

ETUDE DE L'INFLUENCE DE LA DEPENDANCE DES TESTS SUR LES VALEURS PREDICTIVES D'UNE COMBINAISON DE DEUX TESTS.

APPLICATION A L'EXEMPLE DE L'IBR*

Etienne Petit¹ et Régis Pouillot²

RESUME : Dans le cadre du dépistage des maladies infectieuses, le recours à un second test lors de réaction positive à un premier test est souvent utilisé pour confirmer ou infirmer la première réaction. Or, l'aptitude à établir un diagnostic biologique juste dépend des valeurs classiques de sensibilité et de spécificité des tests, mais également de la dépendance entre les tests. Cette notion essentielle est le plus souvent négligée.

Dans ce contexte, l'influence de la dépendance de deux tests utilisés est étudiée, dans un premier temps sous un angle théorique et dans un second temps sur un exemple numérique inspiré des tests sérologiques ELISA utilisés pour le dépistage de l'IBR.

Plus deux tests sont dépendants, plus ils auront tendance à donner le même résultat. La dépendance des réactifs peut se décomposer en :

1. une dépendance de leur sensibilité : si la dépendance est importante, le second test aura tendance à confirmer le premier test sur un animal infecté. Cette dépendance en sensibilité se reflète également dans la proportion de bovins infectés et non dépistés par les deux tests ;
2. une dépendance de spécificité : si la dépendance est importante, le second test aura tendance à confirmer un premier test anormalement positif sur un animal indemne. Cette dépendance en spécificité se reflète par la proportion de bovins non infectés et classés faussement positifs par les deux tests

La connaissance de ces facteurs de dépendance, exprimés ici de manière originale, permet de calculer les valeurs prédictives de chaque combinaison des deux tests. L'application numérique sur l'exemple de l'IBR révèle essentiellement la forte influence de la dépendance des spécificités sur le risque de classer infecté à tort un animal sain, en situation de très faible prévalence (situation des élevages qualifiés). Dans l'exemple étudié, en situation de prévalence de 0,1% d'animaux infectés, la probabilité qu'un animal fournissant une réponse positive aux deux tests soit réellement infecté passe de 90% en cas d'indépendance des tests à moins de 20% (facteurs de dépendance estimée à partir d'une population de 5 221 bovins analysés conjointement avec les deux tests). Ce constat semble rejoindre les observations de terrain, qui déroutent parfois les gestionnaires du dépistage de l'IBR, comme autrefois ceux de la brucellose. La signification des dépendances des tests et leur appréciation quantitative sont également discutées.

Mots-clés : Tests multiples, dépendance des tests, IBR, sensibilité, spécificité, valeurs prédictives.

* Texte de la communication présentée à la Journée AEEMA, 20 mai 2005

¹ FRGDS Bourgogne, 42 rue de Mulhouse, 21000 Dijon, France

² Centre Pasteur du Cameroun, Yaoundé, Cameroun

SUMMARY : For the screening of infectious diseases, a second test is often used to confirm or refute an initial positive test. However, the aptitude to establish a correct biological diagnosis depends on the well known sensitivity and specificity values, but also on the dependence between the tests. This essential concept is generally ignored. In this context, the influence of the dependence of two tests used is studied, theoretically and on a numerical example from IBR ELISA tests used for screening. The more dependent two tests are, the more they will tend to give the same result. The dependence may be split into :

1. a dependence on sensitivities: if this dependence is high, the second test will tend to confirm the first test on an infected animal. However, this dependence on sensitivity is also reflected in the proportion of infected cattle not detected by both reagents;
2. a dependence on specificities: if the dependence is high, the second test will tend to confirm a first false positive test on a disease free animal. This dependence in specificity is reflected by the proportion of disease free cattle classified erroneously positive by the two tests.

The knowledge of these factors of dependence, expressed here in an original way, allows calculating the predictive values of each combination of the two tests. The numerical application on the example of the IBR reveals the strong influence of the dependence of specificities on the risk to falsely classify infected an healthy animal, in situation of very weak prevalence (situation of the qualified herds). In the studied example, in a 0,1% prevalence situation, the probability that a two tests positive animal is really infected decreases from 90% in the case of independence of the tests to less than 20% (factors of dependence estimated from a population of 5 221 bovines analyzed jointly with the two tests). This analysis is in accordance with field observations, where managers of the IBR screening are sometimes confounded, as formerly those of Brucellosis. The significance of the dependences of the tests and their quantitative appreciation are also discussed.

Keywords : Multiple tests, dependence between tests, IBR, sensitivity, specificity, predictive value.



I - INTRODUCTION - OBJECTIFS

Dans le cadre du dépistage des maladies infectieuses, les gestionnaires des plans de lutte sont souvent amenés à utiliser des tests complémentaires pour confirmer ou infirmer des résultats obtenus par un premier test de dépistage. Cette procédure est surtout utilisée pour les résultats douteux ou suspects d'être faussement positifs dans des circonstances épidémiologiques particulièrement favorables et écartant *a priori* la possibilité d'une contamination. En pratique, on utilise souvent un test sensible en première intention, puis une confirmation des individus positifs par un test plus spécifique pour lever les éventuelles suspicions d'individus faussement positifs. Cette pratique se justifie tout à fait lorsque les caractéristiques des tests, telles que la

sensibilité et la spécificité, sont connues, et dans l'hypothèse d'une indépendance des tests. Cependant, la notion de dépendance entre les tests utilisés est rarement prise en compte, en particulier sur un plan quantitatif. Un exemple notoire en santé animale a cependant montré l'importance de cette dépendance : les réactions non spécifiques à la brucellose bovine ont nécessité le recours à un test indépendant des tests sérologiques, à savoir le test à la brucelline. Cette étude présente l'importance théorique de ces facteurs de dépendance et applique cette réflexion à l'exemple du dépistage de l'IBR. Elle propose également quelques pistes de réflexion pour apprécier les dépendances pouvant exister entre divers tests.

II - APPROCHE THEORIQUE SUR DEUX TESTS

Les développements mathématiques sont présentés en annexe.

