

# Evaluation d'une méthode diagnostique (1)

# Plan du cours

- Évaluation d'une méthode diagnostique
  - Phases
    - I: validité, fiabilité: coefficient Kappa (K)
    - II: intrinsèque: Se, Sp, LR+, LR-
      - Courbe ROC (voir le cours prochain)
    - III: extrinsèque: VPP, VPN
    - IV: cout – bénéfice
- Exemples des études
- Exemples des exercices

# Évaluation d'une méthode diagnostique

- **Définition:** des études qui permet l'évaluation
  - des **avantages** (les **qualités diagnostiques**: Se, Sp, ...) et **inconvénients**
    - offerts par un nouveau test de diagnostic, ou plusieurs
      - par rapport aux méthodes standard de diagnostic d'une maladie
- **Utilité:**
  - réaliser le **diagnostique**
  - identifier la gravité/stade d'un maladie afin de **choisir le meilleur traitement**
- **Exemples** des méthodes diagnostiques: signes (fièvre), symptômes (douleur des dents), imagerie médicale (radiographie du crane, tomодensitométrie, radiographie dentaire), examen anatomopathologique, biochimie du sang/ urine... (glycémie, glycosurie, ...)

# Évaluation d'une méthode diagnostique

- On doit préciser dans le titre, fin d'Introduction (but/objective), résumé, chapitre méthodes de l'article/thèse qu'on étudie les qualités diagnostiques.
- Recherché **analytique** – on compare deux tests diagnostiques

# Phase I:

Le **but** - de déterminer:

- **La Validité** de la méthode (qualité technique)
- **La Fiabilité** de la méthode (**ex.** la lecture d'une radiographie du crane):
  - **Reproductibilité inter-observationnelle/inter-opérateur**
    - Définition: désigne les différences entre les observations du même phénomène par les différents chercheurs
  - **Reproductibilité intra-observationnelle/intra-opérateur**
    - Définition: désigne les différences entre observations du même phénomène par un même chercheur.
  - Mesure avec le **coefficient/test Kappa de Cohen** de reproductibilité
    - $>0,8$ : parfait reproductibilité
    - $0,6 - 0,8$ : bonne reproductibilité
    - $0,4 - 0,6$ : modérée reproductibilité
    - $0,2 - 0,4$ : reproductibilité réduite
    - $0 - 0,2$ : reproductibilité presque absente
    - $< 0$ : mauvaise reproductibilité

## Phase II: qualités intrinsèques du test

- Le nouveau test est appliqué chez des patients pour lesquels le **statut malade ou non-malade** est clairement précisé avec un **test de référence** = "Norme d'or" ou "Gold standard", ou « Standard d'or".

		Test standard		Total
		Maladie présente	Maladie absente	
Nouveau test	Test positif	VP a Vrai Positif	FP b Faux Positif	a+b
	Test négatif	FN c Faux Négatif	VN d Vrai Négatif	c+d
Total		a+c	b+d	a+b+c+d

- Les résultats sont généralisables à la population seulement si il est représentatif à la population cible

# Exemple...

European Annals of Otorhinolaryngology, Head and Neck diseases 136 (2019) 361–366



Available online at  
**ScienceDirect**  
www.sciencedirect.com

Elsevier Masson France  
**EM|consulte**  
www.em-consulte.com/en



<https://www.sciencedirect.com/science/article/pii/S187972961930095X>

Original article

## Radiological assessment of mandibular invasion in squamous cell carcinoma of the oral cavity and oropharynx



S. Bouhir<sup>a</sup>, G. Mortuaire<sup>a</sup>, F. Dubrulle-Berthelot<sup>b</sup>, X. Leroy<sup>c</sup>, V. Deken-Delannoy<sup>d</sup>,  
B. Rysman<sup>a</sup>, D. Chevalier<sup>a</sup>, F. Mouawad<sup>a,e,\*</sup>

### ABSTRACT

**Background:** Preoperative assessment of mandibular bone invasion in squamous cell carcinoma of the oral cavity and oropharynx is crucial for optimizing bone resection. The principal aim of this study was to evaluate the diagnostic value of CT and MR imaging for the diagnosis of mandibular bone invasion compared to the histological reference. In addition, we assessed the survival impact of bone invasion.

**Patients and methods:** A single-center retrospective study included all consecutive patients treated by mandibular bone interruption for squamous cell carcinoma of the oral cavity and/or oropharynx.

**Table 2**

Diagnostic values of CT, MRI and CT + MRI versus histology.

Imaging	Se 95% CI	Sp 95% CI	PPV95% CI	NPV95% CI	Kappa concordance95% CI
CT	0.70 0.49–0.86	0.71 0.53–0.85	0.66 0.45–0.82	0.76 0.57–0.89	0.69 0.49–0.89 $P < 0.0001$
MRI	0.83 0.68–0.98	0.50 0.31–0.68	0.59 0.40–0.75	0.78 0.52–0.93	
CT + MRI	0.83 0.69–0.96	0.62 0.46–0.77	0.62 0.46–0.77	0.83 0.69–0.96	–

CT: computed tomography; MRI: magnetic resonance imaging; Se: sensitivity; Sp: specificity; PPV: positive predictive value; NPV: negative predictive value.

# Phase II: Indicateurs évaluable

## Definitions, formules et interpretations statistique de la Sensibilité, Specificité:

- **Sensibilité (Se) - anglais – sensitivity**

- Probabilité d'un sujet malade d'avoir un test positif ( $\approx$  Probabilité d'un sujet d'avoir un test positif sachant qu'il est malade)
  - Ex. La probabilité d'un sujet d'avoir un test Covid 19 rapide positif sachant qu'il a réellement Covid 19 est de 0.70 (70%)
- Proportion de sujets qui ont été testés positifs **parmi les malades:**

$$Se = VP / (VP + FN) = \frac{a}{a+c}$$

- **Spécificité (Sp) - anglais – specificity**

- Probabilité d'un sujet non-malade d'avoir un test négatif
  - Ex. La probabilité d'un sujet d'avoir un test Covid 19 rapide négatif sachant qu'il n'a pas réellement Covid 19 est de 0.98 (98%)
- Proportion de sujets ayant des résultats négatifs chez les non-malades:

$$Sp = VN / (VN + FP) = \frac{d}{b+d}$$

		Test standard		Total
		Maladie présente	Maladie absente	
Nouveau test	Test positif	VP a Vrai Positif	FP b Faux Positif	a+b
	Test négatif	FN c Faux Negatif	VN d Vrai Négatif	c+d
Total		a+c	b+d	a+b+c+d



## Phase II:

		Test standard		Total
		Maladie présente	Maladie absente	
Nouveau test	Test positif	VP a Vrai Positif	FP b Faux Positif	a+b
	Test négatif	FN c Faux Négatif	VN d Vrai Négatif	c+d
Total		a+c	b+d	a+b+c+d

