

## **SENSIBILITE, SPECIFICITE, VALEURS PREDICTIVES : IMPACT de la QUALITE des ANALYSES en EPIDEMIOLOGIE\***

Geneviève ANDRE-FONTAINE<sup>[1]</sup>

---

**RESUME** : Dans cet article sont tout d'abord définis les concepts de sensibilité et de spécificité d'un test, établis dans des conditions expérimentales. Ensuite, sont évoquées la valeur prédictive d'un résultat positif et la valeur prédictive d'un résultat négatif employées en épidémiologie, et leur évolution en fonction des valeurs intrinsèques du test. Enfin, l'attention est attirée sur l'importance des qualités d'une technique en fonction de l'objectif de qualification d'un individu ou d'un cheptel.

**SUMMARY** : In this article, sensitivity and specificity of an assay are first defined in experimental conditions. Then positive and negative predictive values are presented as used in epidemiology. The way they evolve following the intrinsic values of the test is shown. Then it is explained how important is the quality of the technique used following the aim of qualifying an individual or a herd.

\*  
\* \*

L'utilisateur de résultats d'analyses attend de ceux-ci une estimation la plus proche possible de la réalité. Cette estimation est dépendante de nombreux facteurs (qualité des prélèvements et des commémoratifs, qualité de la technique utilisée, qualité de l'interprétation). Parmi ces facteurs multiples, seront seules prises en compte dans ce texte, la sensibilité (sensitivity en anglais) et la spécificité, essentiellement définies dans des conditions expérimentales, puis les valeurs prédictives qui sont des qualités opérationnelles de terrain, dépendantes des premières.

### **I - CONDITIONS EXPERIMENTALES**

Lorsqu'on étudie une nouvelle méthode d'analyse, on se place dans des conditions expérimentales. Cette notion de conditions expérimentales sous-entend généralement que l'on travaille sur deux lots d'animaux, aussi homogènes que possible et ne différant que par le facteur cible de l'analyse étudiée, par exemple l'existence d'une infection, d'une maladie, etc., provoquées expérimentalement.

Ainsi, deux lots d'expérience, de caractéristiques communes et d'effectifs définis sont utilisés, par exemple le lot malade, infecté etc., appartenant à la population dite M dans la suite du texte, et le lot témoin, semblable en tous points au lot M, sauf le facteur maladie ou infection et appartenant à la population dite  $\bar{M}$  pour la suite.

\* Texte de l'exposé présenté le 7 décembre 1989.

[1] E.N.V.N. - Maladies contagieuses - B.P. 527 - 44026 Nantes Cedex 03

La technique étudiée sur ces deux lots M et  $\bar{M}$  est donc supposée réaliser une discrimination parfaite des deux populations M et  $\bar{M}$ . Ainsi, par exemple, tout résultat positif de l'analyse sera caractéristique d'un individu appartenant à M et tout résultat négatif d'un individu appartenant à  $\bar{M}$ .

Malheureusement, cet aspect schématique n'est quasiment jamais réalisé pour des raisons physiologiques, techniques ou autres et la technique employée donne des résultats négatifs même pour certains individus de la population M et inversement des résultats positifs pour certains individus de la population  $\bar{M}$ .

Ainsi, on peut construire le tableau de contingence (tableau I), faisant apparaître les deux lots et les résultats positifs et négatifs obtenus dans chacun des deux lots. C'est ainsi que l'on peut définir :

**. Au sein de la population M**

- **Les vrais positifs (VP)** répondant positivement à l'analyse, et
- **Les faux négatifs (FN)** répondant négativement à l'analyse alors que l'on sait qu'ils appartiennent pourtant à la population M (car ayant, par exemple, reçu une dose infectieuse suffisante).

**. Au sein de la population  $\bar{M}$**

- **Les vrais négatifs (VN)** ne répondant pas à l'analyse étudiée, et
- **Les faux positifs (FP)** qui donnent un résultat positif alors que l'on sait pertinemment qu'ils font partie du lot témoin  $\bar{M}$  non exposé au risque considéré.