1. DEPENDANCE ET COMBINAISON DE DEUX TESTS

Les résultats de deux tests appliqués à un même individu vont dépendre de son statut réel (infecté ou non) et des probabilités de réponse de chacun des tests, conditionnellement à ce statut. Ces probabilités s'expriment pour chaque test, d'une part, en terme de **sensibilité**, qui correspond à la probabilité pour un individu réellement infecté d'être donné positif par le test et, d'autre part, en terme de **spécificité** qui correspond à la probabilité pour un individu réellement sain d'être donné négatif par le test (cf. annexe). Cependant, la connaissance de la sensibilité des deux tests ne permet pas de connaître directement la probabilité pour un individu réellement infecté d'être donné positif **conjointement** par les deux tests.

Classiquement, il est supposé une indépendance des tests. Dans ce cas, la probabilité pour un individu réellement infecté d'être donné positif conjointement par les deux tests est égale au produit de la sensibilité de chaque test. De même, la probabilité pour un individu réellement sain d'être donné conjointement négatif par les deux tests est égale au produit de la spécificité de chaque test.

Or, la probabilité de réponse positive du second test peut être différente selon la réponse du premier test, et ceci dans les sous-populations infectées ou non infectées : les réponses peuvent être liées par une relation de dépendance au niveau de leur sensibilité et de leur spécificité [Pouillot, 2001].

L'estimation des performances de la combinaison de deux tests T1 et T2 nécessite donc la connaissance de deux paramètres complémentaires qui traduisent la relation de dépendance qui peut exister entre les deux tests T1 et T2, tant pour la sensibilité que pour la spécificité.

Ces paramètres peuvent être représentés en terme probabiliste, d'une part, par un facteur multiplicatif (K_{se}) de la sensibilité du test 2, alors que le test 1 s'est avéré positif sur un animal infecté et, d'autre part, par un autre facteur multiplicatif (K_{sp}) de la spécificité du test 2 alors que le test s'est avéré négatif sur

un animal non infecté [Petit, 2002] (cf. annexe).

Alors, la probabilité qu'un test T2 soit positif alors que T1 est positif en population infectée vaut $K_{se} \times Se_2$, où Se_2 représente la sensibilité du test T2. De même, la probabilité qu'un test T2 soit négatif alors que T1 est négatif en population non infectée vaut $K_{sp} \times Sp_2$, où Sp_2 représente la spécificité du test T2.

Si ces facteurs sont égaux à 1, ils traduisent une indépendance des tests.

S'ils sont supérieurs à 1, ils traduisent un degré de dépendance qui s'élève avec la valeur de ces facteurs.

S'ils sont inférieurs à 1 ils traduisent une dépendance « négative » des tests. Par exemple, dans le cas d'une dépendance « négative » des sensibilités, la réponse positive au premier test sur un animal infecté tend à exclure une réponse positive sur le second test.

Par construction et pour préserver des probabilités inférieures à 1, ces paramètres connaissent une limite supérieure et inférieure que l'on peut calculer pour établir l'étendue de leur variation (cf. annexe).

On montre que ces facteurs sont liés à la proportion d'individus faussement classés par le 1^{er} test et confirmés par le second. Ainsi, la dépendance entre sensibilités (K_{se}) est liée linéairement à la proportion de faux négatifs au premier test, confirmés négatifs par le second test. Réciproquement, la dépendance entre spécificités (K_{sp}) est liée linéairement à la proportion de faux positifs au premier test, confirmés positifs par le second test.

Ces proportions sont plus faciles à appréhender que les paramètres multiplicatifs K_{se} et K_{sp} , car elles représentent une donnée observable d'individus faussement classés par le premier test et confirmés par le second test. Ce sont ces valeurs qui seront utilisées pour représenter l'importance des dépendances dans l'exemple numérique de cette étude.

2. VALEURS PREDICTIVES D'UNE COMBINAISON DE DEUX TESTS

A partir de ces paramètres et de la connaissance de la prévalence réelle de l'infection, on peut dès lors calculer les valeurs

prédictives (positive, notée « VPP », et négative, notée « VPN ») de chaque combinaison possible de résultats (cf. annexe).

Dans l'usage courant d'une procédure de contrôle des résultats positifs, on s'intéressera surtout à la VPP du résultat « T1 positif confirmé par T2 positif » (confirmation, noté « pp ») et à la VPN du résultat « T1 positif infirmé par T2 négatif » (infirmité, notée « pn »). La VPP d'une combinaison « pp » (notée VPP(pp)) est égale à la proportion d'animaux réellement **infectés** parmi ceux positifs aux deux tests ; la VPN d'une combinaison « pn » (notée VPN(pn)) est égale à la proportion d'animaux **non infectés** parmi les individus classés « pn ».

A partir de leur expression en fonction des différents paramètres de sensibilité, spécificité et de prévalence, l'étude des variations de

VPP(pp) et de VPN(pn) peut être réalisée en fonction des variations de chacun de leurs paramètres constitutifs (tableau I).

On constate qu'une dépendance entre les sensibilités exerce une influence positive sur les deux valeurs prédictives alors qu'une dépendance entre les spécificités exerce une influence négative sur les deux valeurs prédictives. Paradoxalement, une meilleure sensibilité ou une meilleure spécificité du premier test jouent défavorablement sur la VPN du résultat « pn ».

Il faut cependant relativiser le poids de chacun des paramètres dans la variation des valeurs prédictives, comme le montrera l'exemple numérique.

Tableau I

Variations de VPP(pp) et de VPN(pn) en fonction de chaque paramètre

Paramètre	VPP(pp)	VPN(pn)
Se1 ↗	↗	↘
Se2 ↗	↗	↗
Sp1 ↗	↗	↘
Sp2 ↗	↗	↗
Kse ↗	↗	↗
Ksp ↗	↘	↘
Prévalence (F) ↗	↗	↘

III - APPLICATION NUMERIQUE A L'EXEMPLE DE L'IBR

1. ESTIMATION DES SENSIBILITES ET SPECIFICITES DES TESTS

L'exemple de l'IBR avec deux tests sérologiques, le premier basé sur la recherche d'anticorps dirigés contre la protéine gB, le second sur la recherche des anticorps totaux (AcT), est utilisé pour illustrer l'influence quantitative des facteurs de dépendance sur un dépistage s'appuyant sur deux tests. Le test gB étant réputé plus sensible que le test AcT, il sera utilisé en première intention.

En première approche, les caractéristiques de ces deux tests ont été estimées sur un échantillon de 5 221 sérums analysés conjointement avec ces deux tests et un

troisième test sérologique (recherche des anticorps dirigés contre la glycoprotéine gE). L'estimation a été effectuée par la méthode proposée par Hui et Walter [1981]. Cette méthode, fondée sur une estimation du maximum de vraisemblance, permet d'estimer les caractéristiques des tests en l'absence de test de référence, sous l'hypothèse d'une absence de dépendance entre les tests [Gardner et Greiner, 1999].

