

Valeurs discriminantes et prédictives d'un test diagnostique ou de dépistage

Dr Cécile Couchoud

Principe

- Capacité du test à remplir ses objectifs
- Nécessité d'un examen de référence (gold-standard)
- En fonction de l'utilisation envisagée (diagnostic ou dépistage)

Des objectifs à priori différents

■ Dépistage

- ◆ Population à majorité « saine »
- ◆ Un outil simple
- ◆ + Sensibilité (éviter des faux négatifs)
- ◆ Examen à un moment t
- capacité discriminante

■ Diagnostic

- ◆ Un individu ciblé
- ◆ Un outil précis
- ◆ + Spécificité (éviter des faux positifs)
- ◆ Examens répétés
- capacité prédictive

Problématique du dépistage



- « Ne pas rater de malades » : méthode sensible
- « Ne pas être submergé de non-malades » : méthode spécifique

Variable dichotomique

- Test positif ou négatif
- Exemple :
 - ◆ Présence ou absence d'une image suspecte à la radio pulmonaire
 - ◆ Présence ou absence d'anticorps antinucléaire
 - ◆ Présence ou absence d'une douleur de la fosse iliaque gauche

Valeur diagnostique

		Référence (M)	
		+	-
Test (T)	+	VP	FP
	-	FN	VN

Indice de Youden (J)
 $= Se + Sp - 1$

Sensibilité
 $= Pr (T+/M+)$
 $= VP / (VP+FN)$

Spécificité
 $= Pr (T-/M-)$
 $= VN / (VN+FP)$

Interprétation

		Référence (M)	
		+	-
Test (T)	+	90	10
	-	60	40

Examens inutiles
Inquiétudes...

Sensibilité
 $= 90/150 = 60\%$

Spécificité
 $= 40/50 = 80\%$

Maladie non
détectée à
temps...

Erreur dans 70 cas /200 !

