

# Αξιολόγηση εργαστηριακών ευρημάτων

## Γιώτα Τουλούμη

Καθηγήτρια Βιοστατιστικής και Επιδημιολογίας  
Εργ. Υγιεινής, Επιδημιολογίας και Ιατρικής Στατιστικής  
Ιατρική Σχολή, ΕΚΠΑ  
gtouloum@med.uoa.gr

## Βάνα Σύψα

Καθηγήτρια Επιδημιολογίας και Ιατρικής Στατιστικής  
Εργ. Υγιεινής, Επιδημιολογίας και Ιατρικής Στατιστικής  
Ιατρική Σχολή, ΕΚΠΑ  
vsipsa@med.uoa.gr



# Είδη δοκιμασιών

Π.χ. έλεγχος επιπέδων χοληστερόλης, έλεγχος για παρουσία αντισωμάτων ηπατίτιδας C

- **Ποιοτικές (Qualitative)**

- Με 2 επίπεδα (π.χ. +/-)
- Με >2 επίπεδα, συνήθως διατάξιμες (π.χ. Normal, Suspect, Abnormal)

- **Ποσοτικές (Quantitative)**

- Π.χ. επίπεδα χοληστερόλης

# Ποσοτικές δοκιμασίες

# Χαρακτηριστικά επίδοσης (Performance characteristics)

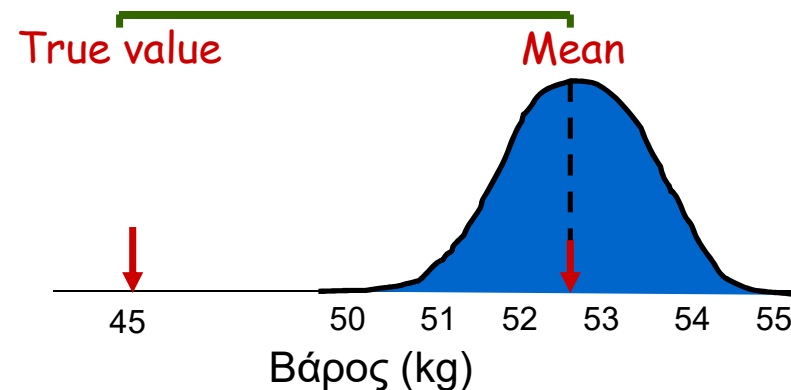
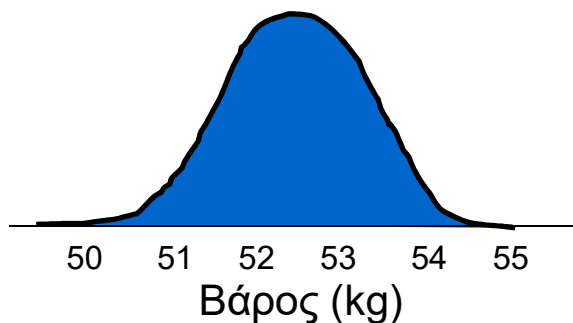
Η μέτρηση εκφράζει την πραγματικότητα;  
**Ακρίβεια/Αυθεντικότητα (Accuracy)**

## Πιστότητα (precision)

Εκφράζει τη διακύμανση στα αποτελέσματα όταν η δοκιμασία επαναλαμβάνεται στο ίδιο δείγμα λόγω **τυχαίου σφάλματος**

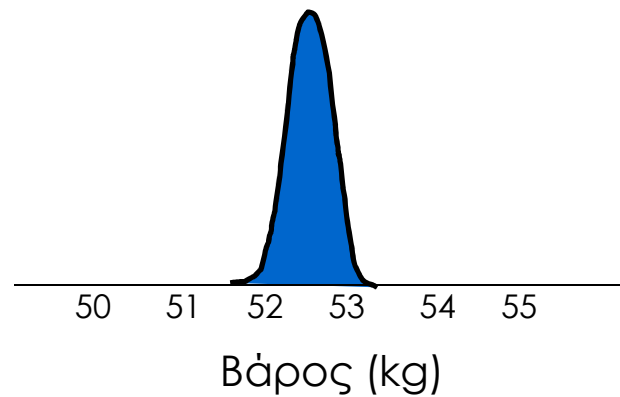
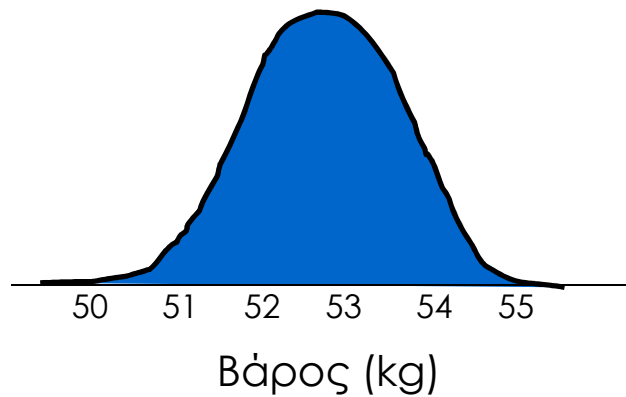
## Ορθότητα (trueness)

Εκφράζει τη διαφορά του μέσου όρου των μετρήσεων από την πραγματική τιμή → Εκφράζει το **συστηματικό σφάλμα**



# Πιστότητα (precision)

- Μία μέθοδος χαρακτηρίζεται από πιστότητα αν η τιμή του μεγέθους είναι σχεδόν ίδια κάθε φορά που μετριέται



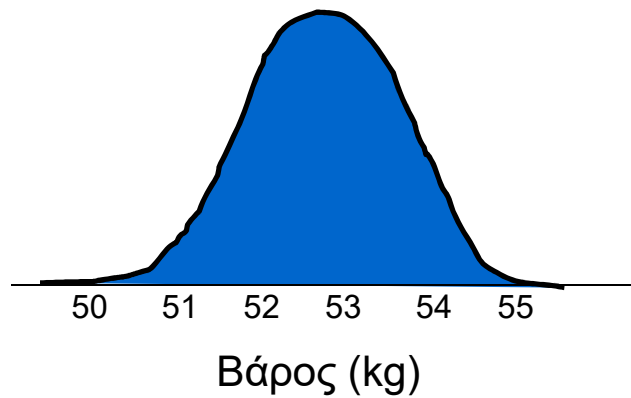
- Η πιστότητα απειλείται από **τυχαία σφάλματα**  
π.χ. διακύμανση στα αποτελέσματα μίας δοκιμασίας που οφείλεται στο ότι την πραγματοποιούν διαφορετικοί εξεταστές (ή μηχανήματα ή μέθοδοι)

# Πως αξιολογείται η πιστότητα μίας ποσοτικής μεθόδου;

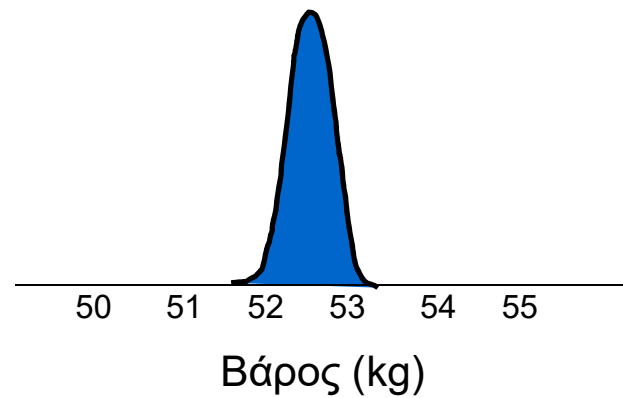
- Η διαπίστωση ενός τυχαίου σφάλματος δεν είναι δυνατή όταν υπάρχει μία μόνο μέτρηση → Η πιστότητα αξιολογείται μέσω **επανελημμένων μετρήσεων**
- Πώς αξιολογούνται τα δεδομένα;

