

Διαγνωστική ποιότητα Πιθανοφάνειες

Ν. Πετρόγλου,
Λ. Σπάρος

Εργαστήριο Κλινικής Επιδημιολογίας,
Τμήμα Νοσηλευτικής, Πανεπιστήμιο
Αθηνών, Αθήνα

Diagnostic quality. Likelihoods

Abstract at the end of the article

1. Εισαγωγή
2. Πιθανοφάνειες
3. Επιλογή των διαγνωστικών δοκιμασιών βάσει των λειτουργικών τους χαρακτηριστικών
4. Μεθοδολογικά σφάλματα κατά την εκτίμηση των πιθανοφανειών
5. Διάστημα εμπιστοσύνης των πιθανοφανειών

Λέξεις ευρετηρίου

Διακριτική ικανότητα
Λειτουργικά χαρακτηριστικά
Νοσολογική ειδικότητα
Νοσολογική ευαισθησία
Πιθανοφάνειες

Υποβλήθηκε 24.3.2005
Εγκρίθηκε 30.3.2005

1. ΕΙΣΑΓΩΓΗ

Κάθε εργαστηριακό αποτέλεσμα αλλά και κάθε κλινικό εύρημα (σύμπτωμα, σημείο), που προκύπτει από το ιστορικό και την αντικειμενική εξέταση, αποτελεί μια διαγνωστική «δοκιμασία», η γνώση του αποτελέσματος της οποίας αποτελεί μια επιπλέον πληροφορία, η οποία τροποποιεί την αρχική πιθανότητα ύπαρξης του νοσήματος.

Η κλινική χρησιμότητα μιας διαγνωστικής δοκιμασίας προσδιορίζεται κατά κύριο λόγο από τη διακριτική της ικανότητα (discriminating ability), δηλαδή από την ακρίβεια με την οποία διακρίνει ασθενείς με ή χωρίς το υπό διερεύνηση νόσημα για το οποίο αυτή επιτελείται.

Οι συχνότερα χρησιμοποιούμενες και αναφερόμενες συνιστώσες της διαγνωστικής ποιότητας (diagnostic quality) μιας δοκιμασίας, που καθορίζουν τη διακριτική της ικανότητα, είναι το ποσοστό των αληθώς θετικών αποτελεσμάτων ή η νοσολογική ευαισθησία της δοκιμασίας και το ποσοστό των αληθώς αρνητικών αποτελεσμάτων ή η νοσολογική ειδικότητα της δοκιμασίας.* Τα ποσοστά αυτά, καθώς και τα συμπληρωματικά τους (ποσοστό ψευδώς αρνητικών και ψευδώς θετικών αποτελεσμάτων, αντίστοιχα), ονομάζονται πιθανοφάνειες** (likelihood) ή, αλλιώς, λειτουργικά χαρακτηριστικά (operating characteristics) της δοκιμασίας.

2. ΠΙΘΑΝΟΦΑΝΕΙΕΣ

Για τον υπολογισμό των πιθανοφανειών μιας διαγνωστικής δοκιμασίας χρησιμοποιείται ο τετράπτυχος πίνακας (ή πίνακας 2×2 ή πίνακας δύο εισόδων) (πίν. 1). Στον πίνακα αυτό φαίνονται οι σχέσεις μεταξύ της πραγ-

* Ο όρος *ευαισθησία* (sensitivity) χρησιμοποιείται και για να ορίσει τη μικρότερη συγκέντρωση μιας ουσίας που μπορεί η μέθοδος να ανιχνεύσει και να μετρήσει (αναλυτική ευαισθησία). Ο όρος *ειδικότητα* (specificity) χρησιμοποιείται και για να ορίσει την ιδιότητα της μεθόδου ν' ανιχνεύει και να μετράει μόνο την ουσία που επιδιώκει να μετρήσει (βιοχημική ειδικότητα). Έτσι, προς αποφυγή σύγχυσης (δεδομένης της πολυσημίας των λέξεων ευαισθησία και ειδικότητα), οι όροι «*νοσολογική ευαισθησία*» και «*νοσολογική ειδικότητα*», που αναφέρονται στο αποτέλεσμα της δοκιμασίας και όχι στην ίδια τη δοκιμασία, είναι «καλύτεροι». Παρόλα αυτά, και οι δύο όροι τείνουν να καταργηθούν και αντικαθίστανται με τους ουδέτερους όρους «ποσοστό των αληθώς θετικών αποτελεσμάτων» και «ποσοστό των αληθώς αρνητικών αποτελεσμάτων», αντίστοιχα.

Βέβαια, και οι τελευταίοι όροι δεν είναι ιδιαίτερα επιτυχείς, δεδομένου ότι το αποτέλεσμα της δοκιμασίας στην πραγματικότητα είναι είτε θετικό είτε αρνητικό και δεν μπορεί να είναι αληθώς ή ψευδώς θετικό ή αρνητικό. Στην περίπτωση αυτή, το «αληθώς» ή «ψευδώς» αναφέρεται στην παρουσία ή απουσία του νοσήματος και όχι στο αποτέλεσμα της δοκιμασίας.^{1,3}

** Ο όρος *πιθανοφάνεια* (likelihood) διατηρείται προς τιμή του διάσημου Άγγλου στατιστικού R.A. Fisher, ο οποίος εισήγαγε τον όρο για τις δεσμευμένες πιθανότητες, όπου στον παρονομαστή βρίσκεται η παράμετρος (δηλαδή η φύση: το νόσημα) και στον αριθμητή η εμπειρία (δοκιμασία). Επισημαίνεται ότι η πιθανοφάνεια δεν είναι πάντοτε πιθανότητα.¹

ματικής κατάστασης του ατόμου (παρουσία ή απουσία του νοσήματος που διερευνάται) και της έκβασης της δοκιμασίας, που εκφράζεται, κατ' ανάγκη, διχοτομικά, δηλαδή θετικό (παθολογικό) ή αρνητικό (φυσιολογικό) αποτέλεσμα.

Η διάκριση των ατόμων σε πάσχοντες και μη πάσχοντες από το συγκεκριμένο νόσημα γίνεται βάσει μιας άλλης δοκιμασίας, που μπορεί (ή θεωρείται ότι μπορεί κατόπιν συμφωνίας) να εκτιμά την πραγματική κατάσταση του ασθενούς. Η δοκιμασία αυτή καλείται μέθοδος αναφοράς (gold standard).