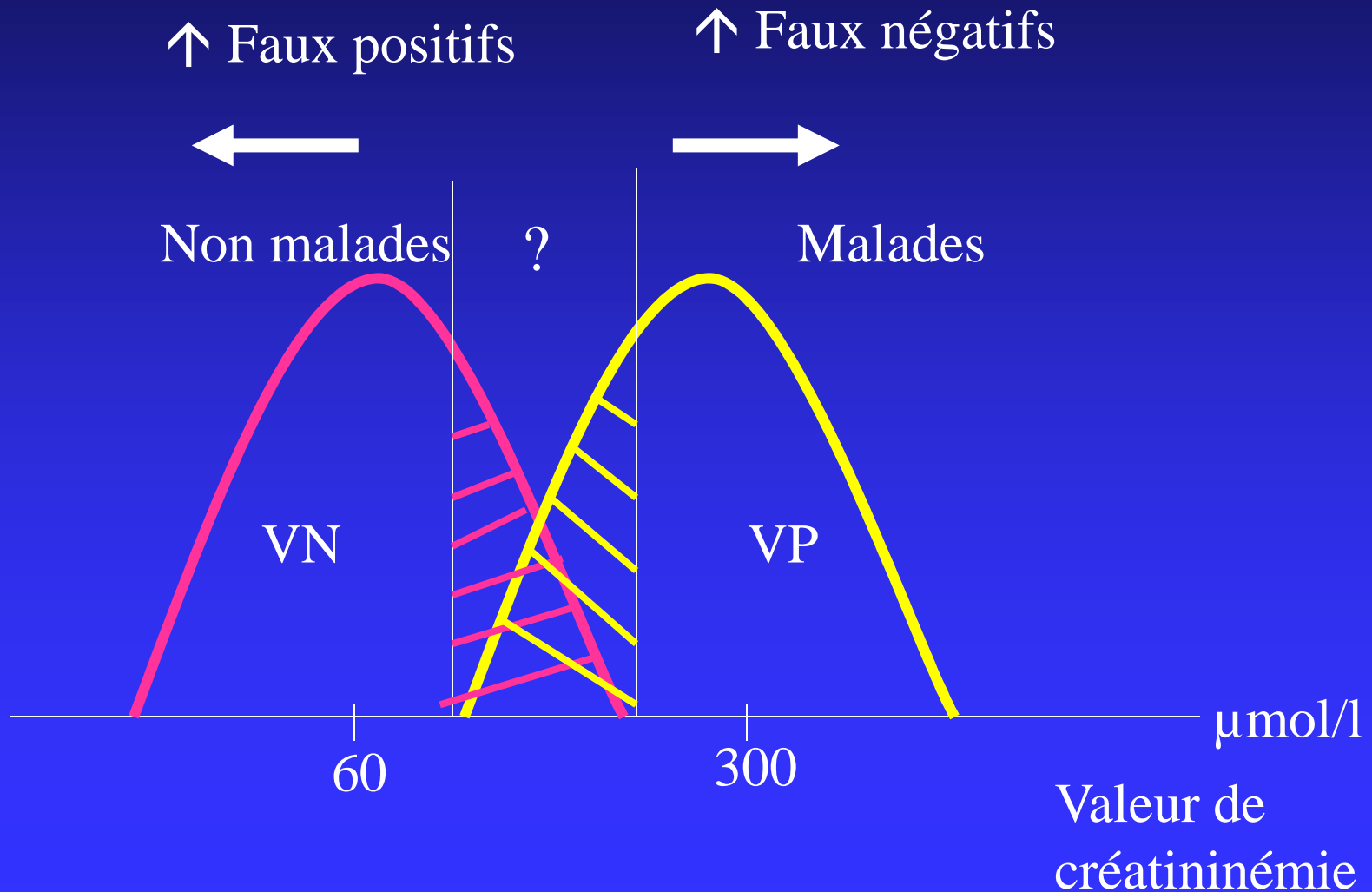
Exactitude : $90+40 / 200 = 75\%$

Tirage au sort : exactitude = 50%

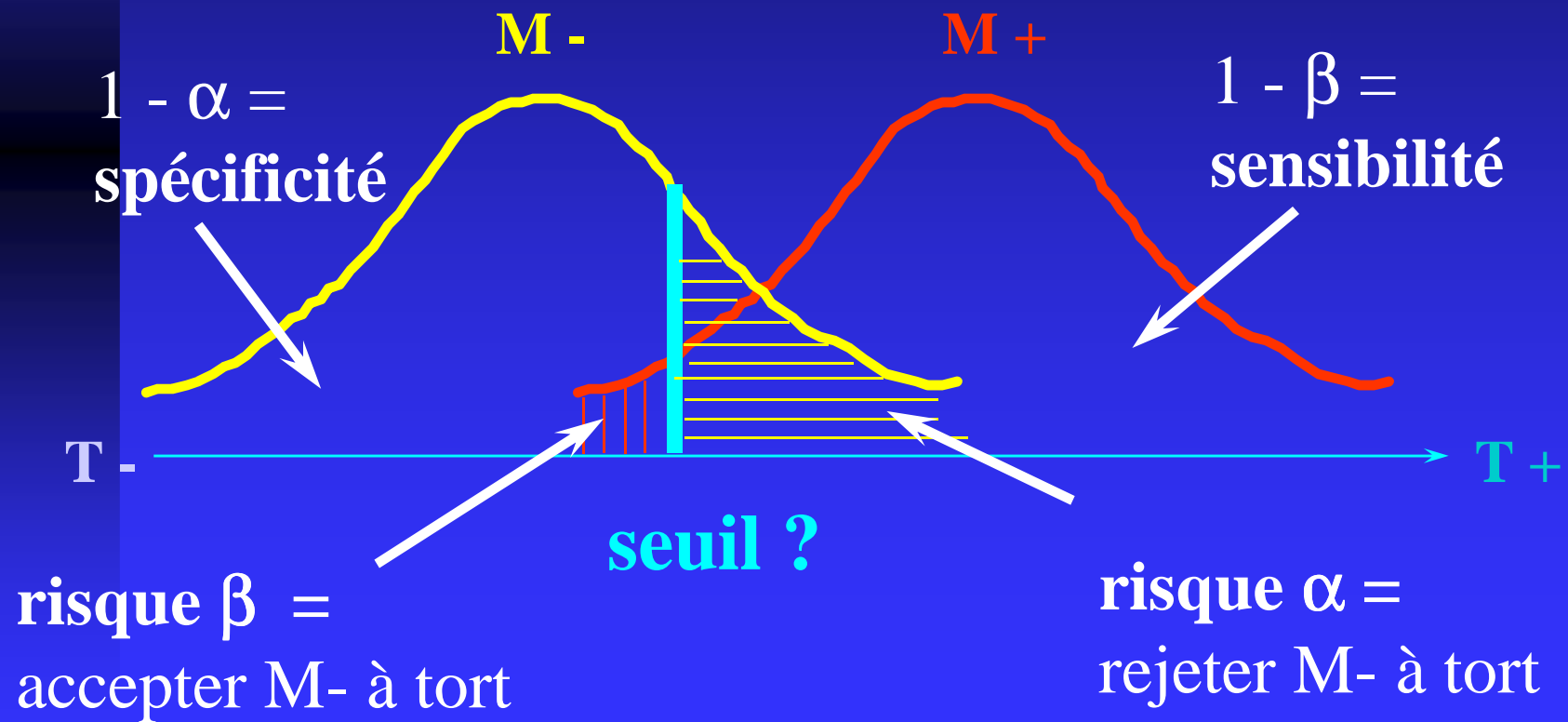
Variable quantitative

- Continuum entre valeurs « basses » et « hautes »
 - ◆ Créatininémie
 - ◆ Pression artérielle
- Continuum entre image « normale » et image « pathologique »
 - ◆ Image suspecte à la radio
 - ◆ Taille d'un anévrisme de l'aorte

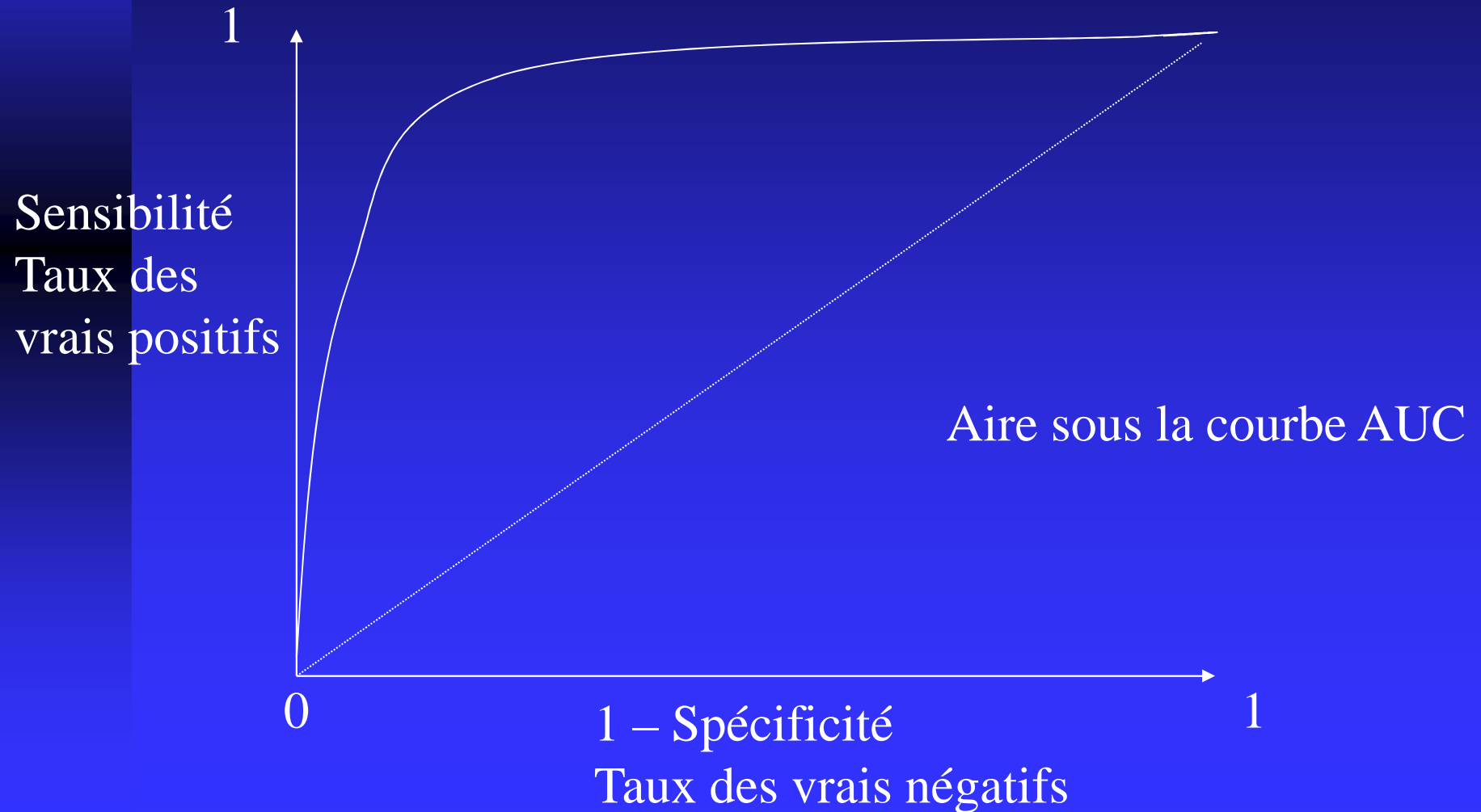
Choix du seuil : ex créatininémie



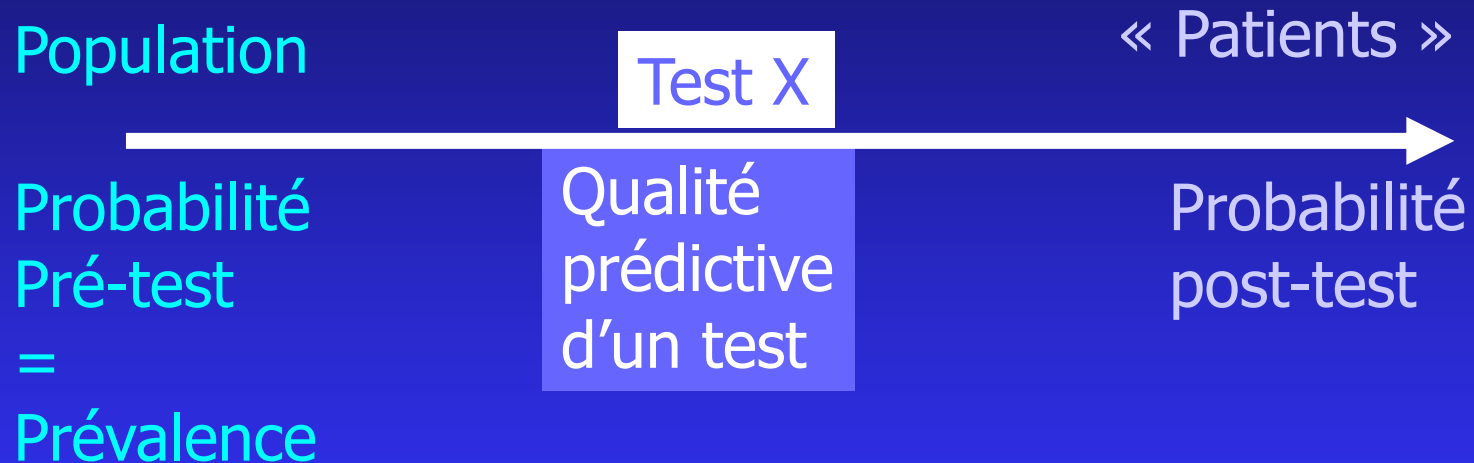
Lien avec les erreurs α et β



Courbe ROC : choix du « meilleur » seuil



Problématique du diagnostique



- Probabilité de maladie si le test est positif ? = Valeur prédictive positive
- Probabilité de non-maladie si le test est négatif ? = Valeur prédictive négative

