

Κατανομές κυτταρικού φθορισμού Cell Fluorescence Distributions

Κατερίνα Ψαρρά

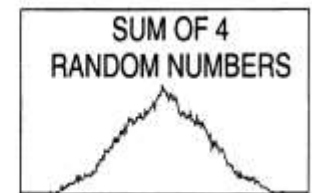
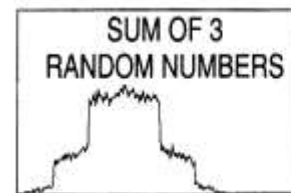
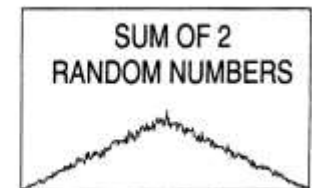
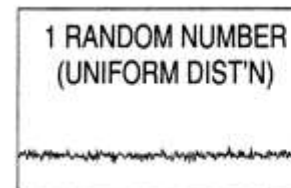
Βασική κυτταρομετρία (μέρος 2^ο)

Σκοπός της απεικόνισης δεδομένων στην Κ.Ρ

- Προσδιορισμός της πραγματικής συχνότητας της παραμέτρου υπό μελέτη
- Ιστόγραμμα κατανομής συχνοτήτων, κατανομής πιθανοτήτων, κατανομής πυκνοτήτων
- Ιστόγραμμα = άμεση κατάταξη σε πίνακα των συχνοτήτων των μετρουμένων τιμών μιας παραμέτρου σε ένα δείγμα
- Ιστόγραμμα: θεωρείται ακριβής οπτική απεικόνιση της κατανομής συχνοτήτων των τιμών της παραμέτρου

Κατανομές στατιστικές

- Κάνουμε πολλές μετρήσεις
- Διάγραμμα τιμών στον άξονα των x και συχνότητες στον άξονα των y .
- Τι περιμένουμε;
- Διάγραμμα αθροισμάτων ανά δύο, τρία, τέσσερα, οκτώ
- Κατανομή Gauss



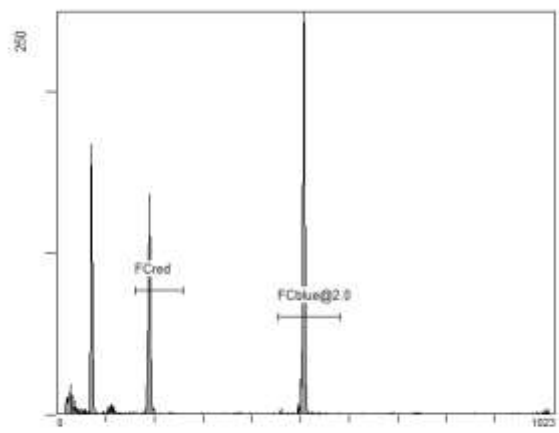
Εξίσωση Gauss

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

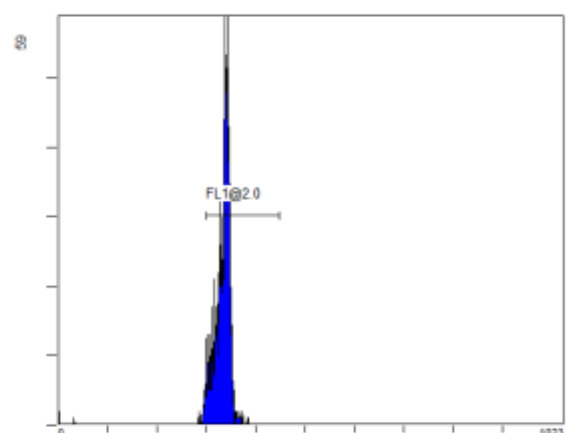
μ =mean, σ =SD

Κατανομές στην κυτταρομετρία I

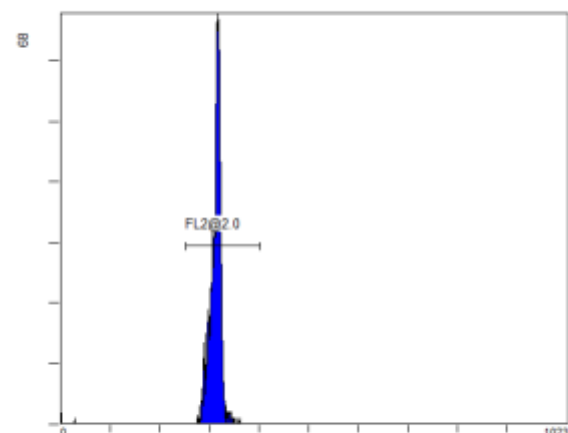
- Σφαιρίδια τύπου I (ευθυγράμμισης): απεικονίζονται σε γραμμική κλίμακα
- DNA ανάλυση: απεικονίζεται σε γραμμική κλίμακα
- Κανονικές κατανομές
- Συνηθίζουμε να χρησιμοποιούμε το CV
- Ζητάμε να έχουμε όσο το δυνατόν μικρότερο CV, (τα σφαιρίδια και οι πυρήνες είναι ίσου μεγέθους και εκπέμπουν σταθερή ένταση φθορισμού)