- **Haute sensibilité** → pour exclure la maladie dans le cas où le test est négatif
- (SnOut = Sensibility rules out the disease )
  - $Se = 100\% + \text{test négatif} \rightarrow \text{certainement le patient n'a pas la maladie}$
  - **$Se = VP / (VP + FN)$**
- **Haute spécificité** → pour confirmer la maladie lorsque le test est positif
- (SpIn = Specificity rules in the disease – la spécificité confirme la maladie)
  - $Sp = 100\% + \text{test positif} \rightarrow \text{le patient a certainement la maladie}$
  - **$Sp = VN / (VN + FP)$**

## Phase II:

**Indicateurs globales** de qualité diagnostique:

- **Valeur globale du test (anglais – accuracy) =  $(VP + VN) / (VP + FP + VN + FN)$** 
  - (proportion des sujets bien classées)
  - De 0% (absente) - 100% (très bien)
  - Le pourcentage des vrai réponses du test
- **Indice Youden =  $Se + Sp - 1$** 
  - De 0 (absente) – 1 (très bien)
- Si ils ne sont proche a 100% ou 1 on ne sait pas si la Se ou la Sp est réduite ou les deux
- On recommande de regarder Se et Sp plus que précision ou l' indice Youden

		Test standard		Total
		Maladie présente	Maladie absente	
Nouveau test	Test positif	VP a Vrai Positif	FP b Faux Positif	a+b
	Test négatif	FN c Faux Négatif	VN d Vrai Négatif	c+d
Total		a+c	b+d	a+b+c+d

## Phase II:

- La sensibilité et la spécificité **ne dépendent pas** de la **prévalence** de la maladie dans l'échantillon étudié.
- Par conséquence, Se et Sp auront les mêmes valeurs, n'importe quelle soit la population à laquelle le test est appliqué.
- Un **test idéal** a la **Se et Sp = 100%**, ce qui est rare dans la pratique:
  - Plus haute la Se, plus basse la Sp, et vice-versa.

# Phase II:

- **Interprétation clinique de la sensibilité, spécificité:**

- Il n'y a pas une règle universelle. Ca dépend de la maladie à évaluer, de la gravité, des possibilités d'être traitée
- Un indicateur de diagnostic pour Se, Sp, précision/Le pourcentage des biens classées (accuracy en anglais)
  - <80% - mauvaise qualité,
  - de 80-90% - modéré
  - > 90% bonne qualité - mais dans certains cas, il doit prendre des valeurs proches de 100% - soit testé pour le virus de l'immunodéficience humaine (VIH)
  - (ex. Se = 95% signifie que le nouveau test de diagnostic a une sensibilité élevée).
  - proche de **100%** signifie un diagnostic de haute qualité

# Phase II:

## Définition, formules, interprétation statistique

du likelihood ratio = rapport de vraisemblance RVP (+LR), RVN (-LR)

		Test standard		Total
		Maladie présente	Maladie absente	
Nouveau test	Test positif	VP a Vrai Positif	FP b Faux Positif	a+b
	Test négatif	FN c Faux Négatif	VN d Vrai Négatif	c+d
Total		a+c	b+d	a+b+c+d

- Rapport de vraisemblance positif (RVP) ou

$$(+LR) = \frac{Se}{(1-Sp)} = \frac{Pr(T \text{ positif} / \text{Malade})}{1 - Pr(T \text{ négatif} / \text{non Malade})} = \frac{Pr(T \text{ positif} / \text{Malade})}{Pr(T \text{ positif} / \text{non Malade})}$$

- Anglais: **+LR – positive likelihood ratio**

- combien de fois un **résultat positif du test** est plus susceptible d'apparaître chez un sujet malade que chez un non-malade.

- $+LR = \frac{a}{b+d} \sim$  il est similaire au Risque Relatif (-> sur les colonnes)

- Rapport de vraisemblance négatif (RVN) ou  $(-LR) = \frac{1-Se}{Sp} = \frac{c}{b+d}$

- Anglais: **-LR – negative likelihood ratio**

- combien de fois un **résultat négatif du test** est plus susceptible d'apparaître chez un sujet malade que chez un non-malade.

# Probabilité pré-test / post-test d'avoir la maladie

## Nomogramme Fagan:

### Utilisation:

On fixe sur l'axe pré-test (avant le test)

la probabilité avant faire une test

On fait le test, et si le résultat du test

est positif on utilise LR +

est négative on utilise LR -

On met une point avec la valeur du LR

sur la deuxième axe (likelihood ratio)

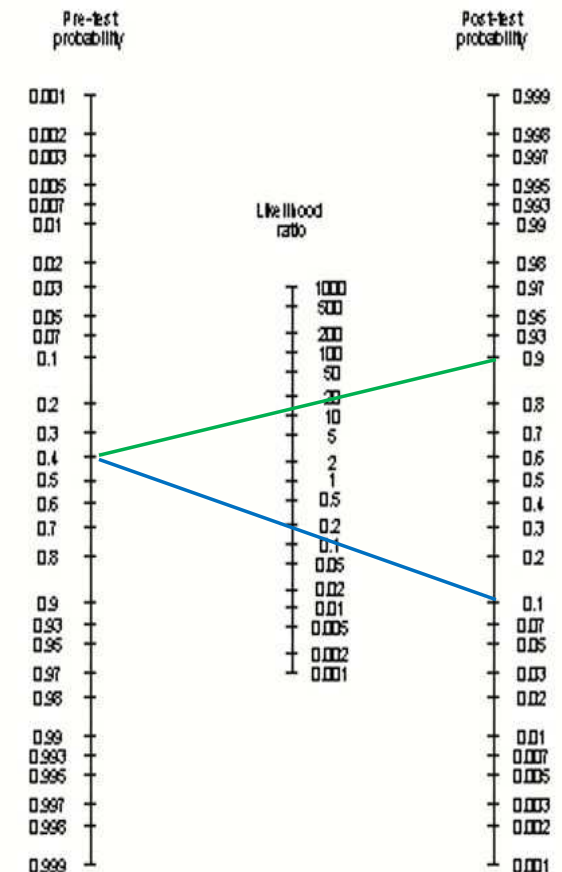
On dessine une ligne entre les deux points

vers la troisième axe et on obtiennent

la probabilité après le test

Ex. P avant=40%, test + => LR+=13 => P après = 90%

P avant=40%, test - => LR-=0.2 => P après = 10%



# Probabilité pré-test / post-test d'avoir la maladie

- **Interprétation clinique des rapports de vraisemblance:**

- certaines valeurs du LR peuvent entraîner des changements plus ou moins importante de la probabilité d'avoir une maladie. Ainsi les valeurs LR dans les gammes:
  - **+LR > 10** ou **-LR < 0,1** entraînent des changements importants et souvent concluante de la probabilité d'un diagnostic (très bon - Diagnostic fiable )
  - LR entre **5 et 10** ou entre **0,1 (=1/10)** et **0,2 (=1/0,5)** entraînent des changements modéré et utile de la probabilité d'un diagnostic (Assez bon - Diagnostic à confirmer avec d'autres tests
  - LR entre **2 et 5** ou entre **0,5(=1/2)** et **0,2(=1/5)** entraînent des changements réduits/faible mais parfois concluante/cliniquement significatif de la probabilité d'un diagnostic (Faible - Autres tests nécessaires )
  - LR entre **1 et 2** ou entre **0,5 et 1** conduit à des modifications de faible probabilité, et rarement cliniquement important (Sans utilité pour le diagnostic )
- LR = 1 – pas de changement
- Donc plus le LR+ est grand plus il permet de confirmer la maladie
- Plus le LR- est petit, plus il permet de exclure la maladie

## Phase III: les qualités extrinsèques du test

- Dans la pratique, le statut malade ou non-malade des sujets soumis au nouveau test n'est pas connue.

