

# Évaluation d'une procédure diagnostique

Prof Emmanuel OGER

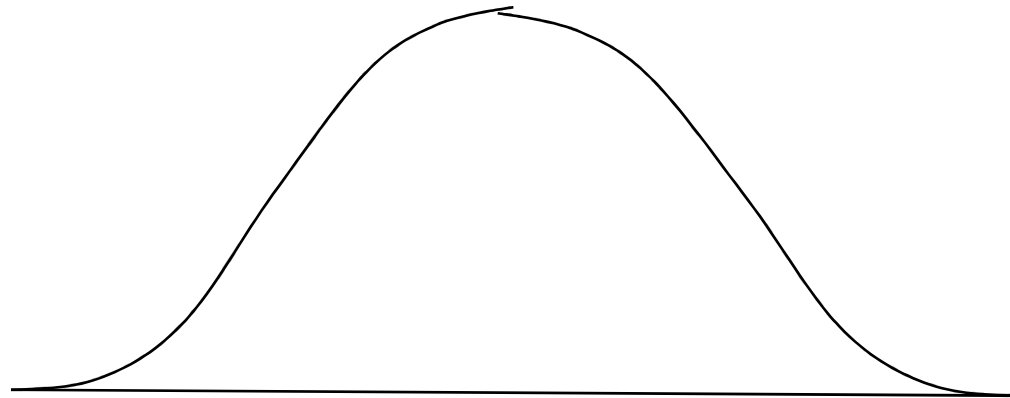
Centre Régional de Pharmacovigilance,  
de Pharmaco-épidémiologie  
et d'Information sur le médicament

# Procédure(s) diagnostique(s)

- Qu'entend-on par procédure diagnostique?
  - Symptôme
  - Signe d'examen
  - Résultat d'examen complémentaire
- Quel est le problème?
  - Ambiguïté des données : notion de concordance, reproductibilité
  - Variabilité des données (variabilité biologique et erreur de mesure)
  - Incertitude sur la relation résultat/diagnostic (modification de la vraisemblance du diagnostic...)

# Résultat de la procédure (1)

- Résultat binaire
  - Échographie-Doppler veineux
- Résultat qualitatif ordonné
  - Scintigraphie pulmonaire de ventilation/perfusion
- Résultat quantitatif continu
  - D-dimères



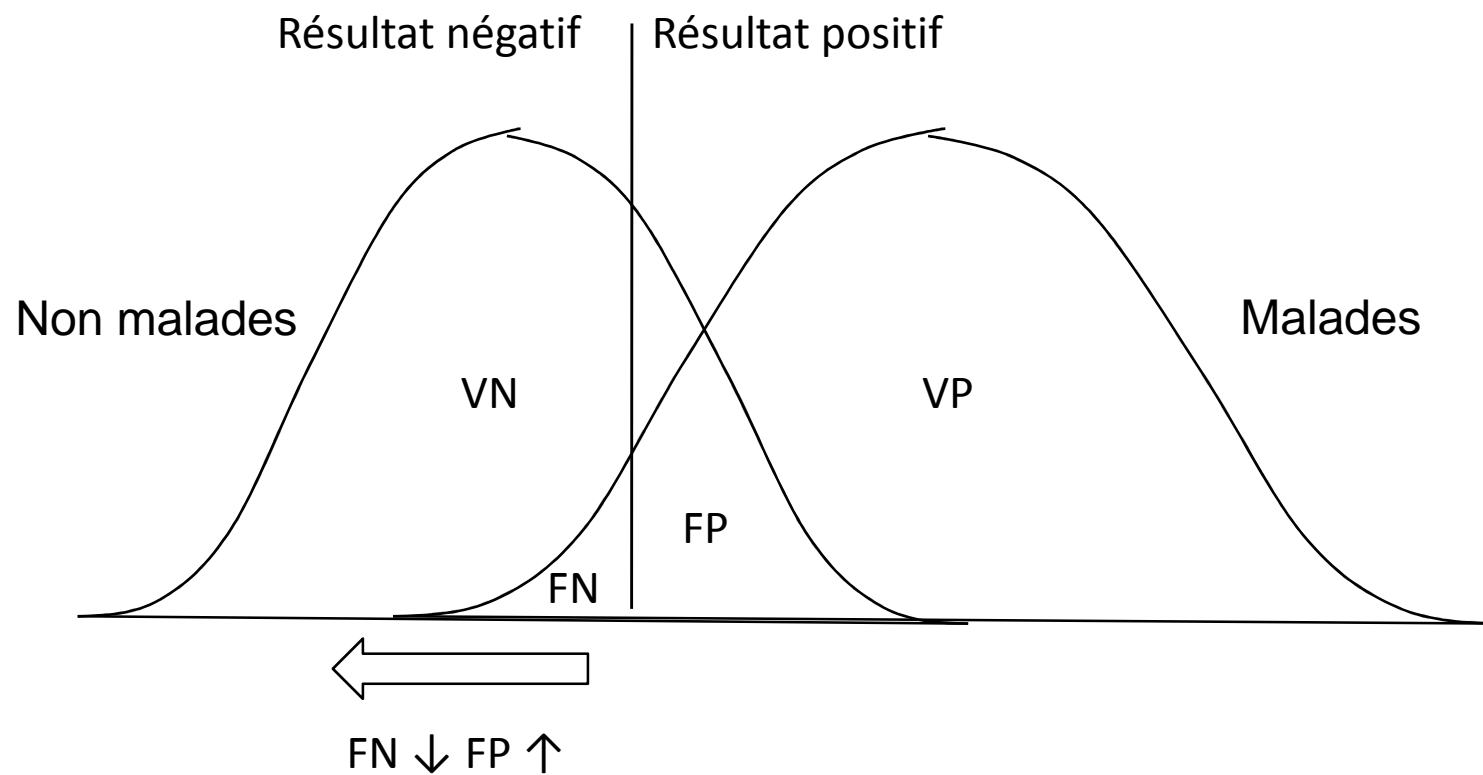
# Résultat de la procédure (2)