## Σταθερή απόκλιση (Standard Deviation)

- Αποτελεί μέτρο της διασποράς των επανειλημμένων μετρήσεων γύρω από τη μέση τιμή τους



↓  
Μεγάλη διασπορά γύρω  
από τη μέση τιμή



↓  
Μικρή διασπορά γύρω από  
τη μέση τιμή

# Συντελεστής μεταβλητότητας (Coefficient of Variation)

- Επανειλημμένες μετρήσεις στην ίδια ερευνητική μονάδα και υπολογισμός μέσης τιμής και σταθερής απόκλισης (SD)

$$CV = \frac{\text{Standard deviation of measurements}}{\text{Mean of measurements}} \times 100$$

- Όσο **μικρότερο** το CV, τόσο μικρότερη η μεταβλητότητα μεταξύ μετρήσεων  
→ **Καλύτερη** πιστότητα



## Συντελεστής μεταβλητότητας (CV): Παράδειγμα

- π.χ. Επανειλημμένες μετρήσεις του ίδιου δείγματος

	113
	99
	118
	120
	132
<b>Μέση τιμή</b>	<b>116.4</b>
<b>Σταθ. Απόκλιση</b>	<b>12.0</b>
<b>CV</b>	<b>10.3%</b>

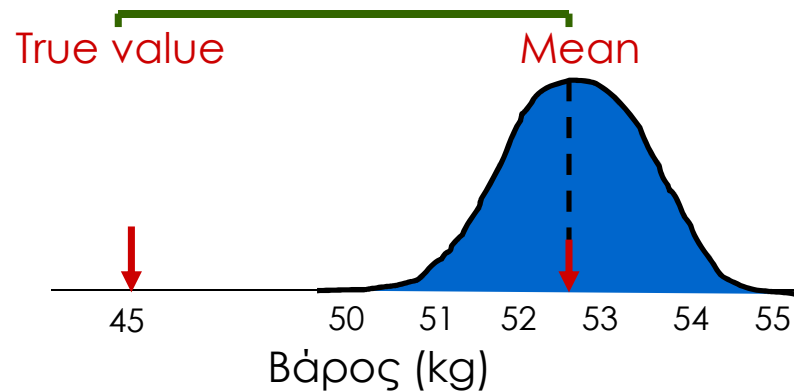
# Συντελεστής μεταβλητότητας (CV): Παράδειγμα

- π.χ. Επανειλημμένες μετρήσεις του ίδιου δείγματος με 2 δοκιμασίες

	<b>A</b>	<b>B</b>
	113	115
	99	116
	118	115
	120	117
	132	113
<b>Μέση τιμή</b>	<b>116.4</b>	<b>115.2</b>
<b>Σταθ. Απόκλιση</b>	<b>12.0</b>	<b>1.5</b>
<b>CV</b>	<b>10.3%</b>	<b>1.3%</b>

# Ορθότητα (Trueness)

- Εκφράζει τη διαφορά του μέσου όρου των μετρήσεων από την πραγματική τιμή



Γιατί του «μέσου όρου» των μετρήσεων;

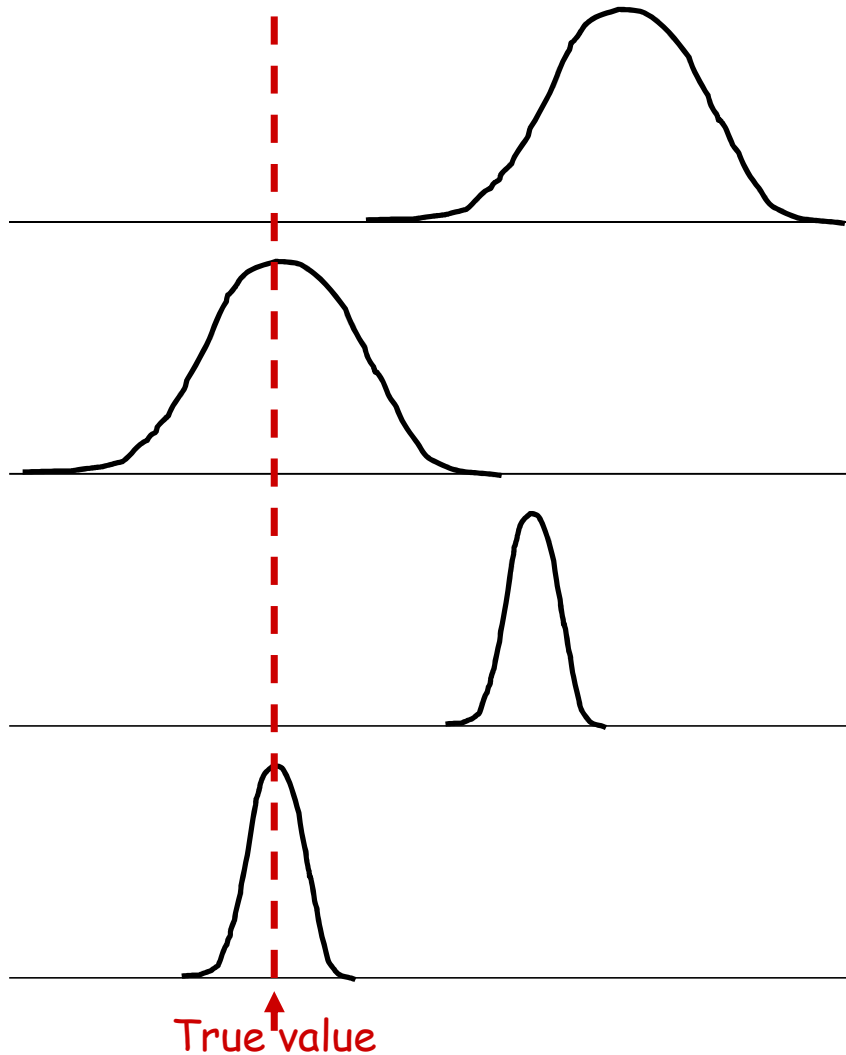
- Για να απαλειφθεί το τυχαίο σφάλμα

→ Η ορθότητα απειλείται από **συστηματικά σφάλματα**

# Πως αξιολογείται η ορθότητα μίας μεθόδου;

- Απαιτείται η ύπαρξη μίας άλλης δοκιμασίας η οποία θεωρείται ότι μετρά την πραγματική τιμή (**reference method, gold standard**)
- Τα δείγματα πρέπει να αξιολογηθούν και με τις δύο δοκιμασίες

Συμπερασματικά σε ποσοτικές δοκιμασίες:  
Ορθότητα (trueness), πιστότητα (precision), ακρίβεια (accuracy)