**Tableau I :** Tableau de contingence général relatif aux résultats positifs et négatifs d'un test appliqué à une population comprenant des malades et des témoins

	M	$\bar{M}$	
+	VP	FP	VP + FP
-	FN	VN	VN + FN
	VP + FN	VN + FP	TOTAL
	↓ SENSIBILITE	↓ SPECIFICITE	

VP = VRAIS POSITIFS  
 FP = FAUX POSITIFS  
 VN = VRAIS NEGATIFS  
 FN = FAUX NEGATIFS

Ce tableau permet donc, à partir de deux lots M et  $\bar{M}$  définis dans le protocole expérimental, d'apprécier la sensibilité et la spécificité de la technique étudiée :

- **La sensibilité** est la probabilité d'obtenir un résultat positif pour un individu appartenant à la population M (composée de VP et FN). Elle peut donc être exprimée de la façon suivante :

$$\text{Sensibilité} = \frac{VP}{VP + FN}$$

- **La spécificité** est la probabilité d'obtenir un résultat négatif (VN) pour un individu appartenant à la population témoin  $\bar{M}$  (composée de VN et FP). Elle peut donc être exprimée de la façon suivante :

$$\text{Spécificité} = \frac{VN}{VN + FP}$$

Par exemple (tableau II) une étude est faite sur deux lots M et  $\bar{M}$  composés chacun de 1000 individus. Les résultats obtenus se répartissent de telle façon que l'on peut estimer, dans ces conditions expérimentales parfaitement définies, la sensibilité à 0,95 et la spécificité à 0,99.

**Tableau II** : Tableau de contingence concernant une population de 2.000 individus.

	M	$\bar{M}$	
+	VP 950	FP 10	960
-	FN 50	VN 990	1 040
	1 000	1 000	2 000

$$\text{SENSIBILITE} = \text{Prob P / M} = \frac{VP}{VP + FN} = 0,95$$

$$\text{SPECIFICITE} : \text{Prob N / } \bar{\text{M}} = \frac{VN}{VN + FP} = 0,99$$

Néanmoins, l'usage courant a consacré les notions d'erreurs par excès ou par défaut, qui concrétisent, dans l'esprit de chacun ces notions de faux positifs dans la population témoin ( $\bar{M}$ ) et de faux négatifs dans la population malade (M).

Ces erreurs sont donc représentées par la probabilité complémentaire de la spécificité et de la sensibilité et peuvent s'exprimer ainsi, selon le tableau I :

$$\text{Erreurs par défaut} = \frac{FN}{VP + FN} = 1 - \text{sensibilité}$$

$$\text{Erreurs par excès} = \frac{FP}{VN + FP} = 1 - \text{spécificité}$$

Par extension, le caractère discriminant global de la technique étudiée encore appelée "efficacité" est la proportion de résultats "vrais" (VP et VN) obtenus sur l'ensemble des résultats :

$$\text{Efficacité} = \frac{VP + VN}{VP + VN + FP + FN}$$

## II - CONDITIONS EPIDEMIOLOGIQUES

Dans les conditions précédentes, les deux lots d'animaux M et  $\overset{\circ}{M}$  étaient parfaitement définis **au départ**, dans leurs effectifs comme dans leurs caractères et plus particulièrement celui d'appartenir ou non à une population infectée.

Lors d'études épidémiologiques, les conditions sont différentes, le but de l'étude étant, à partir d'une population totale analysée, de définir les deux sous-populations M et  $\overset{\circ}{M}$ . La définition des effectifs de ces deux sous-populations permet alors de définir la prévalence de l'infection au sein de la population totale. Pour ce faire, on peut utiliser telle ou telle technique de laboratoire, qui donnera des résultats positifs et négatifs. Ainsi (tableau III), sur une population fixée de 100.000 individus, a été utilisée une technique dont les résultats se sont respectivement répartis en 31.800 positifs et 68.200 négatifs.

Tableau III : Résultats de l'application d'un test sur une population de 100.000 individus.

	M	$\overset{\circ}{M}$	
+			31 800
-			68 200
			100 000

PREVALENCE ?

M ?       $\overset{\circ}{M}$  ?

Très fréquemment, de tels résultats sont interprétés de la façon suivante : la prévalence de l'infection est le rapport de 31.800 à 100.000 soit 0,318.

Or, cette interprétation suppose que les 31.800 résultats positifs sont le fait d'animaux "**vrais positifs**" et donc appartenant exclusivement à la population M. C'est ignorer que l'analyse employée est imparfaite, qu'elle engendre des faux positifs mais aussi des faux négatifs. Ce qui veut dire que pour définir correctement la prévalence, il faut estimer correctement la population M et donc apprécier les effectifs de faux positifs et de faux négatifs.

Si l'on reprend une méthode qui, étudiée dans des conditions expérimentales, avait été caractérisée par une sensibilité de 0,99 et une spécificité de 0,97, le tableau III doit être complété en fonction de ces caractères (tableau IV) et la prévalence devrait alors être estimée par l'effectif de M rapporté à la population totale (30.000/100.000) soit 0,30 et non 0,318.