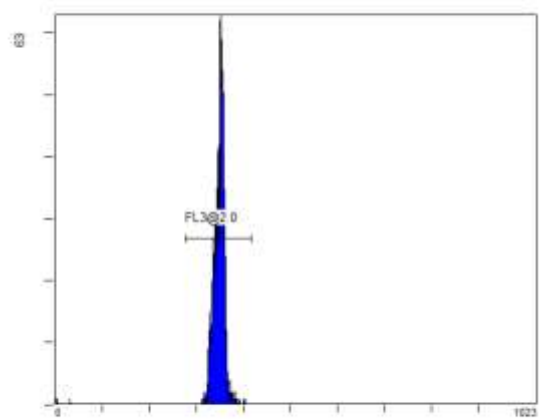
FS NT LN



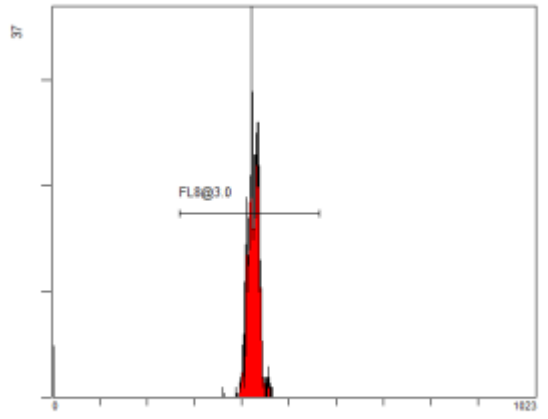
FL1 NT LN



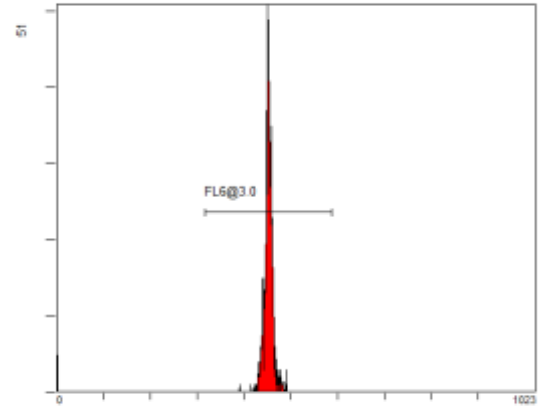
FL2 NT LN



FL3 NT LN



FL4 NT LN



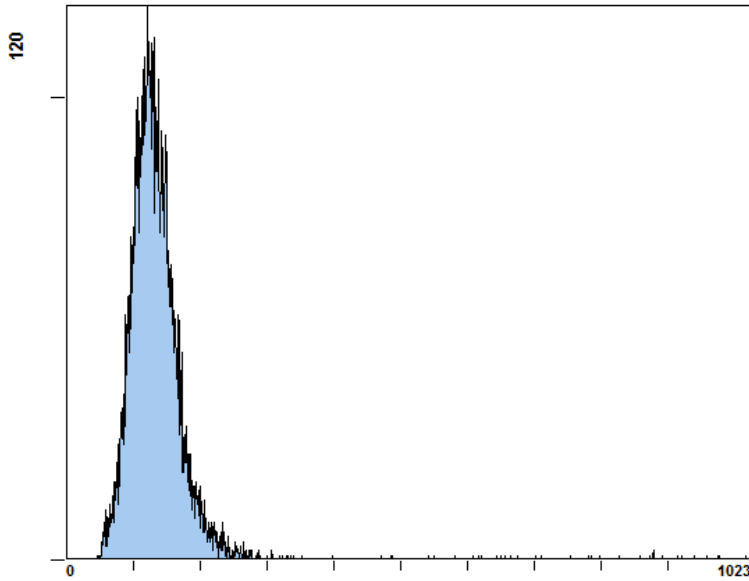
FL6 NT LN

Σφαιρίδια ευθυγράμμισης σε γραμμική κλίμακα – κανονικές κατανομές με χαμηλό CV

Κατανομές στην κυτταρομετρία II

- Οι κατανομές ανοσοφθορισμού και πλάγιας σκέδασης είναι *skewed* (γέρνουν προς τα αριστερά) σε γραμμική κλίμακα.
- Μοιάζουν με κανονική κατανομή όταν μετατραπούν σε λογαριθμική κλίμακα και η κατανομή λέγεται **περίπου lognormal**.
- Το περίπου αφορά ότι στη φύση πάντα υπάρχει κάτι τελείως εκτός του κανονικού π.χ. λεμφοκύτταρα μικρά σαν νετρόνια ή μεγάλα σαν φουντούκια ($1/10^{15}$) (έξω από 8 SD - $1/10^{12}$ 7SD - $1/10^9$ 6SD rare events)
- «κολοβές» κατανομές : δεν εκτείνονται από – άπειρο μέχρι +άπειρο

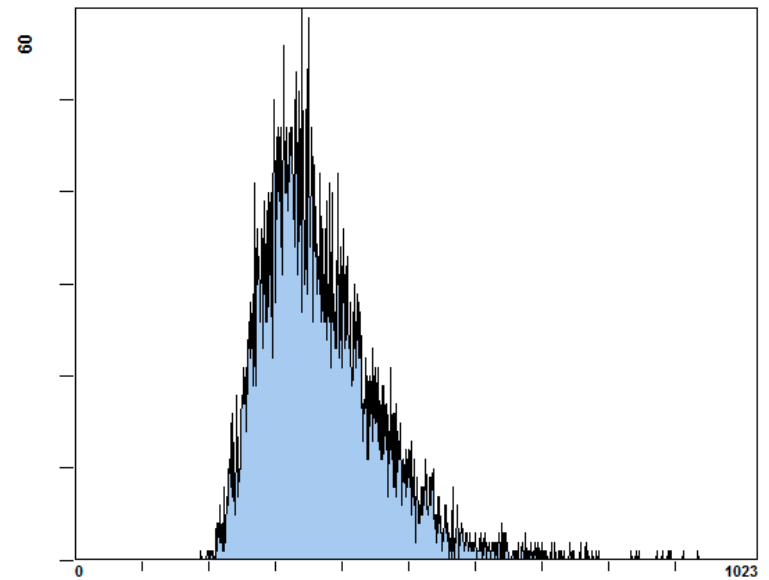
(10000) [A] SS INT LIN



SS INT LIN

Κατανομές που γέρνουν
(skewed)
FS, SS σε γραμμική κλίμακα

(10000) [A] FS INT LIN



FS INT LIN

Μέτρα κεντρικής τάσης

- Mean = μέτρο κεντρικής τάσης, θέση αντιπροσωπευτικού ποσοστού της κατανομής
- Geometric mean = η Νιοστή ρίζα του προϊόντος όλων των τιμών $(\prod x)^{1/N}$. Γιατί τη χρησιμοποιούμε; Διότι οι τιμές είναι λογαριθμικές και το πιο εύκολο mean είναι το geometric.
- Mode = η πιο συχνή τιμή της κατανομής
- Median = η τιμή, πάνω και κάτω από την οποία υπάρχει το 50% των τιμών
- Τα mean και geometric mean επηρεάζονται πολύ από τιμές outliers, τα mode και median όχι

Μέτρα διασποράς τιμών

- μεταβλητότητα (variation)
- σταθερή απόκλιση (standard deviation – SD) = τετραγωνική ρίζα της μεταβλητότητας
- Συντελεστής μεταβλητότητας (coefficient of variation – CV) = $SD / \text{arithmetic mean}$
- Εύρος (range) = διαφορά μέγιστης ελαχίστης τιμής
- Εύρος τεταρτημορίων (interquartile range – i.q.r) = διαφορά τιμών όπου ανήκει το κεντρικό 50% των τιμών της κατανομής