		Test standard		Total
		Maladie présente	Maladie absente	
Nouveau test	Test positif	VP a <b>Vrai Positif</b>	FP b <b>Faux Positif</b>	a+b
	Test négatif	FN c <b>Faux Négatif</b>	VN d <b>Vrai Négatif</b>	c+d
Total		a+c	b+d	a+b+c+d



## Phase III: Indicateurs

### Definitions et interpretations statistique de la VPP, VPN

		Test standard		Total
		Maladie présente	Maladie absente	
Nouveau test	Test positif	VP a Vrai Positif	FP b Faux Positif	a+b
	Test negatif	FN c Faux Negatif	VN d Vrai Negatif	c+d
	Total	a+c	b+d	a+b+c+d

**Valeur prédictive positive (VPP)** en anglais – positive predictive value - PPV

- la probabilité pour un sujet qui a été testé positif de vraiment avoir la maladie
- la probabilité pour un sujet d'avoir la maladie sachant qu'il a été testé positif
  - Ex. la probabilité pour un sujet d'avoir Covid 19 sachant qu'il a été testé positif à le test de rapide pour Covid 19 est de 0.78 (78%)
- proportion des sujets réellement malades parmi les sujets qui ont été testés positifs

$$\text{VPP} = \text{VP} / (\text{VP} + \text{FP})$$

**Valeur prédictive négative (VPN)** : en anglais – negative predictive value - NPV

- la probabilité pour un sujet ayant un résultat négatif de vraiment n'avoir pas la maladie
- la probabilité pour un sujet n'avoir pas la maladie sachant qu'il a été testé négatif
  - Ex. la probabilité pour un sujet de n'avoir pas Covid 19 sachant qu'il a été testé négatif à le test de rapide pour Covid-19 est de 0.93 (93%)
- proportion des sujets vraiment non-malades parmi les personnes testées négatives:

$$\text{VPN} = \text{VN} / (\text{VN} + \text{FN})$$

## Phase III:

### Relations VPP et VPN avec Se, Sp et la prévalence:

- dépendent de la Se, Sp et de la prévalence (la fréquence de la maladie)
- étant fonction de la prévalence, les valeurs prédictives **ne peuvent être appliquées que dans une population ayant une prévalence identique** à celle dans la population étudiée lors du calcul de ces indicateurs.

## Phase III:

- VPP et VPN peuvent être calculé à partir de: Se, Sp et p (prévalence):

$$VPP = \frac{Se}{Se + (1 - Sp) \times \frac{1 - p}{p}} \quad VPN = \frac{Sp}{Sp + (1 - Se) \times \frac{p}{1 - p}}$$

On doit pas apprendre les deux formules

## Phase III:

### Relations VPP et VPN avec Se et Sp et la prévalence

- $VPP = Se * \frac{\text{prévalence de la maladie}}{\text{prévalence du test positif}}$
- $VPN = Sp * \frac{1 - \text{prévalence de la maladie}}{\text{prévalence du test négatif}}$
- Plus la Sp est proche de 1 (100%) → le VPP est proche du 1
- Plus la Se est proche de 1 (100%) → le VPN est proche du 1

# Phase III:

## Relations VPP et VPN avec Prévalence

- $VPP = Se * \frac{\text{prévalence de la maladie}}{\text{prévalence du test positif}}$
- $VPN = Sp * \frac{1 - \text{prévalence de la maladie}}{\text{prévalence du test négatif}}$
- En tenant la:
  - Sensibilité égale –
    - VPP augmente avec une prévalence de la maladie plus grande
    - VPP diminue avec une prévalence du test positif petite
  - Spécificité égale –
    - VPN augmente avec une prévalence de la maladie plus petite
    - VPN diminue avec une prévalence du test positif plus grande (la prévalence du test négatif est petite)

# Phase II et III

- **Méthodologie d'étude diagnostique (pour phase II et III)**
  - **Sujets**
    - **Population étudiée**
      - **Critères d'inclusion/exclusion** – essentiel pour limiter les biais de sélection (évaluation de la validité de l'étude)
      - Depuis où les patients ont été **recrute** - la location, le cadre, ou ont été réalisés les deux tests diagnostiques – essentiel pour évaluer la généralisabilité des résultats
    - **Recrutement** (en fonction des symptômes, résultats d'autres tests ou si le sujet a reçu le nouveau test ou le test standard)

## Phase II et III

- **Méthodologie d'étude diagnostique (pour phase II et III)**
  - **Sujets**
    - Sélection des sujets:
      - **échantillon représentatif** – meilleure / difficile à faire
      - **série consécutive des patients** – l'un des meilleurs si on ne peut pas faire représentatif
        - Les deux permettent que toutes les situations soient prises en compte (malades, non malades, différents stades de la maladie, avec/sans traitement, maladies similaires)
      - **ou série des patients non consécutifs** – on doit préciser comment – possibles biais de sélection
      - **nombre** des sujets suffisamment grand
      - on doit préciser la **période du temps du recrutement**
  - **Variantes des études:**
    - Collection **prospective** des patients sur qui on applique les deux tests diagnostiques
      - idéale
    - Collection **rétrospective** des données enregistrées dans les feuilles d'observation des hôpitaux

# Phase II et III

- Méthodologie d'étude diagnostique (pour phase II et III)
  - Les tests diagnostiques – méthodologie de travail
    - **Le test standard** doit être **acceptable** (le **meilleur** ou proche: un examen histologique, ou plusieurs tests, on doit argumenter qu'il est valide, préciser qu'il représente la référence), doit être bien décrit – critères pour les malades/ non malades.
    - **Le « nouveau » test** est parfois, **moins invasive (moins des risques)**, moins chère, plus simple, mieux acceptée par le patient - en principe
    - **Les tests doivent être reproductible** (intra/ inter observationnel – on doit préciser le coefficient Kappa ou autres mesures de reproductibilité)
    - **Evaluation** des toutes les patients
      - **Indépendante** (deux évaluateurs différents des deux tests)
      - En **aveugle** (les résultats ne sont pas connus pour l'autre évaluateur)
      - Répartition des tests diagnostiques – **randomisée**