- Résultat binaire

	<b>Maladie présente</b>	<b>Maladie absente</b>
Test positif	VP	FP
Test négatif	FN	VN

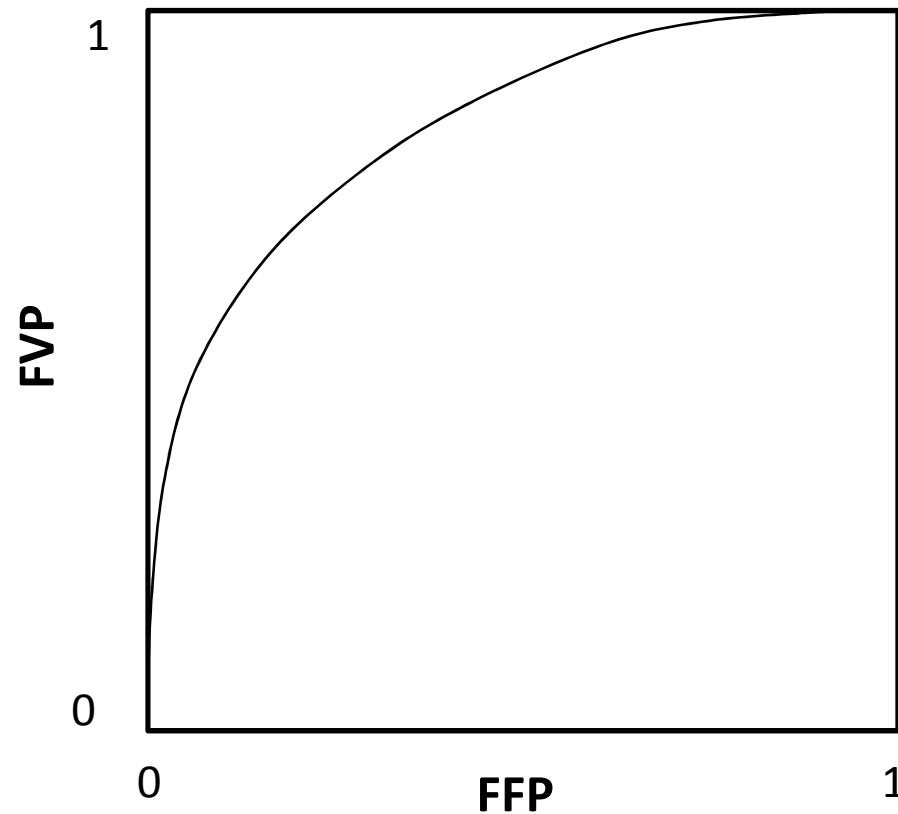
# Résultat de la procédure (3)

- Résultat quantitatif continu



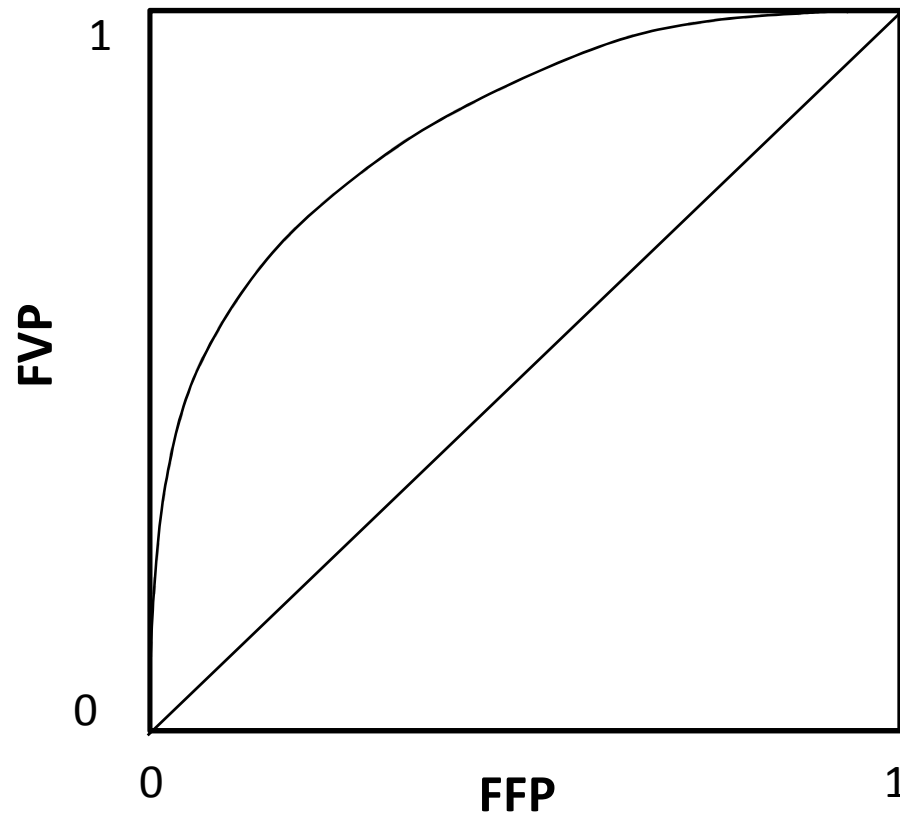
# Résultat de la procédure (4)