Παραμετρικές δοκιμασίες

- Οι παραμετρικές βασίζονται στην υπόθεση ότι η κατανομή είναι κανονική.
- Ανθεκτικές δοκιμασίες είναι αυτές που διατηρούνται ακόμη και όταν οι κατανομές δεν είναι κανονικές.
- Εύρος, mean, geometric mean, SD : όχι ανθεκτικές (επηρεάζονται από outliers)
- Median, interquartile range : ανθεκτικές

Συμμετρική κατανομή μονής κορυφής

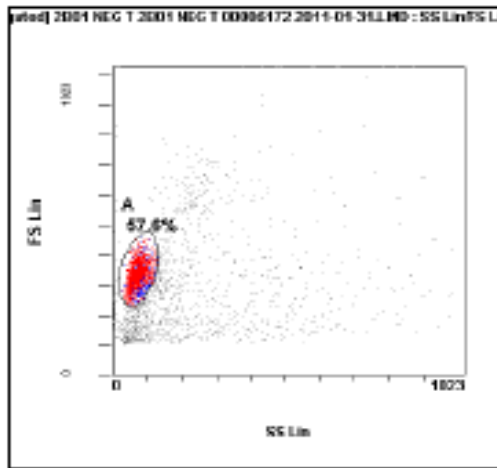
Mean = median = mode

- Οι τιμές σε αληθινή κλίμακα και όχι σε κανάλια

Skewed κατανομή

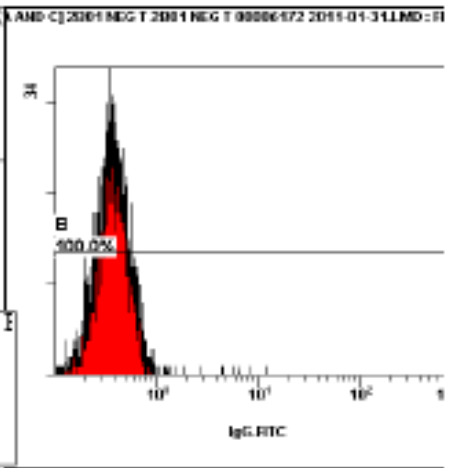
- Mean \neq median \neq mode
- RCV = robust CV

$$\text{RCV} = 0.75 \times (\text{i.q.r.}) / \text{median}$$



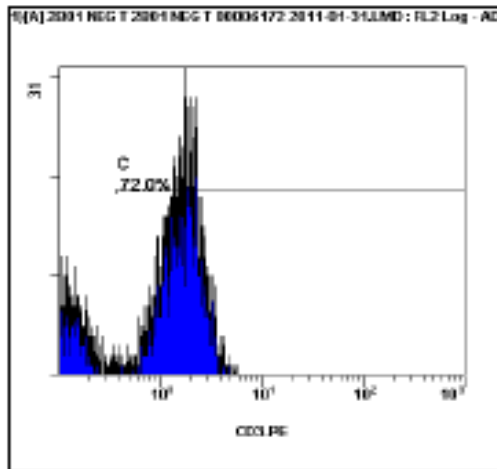
(F1)[Ungated] 2B01 NEG T 2B01 NEG T 00006172 2011-01-31.LMD : S

Region	Number	X-Median	50.0	X-Mean	Y-Mean
ALL	5788	83		160	330
A	3335	72		74.7	337



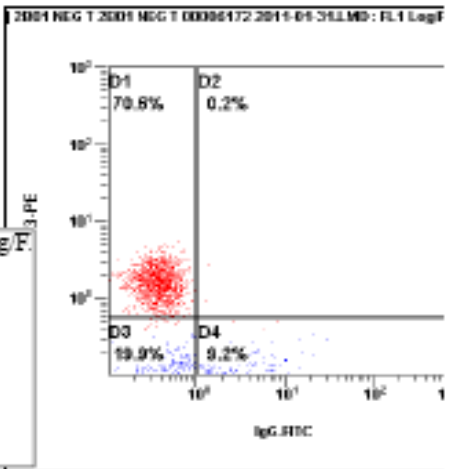
(F1)[A AND C] 2B01 NEG T 2B01 NEG T 00006172 2011-01-31.LMD : F

Region	Number	X-Median	50.0	X-Mean	Y-Mean
ALL	2402	0.367		0.412	###
B	2402	0.367		0.412	###



(F1)[A] 2B01 NEG T 2B01 NEG T 00006172 2011-01-31.LMD : FL

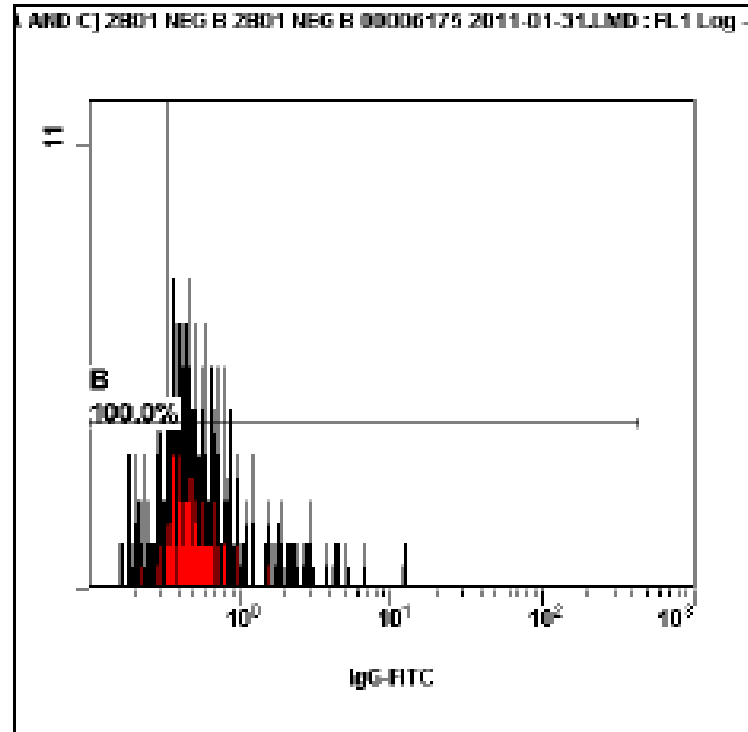
Region	Number	X-Median	50.0	X-Mean	Y-Mean
ALL	3335	1.31		1.31	###
C	2402	1.65		1.77	###