Les valeurs obtenues et leur intervalle de confiance à 95% (IC 95%) figurent dans le tableau II.

Tableau II
Estimations des sensibilités et spécificités de deux tests sérologiques IBR

Paramètre	Valeur estimée	IC95%
Se Ac gB	99,3%	[98,7% ;99,9%]
Sp Ac gB	98,4%	[98,1% ;98,9%]
Se Ac totaux	98,4%	[97,4% ;99,4%]
Sp Ac totaux	99,3%	[99,0% ;99,5%]

L'estimation confirme la plus forte sensibilité du test gB, mais également une spécificité moindre que celle des anticorps totaux.

l'évolution de Kse, (exprimée par le pourcentage de confirmation des faux négatifs) et pour trois valeurs de prévalence : 10%, 1% et 0,1%.

2. EVOLUTIONS DE VPP(PP) ET DE VPN(PN) EN FONCTION DE LA DEPENDANCE DES SENSIBILITES DES DEUX TESTS

La dépendance entre sensibilités n'exerce pratiquement aucune influence sur la VPP(pp), mais elle améliore sensiblement la VPN(pn) d'autant plus que la prévalence est élevée (cf. figure 1). A très faible prévalence, elle n'exerce plus d'influence (cf. figure 3).

Les figures 1 à 3 donnent l'évolution de VPP(pp) et de VPN(pn) en fonction de

Figure 1

Evolution de VPP (pp) et VPN (pn) pour une prévalence de 10%

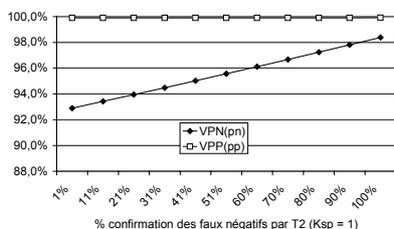


Figure 2

Evolution de VPP (pp) et VPN (pn) pour une prévalence de 1%

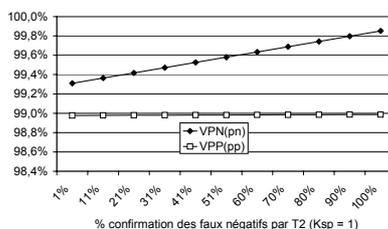
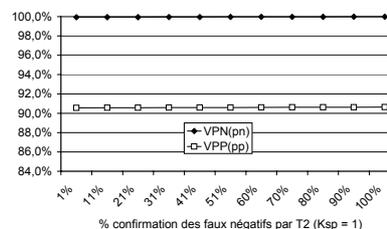


Figure 3

Evolution de VPP (pp) et VPN (pn) pour une prévalence de 0,1%



3. EVOLUTIONS DE VPP(PP) ET DE VPN(PN) EN FONCTION DE LA DEPENDANCE DES SPECIFICITES DES DEUX TESTS

VPN(pn). En situation faiblement infectée (1%, cf. figure 5) la dégradation ne devient importante (VPN inférieure à 99%) que pour une dépendance importante (plus de 30% de faux positifs au test 1 confirmés par le test 2). Pour les faibles prévalences, la VPP(pp) se dégrade très fortement avec la dépendance des spécificités. Ainsi, si la prévalence est de 0,1%, une dépendance qui amènerait une confirmation par le test 2 de seulement 8% de faux positifs au test 1, amènerait une VPP de moins de 50%, contre 90% en cas d'indépendance entre les tests (cf. figure 6).

Les figures 4 à 6 donnent l'évolution de VPP(pp) et de VPN(pn) en fonction de l'évolution de Ksp, (exprimée par le pourcentage de confirmation des faux positifs) et pour trois valeurs de prévalence : 10%, 1% et 0,1%.

Plus la prévalence est élevée, plus la dépendance entre spécificités dégrade la

Figure 4
Evolution de VPP (pp) et VPN (pn) pour une prévalence de 10%

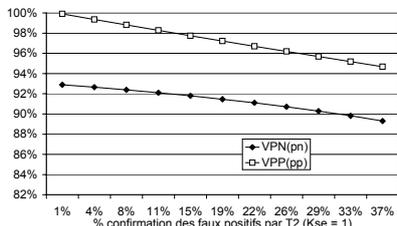


Figure 5
Evolution de VPP (pp) et VPN (pn) pour une prévalence de 1%

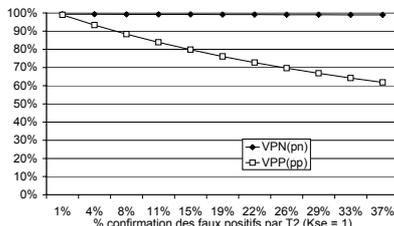
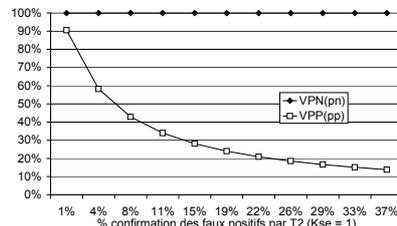


Figure 6
Evolution de VPP (pp) et VPN (pn) pour une prévalence de 0,1%



4. INFLUENCE DES CARACTERISTIQUES DU SECOND TEST

Il est généralement d'usage de privilégier la spécificité du second test, surtout lorsque la sensibilité a été privilégiée sur le premier. Mais il faut cependant maintenir une bonne sensibilité du second test pour préserver une bonne VPN(pn). Les figures 7 à 15 montrent l'influence de la sensibilité du second test sur VPP(pp) et sur VPN(pn) pour trois niveaux de prévalence et trois niveaux de dépendance de la spécificité. La dépendance entre sensibilités est supposée nulle.

Le niveau de sensibilité du second test a un effet très positif sur la VPN(pn) pour les prévalences fortes et moyennes, et ce quel que soit le niveau de dépendance de la spécificité. Cet effet est beaucoup plus réduit sur les faibles prévalences. Ainsi, pour une prévalence de 0,1%, la VPN(pn) varie de 98,4% à 99,8% pour une sensibilité du test 2 variant de 80 à 98%. La dépendance entre les spécificités dégrade légèrement ces valeurs.