Valeur prédictive

Référence (M)

+

-

Test (T)

+

VP

FP

-

FN

VN

VP positive
= $\text{Pr} (M+/T+)$

VP négative
= $\text{Pr} (M-/T-)$

Valeur prédictive

- Sur échantillon représentatif :

$$VPP = \Pr (M+/T+) = VP / (VP+FP)$$

$$VPN = \Pr (M-/T-) = VN / (VN+FN)$$

- Selon Bayes

$$VPP = Se.P / (Se.P + (1-Sp).(1-P))$$

$$VPN = Sp.(1-P) / (Sp.(1-P) + (1-Se).P)$$

P = prévalence de la maladie

Impact de la prévalence

- *Ex : homme, DFG < 60 ml/min, seuil creat = 137 $\mu\text{mol/l}$*
- *Se = 0.9, Sp = 0.93*
- *Prévalence IRC population ?*
 - *Si prévalence = 0.01, VPP = 0.12, VPN = 0.99*
 - *Si prévalence = 0.03, VPP = 0.3, VPN = 0.99*
 - *Si prévalence = 0.1, VPP = 0.6, VPN = 0.83*
 - *Si prévalence = 0.5, VPP = 0.93, VPN = 0.83*
 - *Si prévalence = 0.7, VPP = 0.97, VPN = 0.80*

La probabilité est de 60% d'être malade quand creat > 137 $\mu\text{mol/l}$

la probabilité est de 83% de ne pas être malade quand la creat < 137 $\mu\text{mol/l}$

Intervalle de sécurité du test

Intervalle où la probabilité post test de la maladie est supérieur à 0.5, si le test est positif ($VPP > 0.5$) et où la probabilité post test de la maladie est inférieur à 0.5, si le test est négatif ($VPN > 0.5$)

- *Ex : homme, < 60 ml/min, seuil = $137 \mu\text{mol/l}$*
- $Se = 0.9, Sp = 0.93$

→ Probabilité pré-test (prévalence) entre 7% et 90%

Rapport de vraisemblance (LR)

■ Rapport de vraisemblance positif :

- ◆ Rapport de la probabilité d'être malade / probabilité de ne pas l'être, quand test est positif
- ◆ Taux des VP / taux des FP = $Se / (1 - Sp)$

■ Rapport de vraisemblance négatif

- ◆ Rapport de la probabilité d'être malade / probabilité de ne pas l'être quand test est négatif
- ◆ Taux des FN / taux des VN = $(1 - Se) / Sp$

Rapport de vraisemblance (LR)

- *Ex : homme, < 60 ml/min, seuil creat = 137 $\mu\text{mol/l}$*
 - $\text{Se} = 0.9, \text{Sp} = 0.93$
 - $\text{RVP} = 13.5$: les malades ont 13.5 fois plus de chance d'avoir une creat $> 137 \mu\text{mol/l}$ que les non-malades
 - $\text{RVN} = 0.11$: les malades ont une probabilité de 0.11 fois /non malades d'avoir une creat $< 137 \mu\text{mol/l}$

Changer de seuil ?

- Si prévalence 10%
- homme, < 60 ml/min
- seuil = 137 $\mu\text{mol/l}$, Se= 0.9, Sp= 0.93, VPP = 0.59, VPN = 0.45
- Changement du seuil :
 - Se= 0.95, Sp= 0.81, creat = 121, VPP = 0.36 , VPN = 0.59
 - Se= 0.87, Sp= 0.95, creat = 144, VPP = 0.66, VPN = 0.40
 - Se= 0.84, Sp= 0.98, creat = 150, VPP = 0.82, VPN = 0.35

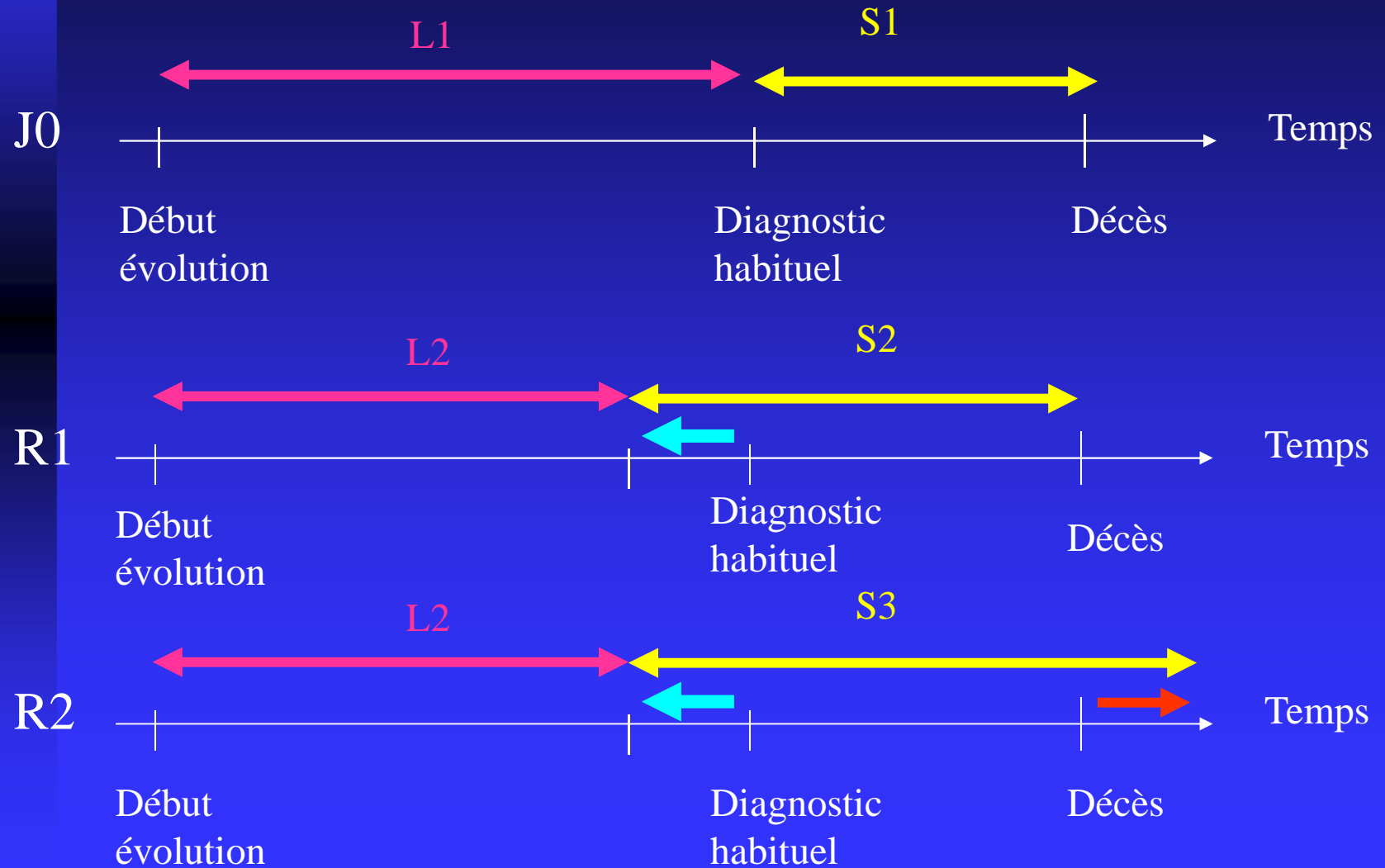
Augmenter la probabilité pré-test ?