Σύμφωνα με τον πίνακα 1, ο συνολικός αριθμός των ατόμων που υποβάλλονται στη δοκιμασία είναι $a+b+c+d$. Το σύνολο των πασχόντων και μη πασχόντων από το νόσημα ατόμων είναι $(a+c)$ και $(b+d)$, αντίστοιχα, ενώ το σύνολο των ατόμων με θετικό και αρνητικό το αποτέλεσμα της δοκιμασίας είναι $(a+b)$ και $(c+d)$, αντίστοιχα.

Κάθε άτομο, ανάλογα με την πραγματική κατάσταση (παρουσία ή απουσία του νοσήματος) και την έκβαση της δοκιμασίας, μπορεί να ανήκει σε μία από τις τέσσερις κατηγορίες:

Να *πάσχει* από το συγκεκριμένο νόσημα και

(a): να έχει θετικό το αποτέλεσμα της δοκιμασίας (αληθώς θετικό: ΑΘ) ή

(c): να έχει αρνητικό το αποτέλεσμα της δοκιμασίας (ψευδώς αρνητικό: ΨΑ)

ή να *μην πάσχει* από το συγκεκριμένο νόσημα και

(b): να έχει θετικό το αποτέλεσμα της δοκιμασίας (ψευδώς θετικό: ΨΘ) ή

(d): να έχει αρνητικό το αποτέλεσμα της δοκιμασίας (αληθώς αρνητικό: ΑΑ).

Από το «κάθετο διάγραμμα» του τετράπτυχου πίνακα προκύπτουν τα αντίστοιχα ποσοστά ή *οι πιθανοφάνειες*.

Ορισμοί¹⁻³

Ποσοστό των αληθώς θετικών αποτελεσμάτων, %ΑΘ (true positive rate, TPR) ή νοσολογική ευαισθησία (sensitivity, Se) μιας δοκιμασίας ονομάζεται το ποσοστό των πασχόντων από ένα συγκεκριμένο νόσημα που έχουν θετικό το αποτέλεσμα της δοκιμασίας:

$$\%ΑΘ = \frac{ΑΑ}{ΑΘ+ΨΑ} \quad \text{ή} \quad \left[\frac{a}{a+c} \right] \quad (1)$$

Το ποσοστό των αληθώς θετικών αποτελεσμάτων είναι, ουσιαστικά, μια δεσμευμένη πιθανότητα,* που εκφράζει την πιθανότητα του θετικού αποτελέσματος της δοκιμασίας (Δ^+), δεδομένου ότι το άτομο πάσχει από το συγκεκριμένο νόσημα (N^+) που διερευνάται, και διατυπώνεται συμβολικά ως:

$$\%ΑΘ = p(\Delta^+/N^+)$$

Συμπληρωματική έννοια του ποσοστού των αληθώς θετικών αποτελεσμάτων ($1-\%ΑΘ$) είναι το ποσοστό των ψευδώς αρνητικών αποτελεσμάτων, %ΨΑ (false negative rate, FNR), που εκφράζει το ποσοστό των πασχόντων από ένα συγκεκριμένο νόσημα που έχουν αρνητικό το αποτέλεσμα της δοκιμασίας:

$$\%ΨΑ = \frac{ΨΑ}{ΑΘ+ΨΑ} \quad \text{ή} \quad \left[\frac{c}{a+c} \right]$$

Ως δεσμευμένη πιθανότητα, το %ΨΑ εκφράζει την πιθανότητα του αρνητικού αποτελέσματος της δοκιμασίας (Δ^-), δεδομένου ότι το άτομο πάσχει από το συγκεκριμένο νόσημα (N^+) που διερευνάται, και διατυπώνεται συμβολικά ως:

$$\%ΨΑ = p(\Delta^-/N^+)$$

Ποσοστό των αληθώς αρνητικών αποτελεσμάτων, %ΑΑ (false positive rate, FPR) ή νοσολογική ειδικότητα (specificity, Sp) μιας δοκιμασίας ονομάζεται το ποσοστό των μη πασχόντων από ένα συγκεκριμένο νόσημα που έχουν αρνητικό το αποτέλεσμα της δοκιμασίας:

Πίνακας 1. Πιθανοφάνειες.

		Νόσημα		Σύνολο
		Ναι	Όχι	
Αποτέλεσμα δοκιμασίας	Θετικό	ΑΘ a	ΨΘ b	a+b
	Αρνητικό	ΨΑ c	ΑΑ d	c+d
Σύνολο		ΑΘ+ΨΑ a+c	ΑΑ+ΨΘ b+d	a+b+c+d
		$\%ΑΘ = \frac{ΑΘ}{ΑΘ+ΨΑ}$	$\%ΨΘ = \frac{ΨΘ}{ΑΑ+ΨΘ}$	
		$\%ΨΑ = \frac{ΨΑ}{ΑΘ+ΨΑ}$	$\%ΑΑ = \frac{ΑΑ}{ΑΑ+ΨΘ}$	

ΑΘ: Αληθώς θετικά, ΨΘ: Ψευδώς θετικά, ΨΑ: Ψευδώς αρνητικά, ΑΑ: Αληθώς αρνητικά αποτελέσματα, %ΑΘ: Ποσοστό αληθώς θετικών, %ΨΘ: Ποσοστό ψευδώς θετικών, %ΨΑ: Ποσοστό ψευδώς αρνητικών, %ΑΑ: Ποσοστό αληθώς αρνητικών αποτελεσμάτων

* Δεσμευμένη πιθανότητα είναι η πιθανότητα να εμφανιστεί ένα ενδεχόμενο Γ, δεδομένου ότι έχει ήδη εμφανιστεί (είναι γνωστό) ένα ενδεχόμενο Δ.

$$\%AA = \frac{AA}{AA+\Psi\Theta} \quad \text{ή} \quad \left[\frac{d}{b+d} \right] \quad (2)$$