## Phase II et III

- **Méthodologie d'étude diagnostique** (pour phase II et III)
  - Pour certaines tests – imagerie, anatomopathologique – le nombre et l' **expérience, formation** des médecins doit être précisée pour les deux tests diagnostiques (années d' expérience, spécialistes/primaire/ résident/ étudiant)
  - Tous les sujets doit être **évaluées avec les deux tests** diagnostiques (si un **grand nombre** des sujets n'est pas évaluées avec un test ca peut poser des problèmes. Si l' étude est prospective et la perte de vue est importante ca peut poser des problèmes - biais d'attrition)
  - Le **temps passée entre les deux tests** (doit être **précises** avec date du debout et du fin – il doit être **court**, sinon il y a un possible biais – modification de l' état du patient)
  - **Préciser** si les patients ont pris des **médicaments** pendant la période de l' étude – qui peuvent modifier les résultats des tests - biais

# Phase II et III

- Méthodologie d'étude diagnostique (pour phase II et III)
  - Les **résultats** doit préciser:
    - Description des **caractéristiques démographiques et cliniques** de l'échantillon
      - age, sexe, ...
      - pour évaluer la generalisabilite des résultats
    - Préciser
      - le **nombre de sujets qui n'ont pas eu le test** de référence, et le nouveau test
        - dans le texte ou diagramme du flux (en: flow-chart)
        - pour évaluer des biais
      - préciser des **données manquantes**
      - préciser si il y a des **situation pas claires** (pour le test standard et nouveau test), comme catégorie séparé, et de ne pas les exclure (biais de pureté diagnostique). Préciser diverses méthodes de tenir compte de ces situations.

# Phase II et III

- Méthodologie d'étude diagnostique (pour phase II et III)
  - Les **résultats** doit préciser:
    - Description des **caractéristiques démographiques et cliniques** de l'échantillon

**Table 1. Characteristics of the included 1014 patients.**

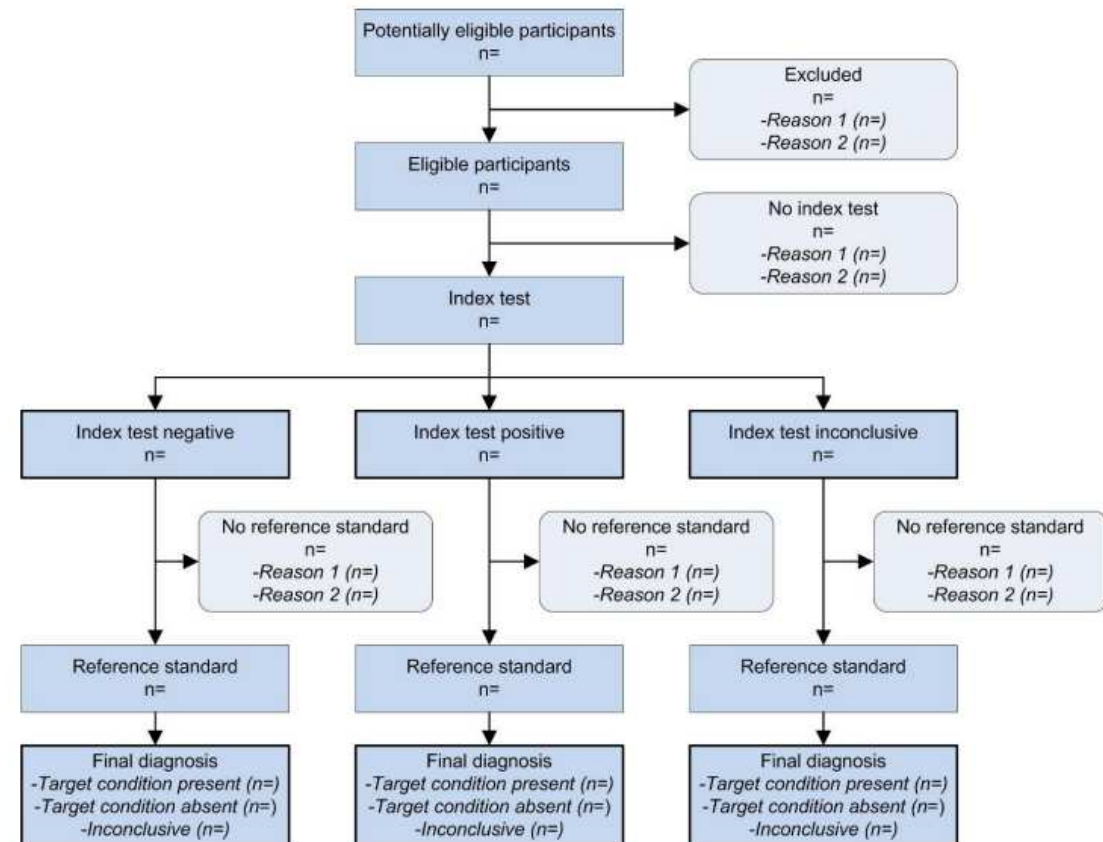
Characteristic	Results
<b>Age (years)</b>	
Mean age	51 ± 15, Range of 2 to 95
< 20	7 (1)
20-39	267 (26)
40-59	409 (40)
≥ 60	331 (33)
<b>Male</b>	467 (46)
<b>Median time-interval between chest CT scan and RT-PCR assay (days)</b>	1, Range of 0 to 7
<b>Results of RT-PCR assay</b>	
Positive	601 (59)
Negative	413 (41)
<b>Findings and manifestations of chest CT</b>	
Consistent with viral pneumonia (positive)	888 (88)
Ground-glass opacity	409/888 (46)
Consolidation	447/888 (50)
Reticulation/thickened interlobular septa	8/888 (1)
Nodular lesions	24/888 (3)
No CT findings of viral pneumonia	126 (12)

Note.-Data are patients (n =7) with percentages in parentheses. Age is mean ± standard deviation. Time-interval is shown as median.

RT-PCR= reverse transcription polymerase chain reaction.

# Phase II et III

- Méthodologie d'étude diagnostique (phase II et III)
  - Les **résultats** doit préciser:
    - dans le texte ou diagramme du flux
    - (en: flow-chart)



# Phase II et III

- Méthodologie d'étude diagnostique (pour phase II et III)
  - Les **résultats** doit préciser:
    - **tableau de contingence**
    - présentation des **Se, Sp, +LR, -LR, VPP, VPN, prévalence, précision**, avec des intervalles de confiance

**Table 2: The performance of chest CT for COVID-19 infection with RT-PCR result as reference.**

Results (n)				Test performance (%)					
TP	TN	FP	FN	Sensitivity [95% CI]	Specificity [95% CI]	PPV [95% CI]	NPV [95% CI]	Accuracy [95% CI]	
Overall									
580	105	308	21	97 (580/601) [95-98]	25 (105/413) [22-30]	65 (580/888) [62-68]	83 (105/126) [76-89]	68 (685/1014) [65-70]	

## Phase IV:

- **Evaluation du test en termes de:**
  - Coût - Bénéfice
  - Coût - Efficacité
- **Réponses à des questions telles que:**
  - Est-ce que l'application du test va modifier le lieu de la pathologie dans la collectivité?
  - Est-ce que l'application du test va avoir un impact
    - sur le clinicien,
    - sur la morbidité et la mortalité de la population?
  - Quel est l'impact économique du test?