- Résultat quantitatif continu : choix du seuil
  - Courbe ROC (*Receiver Operating Characteristics*)



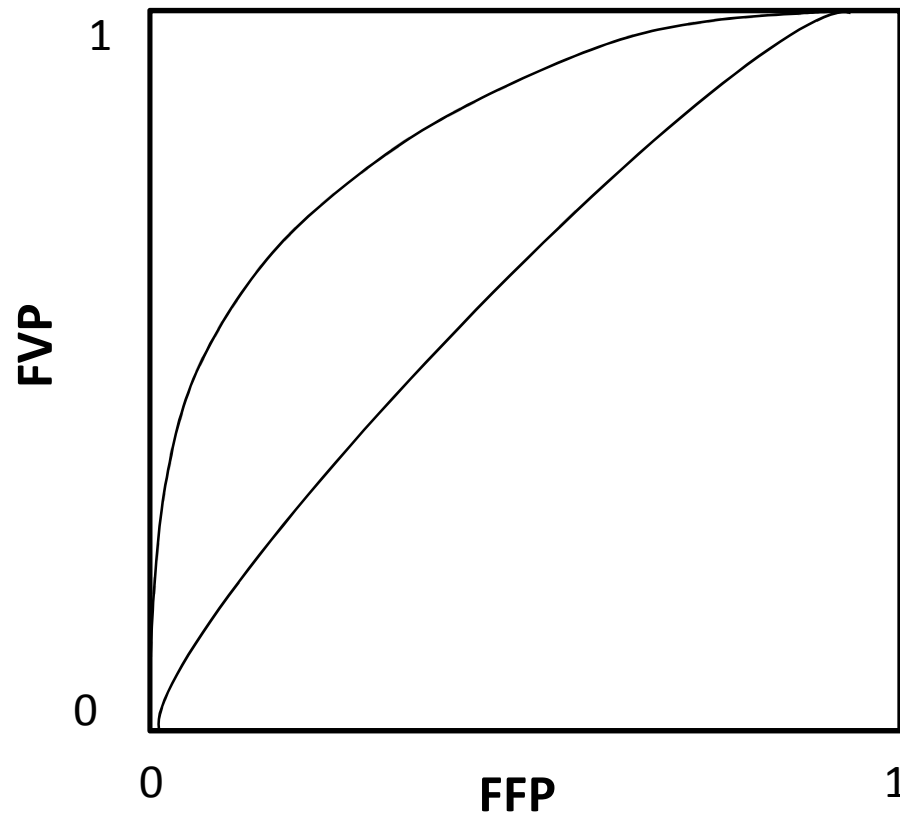
# Résultat de la procédure (5)

- Autre intérêt de la courbe ROC (1)
  - Évaluer la performance global d'un test (*AUC*)



# Résultat de la procédure (6)

- Autre intérêt de la courbe ROC (2)
  - Comparer les performances de plusieurs test (*AUC*)



# Les indices de performances (1)

- La sensibilité est la probabilité d'avoir un test positif quand on est malade

$$SE = VP / (VP + FN)$$

- La spécificité est la probabilité d'avoir un test négatif quand on n'est pas malade

$$SP = VN / (VN + FP)$$

- Qualités intrinsèques du test

# Les indices de performances (2)

- Ces indices sont constants lorsqu'il y a une relation causale entre le nouveau test et le test de référence (vrai test diagnostique).
- La constance est théorique!
- La variation est en grande partie due aux biais de sélection, d'interprétation, de vérification et à l'imperfection du test de référence!

# Les indices de performances (3)

- La valeur prédictive positive est la probabilité d'être malade quand le test est positif

$$VPP = VP / (VP + FP)$$

- La valeur prédictive négative est celle de ne pas être malade quand le test est négatif

$$VPN = VN / (VN + FN)$$

- Valeurs d'usage du test

# Les indices de performances (4)

- Les valeurs prédictives dépendent de la probabilité a priori de la maladie (prévalence)

$$VPP = SE \times P / (SE \times P) + (1 - SP) \times (1 - P)$$

$$VPN = SP \times (1 - P) / SP \times (1 - P) + (1 - SE) \times P$$

- Valeurs d'usage du test varient donc selon le contexte étudié

quand  $P \uparrow$  alors  $VPP \uparrow$  et  $VPN \downarrow$

Si  $P$  constant,      plus  $SP \uparrow$  , plus  $VPP \uparrow$   
plus  $SE \uparrow$  , plus  $VPN \uparrow$

# Les indices de performances (5)

- Les rapports de vraisemblance (*Likelihood Ratio*, LR) estiment le rapport entre la probabilité d'avoir un test positif (ou négatif) chez les sujets malades à celle d'avoir un test positif (ou négatif) chez les sujets sains.

$$LR(+) = SE / 1 - SP$$

$$LR(-) = (1 - SE) / SP$$

# Les indices de performances (6)

- Le rapport de vraisemblance

est le rapport de la probabilité d'avoir un résultat R du test chez les sujets malades sur la probabilité d'avoir ce même résultat R chez les sujets sains.

$$LR = \Pr (T = R \mid M = 1) / \Pr (T = R \mid M = 0)$$

- Généralisation au test non binaire

- Signification connue = risque relatif

LR (R) = 5 : un résultat R a 5 fois plus de chance de provenir d'une population de malade que d'une population de non malade.