La sensibilité du deuxième test influence assez peu la VPP(pp) mais la dépendance entre les spécificités l'influence négativement, et de façon très sensible, notamment en situation de très faible prévalence.

Figure 7
Evolution de VPP (pp) et VPN (pn) pour une prévalence de 10%. Dépendance des spécificités nulle.

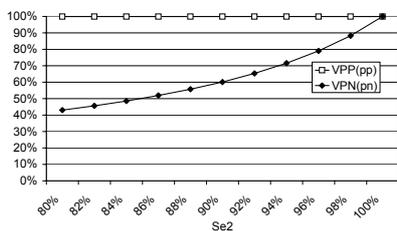


Figure 8
Evolution de VPP (pp) et VPN (pn) pour une prévalence de 1%. Dépendance des spécificités nulle.

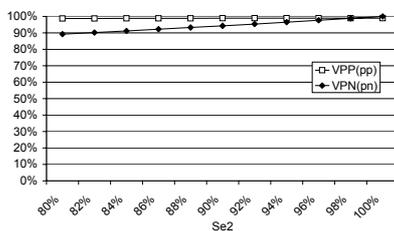


Figure 9
Evolution de VPP (pp) et VPN (pn) pour une prévalence de 0,1%. Dépendance des spécificités nulle.

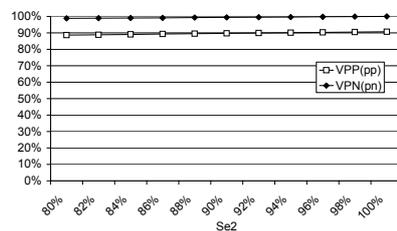


Figure 10

Evolution de VPP (pp) et VPN (pn) pour une prévalence de 10%.

Dépendance des spécificités faible $p(T2+|T1+) = 10\%$ chez des non infectés.

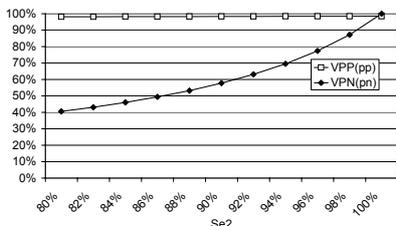


Figure 11

Evolution de VPP (pp) et VPN (pn) pour une prévalence de 1%.

Dépendance des spécificités faible $p(T2+|T1+) = 10\%$ chez des non infectés.

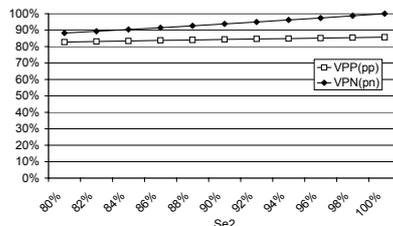


Figure 12

Evolution de VPP (pp) et VPN (pn) pour une prévalence de 0,1%.

Dépendance des spécificités faible $p(T2+|T1+) = 10\%$ chez des non infectés.

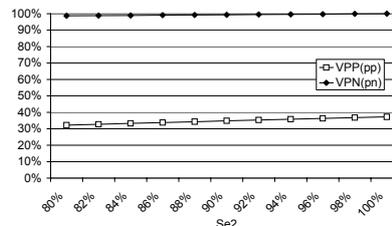


Figure 13

Evolution de VPP (pp) et VPN (pn) pour une prévalence de 10%.

Dépendance des spécificités forte $p(T2+|T1+) = 30\%$ chez des non infectés.

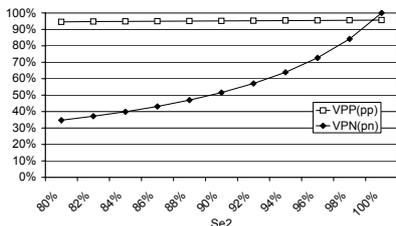


Figure 14

Evolution de VPP (pp) et VPN (pn) pour une prévalence de 1%.

Dépendance des spécificités forte $p(T2+|T1+) = 30\%$ chez des non infectés.

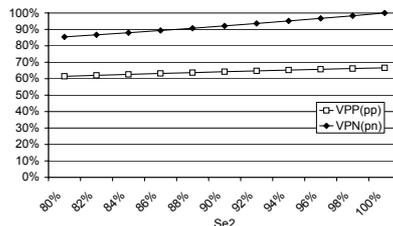
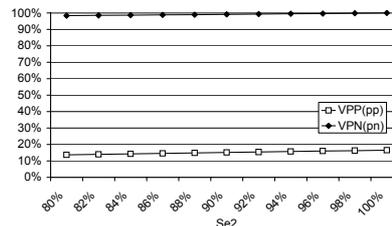


Figure 15

Evolution de VPP (pp) et VPN (pn) pour une prévalence de 0,1%.

Dépendance des spécificités forte $p(T2+|T1+) = 30\%$ chez des non infectés.



IV - APPLICATION NUMERIQUE

AVEC LES FACTEURS DE DEPENDANCE ESTIMES

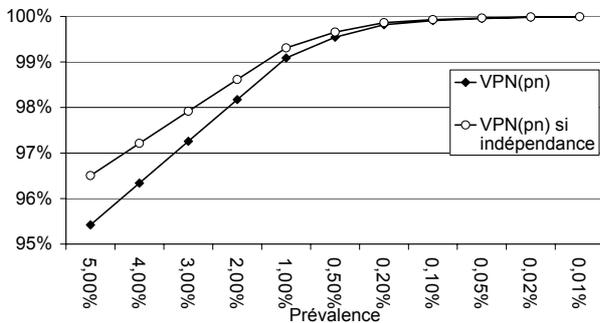
A l'aide d'une méthode inspirée par celle de Hui et Walter (1981), mais n'utilisant pas l'hypothèse d'indépendance entre les deux tests (nos résultats non publiés), l'estimation des sensibilités et spécificités des tests sérologiques de l'IBR a pu s'accompagner de celle des facteurs de dépendance à partir de quatre sous-populations issues de la population des 5 221 bovins analysés avec deux tests individuels.

Le facteur de dépendance des sensibilités est estimé à 1,0001 avec un intervalle de confiance à 95% : [0,9951 ; 1,0051]. Cette valeur traduit une vraisemblable indépendance de la sensibilité des réactifs.

Le facteur de dépendance des spécificités est estimé à 1,0043 avec un intervalle de confiance à 95% : [1,0025 ; 1,0061]. Compte tenu des spécificités estimées des réactifs, cette valeur correspond à une confirmation par le test aux anticorps totaux de 26% des faux positifs décelés par le test aux anticorps gB.