Dans cet exemple, la distorsion entre les deux estimations est limitée car le nombre de "vrais positifs" est important compte tenu de la prévalence réelle élevée et des qualités satisfaisantes de sensibilité et de spécificité. Néanmoins, dans d'autres conditions épidémiologiques et analytiques, l'erreur effectuée sur l'estimation de la prévalence peut être considérable et très lourde de conséquences.

Il devient donc indispensable, pour des estimations épidémiologiques d'apprécier la confiance que l'on peut accorder à un résultat, fût-il positif ou négatif. Ceci correspond à l'appréciation des **valeurs prédictives** des résultats obtenus par une analyse.

Tableau IV : Complément du tableau III.

	M	<sup>o</sup> M	
+	VP 29 700	FP 2 100	31 800
-	FN 300	VN 67 900	68 200
	30 000	70 000	100 000

Si sensibilité = 0,99  
spécificité = 0,97

$$\text{PREVALENCE ESTIMÉE} = \frac{30\,000}{100\,000} = 0,30$$

La **valeur prédictive d'un résultat positif (VPrP)** est donc la probabilité pour que le résultat positif obtenu soit un "vrai positif", autrement dit c'est la probabilité pour qu'un individu qui fournit une réponse positive appartienne effectivement à la population malade. Si l'on reprend le tableau I, cette qualité opérationnelle de l'analyse peut s'exprimer ainsi :

$$\text{VPrP} = \frac{\text{VP}}{\text{VP} + \text{FP}}$$

La **valeur prédictive d'un résultat négatif (VPrN)** est donc la probabilité pour que le résultat négatif soit un "vrai négatif", autrement dit c'est la probabilité pour qu'un individu fournissant une réponse négative appartienne effectivement à la population indemne. Soit d'après le tableau I :

$$\text{VPrN} = \frac{\text{VN}}{\text{VN} + \text{FN}}$$

Si l'on reprend les résultats obtenus dans le tableau IV on peut estimer les valeurs prédictives des résultats obtenus (la valeur prédictive d'un résultat positif étant le nombre de "vrais positifs" rapporté au nombre total de positifs et la valeur prédictive d'un résultat négatif étant le nombre de "vrais négatifs" rapporté au nombre total de négatifs) soit respectivement 93,40 p. cent et 99,56 p. cent.

Dans cet exemple, la prévalence est donc de 30 p. cent. Imaginons qu'à la suite d'une action sanitaire efficace, en quelques années la prévalence réelle tombe à 0,1 p. cent. Une nouvelle enquête (ou contrôle) mise en oeuvre toujours sur 100.000 individus à l'aide de la même technique (sensibilité 0,99 et spécificité 0,97), donnerait alors les résultats présentés dans le tableau V. On constate alors que la valeur prédictive d'un résultat positif passe de 93,40 p. cent à 3,2 p. cent quand la prévalence évolue de 30 p. cent à 0,1 p. cent. Ce qui signifie que dans la première situation épidémiologique la probabilité pour qu'un positif soit un "vrai positif" était de 93,40 p. cent, ce qui est correct, alors que dans la deuxième situation, cette probabilité ne se réalise que dans 3,20 p. cent des cas, ce qui est inadmissible.

Tableau V : Comparaison des résultats obtenus avec le même test dans deux populations de prévalence de maladie 30 p. cent et 0,1 p. cent.