# Les indices de performances (7)

- Le rapport de vraisemblance : utilisation en pratique

Pr pré-test → Pr post-test

Odds post-test = LR x odds pré-test

$$\text{Odds} = p / (1 - p)$$

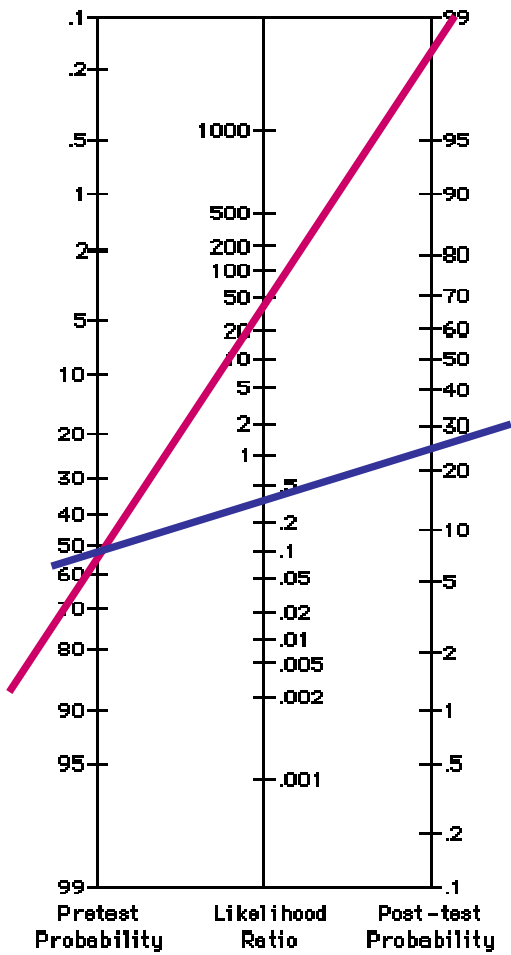
Nomogramme de Fagan

Intérêt pratique si LR(+) élevé (> 10)

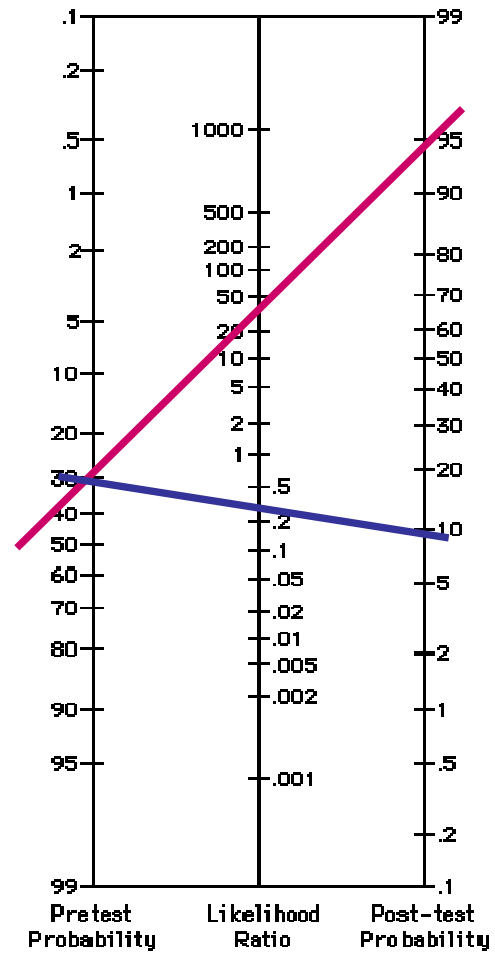
ou LR(-) faible (< 0,1)

# Nomogramme de Fagan

Ex1: prévalence = 60%



Ex2: prévalence = 30%



**AngioIRM:**  
 RV+ = 38  
 RV- = 0,3

# Les indices de performances (8)

- Intervalles de confiance
  - Estimations sur des échantillons des « vraies valeurs »
  - Exprimer la précision des estimations (IC95%)  
méthode exacte ou asymptotique  
région à 95% pour une paire d'estimateurs (SE, SP) plutôt  
que des intervalles individuels!

# Reproductibilité (1)

- Test qualitatif : concordance intra-observateur et inter-observateurs
  - Coefficient Kappa de Cohen

$$\kappa = \frac{\% \text{ concordance observée} - \% \text{ concordance due au hasard}}{1 - \% \text{ concordance due au hasard}}$$

Dépend de la prévalence de diagnostic positif !