## Évaluation d'une méthode diagnostique Collection des données – un seul group

(Échantillon représentatif, cohorte, série consécutive des cas)

- la seule chose fixée par le chercheur est la **taille totale** de l'échantillon
- le résultat du test et le statut de la maladie résultent en manière aléatoire
- en utilisant la collection des données – avec un seul group de sujets (Échantillon représentatif, série consécutive des cas, série des cas, ...), on peut estimer/calculer:
  - Se et Sp
  - +LR (RVP), -LR (RVN)
  - Le pourcentage des biens classées (accuracy)
  - L' indice de Youden
  - VPP et VPN
  - Prévalence (frequence des toutes les malades) de la maladie dans la population cible – pour échantillon représentatif/serie consécutive des cas – dépends ...

# Évaluation d'une méthode diagnostique

## Collection des données - exposés-non exposés

i.e. résultat positif vs. négatif

- approche prédictive
- **deux groupes de taille prédéterminée**
  - pas nécessairement égaux
  - sont constitués selon les résultats du test : (+) et (-)
- **l'état réel de la maladie** est le facteur aléatoire
- En utilisant la collection des données - exposés-non exposés, on peut estimer/calculer:
  - VPP et VPN



# Évaluation d'une méthode diagnostique

## Collection des données - cas-témoins

- **deux groupes de taille prédéterminée**
  - (pas nécessairement égaux → ils ne reflètent pas la prévalence de la maladie)
  - sont constitués, basée sur la présence ou l'absence de la maladie étudiée.
- **le résultat du test** est le facteur aléatoire
- En utilisant **la collection des données - cas-témoins on peut estimer/calculer:**
  - Se et Sp
  - +LR (RVP), -LR (RVN)
  - L'indice de Youden

# Le choix des tests en fonction de Se et Sp

- Un **test sensible** doit être préférée pour:
  - Pour des **maladies graves** mais **traitables**
  - Si le résultats sont faux positifs les répercussions psychologiques et financières sur les sujets ne sont pas très graves
  - Ex. maladies infectieuses, a transmission sexuelle.
- Un **test spécifique** doit être préférée pour:
  - Maladie grave **non traitable** ou **difficile a traiter**
  - Si on connait le fait qu'on n'a pas la maladie a une utilité psychologique et sanitaire
  - Si le résultats sont faux positifs les répercussions psychologiques et financières sur les sujets sont très importantes
  - Ex. des cancers

## Le choix des tests en fonction de VPP et VPN

- Un **test avec bon VPP** doit être préférée pour:
  - Si le résultats sont faux positifs les traitement peut avoir des conséquences graves
  - Ex. radiothérapie/ chimiothérapie pour une cancer mal diagnostiqué
- Un **test avec un bon VPN** doit être préférée pour:
  - Maladie grave mais bien traitable
  - Si on connait le fait qu'on n'a pas la maladie a une utilité psychologique et sanitaire
  - Si le résultats faux positifs et faux négatifs ont les répercussions psychologiques et financières sur les sujets sont très importantes
  - Ex. certaines formes des cancers bien traitables, maladies cardiovasculaires graves traitables.

# Evaluation de la validité des études diagnostiques

1. Le **spectre** des patients était-il **représentatif**? (échantillon consecutive ou representative - choisi par hasard)
2. La **test de référence acceptable** (identifie bien la présence/absence de la maladie)
3. pour être raisonnablement sûr que la condition cible n'a pas La période de **temps** entre le test de référence et le test d'index est-elle **suffisamment courte** hangé entre les deux tests?
4. **L'ensemble de l'échantillon** ou une sélection aléatoire de l'échantillon a-t-il reçu une **vérification à l'aide du test de référence** prévue? (vérification partielle évitée)
5. Les patients ont-ils reçu **le même test de référence** quel que soit le résultat du test d'index?
6. **Le test de référence était-il indépendant du test d'indice** (c'est-à-dire que le test d'index n'a pas partie de la norme de référence)? (incorporation évitée)
7. Les résultats du test de référence ont-ils été interprétés sans connaissance des résultats de la test d'index? (**résultats du test d'index en aveugle**)
8. Les résultats du test d'index ont-ils été interprétés sans connaissance des résultats du test de référence? (**résultats standard de référence aveuglés**)
9. **Les mêmes données cliniques étaient-elles disponibles** lorsque les résultats des tests ont été interprétés comme le seraient disponible lorsque le test est utilisé dans la pratique?
10. Les **résultats des tests non interprétables / intermédiaires ont-ils été rapportés**?
11. Les **retraits** de l'étude ont-ils été **expliqués**? (retraits expliqués)

[https://methods.cochrane.org/sites/methods.cochrane.org.sdt/files/public/uploads/ch09\\_Oct09.pdf](https://methods.cochrane.org/sites/methods.cochrane.org.sdt/files/public/uploads/ch09_Oct09.pdf)

## Exemple article - MD

[Br Dent J](http://www.nature.com/bdj/journal/v190/n8/full/4800998a.html). 2001 Apr 28;190(8):440-3. Available from: <http://www.nature.com/bdj/journal/v190/n8/full/4800998a.html>

### Occlusal caries detection in primary teeth: a comparison of DIAGNOdent with conventional methods.

[Attrill DC](#)<sup>1</sup>, [Ashley PF](#).



**OBJECTIVE:** To compare the accuracy and repeatability of three diagnostic systems (DIAGNOdent, visual and radiographic) for occlusal caries diagnosis in primary molars.

**DESIGN:** Two examiner, in vitro, blinded study. Histological gold standard.

**MATERIALS AND METHODS:** 58 occlusal surfaces of primary molars examined in turn by two examiners using each of three diagnostic systems (DIAGNOdent, visual and radiographic). These results were compared with a histological gold standard. Sensitivity and specificity were calculated for each diagnostic system at a range of thresholds. Inter- and intra- examiner repeatability were calculated for each diagnostic system using the kappa statistic.

**MAIN OUTCOME MEASURES:** Sensitivity, specificity, inter and intra examiner repeatability for each diagnostic system.

**RESULTS:** The highest **sensitivity** values were provided by DIAGNOdent (**0.77 and 0.80**, examiners 1 and 2 respectively) however this was offset by a lower **specificity (0.82 and 0.85)** than all other systems with the exception of examiner 1 at V1 visual threshold. The DIAGNOdent gave the highest values of **kappa for intra- and inter-examiner repeatability** with the exception of intra-examiner repeatability for examiner 2 where visual diagnosis had the highest value of kappa.

**CONCLUSION:** The DIAGNOdent was the most accurate system tested for the detection of occlusal dentine caries in primary molars. The performance of the DIAGNOdent systems was not statistically significantly better than that achieved using visual examination for non-cavitated teeth (V1 and V2 threshold). DIAGNOdent may prove useful as a predictive clinical tool, however with appropriate training visual examination may offer similar results without the need for additional equipment.

## Exemple article - MG

[BMC Res Notes](#). 2011 Mar 24;4:79. doi: 10.1186/1756-0500-4-79. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/pmid/21429231/>

### **MPT 64 Antigen detection for Rapid confirmation of M.tuberculosis isolates.**

[Kumar VG](#)<sup>1</sup>, [Urs TA](#), [Ranganath RR](#).