Το ποσοστό των αληθώς αρνητικών αποτελεσμάτων είναι επίσης δεσμευμένη πιθανότητα, που εκφράζει την πιθανότητα του αρνητικού αποτελέσματος της δοκιμασίας (Δ^-), δεδομένου ότι το άτομο δεν πάσχει από το συγκεκριμένο νόσημα (N^-) που διερευνάται, και διατυπώνεται συμβολικά ως:

$$\%AA = p(\Delta^-/N^-)$$

Συμπληρωματική έννοια του ποσοστού των αληθώς αρνητικών αποτελεσμάτων ($1-\%AA$) είναι το ποσοστό των ψευδώς θετικών αποτελεσμάτων, $\%Ψ\Theta$ (false positive rate, FPR), που εκφράζει το ποσοστό των μη πασχόντων από ένα συγκεκριμένο νόσημα που έχουν θετικό το αποτέλεσμα της δοκιμασίας:

$$\%Ψ\Theta = \frac{\Psi\Theta}{AA+\Psi\Theta} \quad \text{ή} \quad \left[\frac{b}{b+d} \right]$$

Ως δεσμευμένη πιθανότητα, το $\%Ψ\Theta$ εκφράζει την πιθανότητα του θετικού αποτελέσματος της δοκιμασίας (Δ^+), δεδομένου ότι το άτομο δεν πάσχει από το συγκεκριμένο νόσημα (N^-) που διερευνάται, και διατυπώνεται συμβολικά ως:

$$\%Ψ\Theta = p(\Delta^+/N^-)$$

Παράδειγμα: Από μελέτη 96 ασθενών με στρεπτοκοκκική (β-αιμολυτικό στρεπτόκοκκο της ομάδας Α) ή μη κυνάγχη, με σκοπό την εκτίμηση της διαγνωστικής ποιότητας της καλλιέργειας του φαρυγγικού επιχρίσματος στη διάγνωση της στρεπτοκοκκικής κυνάγχης, προέκυψαν τα αποτελέσματα των καλλιιεργειών και τα αντίστοιχα ποσοστά αυτών, που περιλαμβάνονται στον πίνακα 2.⁴ Ως μέθοδος αναφοράς για τη διάκριση των ατόμων σε πάσχοντες και μη από στρεπτοκοκκική κυνάγχη χρησιμοποιήθηκε ο τίτλος των αντιστρεπτολυσινών του ορού.

Η εκτίμηση των πιθανοφανειών προϋποθέτει ότι το κλινικό ή το εργαστηριακό εύρημα δεν περιλαμβάνεται ούτε στον ορισμό του νοσήματος που διερευνάται ούτε στον ορισμό των «μη πασχόντων από το νόσημα».^{3,5} Έτσι, στο παράδειγμα του πίνακα 2, ο ορισμός της στρεπτοκοκκικής κυνάγχης βασίζεται στον τίτλο των αντιστρεπτολυσινών του ορού και όχι στο θετικό ή το αρνητικό αποτέλεσμα της καλλιέργειας.

Η ακρίβεια του ποσοστού των αληθώς θετικών και αληθώς αρνητικών αποτελεσμάτων ενός κλινικού ή εργαστηριακού ευρήματος εξαρτάται συγχρόνως από τον ορισμό του νοσήματος και τον ορισμό του ευρήματος (ορισμός του θετικού και του αρνητικού αποτελέσματος).^{1,3}

Πίνακας 2. Ποσοστό των αληθώς θετικών ($\%A\Theta$), αληθώς αρνητικών ($\%AA$), ψευδώς θετικών ($\%Ψ\Theta$) και ψευδώς αρνητικών ($\%ΨA$) αποτελεσμάτων της καλλιέργειας φαρυγγικού επιχρίσματος σε πάσχοντες από στρεπτοκοκκική ή μη στρεπτοκοκκική κυνάγχη.

		Στρεπτοκοκκική κυνάγχη		Σύνολο
		Ναι	Όχι	
Αποτέλεσμα καλλιέργειας	Θετικό	27 a	20 b	47 a+b
	Αρνητικό	3 c	46 d	49 c+d
	Σύνολο	30 a+c	66 b+d	96 a+b+c+d
		$\%A\Theta = \frac{27}{30} = 0,90$	$\%Ψ\Theta = \frac{20}{66} = 0,30$	
		$\%ΨA = \frac{3}{30} = 0,10$	$\%AA = \frac{46}{66} = 0,70$	

Όταν ένα κλινικό εύρημα και κατά κανόνα ένα εργαστηριακό αποτέλεσμα μπορεί να μετρηθεί (ποσοτική έννοια), τότε το ποσοστό των αληθώς θετικών και αληθώς αρνητικών αποτελεσμάτων είναι συνάρτηση της τιμής του διαχωριστικού ορίου (ΔO), της τιμής, δηλαδή, πάνω από την οποία ένα εύρημα χαρακτηρίζεται θετικό (παθολογικό) και κάτω από την οποία το εύρημα είναι αρνητικό (φυσιολογικό).

Στον πίνακα 3^{1,6} φαίνεται η μεταβολή του ποσοστού των αληθώς θετικών αποτελεσμάτων του καρκινοεμβρυϊκού αντιγόνου (CEA) στα διάφορα στάδια του καρκίνου του παχέος εντέρου, ανάλογα με την τιμή του ΔO . Όσο η τιμή του ΔO αυξάνει, τόσο μειώνεται το ποσοστό των αληθώς θετικών αποτελεσμάτων σε κάθε στάδιο του νοσήματος.

Για τον υπολογισμό του ποσοστού των αληθώς αρνητικών αποτελεσμάτων ενός κλινικού ή εργαστηριακού ευρήματος, στην κατηγορία των «μη πασχόντων» δεν συμπεριλαμβάνονται πάντοτε τα υγιή άτομα. Ο πληθυσμός των «μη πασχόντων» καθορίζεται από τη φύση του διαγνωστικού προβλήματος, του οποίου επιζητείται η λύση βάσει του ευρήματος.

Όταν πρόκειται για μαζική ή επιλεκτική διαλογή (προσυμπτωματικό έλεγχο), τότε ο πληθυσμός, στον οποίο προσδιορίζεται το ποσοστό των αληθώς αρνητικών αποτελεσμάτων, είναι αυτός των κλινικά υγιών ατόμων. Στις συνθήκες, όμως, κλινικές συνθήκες, ο πληθυσμός στον οποίο προσδιορίζεται το ποσοστό των αληθώς αρνητικών αποτελεσμάτων συνίσταται από ασθενείς με διαφοροδιαγνωστικά συναφή νοσήματα, οι οποίοι δεν πάσχουν από το συγκεκριμένο νόσημα που διερευνάται.^{1,7}

Εάν, για παράδειγμα, η παρουσία αίματος στα κόπρανα χρησιμοποιηθεί στα πλαίσια προσυμπτωματικού

Πίνακας 3. Ποσοστό των αληθώς θετικών αποτελεσμάτων του καρκινοεμβρυϊκού αντιγόνου (CEA) στα διάφορα στάδια του καρκίνου του παχέος εντέρου (ταξινόμηση Dukes), ανάλογα με την τιμή του διαχωριστικού ορίου.