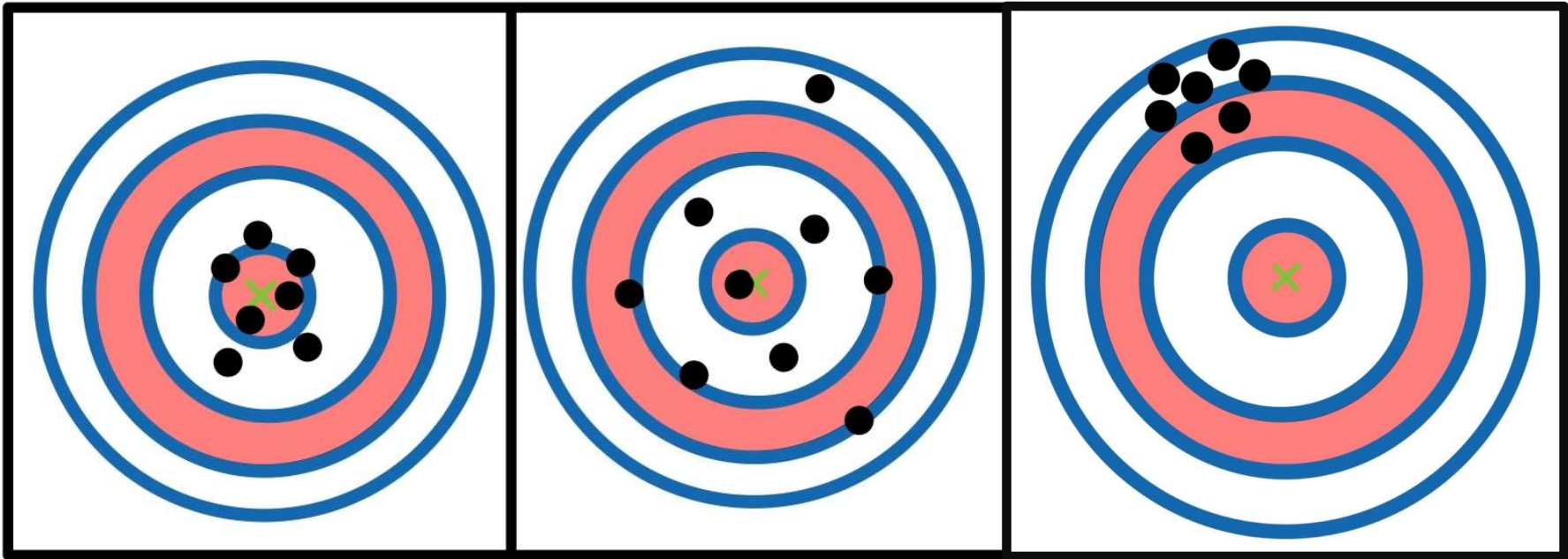


Κακή πιστότητα  
Κακή ορθότητα  
**Κακή ακρίβεια**

Κακή πιστότητα  
Καλή ορθότητα  
**Κακή ακρίβεια**

Καλή πιστότητα  
Κακή ορθότητα  
**Κακή ακρίβεια**

Καλή πιστότητα  
Καλή ορθότητα  
**Καλή ακρίβεια**



# Ποιοτικές δοκιμασίες

# Ερωτήματα

- Αν η μέτρηση επαναληφθεί, θα υπάρχει διαφορά στα ευρήματα;
- Η ποιοτική μέθοδος ανιχνεύει ικανοποιητικά τους ασθενείς – υγιείς;
- Αν έχω μία ποσοτική μέθοδο και θέλω να διαχωρίσω με βάση τα αποτελέσματά της τα άτομα σε ασθενείς-υγιείς, ποιο cut-off είναι καταλληλότερο;



# Ερωτήματα

- Αν η μέτρηση επαναληφθεί, θα υπάρχει διαφορά στα ευρήματα; (μέτρηση στο ίδιο δείγμα αλλά σε άλλη χρονική στιγμή ή από άλλο ερευνητή ή με άλλη μέθοδο)
- Η ποιοτική μέθοδος ανιχνεύει ικανοποιητικά τους ασθενείς – υγιείς;
- Αν έχω μία ποσοτική μέθοδο και θέλω να διαχωρίσω με βάση τα αποτελέσματά της τα άτομα σε ασθενείς-υγιείς, ποιο cut-off είναι καταλληλότερο;

# Ποσοστό συμφωνίας (% agreement)

- Αξιολόγηση των ίδιων δεδομένων από δύο ή περισσότερες μέθοδους (ή ερευνητές).

	Μέθοδος Β	
	+	-
Μέθοδος Α	+	-
	α	β
	-	δ
	γ	δ

- Τέλεια συμφωνία:  $\beta=0$  &  $\gamma=0$

- Παρατηρηθέν ποσοστό συμφωνίας =  $\frac{a + \delta}{a + \beta + \gamma + \delta} \times 100$

# Ποσοστό συμφωνίας: Παράδειγμα

		Ακτινολόγος 2	
		+	-
Ακτινολόγος 1	+	4	5
	-	2	6

$$\% \text{ συμφωνίας} = (4+6)/(4+5+2+6) = 58.8\%$$

# Μειονέκτημα του % συμφωνίας

- Συμφωνία μπορεί να υπάρξει και κατά τύχη
  - Π.χ. Αν δύο εξεταστές αξιολογούσαν την ίδια σειρά δειγμάτων ως θετικά ή αρνητικά εντελώς τυχαία, σε κάποιες περιπτώσεις τα αποτελέσματά τους θα συμφωνούσαν
- Δείκτης που διορθώνει για την ύπαρξη τυχαίας συμφωνίας → **kappa statistic**

# Kappa statistic

	Μέθοδος Β		
	+	-	
Μέθοδος Α	+	α	β
	-	γ	δ

$$\text{Kappa} = \frac{\% \text{ συμφωνίας που παρατηρήθηκε}(I_o) - \% \text{ τυχαίας συμφωνίας}(I_e)}{1 - \% \text{ τυχαίας συμφωνίας}(I_e)}$$

- Παίρνει τιμές από  $<0$  έως  $1$
- Αν η συμφωνία που παρατηρήθηκε μικρότερη από αυτή που θα παρατηρούσαμε κατά τύχη  $\rightarrow \text{kappa} < 0$
- Τιμές κοντά στο  $1$  δείχνουν καλή συμφωνία

# Υπολογισμός τυχαίας (αναμενόμενης) συμφωνίας

Μέθοδος Β

Μέθοδος Α

	+	-	Σύνολο
+	a	β	$r_1$
-	γ	δ	$r_2$
Σύνολο	$c_1$	$c_2$	N

- Αναμενόμενη συχνότητα στα (+):  $\frac{c_1 \cdot r_1}{N}$
- Αναμενόμενη συχνότητα στα (-):  $\frac{c_2 \cdot r_2}{N}$
- Συνολικά  $\frac{c_1 \cdot r_1 + c_2 \cdot r_2}{N}$
- Τελικά, αναμενόμενο ποσοστό συμφωνίας =  $\frac{\frac{c_1 \cdot r_1 + c_2 \cdot r_2}{N}}{N} = (c_1 \cdot r_1 + c_2 \cdot r_2) / N^2$

# Παραδείγματα

Ερευνητής B

	+	-	Σύνολο
Ερευνητής A			
+	5	5	10
-	5	85	90
Σύνολο	10	90	100

$$I_o = (85+5)/100 = 0,90$$

$$I_e = (10*10+90*90)/100^2 = 0,82$$

$$k = \frac{I_o - I_e}{1 - I_e} = \frac{0,90 - 0,82}{1 - 0,82} = 0,44$$

Ερευνητής B

	+	-	Σύνολο
Ερευνητής A			
+	35	5	40
-	5	55	60
Σύνολο	40	60	100

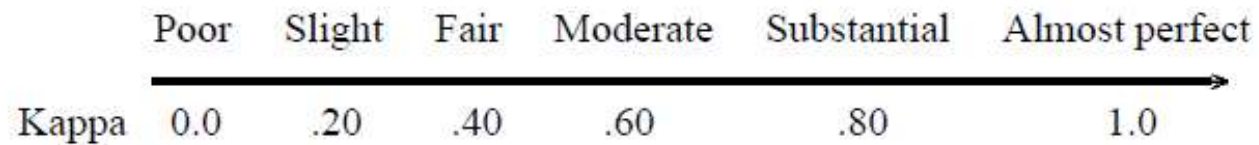
$$I_o = (35+55)/100 = 0,90$$

$$I_e = (40*40+60*60)/100^2 = 0,52$$

$$k = \frac{I_o - I_e}{1 - I_e} = \frac{0,90 - 0,52}{1 - 0,52} = 0,79$$

# Kappa statistic - ερμηνεία

## Interpretation of Kappa



<u>Kappa</u>	<u>Agreement</u>
< 0	Less than chance agreement
0.01–0.20	Slight agreement
0.21– 0.40	Fair agreement
0.41–0.60	Moderate agreement
0.61–0.80	Substantial agreement
0.81–0.99	Almost perfect agreement



