

# ΠΟΛΥΠΑΡΑΓΟΝΤΙΚΗ ΑΝΑΛΥΣΗ

# ΠΟΛΥΠΑΡΑΓΟΝΤΙΚΗ ΚΑΙ ΠΟΛΥΜΕΤΑΒΛΗΤΗ ΑΝΑΛΥΣΗ

Εξαρτημένες μεταβλητές	Ανεξάρτητες μεταβλητές
Μονοπαραγοντική και μονομεταβλητή ανάλυση	
Μια εξαρτημένη	Μια ανεξάρτητη
Πολυπαραγοντική ανάλυση (Multifactor analysis)	
Μια εξαρτημένη	Δύο ή περισσότερες ανεξάρτητες (παράγοντες)
Πολυμεταβλητή ανάλυση (Multivariable or multivariate analysis)	
Δύο ή περισσότερες εξαρτημένες	Μια ανεξάρτητη (Ένας παράγοντας)
Πολυπαραγοντική και πολυμεταβλητή ανάλυση	
Δύο ή περισσότερες εξαρτημένες	Δύο ή περισσότερες ανεξάρτητες (παράγοντες)

Σημείωση: Στις πλείστες των περιπτώσεων, μετά την πραγματοποίηση πολυπαραγοντικών ή/και πολυμεταβλητών αναλύσεων, αν προκύψει οποιαδήποτε στατιστική σημαντικότητα, ακολουθεί απαραίτητα η μονοπαραγοντική ή/και μονομεταβλητή ανάλυση για να αναδείξουμε σε τι οφείλεται η ύπαρξη στατιστικής σημαντικότητας.

## Πολυπαραγοντική ανάλυση

Μια εξαρτημένη

Δύο παράγοντες

P300 amplitudes at electrode Cz  
(Συνεχής μεταβλητή που ακολουθεί την κανονική κατανομή)

1. Group - Controls (n=57) vs Dyslexics (n=114)
2. Gender – Male (n=99) vs Female (n=72)

Αποτελέσματα της μονοπαραγοντικής ανάλυσης – t-test για δύο ανεξάρτητα δείγματα

**Group Statistics**

	Group	N	Mean	Std. Deviation	Std. Error Mean
P300 amplitudes at Cz	Controls	57	4.4334	4.58336	.60708
	Dyslexics	114	3.0829	4.48561	.42012

**Independent Samples Test**

		t-test for Equality of Means		
		t	df	Sig. (2-tailed)
P300 amplitudes at Cz	Equal variances assumed	1.843	169	.067

**Group Statistics**

	Sex	N	Mean	Std. Deviation	Std. Error Mean
P300 amplitudes at Cz	Male	99	3.3122	4.60121	.46244
	Female	72	3.8368	4.49275	.52948

**Independent Samples Test**

		t-test for Equality of Means		
		t	df	Sig. (2-tailed)
P300 amplitudes at Cz	Equal variances assumed	-.743	169	.458