- Cibler les indications de dosage de la créatininémie
 - Population à risque
 - Approche séquentielle : protéinurie puis créatininémie ...

Dépistage : problème « éthique »

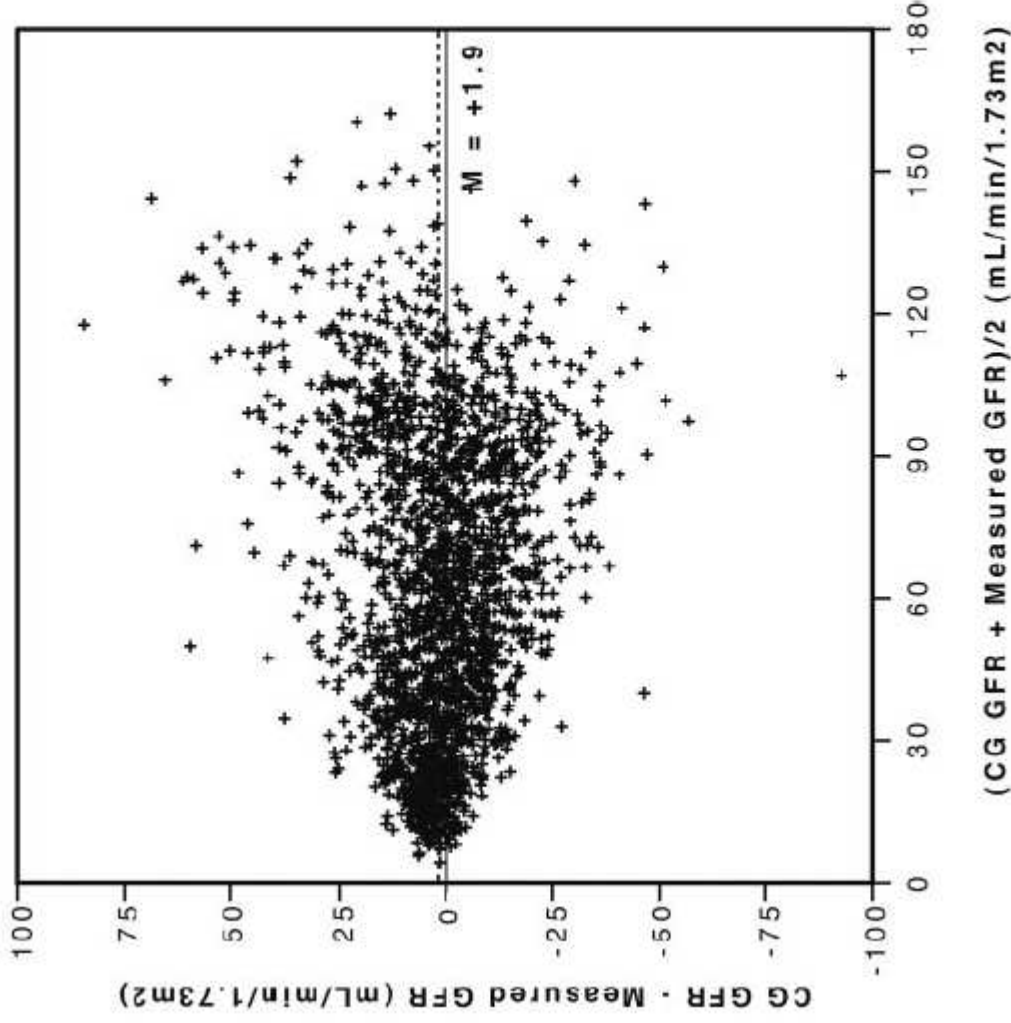
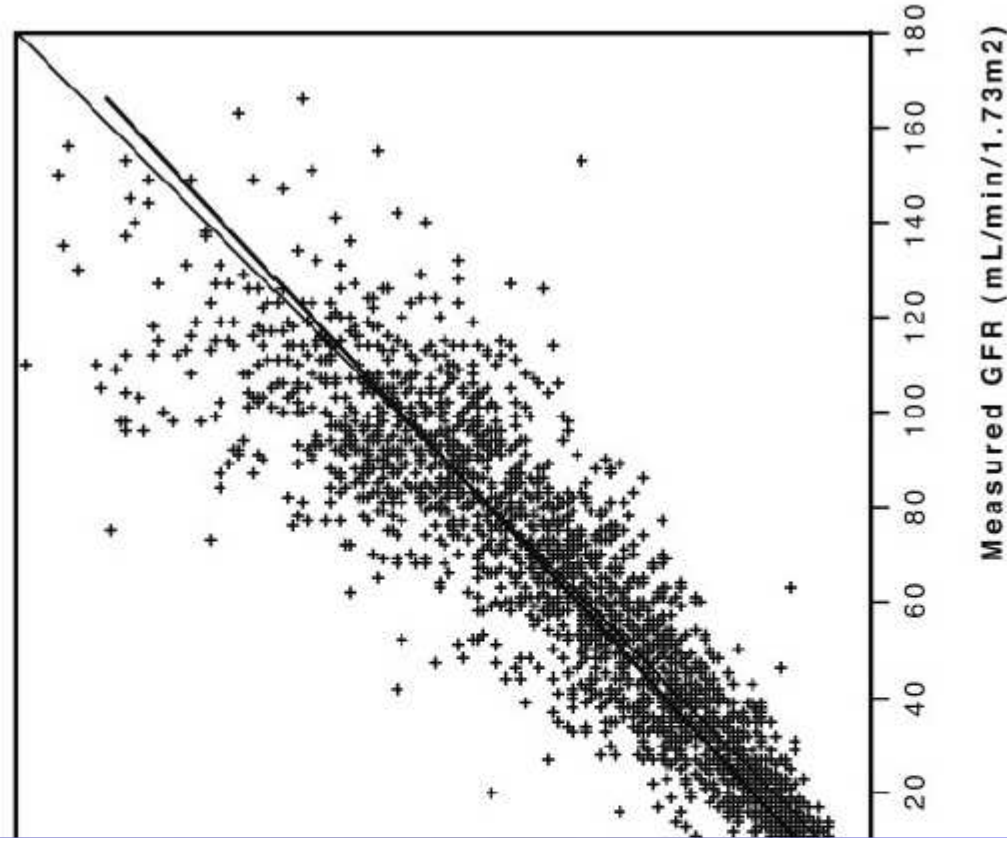
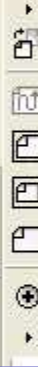
- Population apparemment indemne
 - ◆ Affection fréquente, responsable d'une mortalité ou morbidité importante (menace pour la santé publique)
 - ◆ Histoire naturelle connue +++
 - ◆ Existence d'un test sensible et spécifique
 - ◆ Test simple, sans danger, facilement accepté
 - ◆ Traitement disponible, efficacité démontrée, acceptable pour les patients
- Examens diagnostiques pour vérification

Histoire naturelle



En l'absence de test de référence

- Comparaison de 2 tests « équivalents »
- Ne pas utiliser de test statistique (coefficient de corrélation) !
- Méthode graphique de Bland et Altman
 - ◆ En abscisse : moyenne des 2 valeurs
 - ◆ En ordonnée : différence entre les 2 valeurs
- Biais global : la différence moyenne
- Précision : écart-type de la différence moyenne
- Variation du biais selon les valeurs moyennes



) Relationship between measured GFR and Cockcroft-Gault (CG) GFR. (B) Bland and Altman plot comparing measured and CG GFR. The mean difference (M) is represented by the dashed line.