TEST EMPLOYE : SENSIBILITE = 0,99 SPECIFICITE = 0,97

	M	M̄	
+	29 700	2 100	31 800
-	300	67 900	68 200
	30 000	70 000	100 000

PREVALENCE = 30 %

$$VPrP = \frac{VP}{VP + FP} = 93,40 \%$$

$$VPrN = \frac{VN}{VN + FN} = 99,56 \%$$

	M	M̄	
+	99	2 997	3 096
-	1	96 903	96 904
	100	99 900	100 000

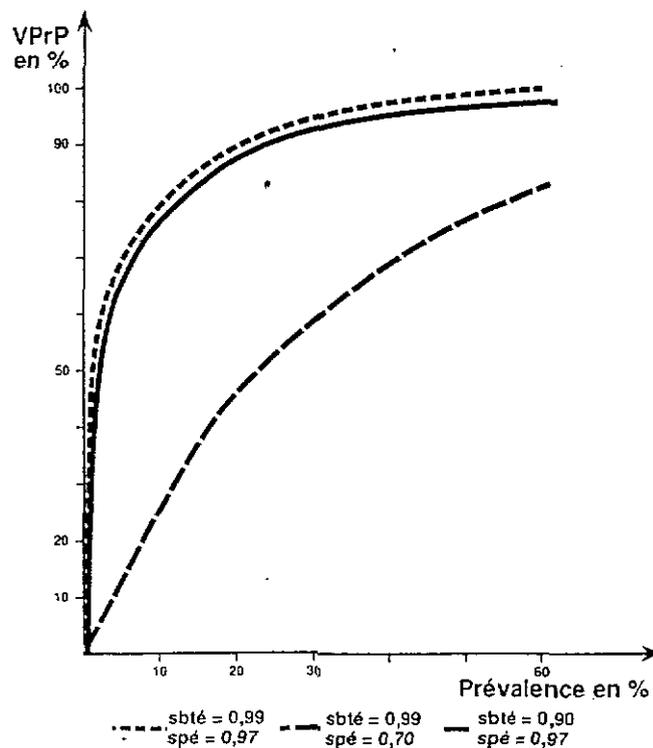
PREVALENCE = 0,1 %

$$VPrP = 3,2 \%$$

$$VPrN = 99,99 \%$$

Ainsi, on voit que pour un test de qualités intrinsèques données (spécificité et sensibilité), les qualités opérationnelles et plus particulièrement la valeur prédictive d'un résultat positif, varie en fonction de la prévalence, ce qui est représenté sur la figure 1.

Figure 1 : Evolution de la valeur prédictive d'un résultat positif en fonction de la prévalence.



On constate que lorsque la prévalence diminue, la valeur prédictive d'un résultat positif diminue de façon importante.

Par ailleurs, on note que pour deux méthodes d'analyse de même spécificité, le gain en valeur prédictive, d'un résultat positif, est limité si l'on améliore la sensibilité et qu'inversement pour deux méthodes de même sensibilité le gain en valeur prédictive d'un résultat positif est important dès que la spécificité est améliorée.

Cette liaison entre les qualités opérationnelles d'une méthode d'analyse, ses qualités intrinsèques et les conditions épidémiologiques de prévalence peut être objectivée par les expressions suivantes :

$$VPrP = \frac{\text{sensibilité} \times \text{prévalence}}{(\text{sensibilité} \times \text{prévalence}) + (1 - \text{prévalence})(1 - \text{spécificité})}$$

$$VPrN = \frac{\text{spécificité}(1 - \text{prévalence})}{[\text{spécificité}(1 - \text{prévalence})] + \text{prévalence}(1 - \text{sensibilité})}$$

Ainsi, dans ces deux expressions on constate que :

- La valeur prédictive d'un résultat positif tend vers 1
  - . si la prévalence tend vers 100 p. cent (ce qui n'est pas le cas le plus fréquent), ou
  - . si la spécificité tend vers 100 p. cent (ce qui doit donc motiver les choix entre techniques différentes).
- La valeur prédictive d'un résultat négatif tend vers 1
  - . si la prévalence devient faible (but recherché dans les politiques sanitaires), ou
  - . si la sensibilité tend vers 100 p. cent.

Ces divers aspects doivent donc conduire à des choix techniques raisonnés en fonction d'une situation épidémiologique donnée.

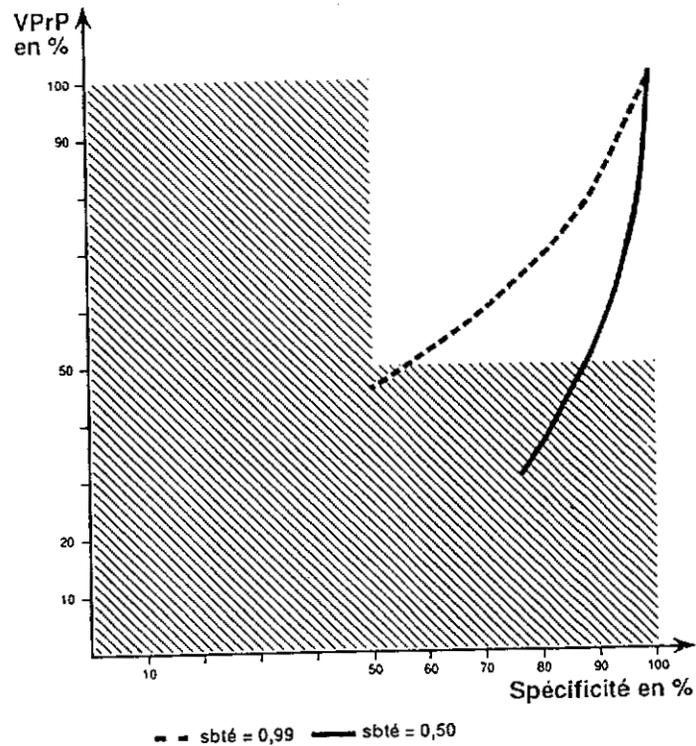
### **III - EVOLUTION DES VALEURS OPERATIONNELLES EN FONCTION DES VALEURS INTRINSEQUES DE SPECIFICITE ET SENSIBILITE**