Les figures 16 et 17 donnent respectivement l'évolution de la VPN(pn) et la VPP(pp) pour diverses situations de prévalence décroissantes et comparent ces valeurs entre une situation d'indépendance des spécificités et la situation de dépendance estimée.

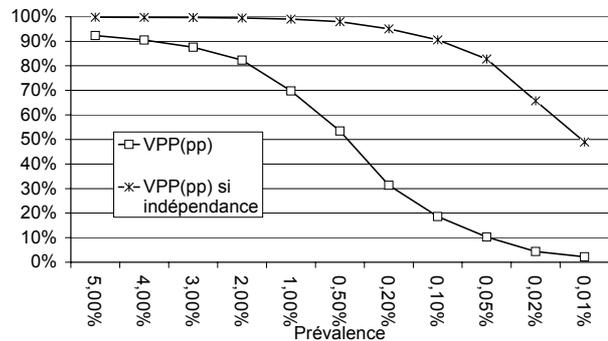
Figure 16
Evolution de la VPN(pn)
en fonction de la prévalence



Dans les situations de très faible prévalence *a priori* (exemple des élevages quasiment indemnes ou considérés comme tels, dont on peut considérer la prévalence comme proche ou inférieure à 0,1%), la VPN(pn) est de l'ordre de 99,9% : la probabilité pour un animal positif au premier test, infirmé par un second test négatif, d'être indemne est équivalent au risque de base de la population, que les spécificités des tests soient dépendantes ou non.

Par contre, dans la situation de dépendance des spécificités, la VPP(pp) est de l'ordre de 20% lorsque la prévalence est de 0,1% : seul un bovin positif sur cinq confirmés par les deux tests correspondrait à un vrai infecté. Cette valeur est bien inférieure à la VPP de 90% que

Figure 17
Evolution de la VPP(pp)
en fonction de la prévalence



l'on évaluerait en l'absence de prise en compte de la dépendance entre les tests. Ce constat rejoint les observations de terrain, où la plupart des résultats positifs confirmés en élevage qualifié ne trouvent aucune explication épidémiologique, et ne sont associés à aucune circulation virale dans le cheptel du bovin concerné. De plus, les investigations complémentaires (sérologies sur de nouveaux prélèvements, recherches virologiques, suivi des cheptels concernés) tendent à confirmer le statut négatif de ces animaux. Cet exemple illustre la nécessité de recourir à des tests - virologie, analyse épidémiologique - indépendants de la réponse sérologique d'un seul bovin.

V - DISCUSSION

1. METHODE

L'approche quantitative des dépendances entre tests est assez peu documentée, mais elle permet de comprendre leur importance. Les méthodes basées sur l'indicateur Kappa permettent une telle approche, mais ce facteur est trop global et résulte de tous les paramètres qui conditionnent les valeurs prédictives, y compris la prévalence [Bonnardel, 1995]. D'où la nécessité de décomposer la concordance des tests en facteurs de dépendance.

La modélisation quantitative des dépendances a été rapportée par Gardner et Greiner sous forme de covariance des sensibilités et spécificités [Gardner et Greiner, 1999]. La transformation de la covariance, qui est

additive, en un facteur multiplicatif facilite le calcul des valeurs prédictives des combinaisons de tests et permet une modélisation au delà de deux tests [Petit, 2002].

Une des difficultés principales est l'estimation de ces nouveaux paramètres K_{sp} et K_{se} . Une possibilité est, à partir d'individus classés réellement infectés sur la base d'un diagnostic clinique, épidémiologique ou bactériologique (ou viral), d'établir la fréquence d'individus confirmés « négatif » par le second test alors qu'ils étaient déjà négatifs au premier test (= $p(T2-|T1-, \text{infecté})$). Parallèlement, sur des individus classés réellement non infectés sur la base d'un diagnostic clinique, épidémiologique ou bactériologique (ou viral), on peut établir la fréquence d'individus confirmés « positif » par

le second test alors qu'ils étaient déjà positifs au premier test (= $p(T2+|T1+, indenne)$). Les paramètres de dépendance peuvent alors être estimés si les caractéristiques « classiques » des tests sont connus avec précision, à l'aide des relations 1) et 2)(Annexe).

Des méthodes statistiques par analyse de classes latentes, bayésiennes ou fréquentistes, peuvent permettre d'estimer les facteurs de dépendance. Elles offrent l'avantage d'être applicables sans test de référence, mais nécessitent de recourir à plusieurs tests simultanés sur plusieurs sous populations d'animaux. De plus, elles imposent d'assumer le paradigme de Hui et Walter, à savoir que la sensibilité et la spécificité des tests soient constants entre les sous populations. En cas de réactions croisées liées à un autre agent infectieux, cette assertion peut ne pas être vérifiée. Une solution consiste alors à intégrer le phénomène des réactions croisées dans le modèle des classes latentes, mais cette pratique multiplie les paramètres à évaluer et impose des sous populations supplémentaires, ce qui complique les calculs à réaliser.

2. SIGNIFICATION BIOLOGIQUE DES DEPENDANCES

Plus les techniques seront proches, par exemple deux techniques sérologiques, plus les dépendances seront importantes.

Pour la sensibilité l'intensité de la dépendance entre deux techniques proches provient essentiellement d'un problème de détectabilité de certains individus malades mais présentant des réponses faibles ou nulles aux tests utilisés. Des techniques recherchant des réponses de nature différente auront plus de chances de repérer ces individus sur au moins un test (par exemple, virologie et sérologie), mais il reste fort probable que certains individus infectés échappent à tout test recherchant une réponse biologique de leur part. Dès lors, ils créent une certaine dépendance entre ces tests.

Pour la spécificité deux cas de figure peuvent se présenter.

Premièrement, les faux positifs sont souvent liés à un « bruit de fond » dû à des individus dépassant légèrement le seuil de positivité fixé souvent pour répondre à des objectifs de détectabilité. Les réponses aux tests de ces individus constituent souvent la partie extrême de la réponse « normale » d'une population d'individus non malades, qui présente une variabilité plus ou moins importante. Cette

variabilité « normale » est liée aux individus qui composent la population, mais également à la variabilité intrinsèque des tests, en particulier leur répétabilité. Il paraît vraisemblable que ces variabilités soient relativement indépendantes d'un test à l'autre. Ces défauts de spécificité s'expriment souvent par des valeurs proches du seuil de positivité retenu pour chaque test considéré.