Διαχωριστικό όριο (ng/mL)	Εξελικτικό στάδιο			
	A (n=58)	B (n=51)	C (n=63)	D (n=31)
>2,5	0,28	0,45	0,75	0,84
>5,0	0,03	0,25	0,45	0,65
>10,0	0,00	0,15	0,31	0,62
>20,0	0,00	0,00	0,18	0,43

ελέγχου για τη διάγνωση καρκίνου του παχέος εντέρου, τότε, για τον προσδιορισμό του ποσοστού των αληθώς αρνητικών αποτελεσμάτων, ο πληθυσμός των «μη πασχόντων» θα αποτελείται από άτομα κλινικά υγιή. Εάν, όμως, χρησιμοποιηθεί για τη διάγνωση του συγκεκριμένου νοσήματος σε ασθενείς με πεπτικά ενοχλήματα, τότε, για τον προσδιορισμό του ποσοστού των αληθώς αρνητικών αποτελεσμάτων, ο πληθυσμός των «μη πασχόντων» θα αποτελείται από ασθενείς με ανάλογα πεπτικά ενοχλήματα (π.χ. φλεγμονώδεις παθήσεις του εντέρου) που δεν πάσχουν από καρκίνο του παχέος εντέρου και όχι από υγιή άτομα.

3. ΕΠΙΛΟΓΗ ΤΩΝ ΔΙΑΓΝΩΣΤΙΚΩΝ ΔΟΚΙΜΑΣΙΩΝ ΒΑΣΕΙ ΤΩΝ ΛΕΙΤΟΥΡΓΙΚΩΝ ΤΟΥΣ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ

Όταν ένα κλινικό ή εργαστηριακό εύρημα έχει υψηλή νοσολογική ευαισθησία, δηλαδή απαντά σε όλους σχεδόν τους πάσχοντες από το νόσημα N, η απουσία του δραστικά αποκλείει την παρουσία του νοσήματος N, ανεξαρτήτως της νοσολογικής ειδικότητάς του.* Παρομοίως, όταν ένα κλινικό ή εργαστηριακό εύρημα έχει υψηλή νοσολογική ειδικότητα, δηλαδή απουσιάζει απ' όλους σχεδόν τους μη πάσχοντες από το νόσημα N, τότε η παρουσία του δραστικά θέτει τη διάγνωση του νοσήματος N, ανεξαρτήτως της νοσολογικής ευαισθησίας του.** Το ιδανικό εύρημα που απουσιάζει από όλους τους μη πάσχοντες από το νόσημα N (νοσολογική ειδικότητα 1) και με την παρουσία του επιβεβαιώνει τη διάγνωση του νοσήματος καλείται παθολογικό του νοσήματος N.^{3,8,9}

* Ο μνημοτεχνικός κανόνας που χρησιμοποιείται στην αγγλική είναι **SnNout** (sufficiently high **S**ensitivity - **N**egative result rules **o**ut the diagnosis),⁴ ο οποίος έχει αποδοθεί στα ελληνικά ως ΕυΑΟχι [υψηλή ευαισθησία - αρνητικό αποτέλεσμα - αποκλείει (όχι) τη διάγνωση].⁹

** Ο μνημοτεχνικός κανόνας που χρησιμοποιείται στην αγγλική είναι **SpPin** (sufficiently high **S**pecificity - **P**ositive result rules **i**n the diagnosis),⁴ ο οποίος έχει αποδοθεί στα ελληνικά ως ΕιΘεΝαι [υψηλή ειδικότητα - θετικό αποτέλεσμα - θέτει (ναι) τη διάγνωση].⁹

Δηλαδή, μια δοκιμασία με υψηλή νοσολογική ευαισθησία (ποσοστό αληθώς θετικών) είναι περισσότερο χρήσιμη στον κλινικό όταν το αποτέλεσμα της είναι αρνητικό (αποκλείει τη διάγνωση), ενώ μια δοκιμασία με υψηλή νοσολογική ειδικότητα (ποσοστό αληθώς αρνητικών) είναι περισσότερο χρήσιμη όταν το αποτέλεσμα της είναι θετικό (θέτει τη διάγνωση).

Είναι βέβαια επιθυμητό, η δοκιμασία που χρησιμοποιείται σε κάθε περίπτωση να έχει ταυτόχρονα υψηλό ποσοστό αληθώς θετικών και αληθώς αρνητικών αποτελεσμάτων. Δεδομένου, όμως, ότι αυτό δεν είναι πάντοτε δυνατό, μια διαγνωστική δοκιμασία θα πρέπει να επιλέγεται βάσει του υψηλού ποσοστού αληθώς θετικών ή αληθώς αρνητικών αποτελεσμάτων της, ανάλογα με το σκοπό για τον οποίο επιτελείται.

Δεδομένου ότι το αρνητικό αποτέλεσμα μιας δοκιμασίας με υψηλό ποσοστό αληθώς θετικών αποτελεσμάτων αποκλείει τη διάγνωση του νοσήματος, οι δοκιμασίες που παρουσιάζουν υψηλό ποσοστό αληθώς θετικών αποτελεσμάτων είναι εξαιρετικά χρήσιμες κατά την αρχική φάση της διαγνωστικής διαδικασίας, όταν υπάρχει ένας αριθμός πιθανών διαγνωστικών υποθέσεων, για τον αποκλεισμό κάποιων από αυτές, ώστε να διευκολυνθεί και να συντομευτεί ο δρόμος προς την τελική διάγνωση.

Επίσης, όταν ο στόχος του κλινικού είναι να μη διαφύγει η διάγνωση σοβαρού θεραπεύσιμου νοσήματος (π.χ. φυματίωση, νόσος Hodgkin) σε δεδομένο ασθενή, τότε επιλέγεται η δοκιμασία που έχει τα λιγότερα ψευδώς αρνητικά αποτελέσματα, δηλαδή αυτή που έχει υψηλότερο ποσοστό αληθώς θετικών αποτελεσμάτων.