#### **BACKGROUND:**

A new rapid Immunochromatographic test kit(SD MPT64TB Ag Kit) for detection of MPT 64 Antigen in M. tuberculosis isolates using mouse monoclonal MPT 64 Antibody developed by SD Bioline, South Korea was evaluated for rapid identification of M. tuberculosis isolates. We also assessed the sensitivity, specificity and predictive values of this kit. The test kit has an excellent sensitivity, specificity, negative predictive value & positive predictive value. This rapid method is found to be a reliable, rapid and cheaper method for confirming MTB culture isolates in resource poor laboratories.

#### **MATERIAL/METHODS:**

54 culture isolates of M. tuberculosis in broth & on LJ medium, 12 Non mycobacterial isolates, 10 Non tubercular (NTM) rapidly growing Mycobacteria isolated from pus & 5 smear positive sputum samples were tested for detection of MPT64 antigen using the SD Bioline immunochromatography (ICT)test kit. H37 RV strain was employed as the positive reference control.

#### **FINDINGS:**

H37 RV strain showed the presence of MPT64 antigen band. Similar band was formed in all the 54 MTB isolates tested **proving 100% sensitivity**. MPT64 band formation was not detected in any of the other test isolates which **proved the 100% specificity** of the test kit. **Both PPV & NPV were 100%.**

#### **CONCLUSION:**

Tuberculosis is a global pandemic. Rapid identification of MTB culture isolate is very important for drug susceptibility testing. MPT 64 TB Ag detection ICT kit is a rapid, reliable method; it can be a substitute for the molecular identification methods.

## Exemple article - MG

Radiology, 200642 2020 Feb 26[Online ahead of print] DOI: [10.1148/radiol.2020200642](https://doi.org/10.1148/radiol.2020200642)

### Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases

[Tao Ai<sup>1</sup>](#), [Zhenlu Yang<sup>1</sup>](#), [Hongyan Hou<sup>1</sup>](#), [Chenao Zhan<sup>1</sup>](#), [Chong Chen<sup>1</sup>](#), [Wenzhi Lv<sup>1</sup>](#), [Qian Tao<sup>1</sup>](#), [Ziyong Sun<sup>1</sup>](#), [Liming Xia<sup>1</sup>](#)

#### Abstract

Background Chest CT is used for diagnosis of 2019 novel coronavirus disease (COVID-19), as an important complement to the reverse-transcription polymerase chain reaction (RT-PCR) tests. Purpose To investigate the diagnostic value and consistency of chest CT as compared with comparison to RT-PCR assay in COVID-19. Methods From January 6 to February 6, 2020, 1014 patients in Wuhan, China who underwent both chest CT and RT-PCR tests were included. With RT-PCR as reference standard, the performance of chest CT in diagnosing COVID-19 was assessed. Besides, for patients with multiple RT-PCR assays, the dynamic conversion of RT-PCR results (negative to positive, positive to negative, respectively) was analyzed as compared with serial chest CT scans for those with time-interval of 4 days or more. Results Of 1014 patients, 59% (601/1014) had positive RT-PCR results, and 88% (888/1014) had positive chest CT scans. **The sensitivity of chest CT** in suggesting COVID-19 **was 97% (95%CI, 95-98%, 580/601 patients)** based on positive RT-PCR results. In patients with negative RT-PCR results, 75% (308/413) had positive chest CT findings; of 308, 48% were considered as highly likely cases, with 33% as probable cases. By analysis of serial RT-PCR assays and CT scans, the mean interval time between the initial negative to positive RT-PCR results was  $5.1 \pm 1.5$  days; the initial positive to subsequent negative RT-PCR result was  $6.9 \pm 2.3$  days). 60% to 93% of cases had initial positive CT consistent with COVID-19 prior (or parallel) to the initial positive RT-PCR results. 42% (24/57) cases showed improvement in follow-up chest CT scans before the RT-PCR results turning negative. Conclusion Chest CT has a high sensitivity for diagnosis of COVID-19. Chest CT may be considered as a primary tool for the current COVID-19 detection in epidemic areas.

**Keywords:** 2019-nCoV pneumonia; chest CT imaging; diagnostic value; positive rate; reverse transcription polymerase chain reaction.

# Exemple article - MG

Radiology, 200642 2020 Feb 26[Online ahead of print] DOI: [10.1148/radiol.2020200642](https://doi.org/10.1148/radiol.2020200642)

## Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases

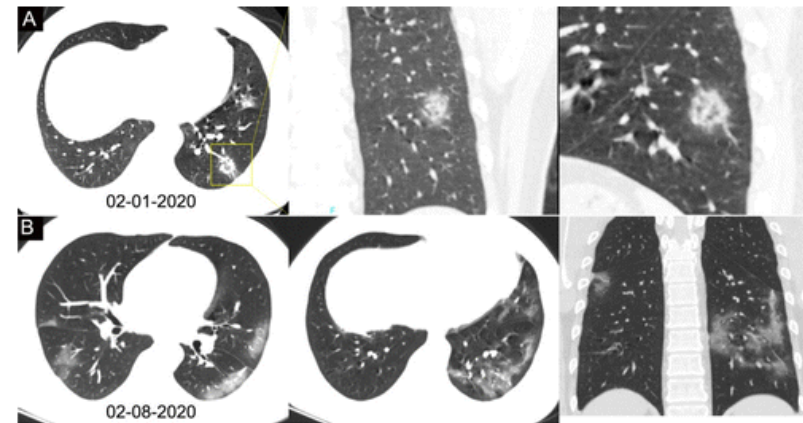
Tao Ai<sup>1</sup>, Zhenlu Yang<sup>1</sup>, Hongyan Hou<sup>1</sup>, Chenao Zhan<sup>1</sup>, Chong Chen<sup>1</sup>, Wenzhi Lv<sup>1</sup>, Qian Tao<sup>1</sup>, Ziyong Sun<sup>1</sup>, Liming Xia<sup>1</sup>

### Abstract

Table 2: The performance of chest CT for COVID-19 infection with RT-PCR result as reference.										
	Results (n)				Test performance (%)					[95% CI]
	TP	TN	FP	FN	Sensitivity	Specificity	PPV	NPV	Accuracy	
					[95% CI]	[95% CI]	[95% CI]	[95% CI]		
Overall	580	105	308	21	97 (580/601) [95-98]	25 (105/413) [22-30]	65 (580/888) [62-68]	83 (105/126) [76-89]	68 (685/1014) [65-70]	
Age										
< 60 years	362	81	225	15	96 (362/377) [94-98]	27 (81/306) [22-32]	62 (362/587) [58-66]	84 (81/96) [76-90]	65 (443/683) [61-68]	
≥ 60 years	218	24	83	6	97 (218/224) [94-99]	22 (24/107) [16-31]	72 (218/301) [67-77]	80 (24/30) [63-91]	73 (242/331) [68-78]	
Sex										
Male	272	35	148	12	96 (272/284) [93-98]	19 (35/183) [14-25]	65 (272/420) [60-69]	75 (35/47) [61-85]	66 (307/467) [61-70]	
Female	308	70	160	9	97 (308/317) [95-99]	30 (70/230) [25-37]	66 (308/468) [61-70]	89 (70/79) [80-94]	69 (378/547) [65-73]	

TP= true positive, TN=true negative, FP=false positive, FN=false negative, PPV= positive predictive value, NPV=negative predictive value, RT-PCR= reverse transcription polymerase chain reaction.