Cette évolution est représentée dans les figures 2, 3, 4 et 5 pour une situation épidémiologique donnée (30 p. cent), choisie arbitrairement. Cette évolution n'est présentée que dans le cadran défini par valeur prédictive (positive ou négative) supérieure à 50 p. cent et par la valeur intrinsèque (spécificité ou sensibilité) supérieure à 50 p. cent. En effet, il ne paraît pas concevable que, dans une situation épidémiologique donnée, on puisse utiliser des méthodes dont l'aspect discriminant soit inférieur à 50 p. cent (probabilité d'un phénomène aléatoire).

#### **1. EVOLUTION DE LA VALEUR PREDICTIVE D'UN RESULTAT POSITIF**

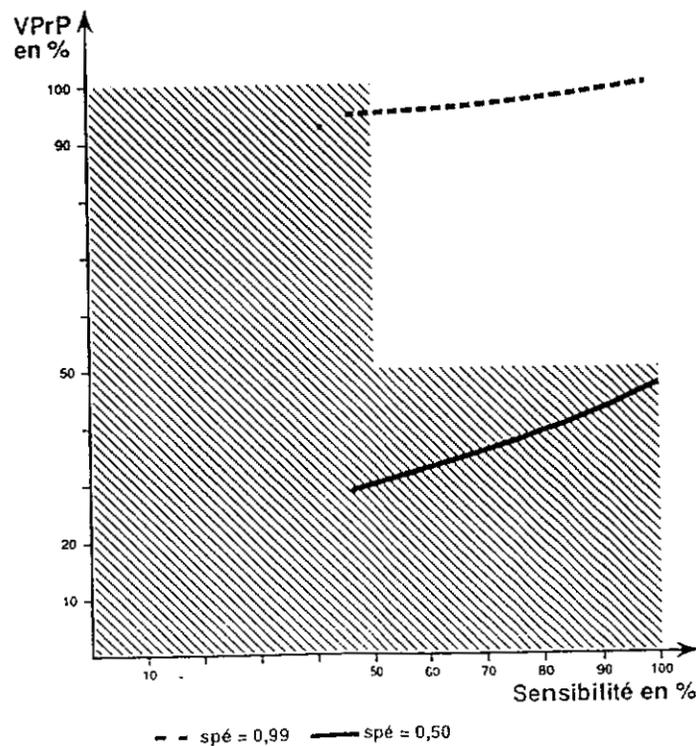
On constate, figure 2, que pour des tests de même sensibilité (0,99 par exemple), un gain considérable peut être obtenu pour la VPrP en jouant sur la spécificité.

**Figure 2 :** Evolution de la VPrP en fonction de la spécificité dans une situation épidémiologique donnée (Prévalence 30 %).



En revanche, des améliorations notables de la sensibilité d'un test de spécificité 0,50 par exemple ne permettent pas un gain satisfaisant de la VPrP (figure 3).

**Figure 3 :** Evolution de la VPrP en fonction de la sensibilité dans une situation épidémiologique donnée (Prévalence 30 %).