Αντίθετα, όταν ο στόχος του κλινικού είναι να αποφευχθεί ο σιγματισμός του ατόμου ως πάσχοντος (ψευδοδιάγνωση) από σοβαρό νόσημα (π.χ. καρκίνος, AIDS), τότε επιλέγεται η δοκιμασία με τα λιγότερα ψευδώς θετικά αποτελέσματα, δηλαδή αυτή που έχει υψηλότερο ποσοστό αληθώς αρνητικών αποτελεσμάτων.

4. ΜΕΘΟΔΟΛΟΓΙΚΑ ΣΦΑΛΜΑΤΑ ΚΑΤΑ ΤΗΝ ΕΚΤΙΜΗΣΗ ΤΩΝ ΠΙΘΑΝΟΦΑΝΕΙΩΝ

Η ακρίβεια με την οποία εκτιμάται η διαγνωστική ποιότητα μιας δοκιμασίας εξαρτάται από την επιλογή της μεθόδου αναφοράς, τη μεροληψία (bias) κατά την εκτίμηση της διακριτικής ικανότητας της δοκιμασίας και την επιλογή του πληθυσμού των ασθενών στον οποίο προσδιορίζεται το ποσοστό των αληθώς θετικών και αληθώς αρνητικών αποτελεσμάτων.^{8,10}

- α. Για την αξιολόγηση της εγκυρότητας μιας νέας δοκιμασίας, όπως προαναφέρθηκε, προϋποτίθεται η γνώση της πραγματικής κατάστασης (παρουσία ή απουσία του νοσήματος) των ατόμων, που επιτυγχάνεται χρησιμοποιώντας μια άλλη δοκιμασία, τη μέθοδο αναφοράς, η οποία στις περισσότερες περιπτώσεις δεν είναι απόλυτα έγκυρη (%ΑΘ και %ΑΑ ίσο με 100%), αλλά θεωρείται κατόπιν συμφωνίας των ειδικών ως η «καλύτερη» διαθέσιμη.

Κατά τη σύγκριση, όμως, μιας νέας δοκιμασίας με μια όχι απόλυτα έγκυρη μέθοδο αναφοράς, η πρώτη δεν μπορεί να φανεί εγκυρότερη από τη δεύτερη ακόμη και αν προσεγγίζει περισσότερο την πραγματική κατάσταση. Εάν, π.χ., το αποτέλεσμα της νέας δοκιμασίας είναι περισσότερο «ευαίσθητο» από αυτό της μεθόδου αναφοράς, τα επιπλέον άτομα που θα διαγνωστούν από αυτή ως πάσχοντες θα θεωρηθεί ότι έχουν ψευδώς θετικό το αποτέλεσμα της δοκιμασίας. Παρομοίως, εάν το αποτέλεσμα της νέας δοκιμασίας είναι περισσότερο «ειδικό» από εκείνο της μεθόδου αναφοράς, τα επιπλέον άτομα που θα διαγνωστούν από αυτή ως μη πάσχοντες θα θεωρηθεί ότι έχουν ψευδώς αρνητικό το αποτέλεσμα της δοκιμασίας.⁸

- β. Μια άλλη μορφή μεθοδολογικού σφάλματος είναι η μεροληψία (bias) κατά την εκτίμηση της διακριτικής ικανότητας της δοκιμασίας, η οποία παρατηρείται στην περίπτωση που η υπό αξιολόγηση δοκιμασία αποτελεί κριτήριο για τη διενέργεια της μεθόδου αναφοράς.

Το θετικό αποτέλεσμα της δοκιμασίας, συχνά, παρακινεί τον κλινικό να συνεχίσει περισσότερο επιμελώς την αναζήτηση του υπό διερεύνηση νοσήματος, αυξάνοντας έτσι την πιθανότητα διάγνωσής του. Αντίθετα, έχοντας ένα αρνητικό αποτέλεσμα, ο κλινικός συχνά εγκαταλείπει την περαιτέρω αναζήτηση του υπό διερεύνηση νοσήματος, ταξινομώντας έτσι λανθασμένα κάποια άτομα στους «μη πάσχοντες». Η ίδια μορφή μεθοδολογικού σφάλματος συμβαίνει όταν τα

αποτελέσματα των δοκιμασιών ερμηνεύονται λαμβάνοντας υπόψη τις λοιπές κλινικές πληροφορίες ή την τελική διάγνωση του ασθενούς, όπως γίνεται συχνά κατά την ανάγνωση ακτινολογικών εξετάσεων.^{8,10,11}

- γ. Η ακρίβεια, τέλος, με την οποία εκτιμάται η ποιότητα μιας διαγνωστικής δοκιμασίας εξαρτάται και από τον πληθυσμό των ασθενών που επιλέγεται για το σκοπό αυτόν.

Όπως φαίνεται από τον τετράπτυχο πίνακα (πίν. 1), το ποσοστό των αληθώς θετικών και αληθώς αρνητικών αποτελεσμάτων της δοκιμασίας δεν εξαρτάται από την αναλογία των πασχόντων προς τους μη πάσχοντες στο δείγμα του πληθυσμού (προκύπτουν από το «κάθετο διάβασμα» του πίνακα). Το γεγονός αυτό οδήγησε στην ευρέως επικρατούσα αντίληψη ότι τα δύο αυτά βασικά μέτρα εγκυρότητας μιας δοκιμασίας αποτελούν εγγενή χαρακτηριστικά της δοκιμασίας και επομένως δεν επηρεάζονται από τον επιπολασμό του νοσήματος, δηλαδή από το ποσοστό των πασχόντων από το νόσημα στον πληθυσμό.