Deuxièmement, les réactions faussement positives peuvent être provoquées par des réactions dites croisées qui génèrent des réponses biologiques voisines de celles provoquées par la maladie recherchée. Selon l'affinité de chacun des tests vis-à-vis de l'agent croisant, le niveau de dépendance de leur spécificité sera plus ou moins élevé. De même, l'intensité des réactions faussement positives pourra être très variable. Si l'agent croisant est de nature contagieuse, les réactions faussement positives peuvent suggérer un profil épidémiologique de type contagieux : concentration des réactions dans certains élevages, voire certains lots ou catégories d'animaux. Dans ces situations, la distinction avec la maladie recherchée est souvent difficile, et s'appuie souvent sur les circonstances épidémiologiques locales. Dans l'exemple de la brucellose bovine, les réactions faussement positives apparues dans le début des années 1990 sont très vraisemblablement liées à une infection par *Yersinia enterocolitica* O:9, présentant un lipopolysaccharide (S-LPS) de structure très proche de celle de *Brucella*. Tous les tests sérologiques fondés sur la recherche de la réponse antigénique contre ce S-LPS présentaient une forte dépendance en spécificité, vis-à-vis de *Brucella* (et une forte dépendance en sensibilité vis-à-vis du S-LPS). Il a fallu recourir à un test de nature complètement différente, la brucellination, pour distinguer les réactions dites « atypiques » des infections réelles à *Brucella* [Pouillot *et al.*, 1999].

Remarque : seul le cas des dépendances « positives » ($K > 1$) a été envisagé dans cet article. Il est cependant possible d'imaginer une dépendance « négative » ($K < 1$) dans les cas où les tests font appel à des réponses biologiques qui sont relativement incompatibles sur un même individu au même moment. Par exemple, en paratuberculose bovine, la sérologie et la coproculture semblent présenter une telle dépendance « négative ».

3. APPLICATION NUMERIQUE A L'IBR

Les caractéristiques des réactifs sérologiques de l'IBR et leurs dépendances ont été estimées à partir d'une population de 5 221 bovins répartis dans 47 cheptels de l'Yonne. Ces bovins ont été analysés avec trois réactifs (anticorps totaux, gB, gE). Pour estimer l'ensemble des caractéristiques intrinsèques des deux premiers réactifs et leurs dépendances, cette population a été scindée en quatre sous-populations de taille homogène en distinguant les cheptels les plus positifs, les moins positifs et les situations intermédiaires [Petit et Perrin, 2003].

Sur l'exemple numérique de l'IBR, on peut constater que les facteurs de dépendance jouent une influence plus ou moins sensible en fonction du contexte épidémiologique. La dépendance entre sensibilités s'exprime en situation de prévalence moyenne à élevée et améliore la VPN(pn). Cependant, dans ces situations, la confirmation d'un résultat positif présente un intérêt assez relatif. En situation de faible prévalence (0,1%), la VPN(pn) est égale à 99,90% dans l'hypothèse d'une indépendance des réactifs et à 99,87% avec les facteurs de dépendances estimés. Ainsi, le risque associé à un tel résultat est proche du risque de base. Ce constat a été intégré dans le cahier des charges des appellations IBR qui a validé le simple recontrôle par un second test pour infirmer un premier résultat positif observé dans un élevage qualifié.

La dépendance entre spécificités joue un rôle plus déterminant sur la VPP(pp), d'autant plus que la prévalence est basse. C'est malheureusement la situation dans laquelle travaillent les gestionnaires des plans de qualification des élevages. Une procédure complémentaire au cahier des charges national est en cours de rédaction pour proposer un protocole de confirmation des bovins à réponse positive confirmée issus des cheptels indemnes. Ce protocole s'appuiera sur des nouveaux contrôles sérologiques du bovin concerné, complétés si besoin est par une recherche virologique, et une analyse du lot de bovins en contact avec le bovin positif.

Sur le plan pratique, il est très difficile de chiffrer la valeur de ces dépendances, notamment en l'absence de méthode de référence absolue. On peut cependant proposer plusieurs méthodes, imparfaites mais complémentaires, pour estimer ces dépendances, en particulier pour des tests sérologiques comme les tests IBR.

Sur le terrain, des sérums individuels sont souvent repérés car ils posent des problèmes d'interprétation. Il s'agit généralement de sérums positifs « discordants » (résultats contradictoires selon les tests utilisés) ou « aberrants » (résultats positifs généralement confirmés par plusieurs tests mais inexplicables sur un plan épidémiologique). Ces sérums sont fréquemment adressés au laboratoire de référence pour une contre-expertise. Il serait utile de conserver ces sérums et de confirmer *a posteriori* le statut des bovins dont ils sont issus. Cette confirmation peut s'établir à partir d'un suivi dans le temps des individus concernés (absence de séroconversion pendant trois ans par exemple) ou d'une recherche plus approfondie (virologie PCR sur ganglions trijumeaux ou protocole de réactivation virale par exemple). Sur les individus ainsi classés réellement infectés, on peut établir la fréquence d'individus observés « négatif » par le second test alors qu'ils étaient déjà négatifs au premier test. Parallèlement, sur les individus ainsi classés réellement non infectés, on peut établir la fréquence d'individus confirmés « positif » par le second test alors qu'ils étaient déjà positifs au premier test. Les paramètres de dépendance peuvent alors être estimés si les caractéristiques « classiques » des tests sont connues avec précision, à l'aide des relations 1) et 2) (Annexe). Ce travail nécessite de collecter ces sérums et devrait revenir à un laboratoire de référence.

Une analyse statistique des combinaisons de résultats individuels observés sur le terrain permet une approche complémentaire à ces observations. En effet, à partir de plusieurs populations de prévalence *a priori* différentes (par exemple, des départements plus ou moins avancés dans le programme d'élimination), utilisant les mêmes réactifs et les mêmes protocoles d'utilisation de ces réactifs, il est possible d'approcher les caractéristiques des réactifs et leurs dépendances par la méthode du maximum de vraisemblance. Par exemple, pour estimer les six paramètres (deux sensibilités, deux spécificités et deux facteurs de dépendance) de deux réactifs, il suffirait d'au moins sept départements qui utiliseraient le même protocole de dépistage, à savoir un premier dépistage avec un test complété par un second test de confirmation appliqué aux individus révélés positifs au premier test. Ces informations sont *a priori* disponibles dans des laboratoires et pourraient être utilisées à cette fin également par un laboratoire de référence.