(F1)[A] 2B01 NEG T 2B01 NEG T 00006172 2011-01-31.LMD : FL1 Log F

Region	Number	X-Median	50.0	X-Mean	Y-Mean
ALL	3335	0.398		0.737	1.31
D1	2356	0.364		0.39	1.8
D2	8	1.29		2.36	1.81
D3	665	0.444		0.483	0.144
D4	306	2.07		3.92	0.149

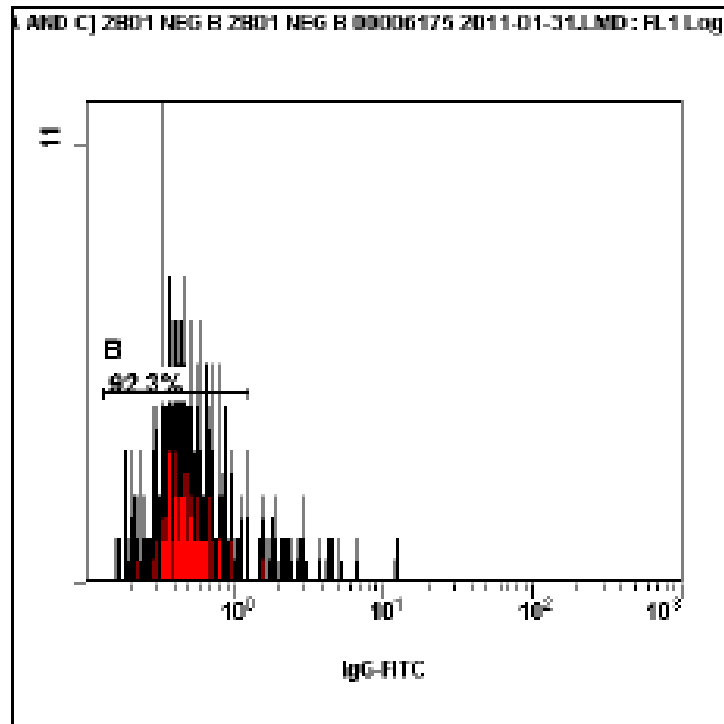
IgG-FITC median \neq mean ($\mu\epsilon$ outliers)
 CD3-PE median = mean ($\chi\omega\rho\acute{\iota}\varsigma$ outliers)



(F1)[A AND C] 2B01 NEG B 2B01 NEG B 00006175 2011-01-31.LMD : FL1 Log

Region	Number	X-Median	50.0	X-Mean	X-Mode	HP X-CV
ALL	428	0.456		0.668	0.324	0.44
B	428	0.456		0.668	0.324	0.44

IgG-FITC median \neq mean \neq mode
 ($\mu\epsilon$ outliers)



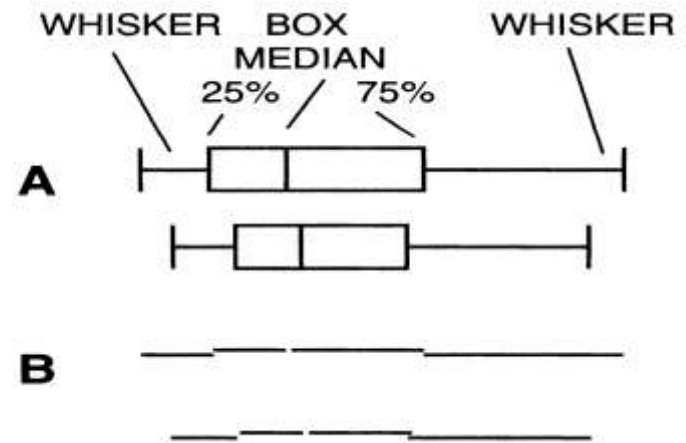
(F1)[A AND C] 2B01 NEG B 2B01 NEG B 00006175 2011-01-31.LMD : FL1 Log

Region	Number	X-Median	50.0	X-Mean	X-Mode	HP	X-CV
ALL	428	0.456		0.668	0.324		0.44
B	395	0.444		0.488	0.324		0.44

IgG-FITC median = mean \neq mode
(χωρίς outliers)

Λογαριθμική κλίμακα

- Median
- i.q.r (στα boxplots το κουτάκι)
- RCV



- Αποθήκευση:

Λογαριθμικές εικόνες και

Γραμμικά δεδομένα? Ρωτούσε ο σπαρτο το ζυζ...

Αυτό ισχύει σήμερα!!!

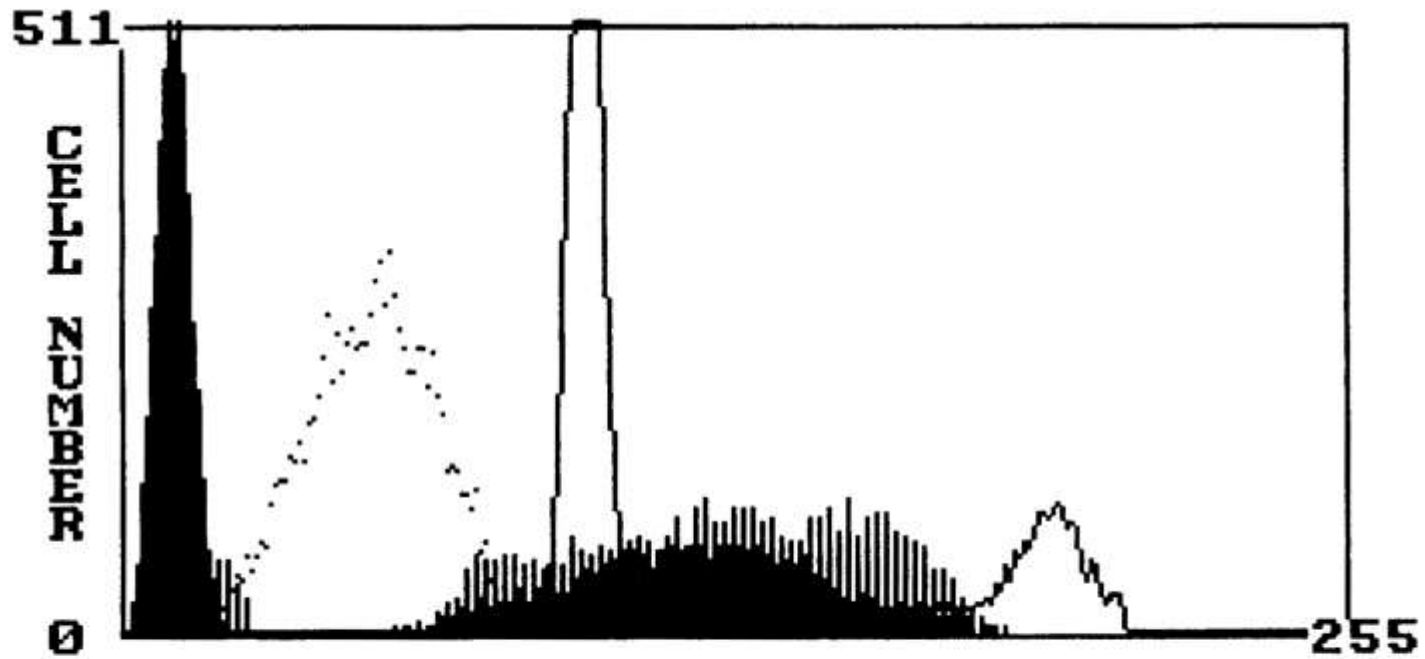
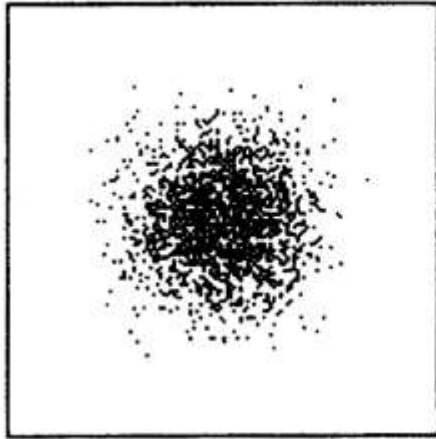