# Καρρα statistic: Παράδειγμα

	Ακτινολόγος 2	
	+	-
Ακτινολόγος 1	+	5
	-	6

% συμφωνίας : 58.8%

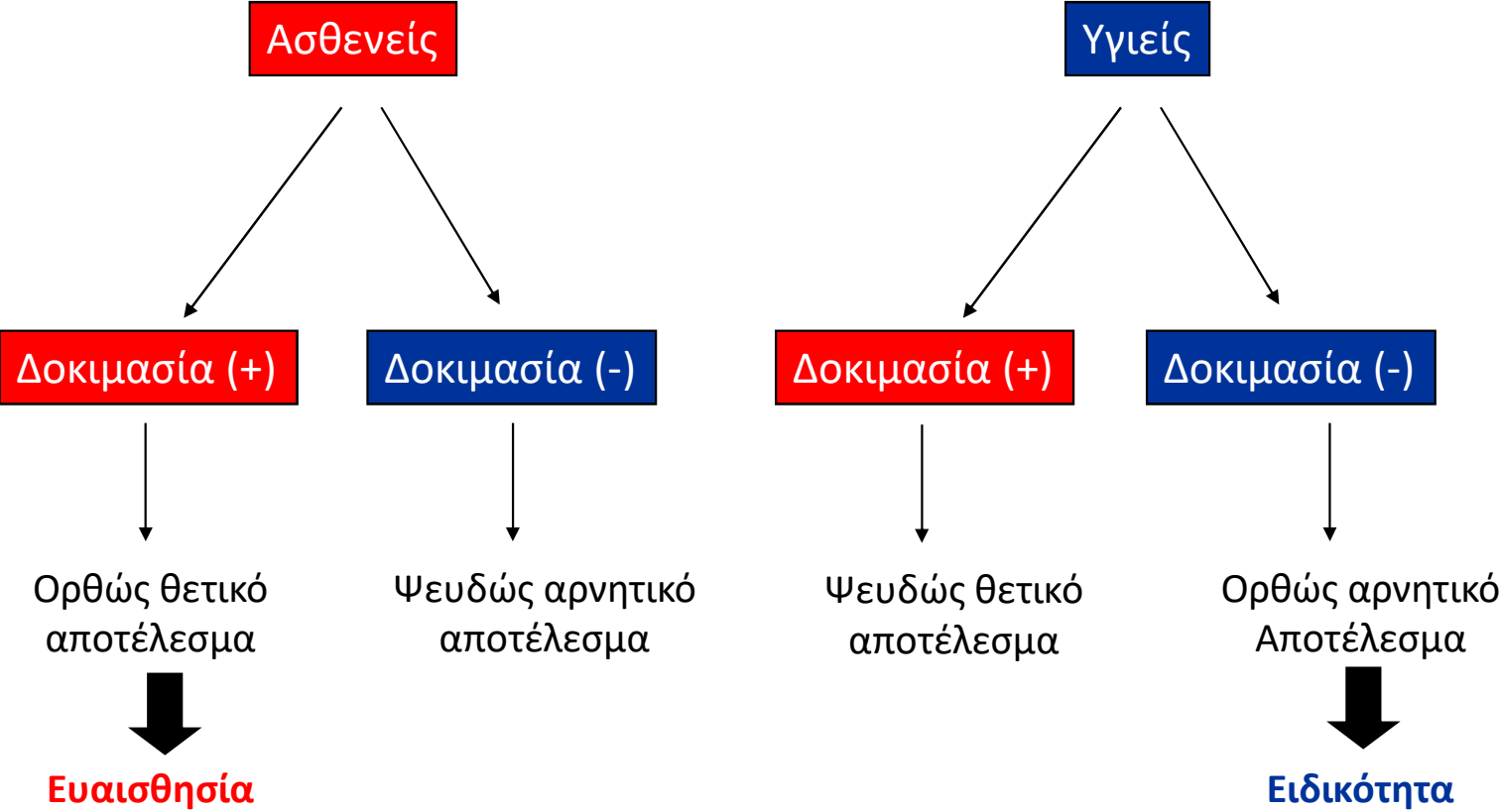
% τυχαίας συμφωνίας: 49.2%

Καρρα = 0.19 (slight agreement)

# Ερωτήματα

- Αν η μέτρηση επαναληφθεί, θα υπάρχει διαφορά στα ευρήματα;
- **Η ποιοτική μέθοδος ανιχνεύει ικανοποιητικά τους ασθενείς – υγιείς;**
- Αν έχω μία ποσοτική μέθοδο και θέλω να διαχωρίσω με βάση τα αποτελέσματά της τα άτομα σε ασθενείς-υγιείς, ποιο cut-off είναι καταλληλότερο;

# Ευαισθησία και ειδικότητα μίας δοκιμασίας



# Ευαισθησία - Ειδικότητα

		Νόσος (D)	
		Ναι	Όχι
Αποτέλεσμα δοκιμασίας (T)	(+)	a ορθώς (+)	b ψευδώς (+)
	(-)	c ψευδώς (-)	d ορθώς(-)
Σύνολο		<b>a+c</b>	<b>b+d</b>

- **Ευαισθησία** =  $a/(a+c)$  ορθώς θετικά
  - Η αναλογία των ασθενών για τους οποίους η δοκιμασία δίνει το σωστό αποτέλεσμα (δηλαδή +)  $\rightarrow P(T+ | D+)$
- **Ειδικότητα** =  $d/(b+d)$  ορθώς αρνητικά
  - Η αναλογία των υγιών για τους οποίους η δοκιμασία δίνει το σωστό αποτέλεσμα (δηλαδή -)  $\rightarrow P(T- | D-)$

# Ευαισθησία

		Σιδηροπενική αναιμία		Σύνολο
		Ναι	Όχι	
Φερριτίνη ορού	(+) ( $<65$ mmol/L)	731 ↓ a	270 b	1001 a+b
	(-) ( $\geq 65$ mmol/L)	78 c	1500 d	1578 c+d
Σύνολο		a+c 809	b+d 1770	2579

Ευαισθησία:

Μεταξύ των ασθενών (a+c), η αναλογία αυτών με θετική δοκιμασία (a)

$$\text{Ευαισθησία} = \frac{a}{a + c} = \frac{731}{809} = 90\%$$

# Ειδικότητα

		Σιδηροπενική αναιμία		Σύνολο
		Ναι	Όχι	
Φερριτίνη ορού	(+) (<65 nmol/L)	731 <i>a</i>	270 <i>b</i>	1001 <i>a+b</i>
	(-) (≥65 nmol/L)	78 <i>c</i>	1500 <i>d</i> ↓	1578 <i>c+d</i>
Σύνολο		809 <i>a+c</i>	1770 <i>b+d</i>	2579

Ειδικότητα:

Μεταξύ των υγιών ( $b+d$ ), η αναλογία αυτών με αρνητική δοκιμασία ( $d$ )

$$\text{Ειδικότητα} = \frac{d}{b+d} = 1500/1770 = 85\%$$

# Ερμηνεία

- Φερριτίνη ορού με cut-off 65 mmol/L:
  - **Ευαισθησία = 90%**  
Δηλαδή από 100 ασθενείς, οι 90 θα διαγνωστούν σωστά ως θετικοί με βάση τη δοκιμασία της φερριτίνης → οι υπόλοιποι 10 **ψευδώς αρνητικοί**
  - **Ειδικότητα = 85%**  
Δηλαδή από 100 υγιείς, οι 85 θα διαγνωστούν σωστά ως αρνητικοί με βάση τη δοκιμασία της φερριτίνης → οι υπόλοιποι 15 **ψευδώς θετικοί**

		Νόσος (D)	
		Ναι	Όχι
Αποτέλεσμα δοκιμασίας (T)	(+)	a ορθώς (+)	b ψευδώς (+)
	(-)	c ψευδώς (-)	d ορθώς(-)
Σύνολο		<b>a+c</b>	<b>b+d</b>