Αυτό, όμως, δεν ισχύει στην πραγματικότητα. Τα διάφορα χαρακτηριστικά των ασθενών (προφίλ*), το εξελικτικό στάδιο και η βαρύτητα (νοσολογικό φάσμα) του νοσήματος σχετίζονται ισχυρά τόσο με τον επιπολασμό του νοσήματος όσο και με το ποσοστό των αληθώς θετικών και αληθώς αρνητικών αποτελεσμάτων της δοκιμασίας. Όσο μεγαλύτερος είναι ο επιπολασμός του νοσήματος, τόσο μεγαλύτερη είναι η πιθανότητα να συμπεριληφθούν στο μελετώμενο πληθυσμό άτομα που βρίσκονται σε περισσότερο προχωρημένα εξελικτικά στάδια της νόσου, με αποτέλεσμα να αυξάνει σημαντικά το ποσοστό των αληθώς θετικών αποτελεσμάτων, ενώ μειώνεται το ποσοστό των αληθώς αρνητικών αποτελεσμάτων.^{5,8,11-14}

Το γεγονός αυτό γίνεται εύκολα κατανοητό βάσει του ορισμού του ποσοστού των αληθώς θετικών και αληθώς αρνητικών αποτελεσμάτων μιας δοκιμασίας (τύποι 1, 2): Όταν η δοκιμασία εφαρμόζεται σε άτομα με προχωρημένο στάδιο της νόσου (όπως αναμένεται στις περιπτώσεις αυξημένου επιπολασμού της νόσου), προκύπτουν περισσότερα θετικά αποτελέσματα (ΑΘ και ΨΘ) και λιγότερα αρνητικά αποτελέσματα (ΑΑ και ΨΑ). Η μείωση των ΨΑ σε συνδυασμό με την αύξηση των ΑΘ αποτελεσμάτων, σύμφωνα με τον τύπο 1, οδηγεί σε

* Από δεκαετίες περίπου έχει εισαχθεί στην ιατρική ορολογία ο όρος προφίλ (client profile), ο οποίος συνίσταται από το προφίλ υγείας (health profile), που περιλαμβάνει τα κλινικά ευρήματα, και από το προφίλ κινδύνου (risk profile), που περιλαμβάνει χαρακτηριστικά του ατόμου που συνδέονται με τη συχνότητα του νοσήματος (π.χ. φύλο, ηλικία, συνυπάρχοντα νοσήματα).^{1,5}

αύξηση του ποσοστού των αληθώς θετικών αποτελεσμάτων της δοκιμασίας. Παρομοίως, από τον τύπο 2 προκύπτει η μείωση του ποσοστού των αληθώς αρνητικών αποτελεσμάτων της δοκιμασίας με την αύξηση των ΨΘ και τη σύγχρονη μείωση των ΑΑ αποτελεσμάτων.¹⁰

Στον πίνακα 3 φαίνεται η αύξηση του ποσοστού των αληθώς θετικών αποτελεσμάτων του καρκινοεμβρυϊκού αντιγόνου (CEA) όσο το στάδιο του καρκίνου του παχέος εντέρου είναι πιο προχωρημένο, για την ίδια τιμή ΔΟ. Για παράδειγμα, για τιμή ΔΟ 2,5 ng/mL το ποσοστό των πασχόντων που ταξινομούνται ως πάσχοντες από τη δοκιμασία είναι μόλις 0,28 στο στάδιο Α, ανέρχεται όμως στο 0,84 στο στάδιο D.

Έτσι, μια δοκιμασία, η εγκυρότητα της οποίας εκτιμάται μεταξύ αρρώστων για τους οποίους υπάρχει αυξημένη πεποίθηση, βάσει άλλων κλινικοεργαστηριακών ευρημάτων, ότι πάσχουν από το συγκεκριμένο νόσημα που διερευνάται (όπως συνήθως συμβαίνει για την αξιολόγηση μιας νέας διαγνωστικής δοκιμασίας), θα παρουσιάζει τυπικά χαμηλότερο ποσοστό αληθώς θετικών αποτελεσμάτων και υψηλότερο ποσοστό αληθώς αρνητικών αποτελεσμάτων όταν εφαρμοστεί στα πλαίσια της πρωτοβάθμιας φροντίδας υγείας ή ως δοκιμασία διαλογής για την ανίχνευση του νοσήματος στο γενικό πληθυσμό, όπου ο επιπολασμός του νοσήματος είναι σαφώς μικρότερος.^{10,15}

Παράδειγμα: Στον πίνακα 4^{10,15} φαίνονται τα αποτελέσματα μελέτης 366 ασθενών, με σκοπό την εκτίμηση της διαγνωστικής ποιότητας μιας ταχείας εργαστηριακής δοκιμασίας, κατά την οποία γίνεται ανίχνευση της εστεράσης των λευκοκυττάρων και νιτροδών των βακτηρίων, με βύθιση ειδικού δείκτη στα ούρα, για τη διάγνωση της ουρολοίμωξης. Οι ασθενείς που υποβλήθηκαν στη δοκιμασία χωρίστηκαν βάσει της κλινικής τους εικόνας σε δύο ομάδες: αυτούς που είχαν τυπικά συμπτώματα και σημεία ουρολοίμωξης και επομένως είχαν αυξημένη πιθανότητα (>50%) να πάσχουν από ουρολοίμωξη και αυτούς που είχαν ελάχιστα κλινικά ευρήματα ουρολοίμωξης και επομένως η πιθανότητα να έχουν ουρολοίμωξη ήταν μικρή (≤50%). Ως μέθοδο αναφοράς για το διαχωρισμό των ατόμων σε πάσχοντες και μη από ουρολοίμωξη χρησιμοποιήθηκε η καλλιέργεια ούρων (≥100.000 μικροοργανισμοί/mL: θετικό αποτέλεσμα δοκιμασίας και, άρα, παρουσία ουρολοίμωξης).

Όπως φαίνεται στον πίνακα 4, στην ομάδα με αυξημένη αρχική πιθανότητα για ουρολοίμωξη βάσει της κλινικής εικόνας, το ποσοστό των αληθώς θετικών αποτελεσμάτων (%ΑΘ=0,92) είναι σαφώς μεγαλύτερο και το ποσοστό των αληθώς αρνητικών αποτελεσμάτων

Πίνακας 4. Ποσοστό των αληθώς θετικών (%ΑΘ) και αληθώς αρνητικών (%ΑΑ) αποτελεσμάτων της ταχείας δοκιμασίας για τη διάγνωση ουρολοίμωξης, στο σύνολο του πληθυσμού που εξετάστηκε και στις δύο ομάδες, υψηλής (>50%) και χαμηλής (≤50%) αρχικής πιθανότητας για ουρολοίμωξη, βάσει των κλινικών ευρημάτων.