Data in parentheses are numbers of patients used to calculate percentages. Data in brackets are 95% confidence intervals.





# Exemples des questions pour l'examen

**Q1.** Vous travaillez dans un hôpital, et vous voulez utiliser un nouveau test diagnostique. Le test a la sensibilité de 0,99, et la spécificité de 0,91. Lesquelles des suivantes affirmations sont correctes ?

- a) le test est meilleur comme test de screening
- b) le test est meilleur comme test de précision
- c) la sensibilité est très bonne
- d) 9% des sujets non malades ont le test positif
- e) 1% des malades ont le test négatif

**Réponse: a, c, d, e**

## Explications

Les tests de screening sont des tests avec une grande sensibilité. Les test de diagnostique de précision sont avec une grande spécificité. Ici la sensibilité est plus grande que la spécificité, donc le réponse a) est correcte

Pour la Se, Sp, les valeurs  $\geq 90\%$ , sont cliniquement bonnes, les valeurs entre  $\geq 80\%$  et  $< 90\%$  sont modérées, les valeurs inférieure a  $< 80\%$  sont cliniquement avec une qualité réduite. Ici Se=99%, donc le réponse c) est correcte.

Dans le tableau de contingence sur la colonne des non malades on peut calculer la spécificité ( $d/(b+d)$ ) – qui représente la probabilité d'avoir une test négative sachant que le sujet est non malade. Dans la même colonne on peut la probabilité d'avoir une test positif sachant que le sujet est non malade ( $b/(b+d)$ ). La somme de ces deux probabilités est 1. Ici la Sp=0,91, donc,  $1-0,91=0,09$ , qui est 9%. Le réponse d) est correcte.

Dans le tableau de contingence sur la colonne des malades on peut calculer la sensibilité ( $a/(a+c)$ ) – qui représente la probabilité d'avoir une test positif sachant que le sujet est malade. Dans la même colonne on peut la probabilité d'avoir une test négatif sachant que le sujet est malade ( $c/(a+c)$ ). La somme de ces deux probabilités est 1. Ici la Se=0,99, donc,  $1-0,99=0,01$ , qui est 1%. Le réponse e) est correcte.

# Exemples des questions pour l'examen

**Q2.** Vous travaillez dans un hôpital, et vous voulez utiliser un nouveau test diagnostique. Le test a la sensibilité de 0,91, et la spécificité de 0,99. Lesquelles des suivantes affirmations sont correctes ?

- a) le test est meilleur comme test de screening
- b) le test est meilleur comme test de précision
- c) la spécificité est très bonne
- d) 99% des sujets non malades ont le test négatif
- e) 99% des sujets non malades ont le test positif

**Réponse: b, c, d**

**Q3.** Dans une collection des données exposée-non exposée pour une test diagnostique on peut calculer les suivantes indicateurs médicales

- a) Se
- b) Sp
- c) VPP
- d) VPN
- e) le pourcentage des bien classes

**Réponse: c, d**

# Exemples des questions pour l'examen

**Q4.** Dans une collection des données cas témoin pour une test diagnostique on peut calculer les suivantes indicateurs médicales

- a) Se
- b) Sp
- c) +LR
- d) -LR
- e) le pourcentage des bien classes

**Réponse: a, b, c, d**

**Q5.** Dans une collection des données échantillon représentatif pour une test diagnostique on peut calculer les suivantes indicateurs médicales

- a) Se
- b) Sp
- c) la prevalence de la maladie
- d) VPN
- e) le pourcentage des bien classes

**Réponse: a, b, c, d, e**

# Exemples des questions pour l'examen

**Q6.** Dans une collection des données cas témoin pour une test diagnostique on peut calculer les suivantes indicateurs médicales

- a) Indice Youden
- b) VPP
- c) VNP
- d) -LR
- e) le pourcentage des bien classes

**Réponse: a, d**

**Q7.** Lesquelles des suivantes affirmations sont correctes ?

- a) Une bonne sensibilité permet de limiter le nombre de faux positifs
- b) Une bonne spécificité permet de limiter le nombre de faux positifs
- c) Une bonne valeur prédictive positive permet de limiter le nombre de faux positifs
- d) Une bonne valeur prédictive négative permet de limiter le nombre de faux positifs
- e) Une bonne sensibilité permet de limiter le nombre de faux négatifs

**Réponse: b, c, e**

# Exemples des questions pour l'examen

**Q8.\*** Dans une étude d'évaluation de la radiographie comme test de diagnostic pour la détection de la tuberculose, réalisé avec un collection cas témoin, parmi les 100 malades 90 ont été détectés comme positifs par le nouveau test et parmi les 100 patients en bonne santé 85 ont été qualifiées comme pas affectée à juste titre par le nouveau test. La VPP et la VPN sont:

- a) 0,857 et 0,895
- b) 0,90 et 0,895
- c) 0,85 et 0,857
- d) 0,857 et 0,857
- e) Ne peuvent pas être calculée

**Réponse:** e

**Q9.** Vous travaillez dans un hôpital, et vous avez a votre disposition trois test diagnostique. Le test A a la sensibilité de 0,55, et la spécificité de 0,85; le test B a la sensibilité de 0,82, et la spécificité de 0,97; le test C a la sensibilité de 0,96, et la spécificité de 0,83. Lesquelles des suivantes affirmations sont correctes ?

- a) le test C est meilleur comme test de screening
- b) le test B est meilleur comme test de précision
- c) la sensibilité du test B est cliniquement modérée
- d) la sensibilité du test C est cliniquement bonne
- e) la sensibilité du test A est cliniquement modérée

**Réponse:** b, c, d

# Exemples des questions pour l'examen

**Q10.** Vous suspectez un de vos patients d'avoir la maladie testée par le nouveau test diagnostique. De l'anamnèse vous assumez une probabilité de 55% pour votre patient d'avoir la maladie.  $LR+ = 3,5$ ,  $LR- = 0,5$ . Vous appliquez le nouveau test et le patient a un résultat positif. Utilisez le nomogramme de Fagan pour trouver la probabilité post-teste que votre patient soit réellement malade. Lesquelles des suivantes affirmations sont correctes ?