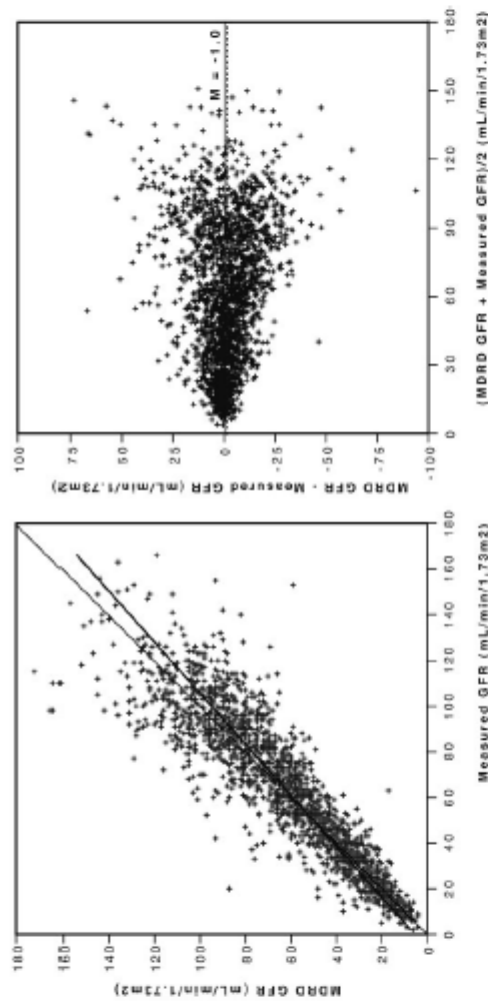


Figure 2. (A) Relationship between measured GFR and Modification of Diet in Renal Disease (MDRD) GFR. (B) Bland and Altman plot comparing measured GFR and MDRD GFR. The mean difference (M) is represented by the dashed line.

ment. A modified kinetic Jaffe colorimetric method was used with a Bayer EA-XI and a Kresch 20 analyzer. A five-point calibration was applied in each assay. Before measurement, ultrafiltration of plasma through a 20-kD cutoff membrane (MPS-1; Amicon, Beverly, MA) was performed to discard chromogens that were linked to albumin and other heavy proteins. In the absence of an international standard for

creatinine assay, the linearity of the measurements was verified by using plasma samples from normal subjects in which increasing amounts of desiccated creatinine hydrochloride (MW 149.6; Sigma Chemicals, Perth, Australia) had been added.

Linear regression analysis showed that the slope of the relationship between measured and expected creatinine concentrations was $1.008 \pm$

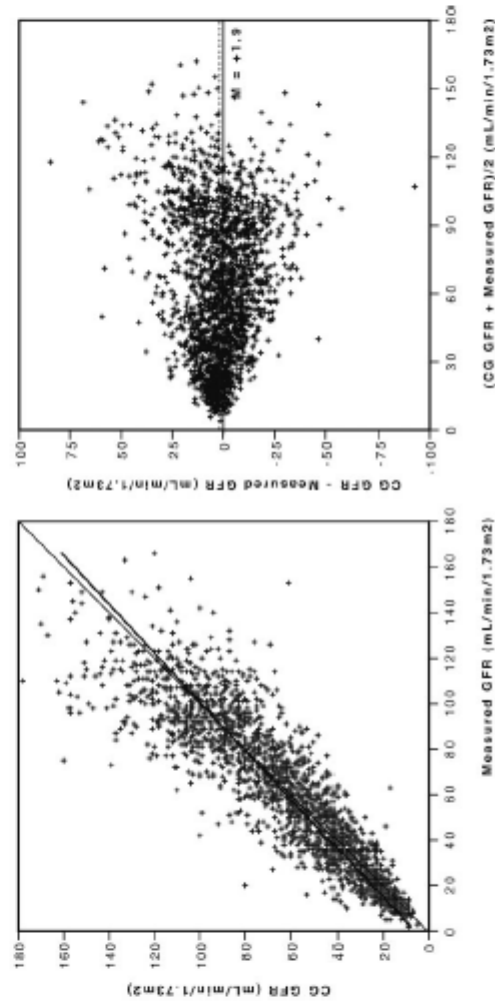


Figure 3. (A) Relationship between measured GFR and Cockcroft-Gault (CG) GFR. (B) Bland and Altman plot comparing measured GFR and CG GFR. The mean difference (M) is represented by the dashed line.