- Αν ευαισθησία=100%
  - $a/(a+c)=100\% \rightarrow c=0$
  - Όχι ψευδώς αρνητικά αποτελέσματα
- Αν ειδικότητα=100%
  - $d/(b+d) =100\% \rightarrow b=0$
  - Όχι ψευδώς θετικά αποτελέσματα
  - Ευρήματα με ειδικότητα κοντά στο 100% → παθογνωμονικά



# Rapid antigen tests για SARS-CoV-2

**Minimum requirements: 80% sensitivity & 97% specificity**

(θεωρώντας σαν gold-standard την PCR)

- Τι σημαίνει αυτό;
  - Ένα 20% των ασθενών θα διαγνωστεί αρνητικό
  - Ένα 3% των υγιών θα διαγνωστεί θετικό
- Η ευαισθησία και η ειδικότητα εξαρτώνται από παράγοντες όπως σε ποια χρονική στιγμή από τη μόλυνση λήφθηκε το δείγμα, το είδος και η ποιότητα του δείγματος

# Θετική & αρνητική διαγνωστική αξία

- Όταν πραγματοποιούμε μία δοκιμασία, λαμβάνουμε αποφάσεις με βάση το αποτέλεσμα της
  - Αν (+), πόσο σίγουροι είμαστε ότι το άτομο **έχει** την ασθένεια;
  - Αν (-), πόσο σίγουροι είμαστε ότι το άτομο **δεν έχει** την ασθένεια;
- **Θετική διαγνωστική αξία (ΘΔΑ)**  
Η πιθανότητα υπάρξεως της νόσου όταν η δοκιμασία είναι (+) →  **$P(D+ | T+)$**
- **Αρνητική διαγνωστική αξία (ΑΔΑ)**  
Η πιθανότητα να μην υπάρχει η νόσος όταν η δοκιμασία είναι (-) →  **$P(D- | T-)$**

## Θετική Διαγνωστική Αξία - ΘΔΑ (Positive predictive value –PPV)

		Σιδηροπενική αναιμία		Σύνολο
		Ναι	Όχι	
Φερριτίνη ορού	(+) ( $<65$ mmol/L)	731 <i>a</i>	270 <i>b</i>	1001 <i>a+b</i>
	(-) ( $\geq 65$ mmol/L)	78 <i>c</i>	1500 <i>d</i>	1578 <i>c+d</i>
	Σύνολο	<i>a+c</i> 809	<i>b+d</i> 1770	2579

PPV:

Μεταξύ των ατόμων με θετική δοκιμασία ( $a+b$ ), η αναλογία αυτών που είναι πραγματικά ασθενείς ( $a$ )

$$PPV = \frac{a}{a+b} = 731/1001 = 73\%$$

## Αρνητική Διαγνωστική Αξία - ΑΔΑ (Negative predictive value – NPV)

		Σιδηροπενική αναιμία		Σύνολο
		Ναι	Όχι	
Φερριτίνη ορού	(+) (<65 mmol/L)	731 <i>a</i>	270 <i>b</i>	1001 <i>a+b</i>
	(-) (≥65 mmol/L)	78 <i>c</i>	1500 <i>d</i>	1578 <i>c+d</i>
Σύνολο		809 <i>a+c</i>	1770 <i>b+d</i>	2579

**NPV:**

Μεταξύ των ατόμων με αρνητική δοκιμασία (*c+d*), η αναλογία αυτών που είναι πραγματικά υγιείς (*d*)

$$\text{NPV} = \frac{d}{c + d} = \frac{1500}{1578} = 95\%$$

# Φερριτίνη ορού με cut-off 65 mmol/L για τη διάγνωση σιδηροπενικής αναιμίας

- **Ευαισθησία: 90%** → από τους 100 ασθενείς, οι 90 διαγιγνώσκονται (+)
  - 90% των ασθενών διαγιγνώσκεται σωστά
  - 10% ψευδώς αρνητικά
- **Ειδικότητα: 85%** → από τους 100 υγιείς, οι 85 διαγιγνώσκονται (-)
  - 85% των υγιών έχει αρνητικό αποτέλεσμα
  - 15% ψευδώς θετικά
- **ΘΔΑ: 73%** → από τους 100 με (+) τεστ, οι 73 είναι ασθενείς
  - Ένα άτομο με θετικό αποτέλεσμα φερριτίνης έχει 73% πιθανότητα να πάσχει από σιδηροπενική αναιμία
- **ΑΔΑ: 95%** → από τους 100 με (-) τεστ, οι 95 είναι υγιείς
  - Ένα άτομο με αρνητικό αποτέλεσμα φερριτίνης έχει 95% πιθανότητα να μην πάσχει από σιδηροπενική αναιμία

## Από τι εξαρτώνται αυτές οι παράμετροι;

- **Ευαισθησία – Ειδικότητα:**

Χαρακτηριστικά της δοκιμασίας

- **ΘΔΑ – ΑΔΑ:**

→ Εξαρτώνται από την ευαισθησία, την ειδικότητα και τον **επιπολασμό\*** του νοσήματος στον υπό μελέτη πληθυσμό

\*επιπολασμός=% ατόμων με το νόσημα στον πληθυσμό

# Θετική διαγνωστική αξία και επιπολασμός

		Διάγνωση		
		Ναι	Όχι	Σύνολο
Εργαστ. εύρημα	+	450	50	500
	-	50	450	500
	Σύνολο	500	500	1000

- Ευαισθησία= $450/500=0,9$
- Ειδικότητα= $450/500=0,9$
- Επιπολασμός= $500/1000=0,5$
- **Θετική Διαγνωστική αξία=**  
**= $450/500=0,9$**

		Διάγνωση		
		Ναι	Όχι	Σύνολο
Εργαστ. εύρημα	+	90	90	180
	-	10	810	820
	Σύνολο	100	900	1000

- Ευαισθησία= $90/100=0,9$
- Ειδικότητα= $810/900=0,9$
- Επιπολασμός= $100/1000=0,1$
- **Θετική Διαγνωστική αξία=**  
**=  $90/180=0,5$**