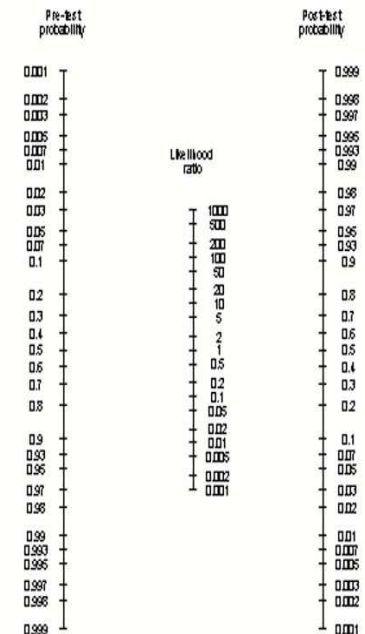
- a) la probabilité d'avoir la maladie post test est  $\sim 80\%$
- b) la probabilité d'avoir la maladie post test est  $\sim 40\%$
- c) le test diagnostique positif entraînent des changements réduits/faible mais parfois concluante/cliniquement significatif de la probabilité d'un diagnostic
- d) le test diagnostique négatif conduit à des modifications de faible probabilité, et rarement cliniquement important
- e) on doit utiliser  $LR+$  sur la deuxième axe pour trouver la probabilité post test

**Réponse: a, c, d, e**

**Q11.** Vous suspectez un de vos patients d'avoir la maladie testée par le nouveau test diagnostique. De l'anamnèse vous assumez une probabilité de 0,7 pour votre patient d'avoir la maladie.  $LR+ = 10,1$ ,  $LR- = 0,09$ . Vous appliquez le nouveau test et le patient a un résultat négatif. Utilisez le nomogramme de Fagan pour trouver la probabilité post-teste que votre patient soit réellement malade. Lesquelles des suivantes affirmations sont correctes ?

- a) la probabilité d'avoir la maladie post test est  $\sim 97\%$
- b) la probabilité d'avoir la maladie post test est  $\sim 0,2$
- c) le test diagnostique positif entraînent des changements importants et souvent concluante de la probabilité d'un diagnostic
- d) le test diagnostique négatif entraînent des changements réduits/faible mais parfois concluante/cliniquement significatif de la probabilité d'un diagnostic
- e) on doit utiliser  $LR+$  sur la deuxième axe pour trouver la probabilité post test

**Réponse: b, c**



# Exemples des questions pour l'examen

## Q12. Explications

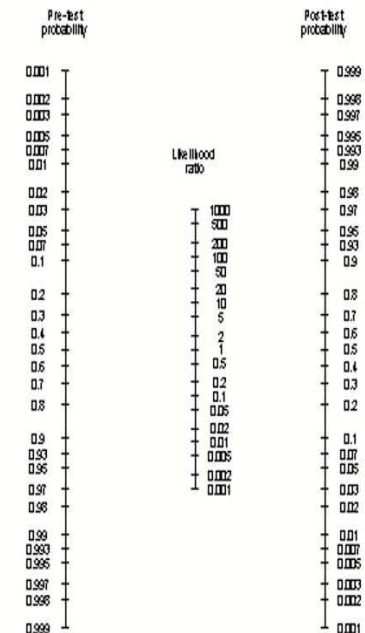
Pour trouver la probabilité post test, sur la première axe on va identifier un point avec la probabilité donnée dans la question: 0,7, puis sur la deuxième axe (Likelihood ratio), si le résultat du test est positif on va identifier un point avec la valeur du +LR (positive likelihood ratio/ rapport de vraisemblance positif), et si le résultat du test est négatif on va identifier un point avec la valeur du -LR (negative likelihood ratio/ rapport de vraisemblance négatif) – ici, dans le scénario il est indiquée que le résultat a été négatif, donc on va utiliser  $-LR=0,09$ , et on va tirer une ligne entre ces deux points (pretest et likelihood ratio), qui on va le prolonger jusqu'à la troisième axe (posttest probability), ou on va trouver la probabilité post test. Donc la réponse b) est correcte, et les réponses a), et e) sont incorrectes.

Concernant le niveau des changements entraînés par des tests positifs et négatifs, on doit regarder les valeurs du LR.

- LR > 10 ou < 0,1 entraînent des changements
  - importants et souvent concluante de la probabilité d'un diagnostic
- LR entre 5 et 10 ou entre 0,1 (=1/10) et 0,2 (=1/5) entraînent des changements
  - modéré et utile de la probabilité d'un diagnostic
- LR entre 2 et 5 ou entre 0,5 (=1/2) et 0,2 (=1/5) entraînent des changements
  - réduits/faible mais parfois concluante/cliniquement significatif de la probabilité d'un diagnostic
- LR entre 1 et 2 ou entre 0,5 et 1 conduit à des modifications
  - de faible probabilité, et rarement cliniquement important.
- LR = 1 – pas de changement

Le LR+ = 10,1 > 10 => changements importants et souvent concluante => réponse c est correcte

Le LR- = 0,09 < 0,1 => changements importants et souvent concluante => réponse e est incorrecte



# Exemples des questions pour l'examen

**Q13.** \* Dans une étude d'évaluation d'une technique radiographique comme test de diagnostic pour la détection de la sarcoïdose thoracique sur un échantillon représentatif, parmi les 200 malades 110 ont été détectés comme positifs par le nouveau test et parmi les 200 patients en bonne santé 105 ont été qualifiées comme pas affectée à juste titre par le nouveau test. La Sensibilité et la VPP sont:

- a) ne peuvent pas être calculées
- b) 0,55 et 0,537
- c) 0,525 et 0,537
- d) 0,55 et 0,538
- e) 0,538 et 0,525

**Réponse:** b

**Q14.** \* Dans une étude d'évaluation d'une technique radiographique comme test de diagnostic pour la détection de la sarcoïdose thoracique sur un échantillon représentatif, parmi les 200 malades 110 ont été détectés comme positifs par le nouveau test et parmi les 200 patients en bonne santé 105 ont été qualifiées comme pas affectée à juste titre par le nouveau test. Les VP, VN, FP, FN sont:

- a) 110, 105, 95, 90
- b) 95, 90, 110, 105
- c) 110, 95, 105, 90
- d) 105, 110, 105, 95
- e) 110, 90, 95, 105

**Réponse:** a



# Exemples des questions pour l'examen

**Q15.\*** Dans une étude d'évaluation d'une technique radiographique comme test de diagnostic pour la détection de la sarcoïdose thoracique sur un échantillon représentatif, parmi les 200 malades 110 ont été détectés comme positifs par le nouveau test et parmi les 200 patients en bonne santé 105 ont été qualifiées comme pas affectée à juste titre par le nouveau test. La Sensibilité et la VPP sont:

- a) la probabilité d'un patient présentant un nouveau test positif d'avoir la maladie testée représente le VPP
- b) la probabilité d'un patient non-malade à tester négatif en utilisant le nouveau test représente la spécificité
- c) la probabilité d'un patient malade à tester positif en utilisant le nouveau test représente la sensibilité
- d) la probabilité d'un patient présentant un nouveau test négatif d'avoir la maladie testée représente la spécificité
- e) la probabilité d'un patient malade à tester négatif en utilisant le nouveau test représente la sensibilité

**Réponse:** a, b, c

# Notions étudiées dans le cours

- Évaluation d'une méthode diagnostique
  - Phases
    - I: validité, fiabilité, le test kappa de Cohen
    - II: intrinsèque: Se, Sp, LR+, LR-
    - III: extrinsèque: VPP, VPN
    - IV: cout – bénéfice
- Exemples des études
- Exemples des exercices

Fin