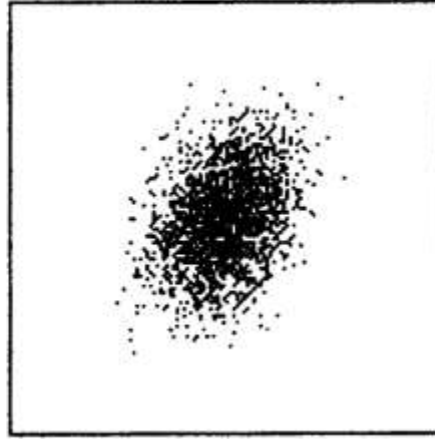
Figure 5-7. Histogram display formats.

Κατά τον Shapiro οι εναλλασσόμενες σκούρες ανοιχτές στήλες καλύτερες για συγκρίσεις

Στο δεξί dotplot χρειάζεται να μελετηθεί μόνο η μία παράμετρος



$r = -.002$



$r = .320$



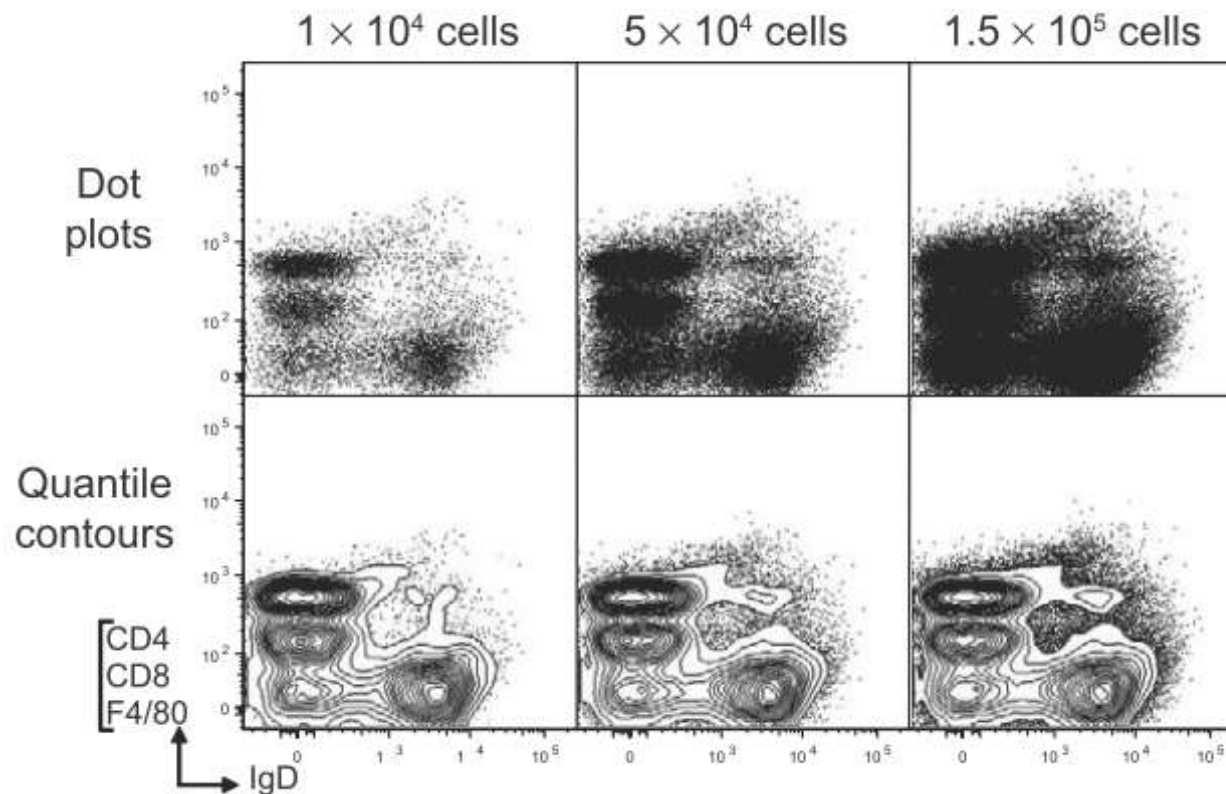
$r = .950$

Figure 5-8: Dot plots of computer-generated data showing various degrees of correlation.