# Ερωτήματα

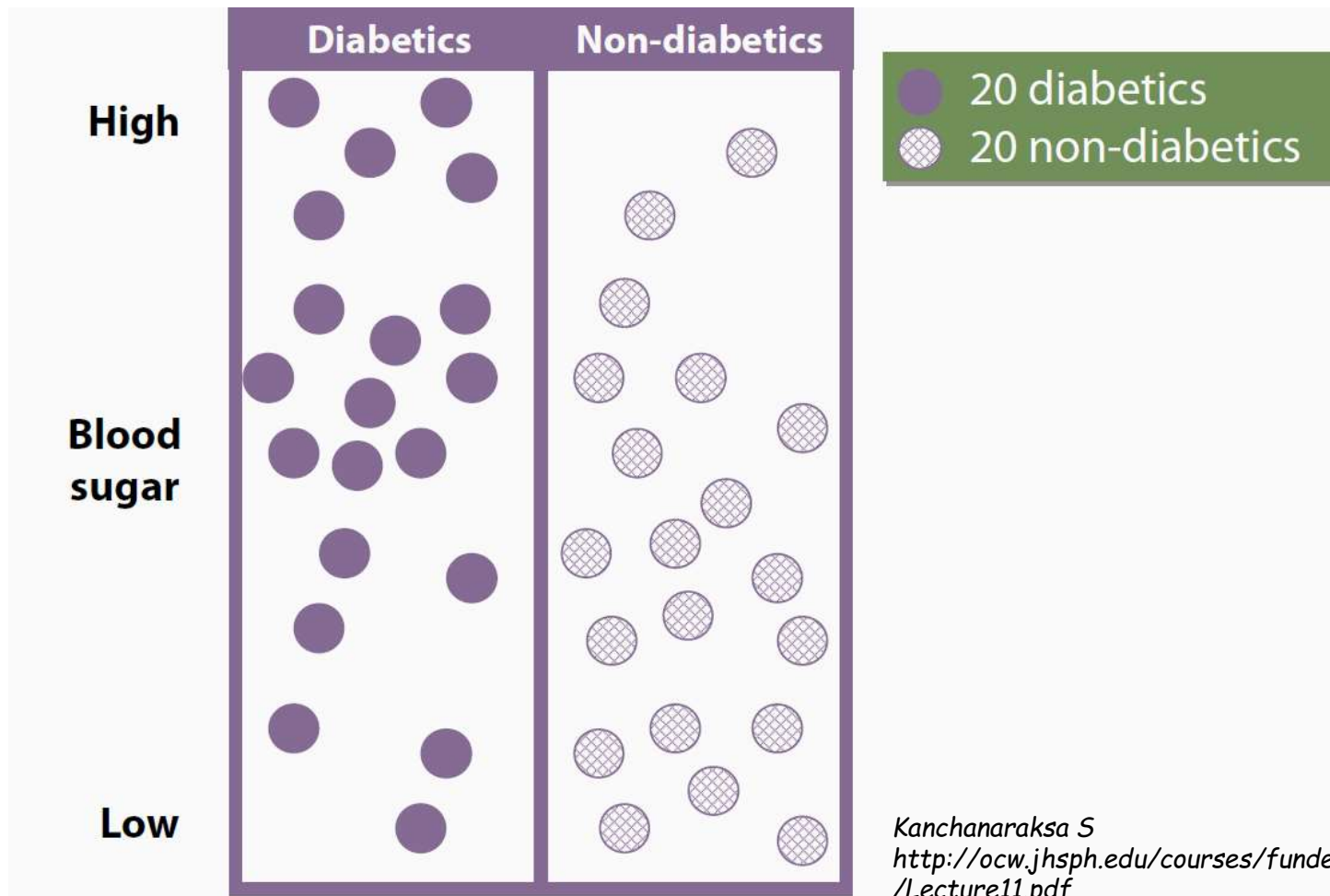
- Αν η μέτρηση επαναληφθεί, θα υπάρχει διαφορά στα ευρήματα;
- Η ποιοτική μέθοδος ανιχνεύει ικανοποιητικά τους ασθενείς – υγιείς;
- Αν έχω μία ποσοτική μέθοδο και θέλω να διαχωρίσω με βάση τα αποτελέσματά της τα άτομα σε ασθενείς-υγιείς, ποιο cut-off είναι καταλληλότερο;



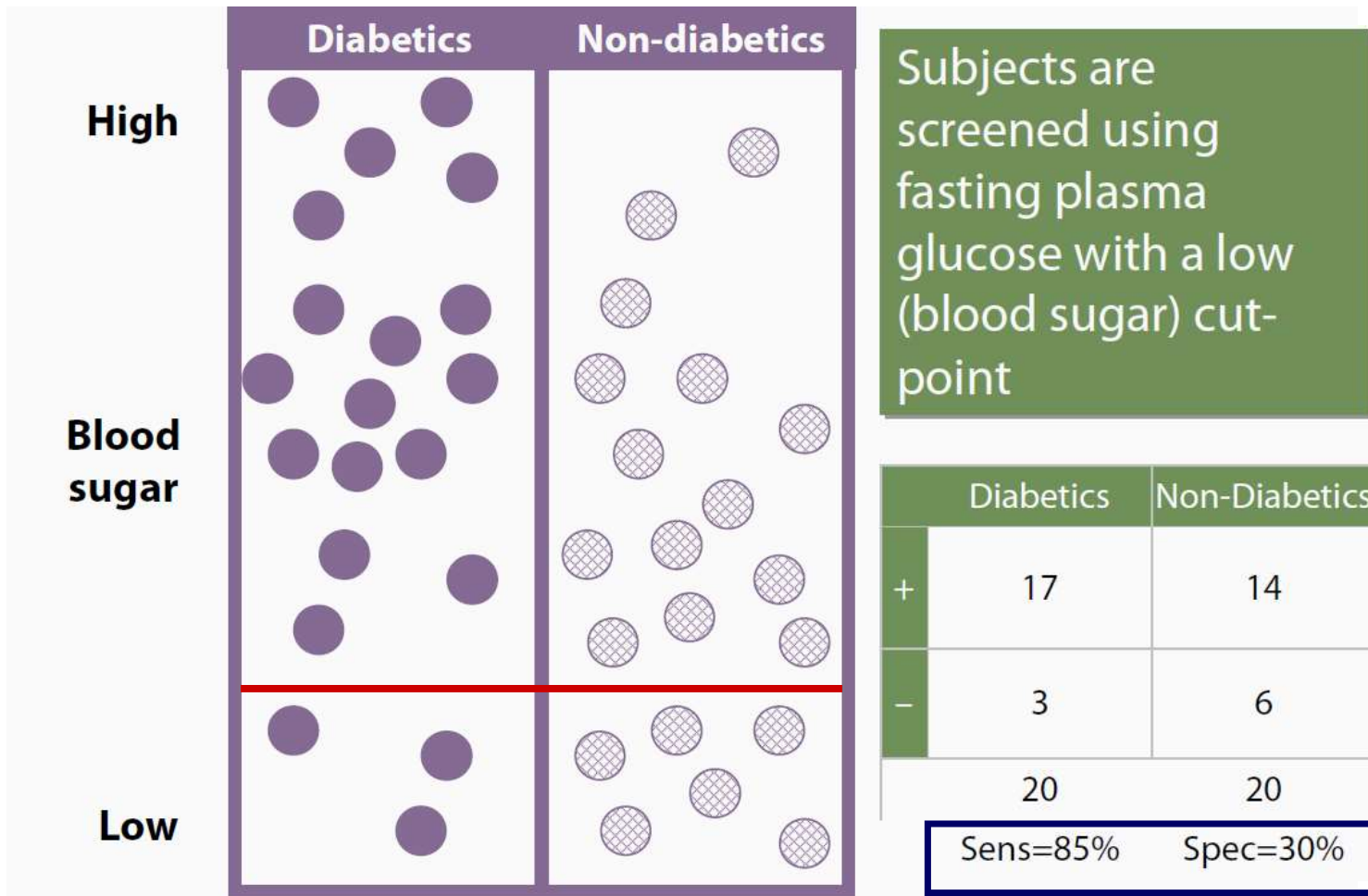
# Επιλογή cut-off

- Στο προηγούμενο παράδειγμα, μία ποσοτική δοκιμασία (επίπεδα φερριτίνης) χρησιμοποιήθηκε για το διαχωρισμό ασθενών-υγιών με βάση το cut-off 65 mmol/L
- Πως θα προσδιορίσω το «ιδανικό» cut-off σε μία **ποσοτική** διαγνωστική δοκιμασία για να διαχωρίζω τους ασθενείς από τους υγιείς;