		Ουρολοίμωξη		Σύνολο
		Ναι	Όχι	
<i>Σύνολο του πληθυσμού που εξετάστηκε</i>				
Αποτέλεσμα ταχείας δοκιμασίας	Θετικό	60	84	144
	Αρνητικό	12	210	222
		72	294	366
		$\%ΑΘ = \frac{60}{72} = 0,83$	$\%ΑΑ = \frac{210}{294} = 0,71$	
<i>Ομάδα χαμηλής αρχικής πιθανότητας για ουρολοίμωξη (≤50%)</i>				
Αποτέλεσμα ταχείας δοκιμασίας	Θετικό	10	53	63
	Αρνητικό	8	188	196
		18	241	259
		$\%ΑΘ = \frac{10}{18} = 0,56$	$\%ΑΑ = \frac{188}{241} = 0,78$	
<i>Ομάδα υψηλής αρχικής πιθανότητας για ουρολοίμωξη (>50%)</i>				
Αποτέλεσμα ταχείας δοκιμασίας	Θετικό	49	29	78
	Αρνητικό	4	21	25
		53	50	103
		$\%ΑΘ = \frac{49}{53} = 0,92$	$\%ΑΑ = \frac{21}{50} = 0,42$	

(%ΑΑ=0,42) μικρότερο απ' ό,τι στην ομάδα με χαμηλή αρχική πιθανότητα για ουρολοίμωξη (%ΑΘ=0,56 και %ΑΑ=0,78) και στο σύνολο του πληθυσμού που εξετάστηκε (%ΑΘ=0,83 και %ΑΑ=0,71).

Η παρουσία τέτοιων μεθοδολογικών σφαλμάτων, κατά την εκτίμηση της εγκυρότητας μιας δοκιμασίας, οδηγεί συχνά σε αύξηση του ποσοστού συμφωνίας ανάμεσα στη δοκιμασία και τη μέθοδο αναφοράς, δηλαδή κάνει τη δοκιμασία να φαίνεται περισσότερο «χρήσιμη» απ' ό,τι είναι στην πραγματικότητα.

Η εκτίμηση, λοιπόν, της διαγνωστικής ποιότητας των δοκιμασιών θα πρέπει να γίνεται χρησιμοποιώντας όσο το δυνατόν πιο αντιπροσωπευτικά δείγματα του πληθυσμού των ασθενών στους οποίους αυτές πρόκειται να εφαρμοστούν στην κλινική πράξη και οι κλινικοί να έχουν πάντοτε υπόψη τους τα μεθοδολογικά σφάλματα που είναι δυνατόν να έχουν συμβεί κατά την αξιολόγηση των δοκιμασιών, όταν επιλέγουν να τις εφαρμόσουν στο δικό τους πληθυσμό ασθενών.

5. ΔΙΑΣΤΗΜΑ ΕΜΠΙΣΤΟΣΥΝΗΣ ΤΩΝ ΠΙΘΑΝΟΦΑΝΕΙΩΝ

Η εκτίμηση του ποσοστού των αληθώς θετικών και αληθώς αρνητικών αποτελεσμάτων μιας δοκιμασίας γίνεται βάσει των παρατηρήσεων από, σχετικώς, μικρά «δείγματα» του πληθυσμού. Το διάστημα εμπιστοσύνης (συνήθως το 95%) παρέχει ένα μέτρο της ακρίβειας (ή αβεβαιότητας) των εκτιμήσεων αυτών, που επιτρέπει τη διατύπωση συμπερασμάτων αναφορικά με τον πληθυσμό από τον οποίο ελήφθη το «δείγμα».

Η εκτίμηση του 95% διαστήματος εμπιστοσύνης, δηλαδή του εύρους των τιμών μέσα στο οποίο με πιθανότητα 95% βρίσκεται η πραγματική τιμή του ποσοστού των αληθώς θετικών και αληθώς αρνητικών αποτελεσμάτων (καθώς και των συμπληρωματικών τους ποσοστών), γίνεται με βάση τον τύπο*^{8,16}

$$p \pm 1,96 \times SE \quad (3)$$

όπου: p = παρατηρούμενο ποσοστό και

SE = το πιθανό σφάλμα (standard error) του p , που δίδεται από τον τύπο:

$$SE = \sqrt{\frac{p(1-p)}{n}} \quad (4)$$

όπου: n = σύνολο των ατόμων με βάση το οποίο υπολογίστηκε το αντίστοιχο ποσοστό.

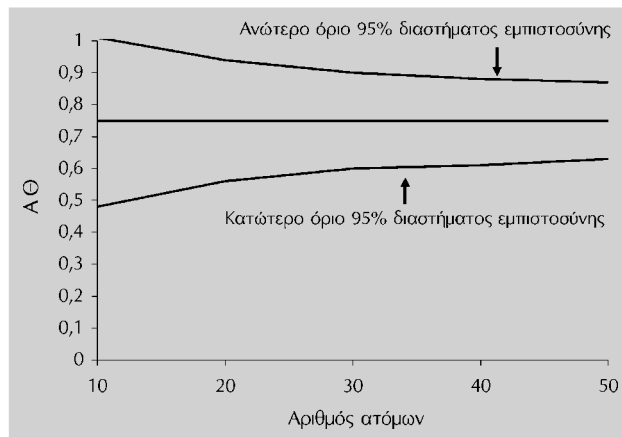
Παράδειγμα: Όσον αφορά στα δεδομένα του παραδείγματος στον πίνακα 2, το πιθανό σφάλμα του ποσοστού των αληθώς θετικών αποτελεσμάτων (0,90) ισούται με:

$$SE = \sqrt{\frac{0,90(1-0,90)}{30}} = 0,05$$

και το 95% διάστημα εμπιστοσύνης είναι $0,90 \pm 1,96 \times 0,05$, δηλαδή 0,79–1.

Με παρόμοιο τρόπο υπολογίζεται το 95% διάστημα εμπιστοσύνης και των υπολοίπων πιθανοφανειών.

Το εύρος του διαστήματος εμπιστοσύνης καθορίζει την ακρίβεια των εκτιμήσεων. Όσο πιο μικρό είναι το διάστημα εμπιστοσύνης, τόσο η ακρίβεια των εκτιμήσεων αυξάνει και αντιστρόφως. Σύμφωνα, όμως, με τους τύπους 3 και 4, το εύρος του διαστήματος εμπιστοσύνης και επομένως η ακρίβεια που αυτό εκφράζει, σε μεγάλο βαθμό, προσδιορίζεται από την τετραγωνική ρίζα του αριθμού των ατόμων που μελετήθηκαν.^{8,10} Όπως φαίνεται στην εικόνα 1, όσο ο αριθμός των ατόμων βάσει του οποίου γίνεται η εκτίμηση του ποσοστού των αληθώς θετικών αποτελεσμάτων (ή του ποσοστού των αληθώς αρνητικών αποτελεσμάτων) αυξάνει, τόσο το εύρος του διαστήματος εμπιστοσύνης μικραίνει, άρα αυξάνει η πιθανότητα η πραγματική τιμή να είναι κοντά στην εκτιμηθείσα τιμή.