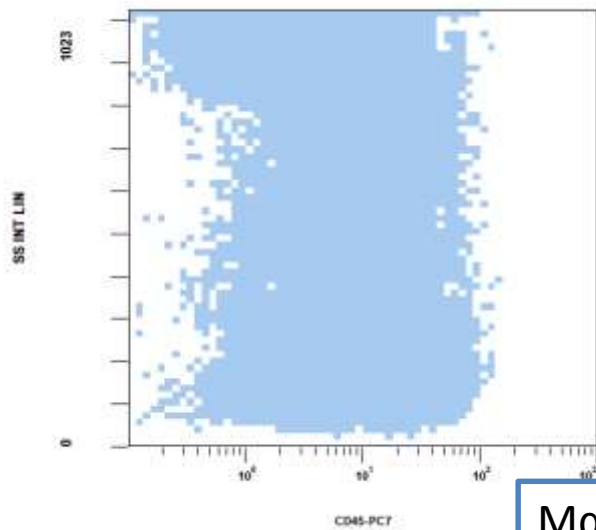
Σχόλια του Shapiro για την προηγούμενη εικόνα (όσο φαίνεται)

- Το dotplot και το density plot για 500 κύτταρα δεν διαφέρουν και πολύ ενώ το contour plot λέει πιο πολλά από τα ίδια τα κύτταρα που υπάρχουν (όχι και τόσο δίκαιο).
- Δείχνουν η δεύτερη και η τρίτη σειρά πόσο πιο πολλά κύτταρα υπάρχουν; (ιδίως στα smoothed – ομαλοποιημένα)
- Τα dotplots όσο πιο πολλά κύτταρα έχουν τόσο πιο «μαυρισμένα» σχήματα. Βλέπουμε όμως τα μεμονωμένα κύτταρα. Τα dotplots λένε την αλήθεια, αλλά και τα μη ομαλοποιημένα density plots. Τα ομαλοποιημένα κρύβουν λίγο την αλήθεια εκτός αν επιτρέπονται τα outliers.

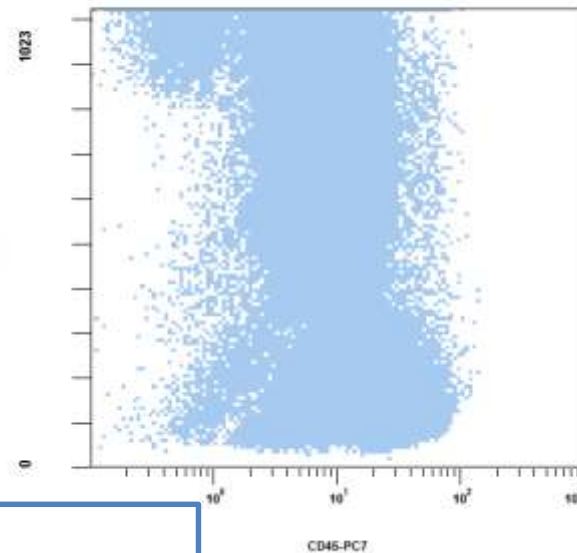
Τα μαυρισμένα dotplots, τα πληροφορικά dotplots, τα contour και τα density plots



Leonore A Herzenberg et al
2006

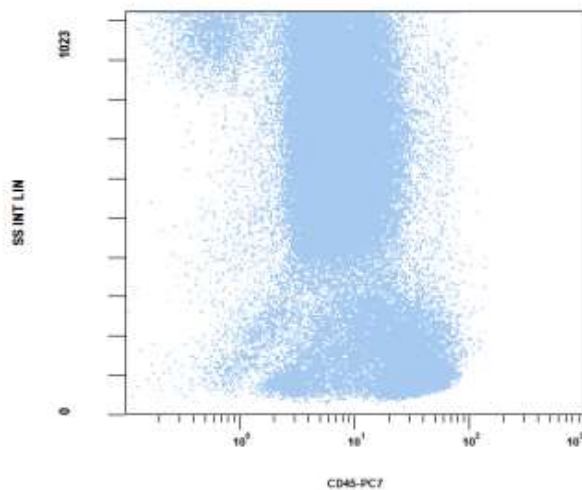


64X64

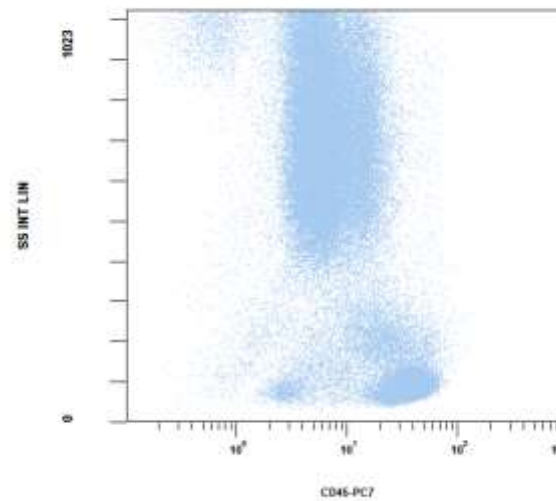


128X128

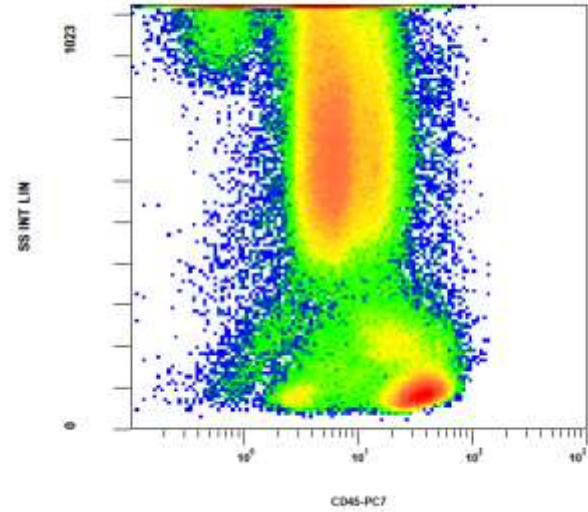
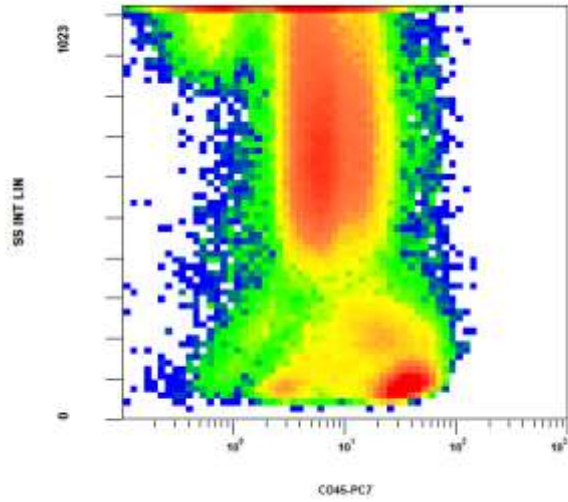
Μαυρισμένα dotplots;
1^η λύση: κάντε τα γαλάζια
2^η λύση: αλλαγή resolution