VI - CONCLUSION

L'approche théorique menée dans cette étude a montré l'influence que pouvait exercer la dépendance des tests sur la valeur prédictive des résultats fournis par l'utilisation de deux tests. L'application numérique a permis de mesurer cette influence dans le cas particulier du dépistage de l'IBR. L'observation la plus significative est la chute importante de la valeur prédictive positive d'un résultat positif aux deux tests dans une situation de très faible prévalence (< 0,1%), lorsque ces deux tests présentent une dépendance de leur spécificité. Ce constat théorique semble s'observer également sur le terrain et doit interroger les utilisateurs des tests sur l'interprétation des résultats en fonction de la dépendance des tests utilisés. Le recours à des tests de confirmation sur la seule base d'une meilleure spécificité n'est pas toujours le plus indiqué s'il existe une dépendance forte avec le premier

test. L'approche quantitative proposée dans cette étude permet d'apprécier le niveau de fiabilité fourni par une combinaison de réactifs aux caractéristiques et à la dépendance connus. En pratique, l'évaluation quantitative des dépendances n'est pas facile, mais peut faire l'objet de différentes approches complémentaires, qui pourraient être menées par des laboratoires de référence. Ceux-ci concentreraient les sérums à problèmes et les observations statistiques des différents réactifs utilisés. Un schéma de « réactovigilance » pourrait ainsi se mettre en place pour assurer un suivi des réactifs et des réactions, et déceler le plus précocement possible l'apparition de problèmes soit au niveau du terrain (apparition de réactions croisées), soit au niveau des réactifs. Une telle approche mériterait d'être discutée et approfondie avec tous les acteurs concernés.

BIBLIOGRAPHIE

Bonnardel P. - Le coefficient Kappa. In site <http://kappa.chez.tiscali.fr>, 1995, page « Limites du test ».

Gardner I.A. et Greiner M. - Advanced Methods for Test Validation and Interpretation in Veterinary Medicine, 78 pages, site www.berlin-info.de, 1999, 44-51.

Hui S.L., Walter S.D. - Estimating the error rates of diagnostic tests. *Biometrics*, 1980, **36**, 167-171.

Petit E. - Etude sur la procédure "résultats aberrants" IBR proposée par l'A.C.E.R.S.A., Ed. A.E.E.M.A., *Epidémiologie et Santé Animale*, 2002, **42**, 133-150.

Petit E. et Perrin M. - Estimation des caractéristiques des réactifs sérologiques IBR à partir de la comparaison de cinq réactifs menée par l'A.F.S.S.A.,

Epidémiologie et Santé Animale, 2003, **43**, 87-96.

Pouillot R., Gerbier G., Garin-Bastuji B. - Synthèse des travaux épidémiologiques sur les réactions sérologiques faussement positives en brucellose bovine, *Epidémiologie et Santé Animale*, 1999, **35**, 1-10.

Pouillot R. - Les tests multiples : effets de la dépendance conditionnelle entre tests sur l'interprétation. In : Méthodes Avancées de Validation et d'Interprétation des Tests de Diagnostic, CDrom des Ateliers de l'AEEMA du 16 mai 2001, 2001.

Toma B. et coll. - Le dépistage des maladies infectieuses animales. In : *Epidémiologie appliquée à la lutte contre les maladies transmissibles majeures*. A.E.E.M.A. (Ed), 2001, 2^{ème} Ed, 696 p.

ANNEXE : DEVELOPPEMENTS MATHÉMATIQUES

1. RAPPELS SUR SENSIBILITE, SPECIFICITE, VALEURS PREDICTIVES
 [Toma *et al.*, 2001]

Pour un test et une maladie donnés, le test peut donner une réponse positive ou négative sur des individus réellement infectés ou réellement sains.

La **sensibilité** (Se) du test correspond à la probabilité pour un individu réellement infecté d'être donné positif par le test (= P(T+ | infecté))

La **spécificité** (Sp) du test correspond à la probabilité pour un individu réellement sain d'être donné négatif par le test (= P(T- | indemne)).

La **valeur prédictive positive** (VPP) correspond à la probabilité d'être réellement infecté alors que le test est positif (= P(infecté | T+)).

La **valeur prédictive négative** (VPN) correspond à la probabilité d'être réellement sain alors que le test est négatif (= P(indemne | T-)).

Ces deux dernières valeurs dépendent à la fois des caractéristiques (sensibilité et spécificité) du test, mais également de la prévalence (F) réelle de la maladie.

Un test idéal doit voir tendre ces valeurs vers 1 (= 100%).

Le tableau III donne les relations entre les probabilités d'obtenir les différentes configurations (test positif ou négatif) et ces différentes valeurs, selon la sous-population considérée (infectée ou non infectée).

Les VPP et VPN s'expriment selon :

$$VPP = F \text{ Se} / [F \text{ Se} + (1 - F) (1 - \text{Sp})]$$

$$VPN = (1 - F) \times \text{Sp} / [F (1 - \text{Se}) + (1 - F) \text{Sp}]$$

Tableau III

Probabilités d'obtenir un résultat, en fonction des caractéristiques du test (Se : Sensibilité, Sp : Spécificité), de la prévalence dans la population (F) et de la sous-population considérée (infectés et non infectés)

	Infectés	Non infectés
Test positif	F x Se	(1 - F) x (1 - Sp)
Test négatif	F x (1 - Se)	(1 - F) x Sp
Total	F	1 - F

2. COMBINAISON DE DEUX TESTS

Classiquement, pour estimer la probabilité d'obtention des différentes combinaisons, on utilise les relations :

$$P(T1+ \text{ et } T2+ | \text{ infecté}) = \text{Se}_1 \times \text{Se}_2$$

$$P(T1- \text{ et } T2- | \text{ indemne}) = \text{Sp}_1 \times \text{Sp}_2$$

Ces relations supposent une indépendance des tests, ce qui se concrétise en terme probabiliste par l'utilisation des relations :

$$P(T1+ | T2+, \text{ infecté}) = P(T1+ | T2-, \text{ infecté}) = P(T1+, \text{ infecté}) = \text{Se}_1$$

$$P(T2+ | T1+, \text{ infecté}) = P(T2+ | T1-, \text{ infecté}) = P(T2+, \text{ infecté}) = \text{Se}_2,$$

$$P(T1- | T2+, \text{ indemne}) = P(T1- | T2-, \text{ indemne}) = P(T1-, \text{ indemne}) = \text{Sp}_1$$

$$P(T2- | T1+, \text{ indemne}) = P(T2- | T1-, \text{ indemne}) = P(T2-, \text{ indemne}) = \text{Sp}_2,$$

Or, la probabilité de réponse du second test peut être différente selon la réponse du premier test, et ceci dans les sous-populations infectées ou non infectées : les réponses peuvent être liées par une relation de dépendance de leurs sensibilités et de leurs spécificités.