# Reproductibilité (2)

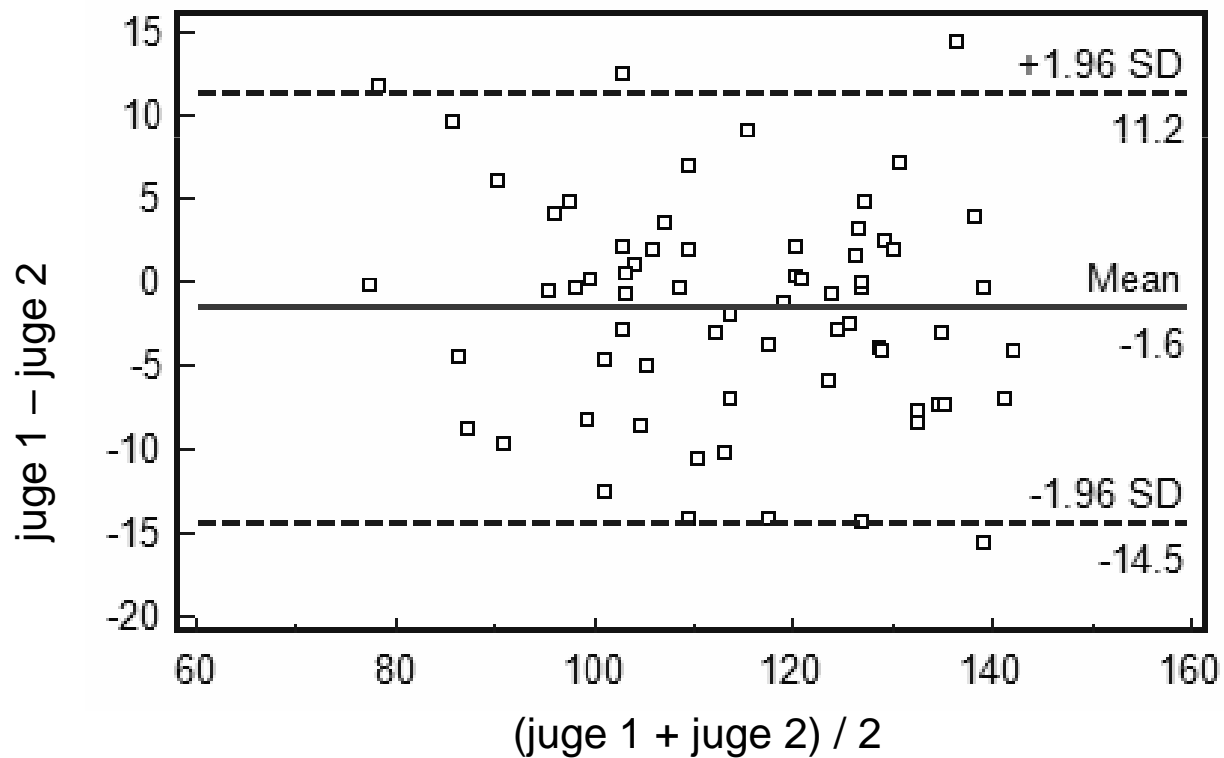
- Test quantitatif (1)
  - Limite du coefficient de corrélation !
  - Coefficient de corrélation intra-classe ++

$$CCI = \frac{\text{variation inter-patients}}{\text{variation inter-patients} + \text{variance inter-juge} + \text{variance résiduelle}}$$

- Graphique de Bland et Altman

# Reproductibilité (3)

- Test quantitatif (2)
  - Graphique de Bland et Altman



# Les phases de développement

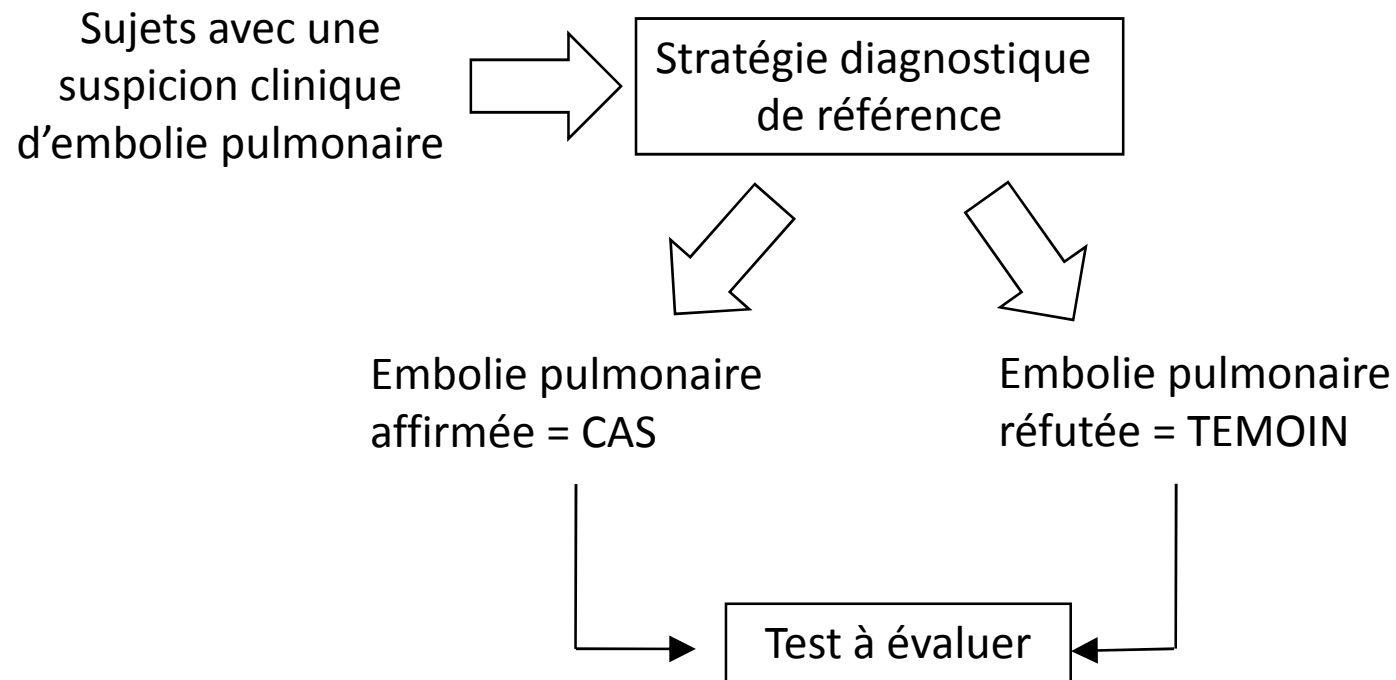
Phase	Description	Objectif	Méthode
1	Investigation exploratoire	Identifier le potentiel et le champ d'application, la reproductibilité	Cas-témoin, large contexte et variétés des malades, contrôles avec maladies proches, échantillonnage aisé
2	Validation rétrospective	Valeurs acceptables de FVP et FFP	Cas-témoin, évaluateurs experts
3	Affinement	Critères de positivité (ROC), impact des covariables	Cas-témoins de grande taille
4	Application	VPP, nature des cas détectés faux positifs en pratique, comparaison de tests	Étude transversale prospective Cohorte
5	Impact	Coût, qualité de vie, mortalité	Essai randomisé / pratique de référence