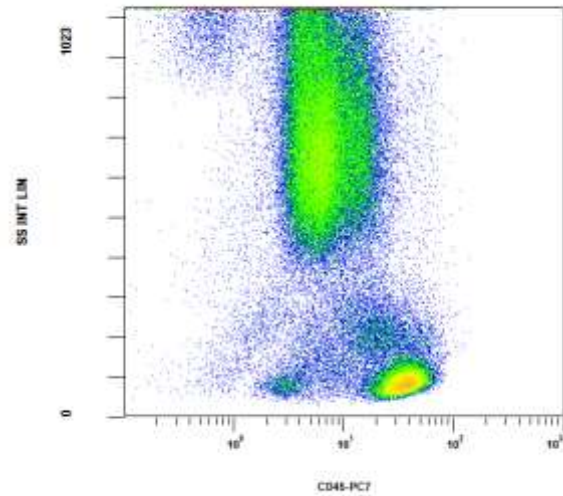
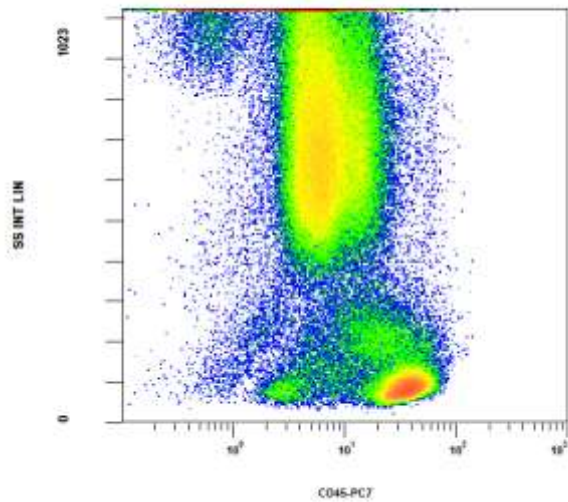
256X256

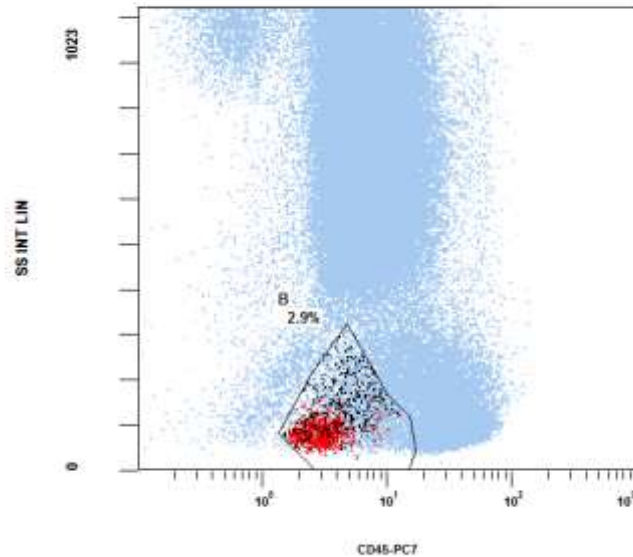
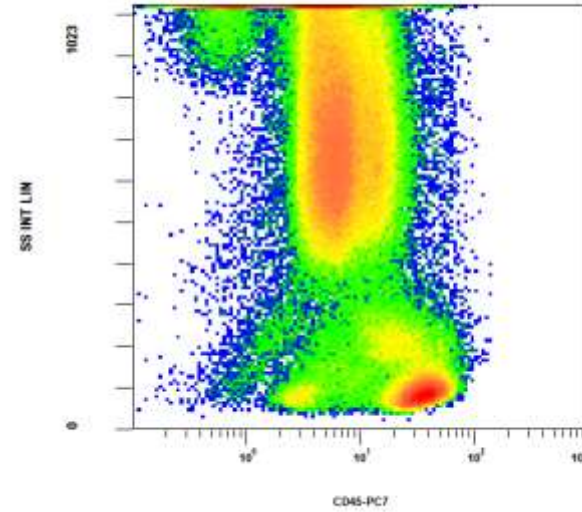
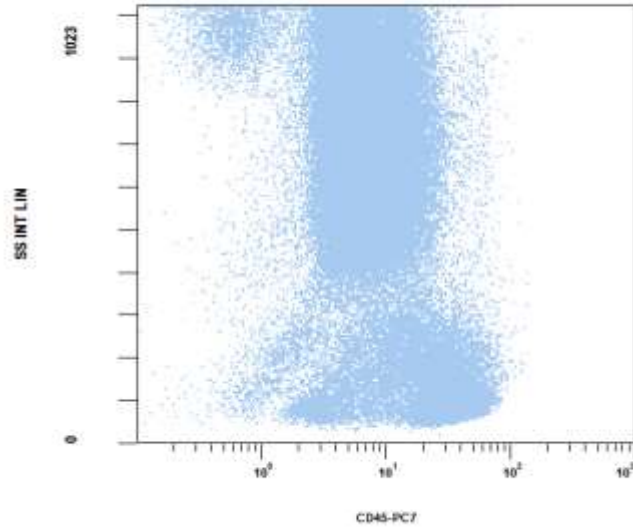