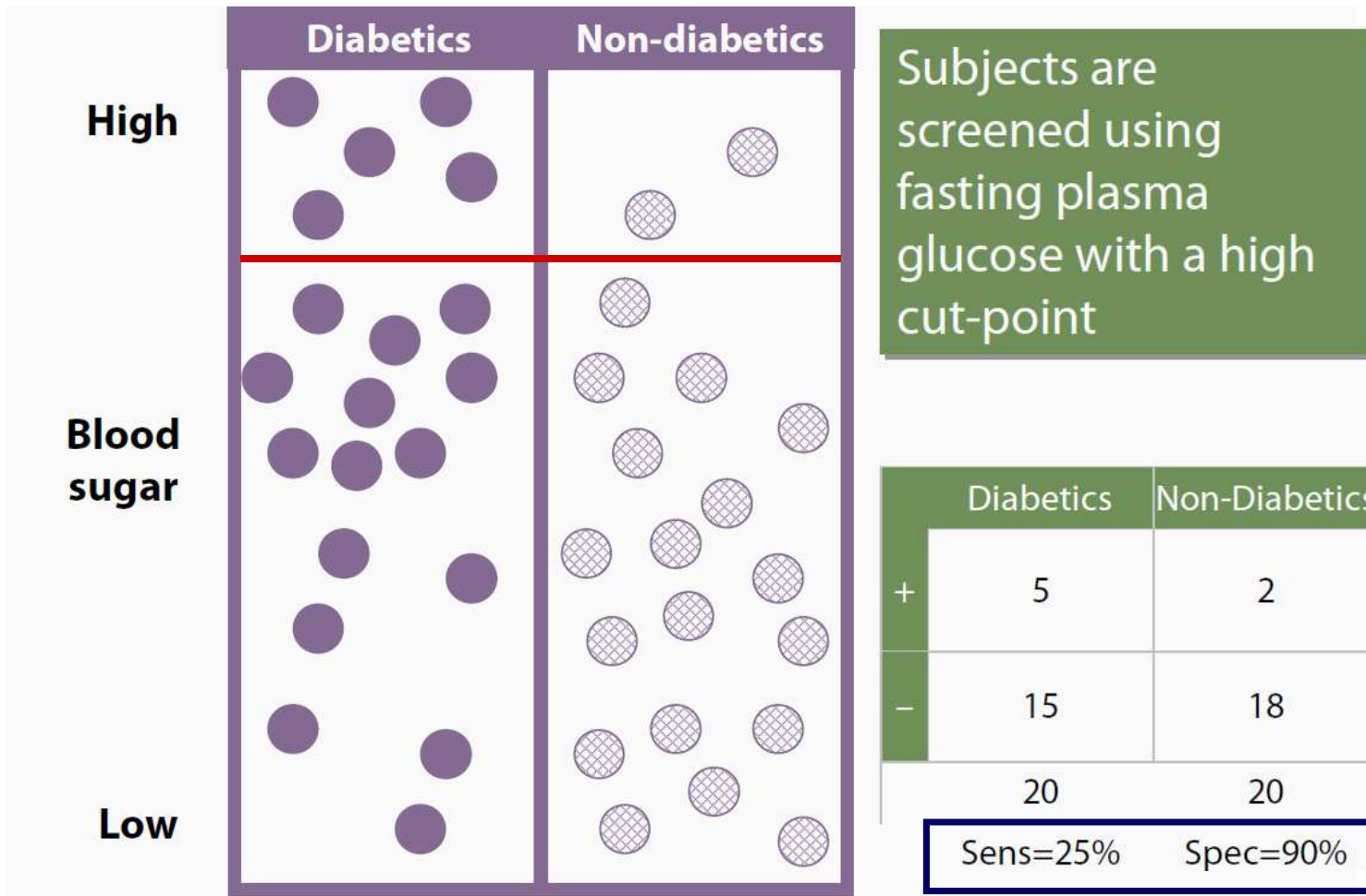
# Επιλογή cut-off



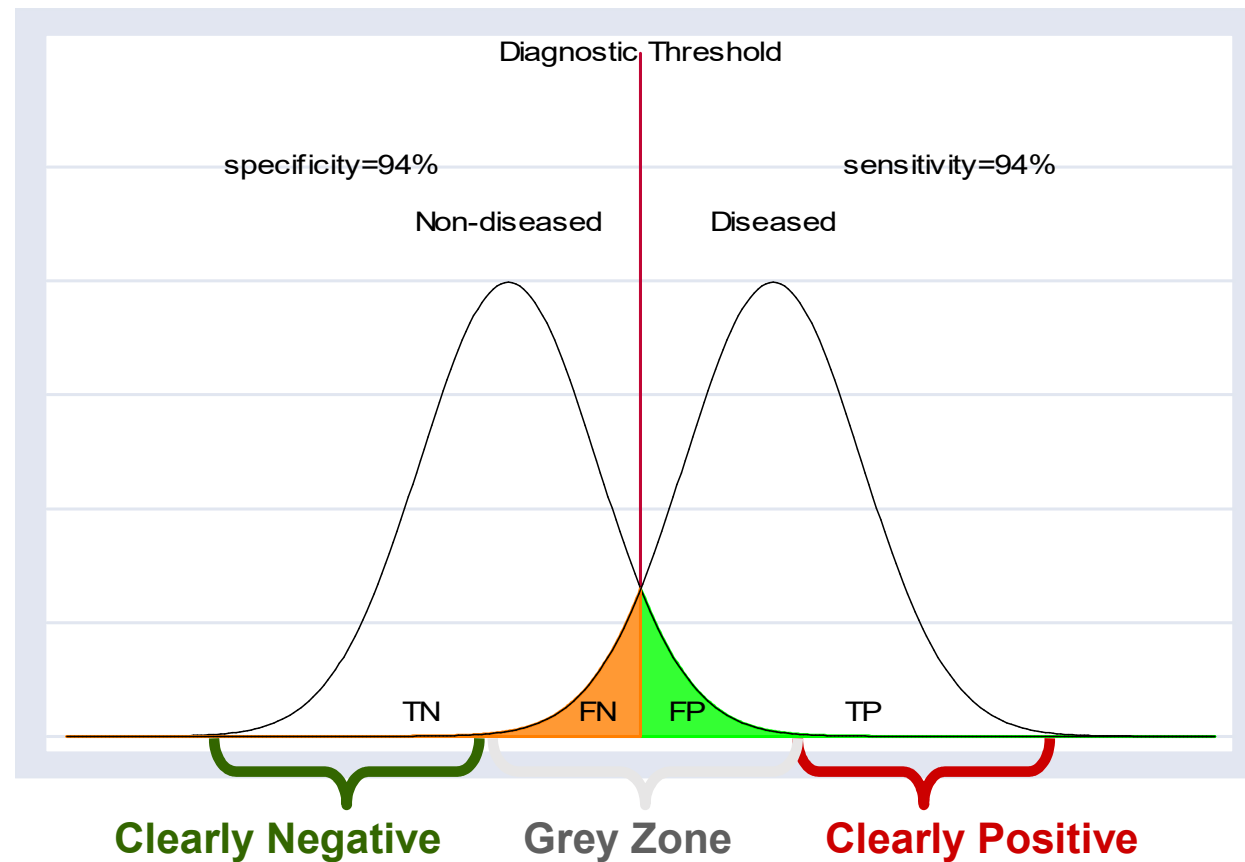
# Επιλογή cut-off



# Επιλογή cut-off

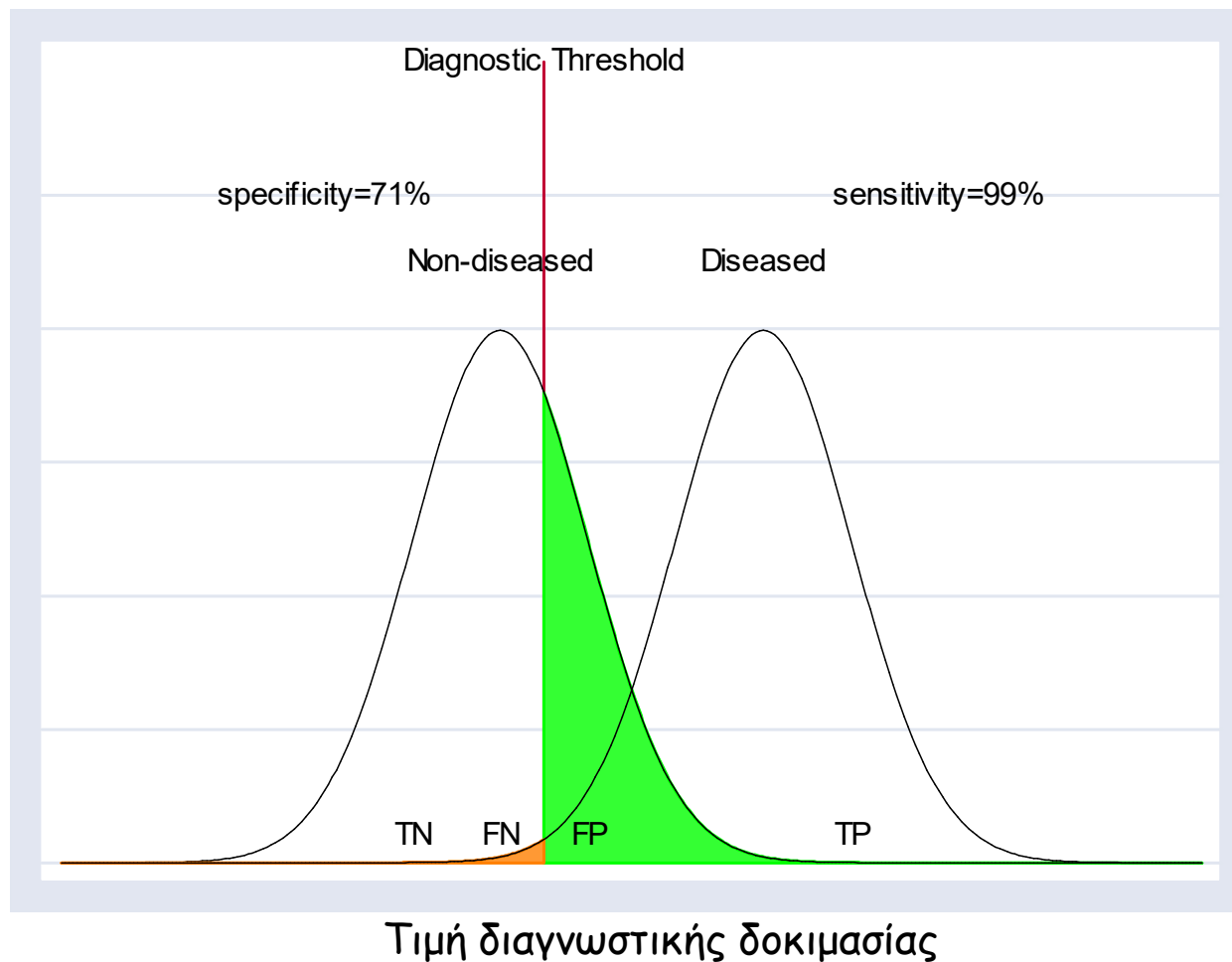


# Επιλογή cut-off

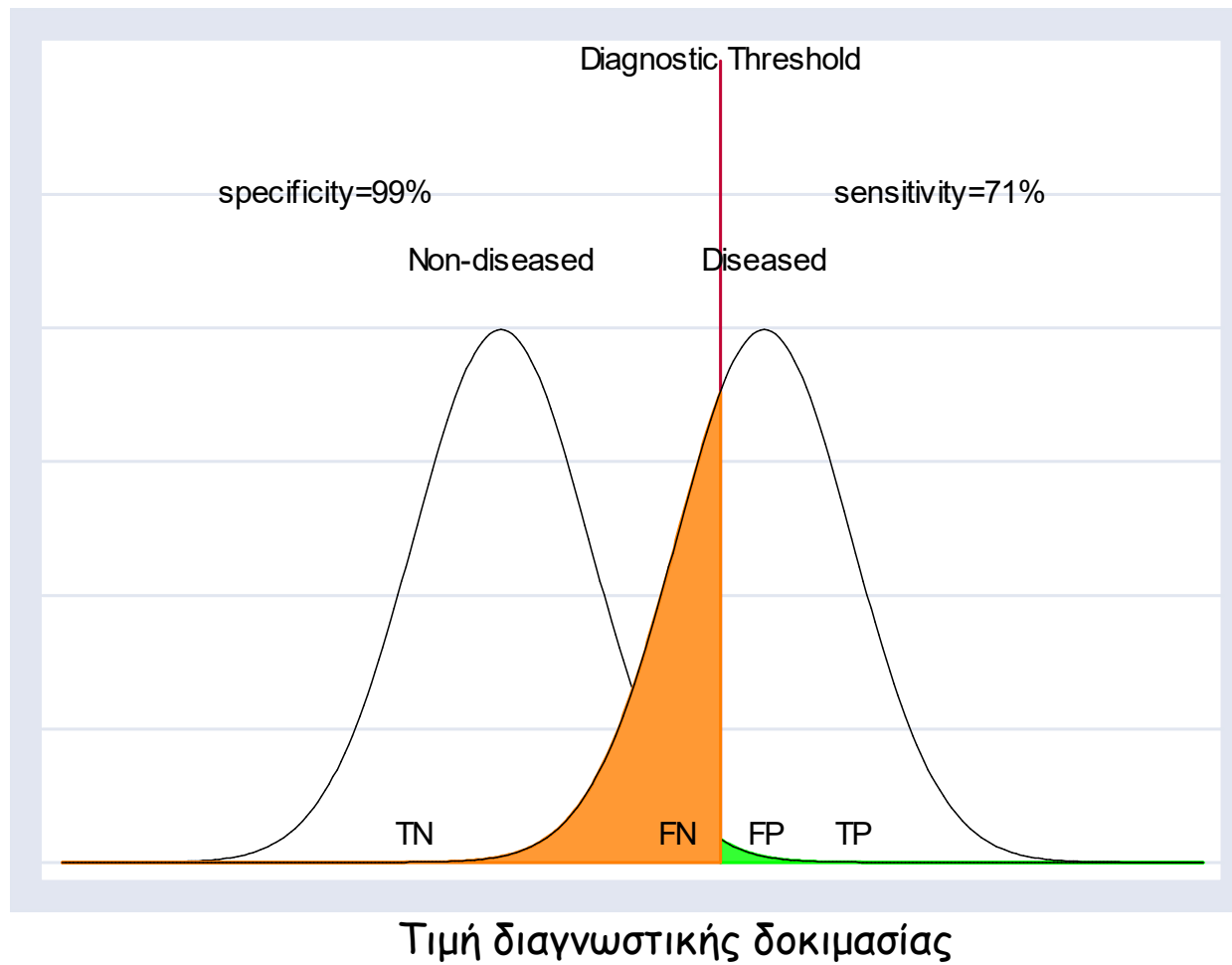


Τιμή διαγνωστικής δοκιμασίας

**Μείωση του cut-off → Λιγότερα ψευδώς (-) (κερδίζω σε ευαισθησία)**

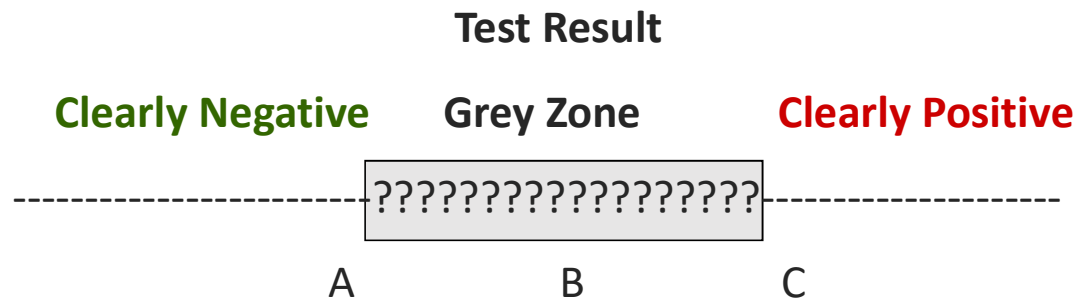


**Αύξηση του cut-off → Λιγότερα ψευδώς (+) (κερδίζω σε ειδικότητα)**



# Κριτήριο θετικότητας (Criterion of Positivity)

Η τιμή της δοκιμασίας στην οποία το αποτέλεσμα θεωρείται (+)



- Η επιλογή του επηρεάζει την ευαισθησία και την ειδικότητα → πρέπει να ζυγίσει κανείς το κόστος των ψευδώς θετικών έναντι του κόστους των ψευδώς αρνητικών
  - **Παράδειγμα:** Αιμοδοσία: HIV
    - Υψηλή ευαισθησία (λίγα ψευδώς αρνητικά)
    - Επιβεβαιωτικό τεστ υψηλής ειδικότητας (λίγα ψευδώς θετικά)