# Question étudiée ?

- Évaluer la reproductibilité d'une procédure
- Évaluer les performances d'une nouvelle procédure par rapport au test de référence
- Comparer les performances de plusieurs procédures
- Comparer les performances d'une procédure sur plusieurs population

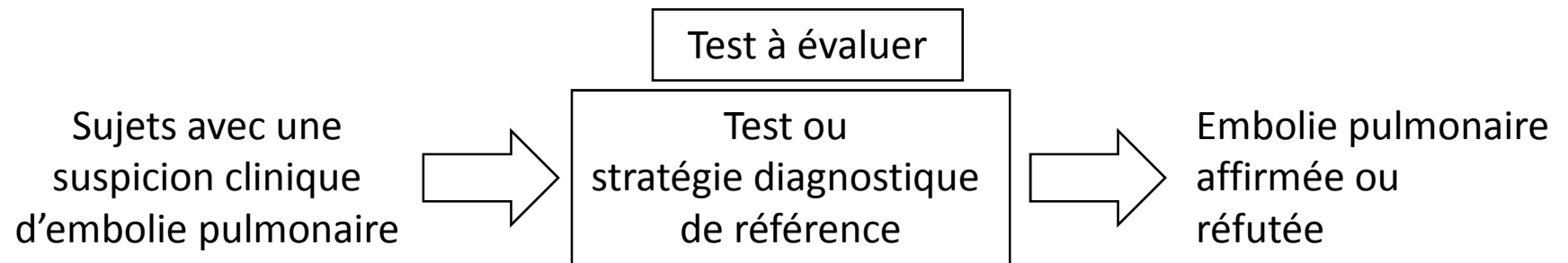
# Méthodologie (1)

- Schéma d'étude (étude cas-témoin)



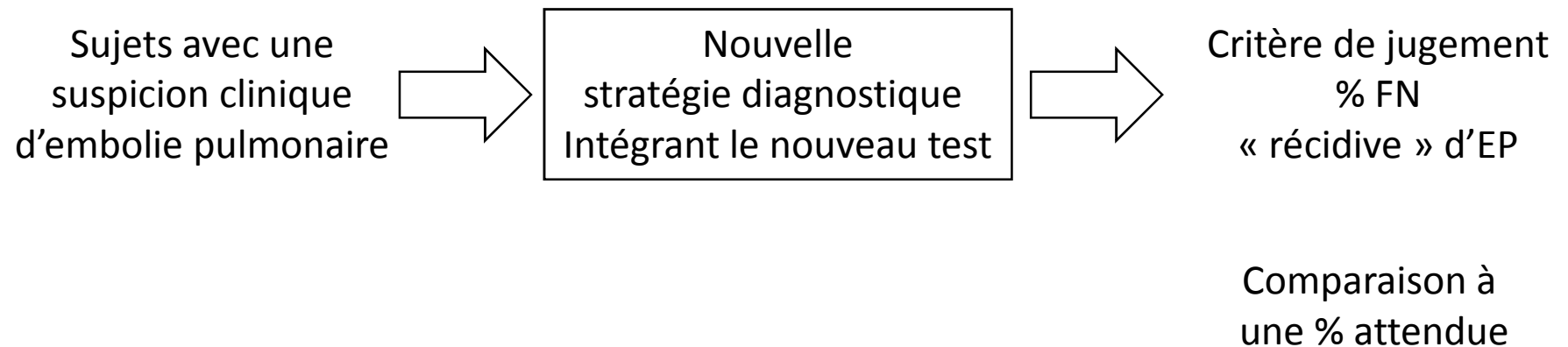
# Méthodologie (2)