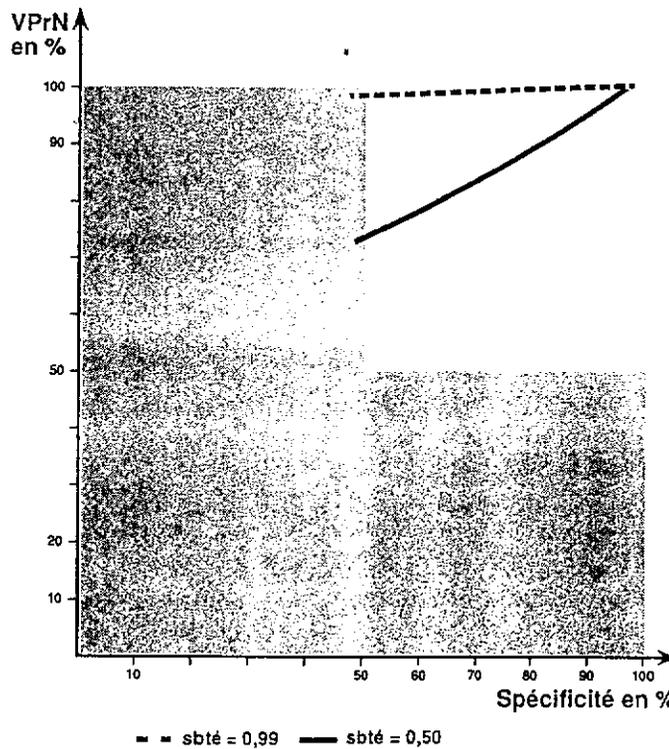


## 2. EVOLUTION DE LA VALEUR PREDICTIVE D'UN RESULTAT NEGATIF

Nous nous sommes placés dans une situation épidémiologique de prévalence plausible (30 p. cent) mais cependant de valeur faible si on se rapporte à l'expression de la valeur prédictive d'un résultat négatif en fonction de la prévalence. Ceci explique, dans ces conditions, l'influence limitée des qualités intrinsèques sur cette qualité opérationnelle (figures 4 et 5). Néanmoins, on constate que pour une sensibilité de 0,99, la spécificité ne modifie pratiquement pas la VPrN (figure 4), inversement pour des tests de spécificité 0,99, une modification limitée mais perceptible peut être obtenue en jouant sur la sensibilité (figure 5).

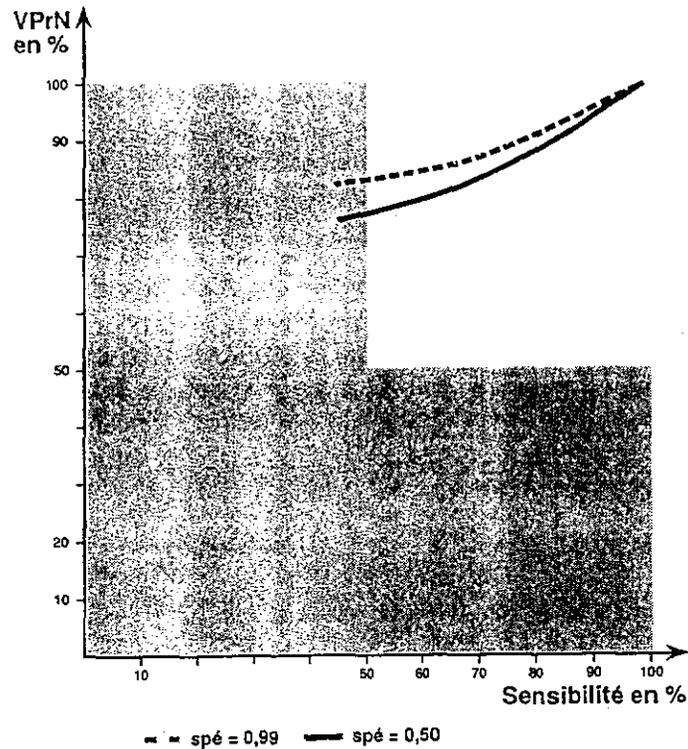
*N.B. : Les altérations de la valeur prédictive d'un résultat négatif seraient plus importantes dans une situation de prévalence beaucoup plus élevée.*

Figure 4 : Evolution de la VPrP en fonction de la spécificité dans une situation épidémiologique donnée (Prévalence 30 %).



Ces différents aspects interviennent donc sur les interprétations données aux résultats de laboratoire et, de la même façon que dans les conditions expérimentales étaient appréciées précédemment les erreurs par excès et les erreurs par défaut, il est important, dans les conditions épidémiologiques, d'apprécier les probabilités de conclure à tort en fonction de l'objectif fixé.

Figure 5 : Evolution de la VPrN en fonction de la sensibilité dans une situation épidémiologique donnée (Prévalence 30 %).



#### IV - IMPORTANCE DES QUALITES D'UNE TECHNIQUE EN FONCTION DE L'OBJECTIF DE QUALIFICATION D'UN INDIVIDU OU D'UN CHEPTEL

##### 1. QUALIFICATION D'UN INDIVIDU

Habituellement, la qualification sanitaire d'un individu est fonction du résultat qu'il donne, en tant qu'individu, à une méthode quelconque de dépistage, complétée par les caractères de son cheptel d'origine.

