur une liste du détenteur

Interprétation IDS

Interprétation IDS	≠ pli de peau (mm)	Signes cliniques*
+	≥ 4	Présence ou absence
+/-	> 2 et < 4	Absence
-	≤ 2	Absence

^{*} œdème diffus ou étendu ; exsudation ou nécrose ; douleur ou réaction inflammatoire des lymphatiques de la région ou des ganglions

*Dans un but didactique:

les résultats + et +/- sont assimilés à des résultats +

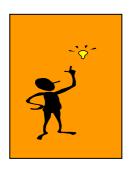
Exercice WinEpiscope (menu « tests », rubrique « agreement »)

 Effectuer les exercices suivants à l'aide du logiciel WinEpiscope.

Table de contingence (2 x 2)

		Maître	Total	
		+	-	
Candidat VT	+	12	4	16
Canadat VI	-	3	190	193
Total		15	194	209

- 1) Que doit vérifier l'étudiant sur la liste du détenteur ?
- 2) Que doit **noter** l'étudiant sur la liste du détenteur ?
- 3) Quel est le taux de concordance entre les résultats du candidat vétérinaire et ceux du maître de stage ?
- 4) Comment peut-on qualifier ce taux de concordance ?



Réponses

- 1) Vérifier que tous les bovins ont été revus (élément clef = identification fiable)
 - Noter les lectures des IDS en regard des numéros d'identification
 - 3) Po = (12 + 190) / 209 = 0,9665 Pc = (1,1483 + 179,1483) / 209 = 0,8627 Kappa = 0,76
 - 4) Bonne (taux # formation théorique et pratique)

Mise en situation (exercices)

Diagnosis of bovine brucellosis by skin test: conditions for the test and evaluation of its performance (Saegerman et al., 1999)

Combina	isons de tests	S	Contexte			
Tests classiques	i-ELISA 1C8	Skin test	Infection chronique	Infection aigue	RSFP	
(seuils légaux)	(2,5 UI/mI)	1,1 mm	922 observations	363 bovins	609 bovins	
+	+	+	46	32	0	
=	+	-	92	33	22	
=	+	+	79	11	0	
=	=	+	50	6	2	
+	+	-	36	6	3	
+	=	-	4	0	0	
+	_	+	2	0	0	
=	_	-	613	275	582	
7	Гotaux		922	363	609	



- A) Quel est le taux de concordance entre les tests classiques de dépistage de la brucellose et le test intradermique (brucellergène®) dans le cas d'infection brucellique aigue ?
- B) Comment qualifiez-vous ce taux de concordance?
- C) Le test intradermique est-il utile ? Pourquoi ?



- A) Quel est le taux de concordance entre les tests classiques de dépistage de la brucellose et le test intradermique (brucellergène®) dans le cas d'infection brucellique chronique ?
- B) Comment qualifiez-vous ce taux de concordance?
- C) Le test intradermique est-il utile ? Pourquoi ?



- A) Quel est le taux de concordance entre les tests classiques de dépistage de la brucellose et le test intradermique dans une zone géographique où l'occurrence des RSFP est élevée ?
- B) Comment qualifiez-vous ce taux de concordance?
- C) Le test intradermique est-il utile dans le contexte des RSFP? Pourquoi?



 Que pouvez-vous inférer des résultats obtenus aux questions 1 à 3 ?

 Quelle est la signification biologique de ce qui précède ?