- Schéma d'étude (étude prospective)



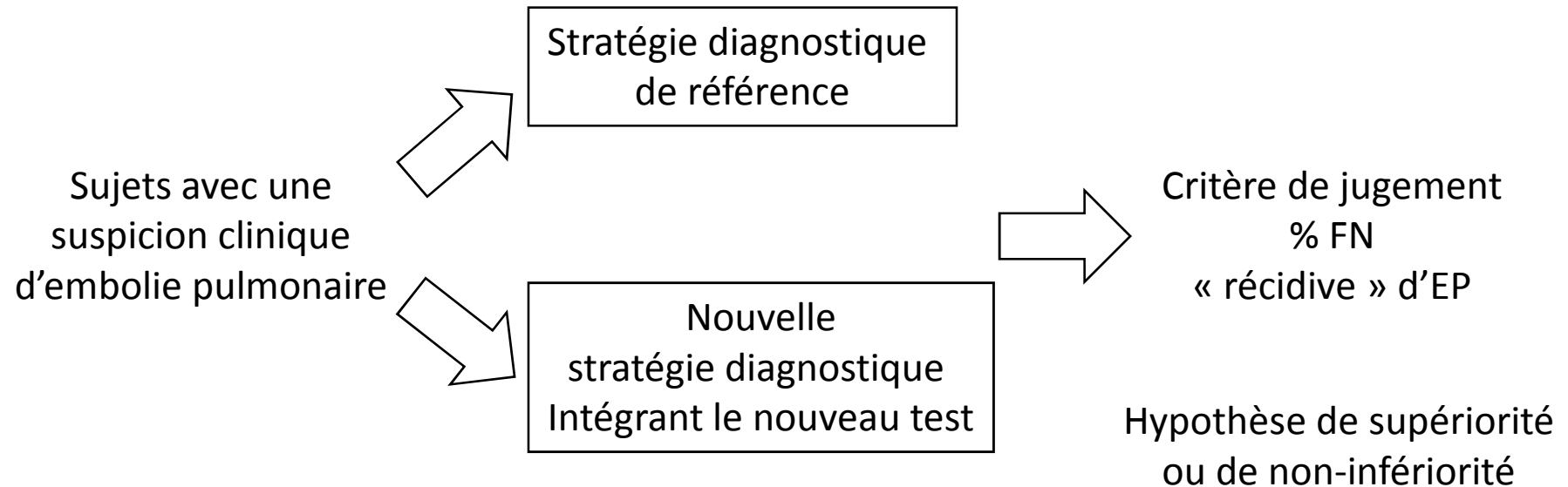
# Méthodologie (3)

- Schéma d'étude (étude prospective)



# Méthodologie (4)

- Schéma d'étude (essai randomisé)



# Méthodologie (5)

- Population

- La population incluse doit être représentative des sujets auxquels on souhaite appliquer le test in fine : sujets suspects de la maladie, à différents stades de gravité et variété de diagnostics différentiels.
- Inclusion consécutive des sujets.
- Modalité et contexte de la sélection...

# Méthodologie (6)

- Procédures diagnostiques (1)
  - Le test admis comme référence doit être justifié et décrit de manière précise.
  - Toutes les techniques utilisées (test de référence et le test étudié), la procédure de leur réalisation doivent être présentées : spécifications techniques, unité de mesure, expérience de l'expérimentateur, moment où les tests sont réalisés...
  - Tous les sujets ont les 2 tests, interprétés de façon indépendante!

# Méthodologie (7)

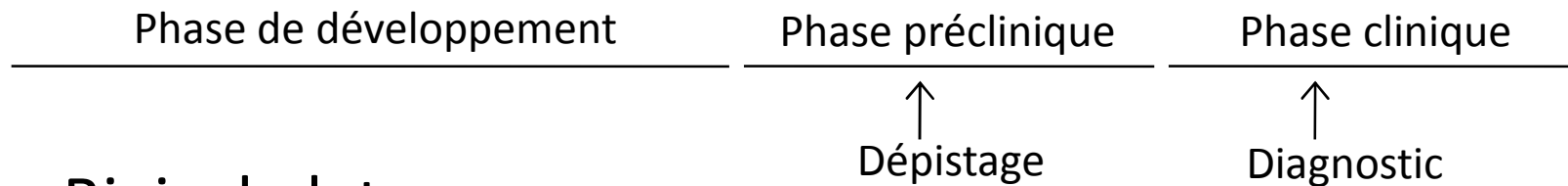
- Procédures diagnostiques (2)
  - Le seuil de positivité doit être justifié et décrit . Si un seuil ne peut pas être défini a priori, l'étude de seuil avec une courbe ROC doit être prévue.
  - La concordance entre plusieurs observateurs doit être décrite si elle n'a pas été évaluée lors des phases plus précoces.
  - Le test de référence n'est jamais parfait et entraîne donc des erreurs de classification de la maladie.