La combinaison de deux tests T1 et T2 nécessite donc la connaissance de deux paramètres complémentaires qui traduisent la relation de dépendance qui peut exister entre

les deux tests T1 et T2, tant de la sensibilité que de la spécificité.

Ces paramètres peuvent être représentés en terme probabiliste, d'une part, par un facteur multiplicatif (Kse) de la sensibilité du test 2, alors que le test 1 s'est avéré positif sur un animal infecté et, d'autre part, par un autre facteur multiplicatif (Ksp) de la spécificité du test 2 alors que le test s'est avéré négatif sur un animal non infecté.

Alors, la probabilité qu'un test T2 soit positif alors que T1 est positif en population infectée vaut $Kse \times Se2$, où $Se2$ représente la sensibilité du test T2 :

$$P(T2+ | T1+, Inf) = Kse P(T2+, Inf) = Kse \times Se2,$$

De même, la probabilité qu'un test T2 soit négatif alors que T1 est négatif en population non infectée vaut $Ksp \times Sp2$, où $Sp2$ représente la spécificité du test T2 :

$$P(T2- | T1-, non infecté) = Ksp P(T2-, non infecté) = Ksp \times Sp2$$

Le tableau IV fournit les probabilités d'obtenir chaque combinaison de résultats avec deux tests, en fonction de leurs caractéristiques et de leurs dépendances.

3. PROPRIETES DES DEPENDANCES

Par construction et pour préserver des probabilités inférieures à 1, Kse (respectivement Ksp) doit être inférieur ou égal au minimum des valeurs $1 / Se1$ et $1 / Se2$ (respectivement $1 / Sp1$ et $1 / Sp2$).

A noter : dans une démarche courante, on utilise un test plus sensible et souvent moins spécifique en première intention et un test plus spécifique en seconde intention. On a donc $Se1 > Se2$ et $Sp1 < Sp2$. Les conditions s'écrivent alors : $Kse \leq 1 / Se1$ et $Ksp \leq 1 / Sp2$.

Tableau IV

Probabilités d'obtenir chaque combinaison de résultats avec deux tests, en fonction de leurs caractéristiques et de leurs dépendances

Combinaison des résultats	Infectés	Non infectés
pp = T1 + , T2 +	F Se1 Se2 Kse	(1 - F) (1 - Sp1 - Sp2 + Ksp Sp1 Sp2)
pn = T1 + , T2 -	F Se1 (1 - Se2 Kse)	(1 - F) (1 - Sp1 Ksp) Sp2
np = T1 - , T2 +	F (1 - Se1 Kse) Se2	(1 - F) (1 - Sp2 Ksp) Sp1
nn = T1 - , T2 -	F (1 - Se1 - Se2 + Se1 Se2 Kse)	(1 - F) Sp1 Sp2 Ksp
Total	F	1 - F

On peut également approcher cette notion par la proportion de faux négatifs au premier test non décelés au second (pour la dépendance entre sensibilités) et par la proportion de faux positifs au premier test confirmés positifs par le second (pour la dépendance entre spécificités). Ces valeurs ont le mérite d'être plus faciles à comprendre et peuvent être estimées expérimentalement.

Une relation existe entre ces proportions et les facteurs de dépendance. On l'obtient en développant sous forme de probabilités conditionnelles l'expression de ces proportions.

Ainsi :

1. sur des animaux infectés, la proportion de faux négatifs au premier test non décelés au second s'écrit :

$$\text{Relation 1 : } P(T2- | T1-, Infecté) = (1 - Se1 - Se2) / (1 - Se1) + [Se1 Se2 / (1 - Se1)] Kse$$

2. Sur des animaux indemnes, la proportion de faux positifs au premier test confirmés positifs par le second s'écrit :

$$\text{Relation 2 : } P(T2+ | T1+, indemne) = (1 - Sp1 - Sp2) / (1 - Sp1) + [Sp1 Sp2 / (1 - Sp1)] Ksp$$

où T2+ (resp. T1+) est l'obtention d'un résultat positif au second test (resp. au premier test).

Dans ces conditions, les exigences posées sur les maxima de Kse et Ksp permettent d'établir une limite pour les valeurs ci-dessus. Ainsi :

1. sur des animaux malades,

$$1 - Se2 \leq P(T2-|T1-) \leq 1$$

2. sur des animaux non malades,

$$1 - Sp2 \leq P(T2+|T1+) \leq (1 - Sp2) / (1 - Sp1)$$

4. VALEURS PREDICTIVES D'UNE COMBINAISON DE DEUX TESTS

Dans l'usage courant d'une procédure de contrôle des résultats positifs, on s'intéressera surtout à la VPP du résultat « pp » (confirmation) et à la VPN du résultat « pn » (infirmité).

Dans ce cas la VPP d'une combinaison « pp » (positif aux deux tests) est égale au rapport du nombre d'infectés positifs aux deux tests sur l'ensemble des positifs aux deux tests, soit :

$$VPP(pp) = F Se1 Se2 Kse / (F Se1 Se2 Kse + (1 - F) (1 - Sp1 - Sp2 + Ksp Sp1 Sp2))$$

$$= 1 / [1 + (1 - F) (1 - Sp1 - Sp2 + Ksp Sp1 Sp2) / (F Se1 Se2 Kse)]$$

De même, la VPN d'une combinaison « pn » (positif au 1^{er} test infirmé au 2^{ème} test) est égale au rapport du nombre de non infectés classés « pn » sur l'ensemble des individus classés « pn », soit :

$$VPN(pn) = (1 - F) (1 - Sp1 Ksp) Sp2 / ((1 - F) (1 - Sp1 Ksp) Sp2 + F Se1 (1 - Se2 Kse))$$

$$= 1 / [1 + F Se1 (1 - Se2 Kse) / (1 - F) (1 - Sp1 Ksp) Sp2].$$