Εικόνα 1. Το 95% διάστημα εμπιστοσύνης για τιμή ποσοστού αληθώς θετικών αποτελεσμάτων (%ΑΘ) ίση με 0,75, σε σχέση με τον αριθμό των ατόμων που μελετήθηκαν.

ABSTRACT

Diagnostic quality. Likelihoods

N. PETROGLOU, L. SPAROS

Laboratory of Clinical Epidemiology, School of Nursing, University of Athens, Athens, Greece

Archives of Hellenic Medicine 2006, 23(1):84–91

The clinical usefulness of a diagnostic test (laboratory test or clinical finding) is determined by the discriminating ability of the test, meaning its ability to correctly classify patients with or without a specified disease. The

* Ο τύπος αυτός χρησιμοποιείται για την εκτίμηση του 95% (1,96) διαστήματος εμπιστοσύνης οποιουδήποτε ποσοστού p .

most commonly used measures of the diagnostic value of a test are the true positive rate or sensitivity (proportion of people with the disease who have a positive test result) and the true negative rate or specificity (proportion of people without the disease who have a negative test result). These and their complementary rates, the false negative rate (proportion of people with the disease who have a negative test result) and the false positive rate (proportion of people without the disease who have a positive test result), respectively, are called the likelihoods or the operating characteristics of the diagnostic test. The diagnostic properties of a test may be inaccurately described because an inappropriate gold standard has been chosen. Also, the spectrum of patients (stage, severity, duration of disease) used to describe the properties of the test and the bias in judging the performance of the test can affect the determination of the true positive and false positive rates. These biases tend to make the test seem more useful than it actually is.

Key words: Discriminating ability, Likelihood, Operating characteristics, Sensitivity, Specificity

Βιβλιογραφία

1. ΣΠΑΡΟΣ Λ. Διαγνωστική ποιότητα των κλινικών και παρακλινικών δεδομένων. Στο: *Μετα-επιδημιολογία ή εφαρμοσμένη ιατρική έρευνα. Αιτιο-γνωστική, δια-γνωστική, προ-γνωστική*. Εκδόσεις ΒΗΤΑ, Αθήνα, 2001:105–115
2. ΑΝΕΥΛΑΒΗΣ Ε. Πιθανολογική λογική. Στο: *Κλινική λογική. Μεθοδολογία λήψης κλινικών αποφάσεων και λύσης κλινικών προβλημάτων*. Επισημονικές εκδόσεις Γ. Παρισιά-νος, Αθήνα, 1993:30–44
3. SACKETT DL, HAYNES RB, GUYATT GH, TUGWELL P. The interpretation of diagnostic data. In: *Clinical epidemiology. A basic science for clinical medicine*. 2nd ed. Little, Brown & Co, Boston, 1991:69–152
4. GRENIER B. Qualites diagnostiques des tests. In: *Evaluation de la decision medicale. Introduction a l'analyse medico-economique*. Masson, Paris, 1990:31–59
5. ΣΠΑΡΟΣ Λ. Διαγνωστική ποιότητα των κλινικών και παρακλινικών δεδομένων. Στο: *Θεωρία της λήψης των κλινικών αποφάσεων*. Εκδόσεις ΒΗΤΑ, Αθήνα, 1999:83–94
6. FLETCHER RH. Carcinoembryonic antigen. *Ann Intern Med* 1986, 104:66–73
7. SCHOUW YT, DIJK R, VERBEEK A. Problems in selecting the adequate patient population from existing data files for assessment studies of new diagnostic tests. *J Clin Epidemiol* 1995, 48:417–422
8. FLETCHER RH, FLETCHER SW, WAGNER EH. Diagnosis. In: *Clinical epidemiology. The essentials*. 2nd ed. Williams & Wilkins, Baltimore, 1988:42–75
9. SACKETT DL, STRAUS SE, RICHARDSON WS, ROSENBERG W, HAYNES RB. Διάγνωση και προσυμπτωματικός έλεγχος. Στο: *Επί ενδείξεων βασισμένη Ιατρική. Πώς να ασκείτε και να διδάσκετε την EBI*. Ιατρικές Εκδόσεις Π.Χ. Πασχαλίδης, Αθήνα, 2002:125–160
10. HAGEN MD. Test characteristics. How good is that test? *Prim Care* 1995, 22:213–233
11. DIAMOND GA. Clinical epistemology of sensitivity and specificity. *J Clin Epidemiol* 1992, 45:9–13
12. BRENNER H, GEFELLER O. Variation of sensitivity, specificity, likelihood ratios and prevalence values with disease prevalence. *Stat Med* 1997, 16:981–991
13. MOONS KGM, VAN ES GA, DECKERS JW, HABBEMA JDF, GROBBEE DE. Limitations of sensitivity, specificity, likelihood ratio, and Bayes' theorem in assessing diagnostic probabilities: A clinical example. *Epidemiology* 1997, 8:12–17
14. MIETTINEN OS, CARO JJ. Foundations of medical diagnosis: What actually are the parameters involved in Bayes' theorem? *Stat Med* 1994, 13:201–209
15. LACHS MS, NACHAMKIN I, EDELSTEIN PH, GOLDMAN J, FEINSTEIN AR, SCHWARTZ JJ. Spectrum bias in the evaluation of diagnostic tests: Lessons from the rapid dipstick test for urinary tract infection. *Ann Intern Med* 1992, 117:135–140
16. SACKETT DL, STRAUS SE, RICHARDSON WS, ROSENBERG W, HAYNES RB. Διαστήματα αξιοπιστίας. Στο: *Επί ενδείξεων βασισμένη Ιατρική. Πώς να ασκείτε και να διδάσκετε την EBI*. Ιατρικές Εκδόσεις Π.Χ. Πασχαλίδης, Αθήνα, 2002:333–343

Corresponding author:

N. Petroglou, 11 Moutalaski street, GR-142 34 Athens, Greece
e-mail: niki.petroglou@yahoo.gr