# Les biais potentiels (1)

- Biais de sélection/spectre biaisé/extrapolation  
Patients non représentatifs...
- Biais de vérification  
Vérification complète des résultats de tous les patients testés avec le test de référence
- Biais d'interprétation  
Interprétation doit être réalisée en insu par deux investigateurs différents
- Biais de classement (test de référence imparfait)
- Test non concluant, ininterprétable

# Les biais potentiels (2)

- Particularités des méthodes de dépistage



- Biais de latence

Le diagnostic précoce (en phase préclinique) indique un bénéfice erroné même si un traitement précoce est totalement inefficace.

- Biais de sélection

Les patients qui acceptent le dépistage sont différents...

# Les biais potentiels (3)

- Biais lié à l'évolution

La phase préclinique étant plus longue dans les maladies à évolution lente, la probabilité de faire un diagnostic (dépistage) à ce stade est plus élevée, et la proportion de ces patients sera plus importante parmi les sujets dépistés....

- Biais de sur-diagnostic

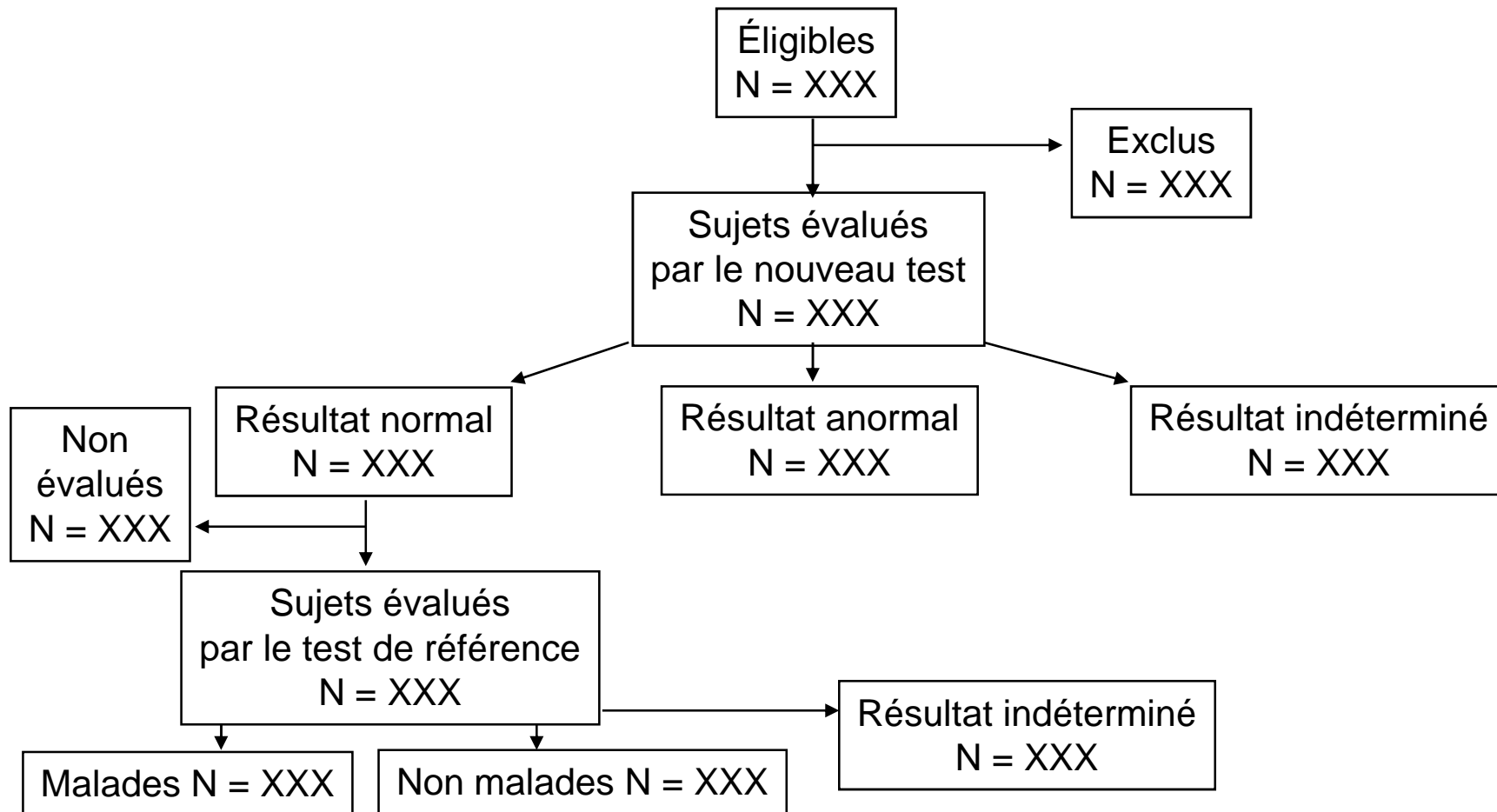
Dépistage de sujet avec une maladie ayant un potentiel à régressé spontanément...

- Biais de détection

Suivi insuffisant pour détecter tous les évènements.

# Présentation des résultats (1)

- Graphique de flux



# Les éléments clés

- Résultats valides (phase 2-4)?

  - Tous les sujets ont les deux tests (dont le test de référence avec une description de la procédure et une définition précise et reconnue de la normalité)

  - Interprétation indépendante

- Résultats utiles?

  - Bonne reproductibilité.

  - Population représentative de la cible.

  - Impact du résultat du test sur la décision (LR)

  - Bénéfice réel pour le patient (phase 5)