Si l'animal donne un résultat positif, il sera qualifié d'infecté. Cette qualification suppose en fait qu'aucun doute n'est porté sur le résultat. Or, nous venons de voir qu'il existe des faux positifs et que la probabilité pour que cet animal appartienne effectivement à la population infectée est appréciée par la valeur prédictive d'un résultat positif.

La probabilité de conclure à tort "animal infecté" peut donc s'exprimer par la probabilité complémentaire :

$$\text{Probabilité de conclure à tort "animal infecté"} : 1 - \text{VPrP}$$

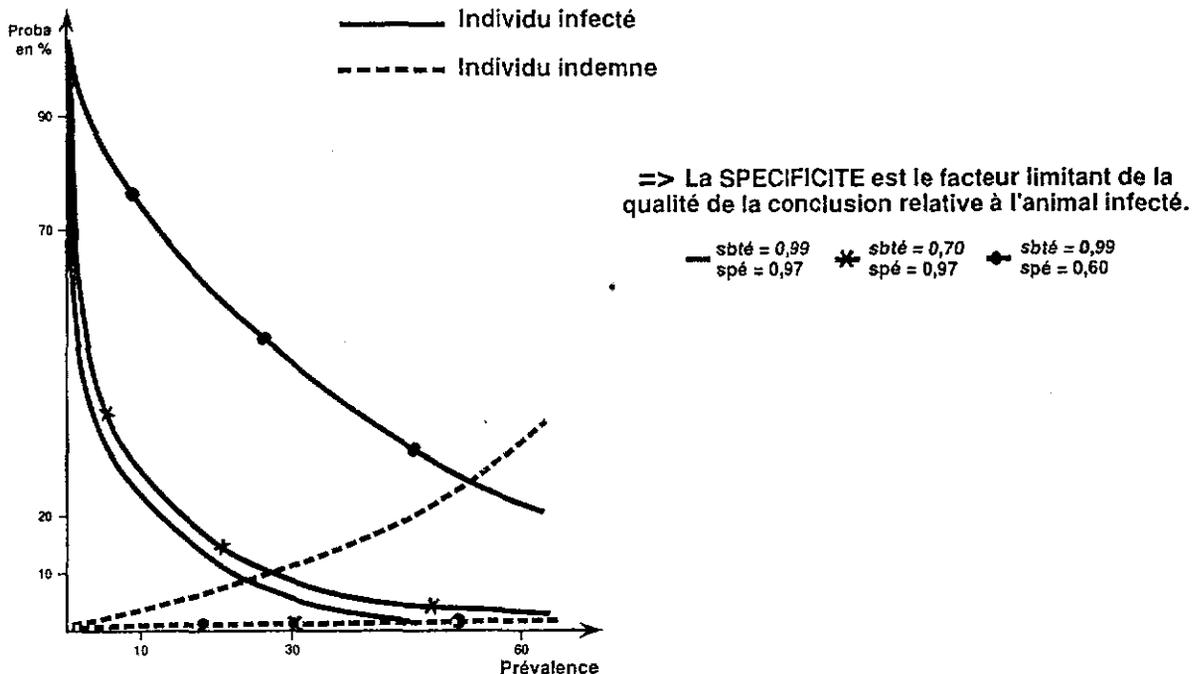
Inversement, la probabilité pour que l'animal qui donne un résultat négatif soit un vrai négatif est appréciée par la valeur prédictive d'un résultat négatif.

La probabilité de conclure à tort "animal indemne" peut donc s'exprimer par la probabilité complémentaire :

$$\text{Probabilité de conclure à tort "animal indemne"} : 1 - \text{VPrN}$$

Nous avons vu précédemment que les valeurs prédictives variaient en fonction de la prévalence (figure 1), il est donc évident que les risques d'attribuer une qualification erronée à un animal varient avec la prévalence (figure 6). On constate sur ce schéma que, pour des qualités intrinsèques données, la probabilité de conclure à faux, en particulier "animal infecté", peut être considérable quand la prévalence diminue.

Figure 6 : Probabilité de conclure à tort.