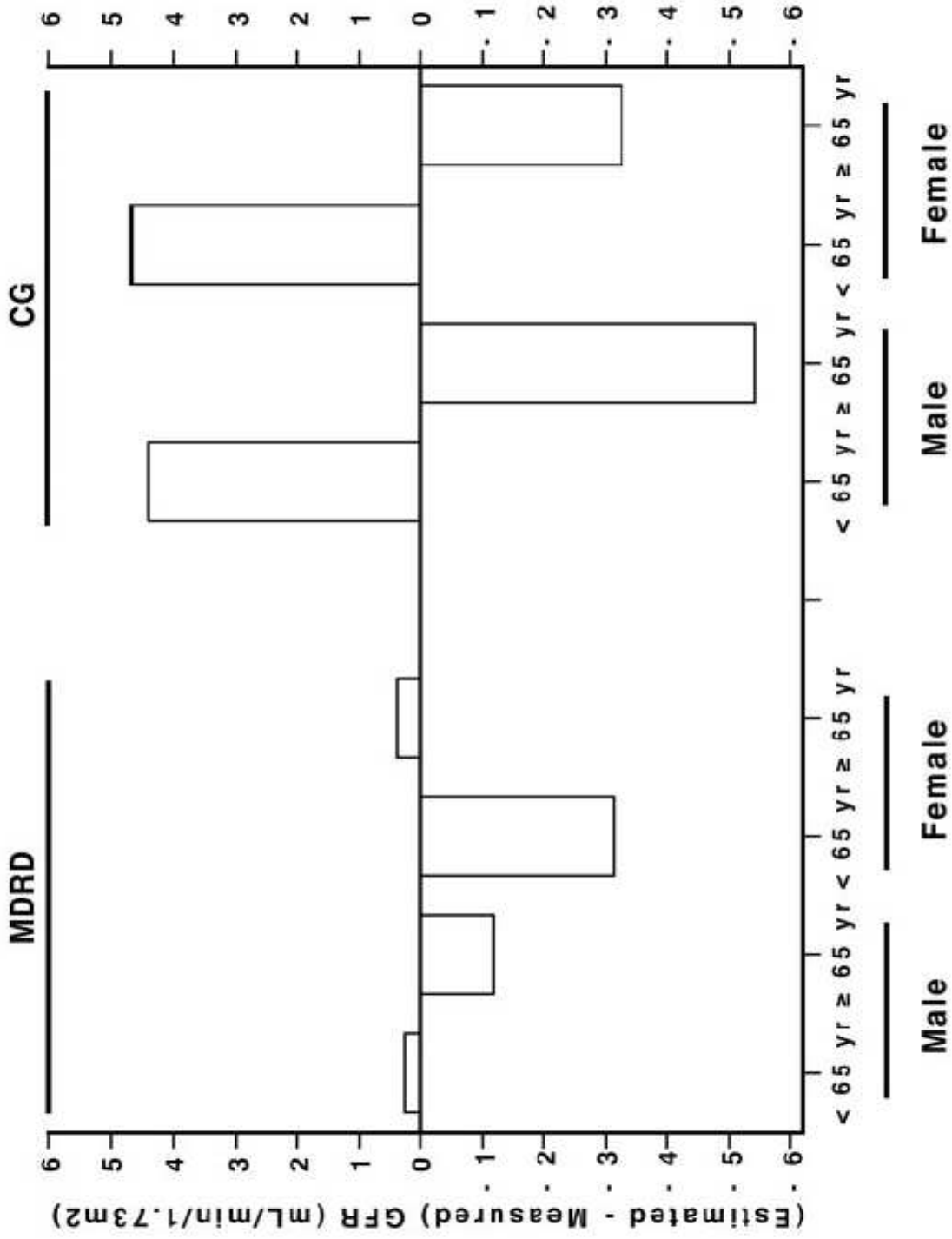
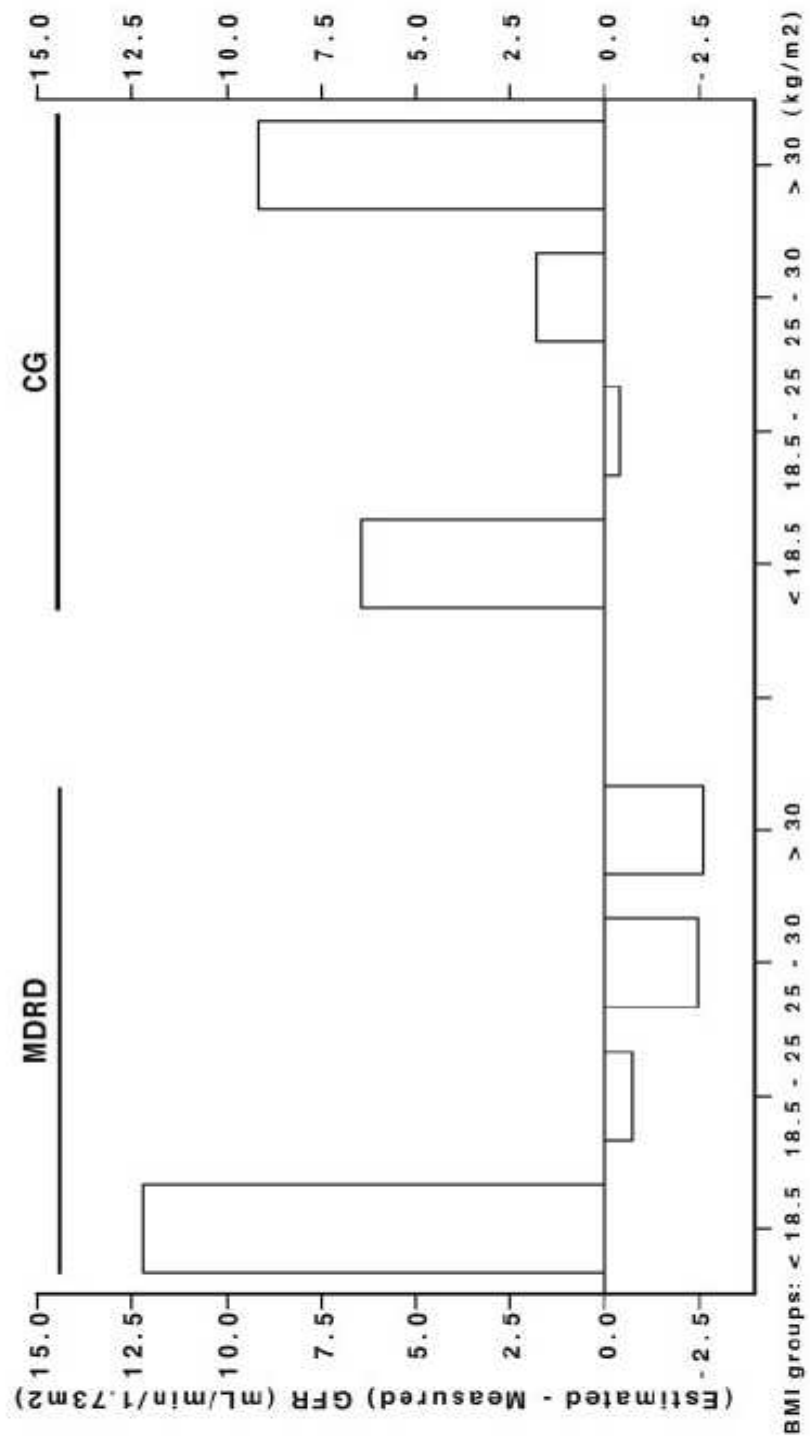


Figure 4. Representation of the mean difference between estimated and measured GFR in the study population. Mean difference shown according to the formula used to estimate GFR and to age groups and gender.



	< 18.5	18.5 - 25	25 - 30	> 30
Overall (N)	94	1,010	712	279
Precision	24.8	13.7	11.4	11.6

	< 18.5	18.5 - 25	25 - 30	> 30
Males (N)	38	550	494	150
Bias	12.1	2.1	-2.7	-2.8
Precision	15.3	12.1	11.5	9.9

Prévalence de l'IRC dans la population générale ?

Quels tests ont été utilisés ?

Nécessité d'un outil simple

■ Créatininémie

- ◆ dosage : variabilité inter-laboratoire (CV 6-8%, écart 15 $\mu\text{mol/l}$), interférences
- ◆ Variabilité inter-individuelle : masse musculaire, alimentation, sécrétion tubulaire...
- ◆ Relation non linéaire entre créatininémie et filtration glomérulaire

■ Formule de Cockcroft et Gault

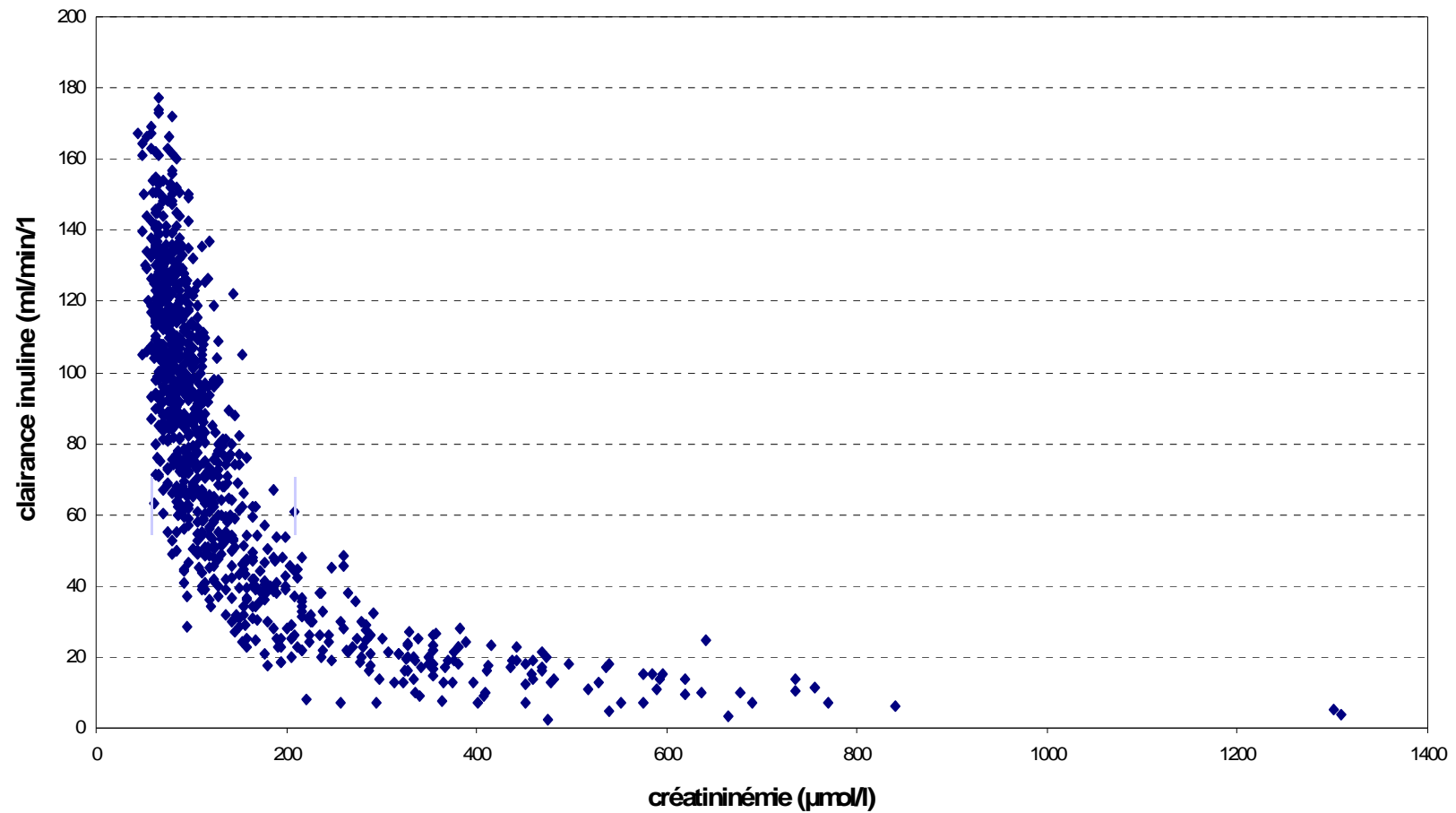
- ◆ Dosage créatininémie
- ◆ Estimation « poids »
- ◆ Recommandations ANAES