512X512



Density plots - resolution



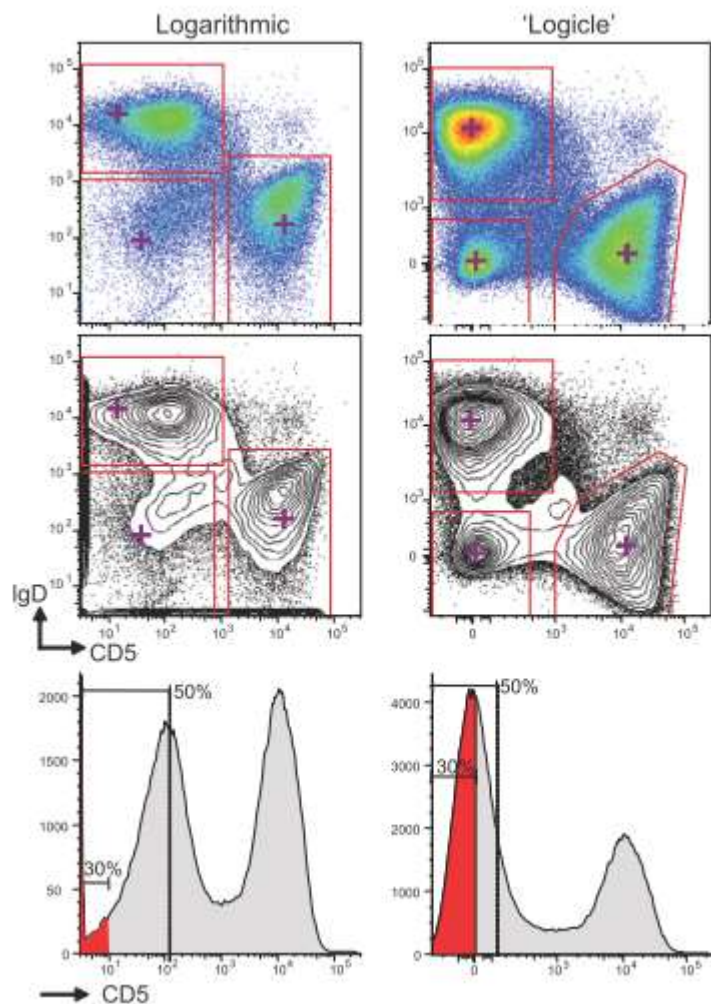


Μερικές φορές το density plot οδηγεί στον εντοπισμό ενός πληθυσμού

Σύγκριση «παλιού» και «νέου»

Τώρα φαίνονται τα διπλά αρνητικά κύτταρα στη logicle κλίμακα.

Στην απεικόνιση παίζει ρόλο ο θόρυβος και η αντιστάθμιση (το πώς το επεξεργάζεται ο κυτταρομετρητής).



Σύγκριση «παλιού» και «νέου»

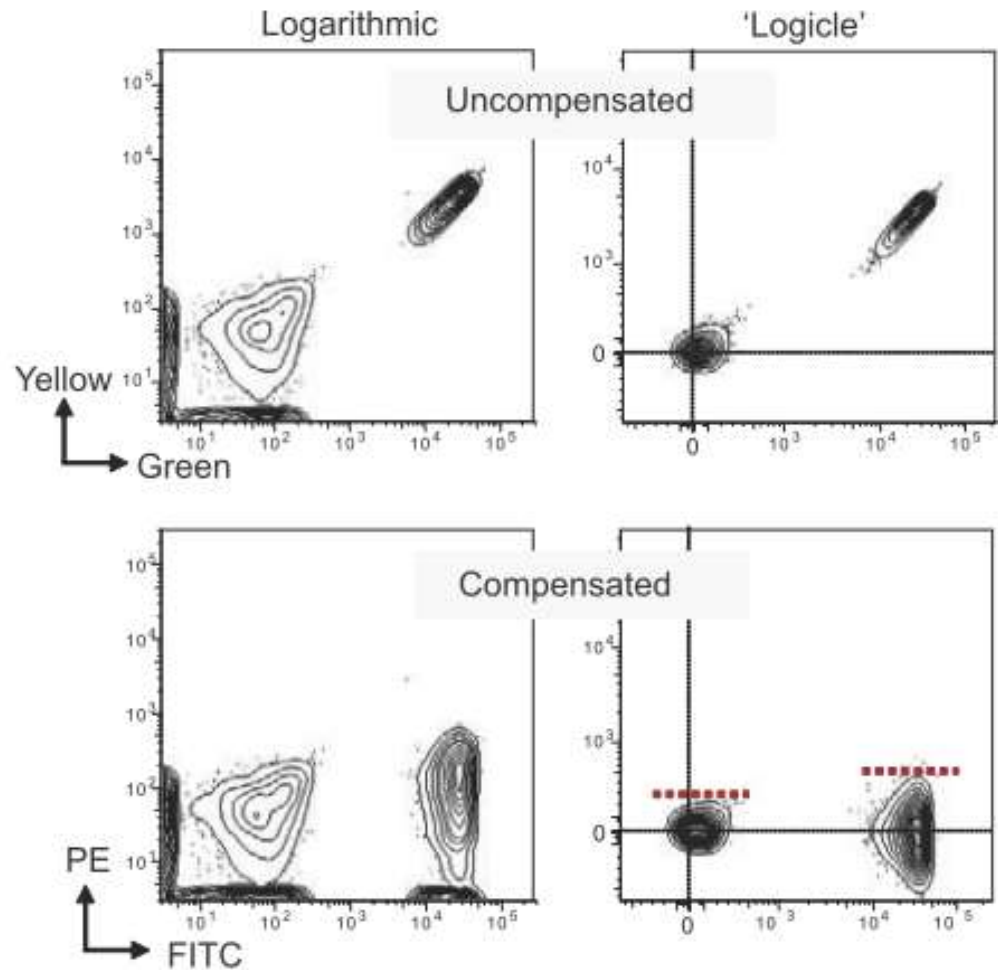
Η αντιστάθμιση στη logicle κλίμακα.

Τα αποτελέσματα πρέπει να συλλέγονται και να αποθηκεύονται χωρίς αντιστάθμιση και αυτή πρέπει να εφαρμόζεται στην ανάλυση.

Αυτό είναι δυνατό στους σύγχρονους κυτταρομετρητές.

Μην ξεχνάτε το FMO!

Τυποποίηση μηχανημάτων με σφαιρίδια!



Leonore A Herzenberg et al
2006



Ι ΚΥΤΤΑΡΟΜΕΤΡΙΚΗ ΒΑΒΕΛ

- Όταν ο άνθρωπος «έπεσε» από τον Παράδεισο έπαψε να μιλάει τη μία «γλώσσα»... δημιουργήθηκε ο Πύργος της Βαβέλ.
- Μήπως σήμερα στην κυτταρομετρία χάθηκε η μία κοινή γλώσσα συνεννόησης;
- Δεδομένα, που έχουν συλλεγεί με τους παλιούς κυτταρομετρητές με λογαριθμικούς ενισχυτές και ηλεκτρονικούς ανιχνευτές της κορυφής μπορούν να συγκριθούν με δεδομένα και εικόνες που έχουν συλλεγεί από μοντέρνους κυτταρομετρητές με ψηφιακή επεξεργασία γραμμικού σήματος; (ιδιαίτερα στα άκρα της κλίμακας)

Εκπαίδευση σαν τη σημερινή
οδηγεί στην κοινή γλώσσα

Ευχαριστώ πολύ