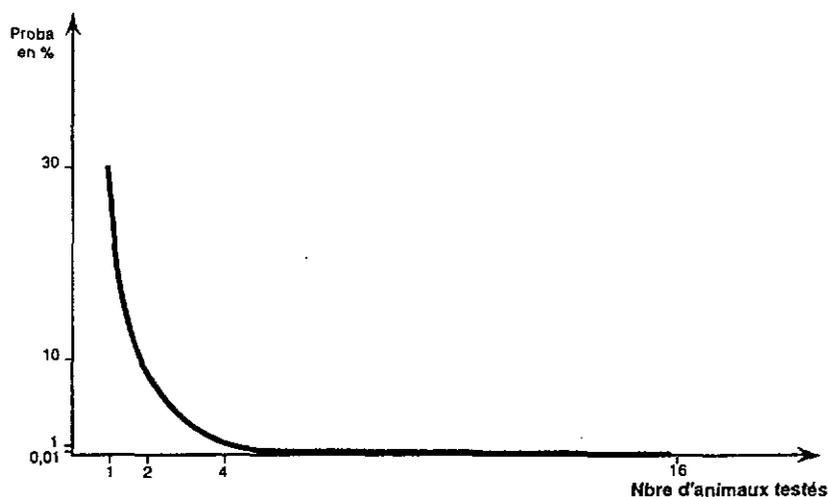
## 2. QUALIFICATION D'UN CHEPTTEL

Néanmoins, la qualification individuelle n'a généralement d'importance que par son impact sur la qualification du cheptel. Ainsi, il est fréquent qu'un cheptel soit déclaré infecté, si un animal et un seul donne un résultat positif au dépistage. On conçoit alors (cf figure 6) qu'en phase terminale d'éradication, les conclusions à tort sont trop nombreuses pour être admissibles. Ainsi, si l'on reprend l'exemple chiffré du tableau V, pour une prévalence de 0,1 p. cent la VPrP n'est que 3,2 P. cent, et donc la probabilité de qualification à tort "cheptel infecté" est de  $1 - 0,032 = 0,968$ , ce qui ne peut être admis ni par le décideur sanitaire ni par l'éleveur.

Dans le cas d'une qualification de cheptel indemne, le problème est tout autre puisque c'est l'ensemble des résultats du troupeau qui sont pris en considération et si la probabilité de conclure à tort "animal indemne" peut être relativement élevée dans certaines circonstances ( $1 - VPrN$ ), la qualification du cheptel est une probabilité conditionnée soit  $(1 - VPrN) (1 - VPrN) (\dots)$ .

Ainsi, figure 7, on constate que, même pour une probabilité individuelle de conclure à faux relativement élevée, soit 30 p. cent par exemple, cette probabilité décroît rapidement avec le nombre d'animaux testés donnant un résultat négatif : 9 p. cent pour 2 animaux simultanément négatifs, 2,7 p. cent pour 3 animaux, etc.

Figure 7 : Probabilité de conclure à tort = cheptel indemne.



Le problème majeur, en politique sanitaire d'éradication, est donc celui de la qualification "infecté" pour un seul résultat positif.

L'attitude "mathématique" serait, là encore, d'utiliser les probabilités conditionnées et de ne qualifier infecté qu'un cheptel qui fournirait simultanément un certain nombre de réponses positives, attitude épidémiologiquement difficile à admettre en terme de risque infectieux. Par ailleurs, cette décision devrait être systématiquement et progressivement adaptée aux conditions épidémiologiques et plus particulièrement à l'évolution de la prévalence, ce qui, pour satisfaisante que soit cette solution pour l'esprit n'en est pas moins d'une lourdeur d'application excessive.

Une autre attitude serait d'utiliser seule ou, mieux, en association, une autre méthode de très bonnes qualités intrinsèques et pour laquelle, surtout, faux positifs et faux négatifs ne seraient pas les mêmes animaux que ceux déterminés dans les lots expérimentaux par la première méthode. L'exclusion d'une erreur par excès ou par défaut d'une technique par la seconde diminuerait de façon considérable l'incertitude de l'interprétation.

### CONCLUSION

Le tableau I résume l'ensemble de ces aspects des qualités des analyses. Dans un cadre expérimental, la lecture du tableau se fait verticalement car les deux sous-populations (infectés et témoins,  $M$  et  $\bar{M}$ ) sont déterminées avant la lecture des résultats ; dans un cadre épidémiologique, les résultats des analyses sont connus avant la répartition en lot infecté et en lot témoin (lecture horizontale) et servent donc de base pour définir a posteriori ces deux sous-populations  $M$  et  $\bar{M}$ .

Il est désormais indispensable que tous les acteurs sanitaires gardent présent à l'esprit l'impact qu'ont les qualités analytiques sur les interprétations et décisions prises.