■ Formule MDRD

- ◆ Dosage créatininémie
- ◆ Recommandations KDIGO

■ Jellife...

Créatininémie et FG



Utilisation d'un seuil « arbitraire »

L'attrait des chiffres « ronds »

Reykjavik Study (Islande)

- 18 912 adultes
- Seuil créatininémie : 1.7 mg/dl = 150 $\mu\text{mol/l}$
- Prévalence créatininémie > seuil :
 - ◆ h \hat{o} : 0.28 %
 - ◆ f \hat{e} : 0.15 %

Utilisation de seuils « hors normes »

95^{ième} percentile

Normes laboratoires

Framingham (USA)

- 6 233 adultes
- Seuil créatininémie (95^{ième} percentile échantillon « sain »):
 - ◆ h^ô : 136 $\mu\text{mol/l}$ = 1.5 mg/dl,
 - ◆ f^ê : 120 $\mu\text{mol/l}$ = 1.4 mg/dl
- Prévalence créatininémie > seuil :
 - ◆ h^ô : 8.9%
 - ◆ f^ê : 8.0%

Bambui Health and Ageing Study (Brésil)

- 818 (18-59) + 1 494 (> 60 ans) adultes
- Seuil créatininémie (normes laboratoires):
 - ◆ h^ô : 1.3 mg/dl = 115 μ mol/l,
 - ◆ f^ê : 1.1 mg/dl = 98 μ mol/l
- Prévalence créatininémie > seuil :
 - ◆ Adultes 18-59 : 0.48%
 - ◆ Adultes > 60 ans : 5.09%
 - ◆ h^ô > 60 ans : 8.0%
 - ◆ f^ê > 60 ans : 3.19%

Utilisation de seuils « réfléchis »

Seuils de créatininémie calculés selon
l'objectif

Formule estimant la FG

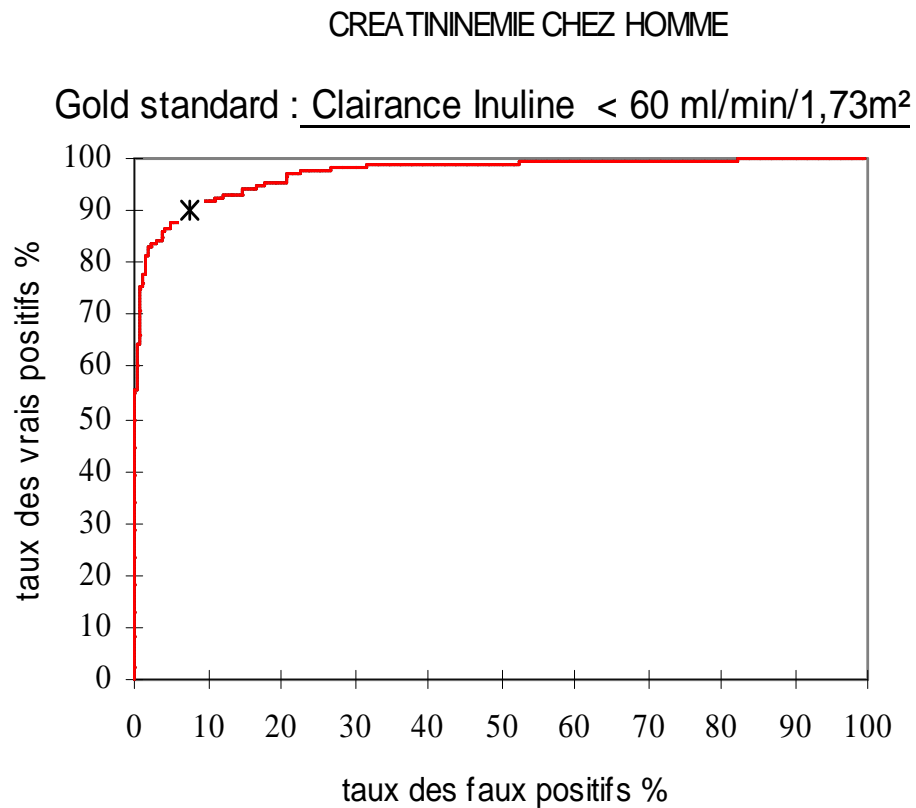
Exemple : la créatininémie

- Connu : répartition des valeurs de créatininémie dans de grands échantillons de population
 - Limites « normales » : IC 95%
- Inconnu : seuil de créatininémie dit « pathologique », fonction d'un objectif de FG
 - Transformation variable continue en une variable qualitative à 2 classes
 - Sensibilité, spécificité, rapport de vraisemblance..

Construction Courbe ROC

- 984 explorations fonctionnelles rénales :
néphropathies, bilan d'HTA, bilan de lithiase,
donneur vivant, volontaires sains ...
- 464 femmes, 520 hommes, âge moyen : 47.6 ans,
cl. Inuline : 81.1 ml/min/1.73m²
- Détermination d'un seuil pour :
 - les 2 sexes
 - créatininémie, Cockcroft
 - 3 degré d'insuffisance rénale (clairance de
l'inuline < 30, 60 et 80 ml/min/1.73m²)
- Bootstrap : 100 * 30% patients

Exemple de courbe ROC



- *Au point d'inflexion
(Indice de Younden) :
- créatininémie : **137 $\mu\text{mol/l}$**
 - sensibilité : 90.0%
 - spécificité : 93.3%
 - rapport de vraisemblance positif : 13.5
 - rapport de vraisemblance négatif : 0.11

Aire sous la courbe : 0.97

Aire sous la courbe ROC

- Burkhardt, 2002 : 30 patients, 70 ml/min/1.73m²
 - ◆ AUC créatininémie = 0.76 (0.58-0.94)
 - ◆ AUC Cockroft = 0.79 (0.61-0.97)
- Couchoud, 1998 : 984 patients, 80-60-30 ml/min
 - ◆ AUC créatininémie chez hô : 0.941, 0.972, 0.991
 - ◆ AUC créatininémie chez fê : 0.914, 0.957, 0.988
 - ◆ AUC Cockroft : 0.936, 0.958, 0.988
 - ◆ L'AUC diminue lorsque le FG cible augmente

NHANES III (USA)

- 13 251 personnes, non diabétiques, >20 ans

Seuils de créatininémie (1.5 mg/dl = 137 μ mol/l et 1.2 mg/dl = 104 μ mol/l),
 Cockcroft, MDRD, MDRD calibré (- 20 μ mol/l)

FG cible	Hommes blancs	Femmes blanches
≥ 80 ml/min/1.73m ²	90.4% 75.1% 45.6% 83.6%	78.8% 46.8% 32% 78.5%
60-79 ml/min/1.73m ²	7.1% 16.7% 45.2% 13.5%	13.4% 32.5% 50.2% 17.4%
30-59 ml/min/1.73m ²	2.2% 7.8% 9% 2.8%	7.4% 19.5% 17.5% 4%
< 30 ml/min/1.73m ²	0.3% 0.4% 0.2% 0.1%	0.4% 1.2% 0.3% 0.2%



-3.1 ± 17.2 ml/min per 1.73 m²), whereas the biases of the CG formula were always significantly larger ($P < 0.0001$).

The precision and the accuracy of the two formulas according to gender and age are reported in Table 6. The MDRD formula was more precise and accurate than the CG one in all subgroups of patients; the only exception was the subgroup of

the drawing of absolute and relative boundaries for the lack of accuracy (Figure 5). In all cases, the MDRD formula was at least as accurate as the CG one. The CG formula principally lacked accuracy in subjects who were younger than 65 yr and had GFR values <60 ml/min per 1.73 m², whereas the accuracy of the MDRD formula was much more uniform (Figure 5B).

Table 7. Classification of the study population according to the MDRD and CG formulas^a

Subjects with Measured GFR (ml/min per 1.73 m ²)	N	Classification Based on the MDRD Formula					Classification Based on the CG Formula				
		Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
		≥ 90 (stage 1)	482	66.8%	32.6%	0.6%	0%	0%	72.2%	27.6%	0.2%
60–89 (stage 2)	576	15.6%	63.7%	20.5%	0.2%	0%	21.7%	58.7%	19.4%	0.2%	0%
30–59 (stage 3)	597	0.5%	11.9%	78.1%	9.5%	0%	0.5%	13.9%	77.9%	7.7%	0%
15–29 (stage 4)	312	0%	0.3%	16.7%	78.8%	4.2%	0%	0.6%	28.8%	67.6%	2.9%
<15 (stage 5)	128	0%	0%	3.1%	32.0%	64.8%	0%	0%	3.1%	53.9%	43.0%

^aMeasured GFR was used to divide the study population into five categories corresponding the five stages of CKD in the K/DOQI CKD classification (5). For each category, the subjects then were reclassified according to the MDRD formula and to the CG formula. Numbers in bold correspond to the percentages of subjects who did not change stage when their GFR level was estimated using a creatinine-based formula. The existence of kidney damage was not taken into account for this analysis.

Conclusion

- Prévalence d'un test positif \neq prévalence de la maladie
- Augmenter la probabilité pré-test : cibler les populations chez qui on effectue le dépistage
- Choisir un seuil selon l'objectif dépistage ou diagnostic, selon la cible = équilibre spécificité/sensibilité à priori

Predictive Performance of the Modification of Diet in Renal Disease and Cockcroft-Gault Equations for Estimating Renal Function

Marc Froissart,^{*†§} Jerome Rossert,^{||} Christian Jacquot,^{‡§} Michel Paillard,^{*†§} and Pascal Houillier^{*†§}

^{*}Department of Physiology and Biophysics, Georges Pompidou Hospital (AP-HP); [†]INSERM U1652 and IFR 58;

[‡]Department of Nephrology, Georges Pompidou Hospital (AP-HP); [§]René Descartes Medical School, Paris V University; and ^{||}Paris VI University, Paris, France

Recent recommendations emphasize the need to assess kidney function using creatinine-based predictive equations to optimize the care of patients with chronic kidney disease. The most widely used equations are the Cockcroft-Gault (CG) and the simplified Modification of Diet in Renal Disease (MDRD) formulas. However, they still need to be validated in large samples of subjects, including large non-U.S. cohorts. Renal clearance of ⁵¹Cr-EDTA was compared with GFR estimated using either the CG equation or the MDRD formula in a cohort of 2095 adult Europeans (863 female and 1232 male; median age, 53.2 yr; median measured GFR, 59.8 ml/min per 1.73 m²). When the entire study population was considered, the CG and MDRD equations showed very limited bias. They overestimated measured GFR by 1.94 ml/min per 1.73 m² and underestimated it by 0.99 ml/min per 1.73 m², respectively. However, analysis of subgroups defined by age, gender, body mass index, and GFR level showed that the biases of the two formulas could be much larger in selected populations. Furthermore, analysis of the SD of the mean difference between estimated and measured GFR showed that both formulas lacked precision; the CG formula was less precise than the MDRD one in most cases. In the whole study population, the SD was 15.1 and 13.5 ml/min per 1.73 m² for the CG and MDRD formulas, respectively. Finally, 29.2 and 32.4% of subjects were misclassified when the CG and MDRD formulas were used to categorize subjects according to the Kidney Disease Outcomes Quality Initiative chronic kidney disease classification, respectively.

J Am Soc Nephrol 16: 763–773, 2005. doi: 10.1681/ASN.2004070549

The prevalent and incident rates of ESRD are continuously increasing in all Western countries. Data from the U.S. Renal Data System predict that the number of patients who were registered with ESRD in 1997 will have doubled in 2010, leading to approximately 700,000 patients with ESRD and 2.2 million patients in 2030 (1), and similar trends are anticipated in other countries (2–4). To level off these incident rates, various initiatives, such as the Kidney Disease Outcomes Quality Initiative (K/DOQI), have provided physicians with guidelines to optimize the care of patients with chronic kidney disease (CKD). These guidelines emphasize the need to assess kidney function using predictive equations rather than serum creatinine alone (5). However, they also highlight that these equations still need to be validated in large samples of subjects, in particular that they should be tested in non-U.S. populations and in individuals with mild decrease in kidney function or normal GFR (5). Validation of the predictive formulas is also particularly important for patients aged 65 and older, who by far have the highest incident rates of ESRD (1,6,7).

The formulas that are most widely used to estimate kidney function and that are recommended in adults by the K/DOQI guidelines (5) are the Cockcroft-Gault (CG) formula (8) and the recently developed (9) and later simplified (10) Modification of Diet in Renal Disease (MDRD) formula. The CG formula is an estimate of creatinine clearance originally developed in a population of 236 Canadian patients, 209 of which were male. The MDRD formulas have been developed as an estimation of ¹²⁵I-iothalamate renal clearance–based GFR measurement in a population of 1628 patients with previously diagnosed CKD (9–11). The mean GFR in this population was 39.8 ± 21.2 ml/min per 1.73 m², and the mean age of the cohort was 50.6 ± 12.7 yr.

The K/DOQI CKD guidelines have established a five-stage classification of patients with CKD that is based solely on kidney function. These stages are defined by GFR ≥ 90 ml/min per 1.73 m² (stage 1), 60 to 89 ml/min per 1.73 m² (stage 2), 30 to 59 ml/min per 1.73 m² (stage 3), 15 to 29 ml/min per 1.73 m² (stage 4), and <15 ml/min per 1.73 m² (stage 5) (5). The guidelines state that the stage of kidney disease should be determined for each CKD patient and that a clinical action plan should be developed on the basis of the stage of disease (5). Thus, inaccurate estimation of kidney function may be responsible for misclassification of some patients and lead to inappropriate evaluation or treatment of these patients (12). However, so far, few studies have assessed the applicability of the MDRD

Received July 13, 2004. Accepted November 24, 2004.

Publication schedule of *Journal of the American Society of Nephrology* at www.jasn.org.

Address correspondence to Dr. Marc Froissart, Department of Physiology and Biophysics, Georges Pompidou Hospital, 20 rue Leblanc, 75015 Paris, France. Phone: +33-1-5603-3073; fax: +33-1-5603-2675; E-mail: marc.froissart@ap-hopital

Copyright © 2005 by the American Society of Nephrology

ISSN: 1046-6673/16(7)763-773