## Αποτελέσματα της πολυπαραγοντικής ανάλυσης

### Tests of Between-Subjects Effects

Dependent Variable: p300a\_13

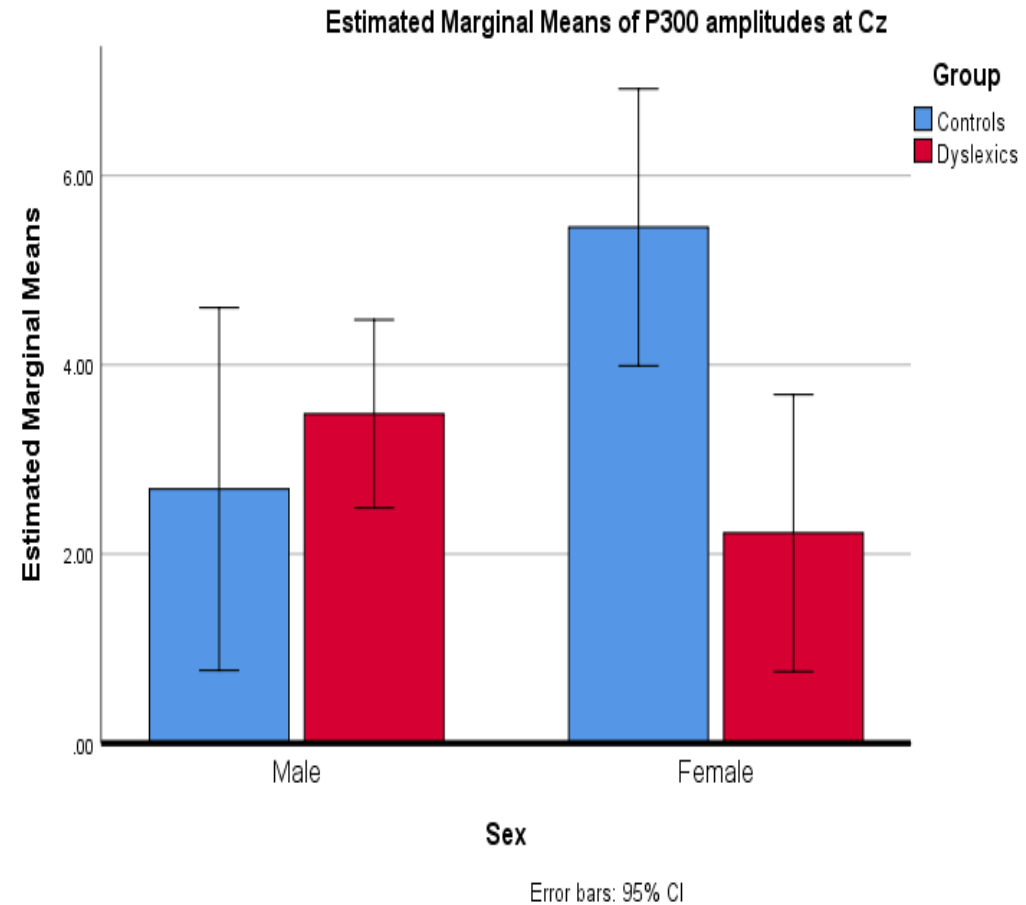
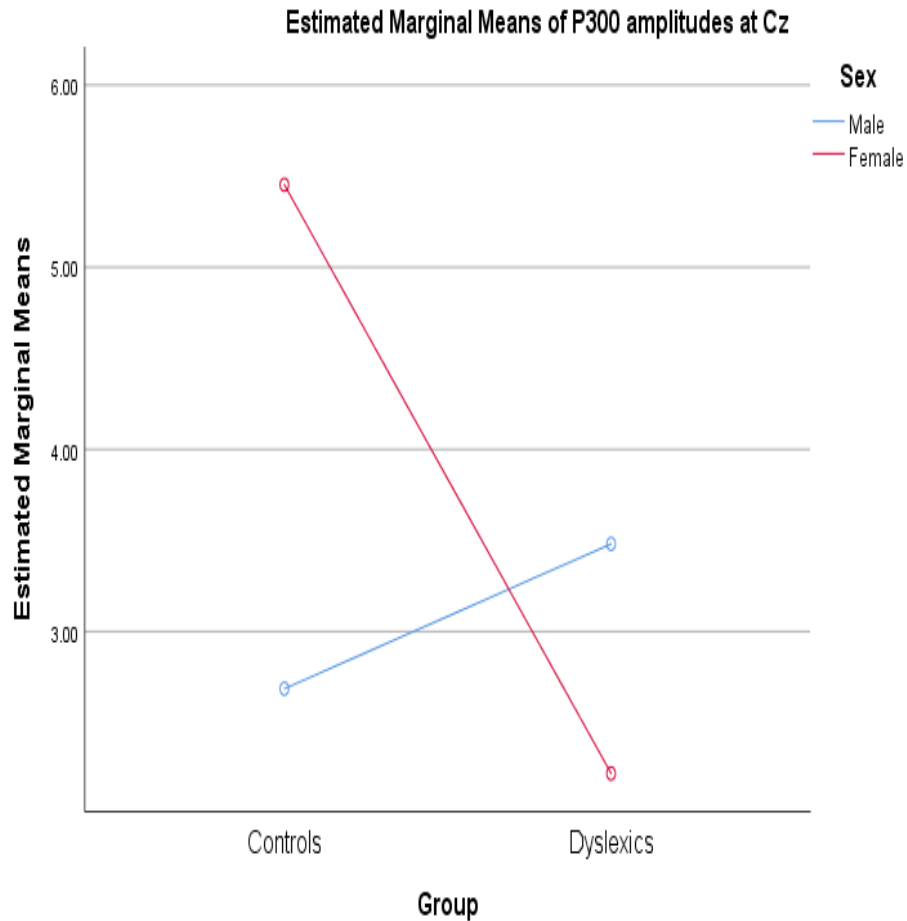
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	209.980	3	69.993	3.532	.016
Intercept	1651.366	1	1651.366	83.332	.000
groupid	51.210	1	51.210	2.584	.110
sexid	19.577	1	19.577	.988	.322
<b>groupid * sexid</b>	139.821	1	139.821	7.056	<b>.009</b>
Error	3309.380	167	19.817		
Total	5653.867	171			
Corrected Total	3519.359	170			

Εδώ βλέπουμε καταρχήν να επαναλαμβάνονται τα αποτελέσματα της μονοπαραγοντικής ανάλυσης, δηλαδή ότι οι παράγοντες Group και Gender από μόνοι τους δεν επηρεάζουν την εξαρτημένη μεταβλητή.

Όμως εδώ εμφανίζεται και ο παράγοντας αλληλεπίδρασης των Group X Gender (two-way interaction effect), ο οποίος δείχνει να έχει στατιστικά σημαντική ( $p=0.009$ ) επίδραση στην εξαρτημένη μεταβλητή.

Η ύπαρξη σημαντικότητας της αλληλεπίδρασης δευτέρου βαθμού αυτόματα καθιστά σημαντικούς και τους πρωτοβάθμιους παράγοντες Group και Gender ανεξάρτητα από τα δικά τους αποτελέσματα.

# Γραφική αναπαράσταση (εξήγηση) της αλληλεπίδρασης (The crossover effect)



Στους μάρτυρες τα κορίτσια είχαν πολύ μεγαλύτερα amplitudes, ενώ αντίστροφα στους δυσλεκτικούς τα κορίτσια είχαν κάπως μικρότερες τιμές.

Τα δυσλεκτικά αγόρια είχαν κάπως μεγαλύτερες τιμές amplitudes από τα αγόρια-μάρτυρες, ενώ αντίστροφα τα δυσλεκτικά κορίτσια είχαν πολύ μικρότερες τιμές από τα κορίτσια-μάρτυρες.

# Πως αυτό θα εκφραστεί στατιστικά σε ένα άρθρο

**Group Statistics**

Sex	Group	N	Mean	Std. Deviation	Std. Error Mean	
Male	P300 amplitudes at Cz	Controls	21	2.6858	4.69377	1.02426
		Dyslexics	78	3.4808	4.59204	.51995
Female	P300 amplitudes at Cz	Controls	36	5.4529	4.25684	.70947
		Dyslexics	36	2.2207	4.17764	.69627

**Independent Samples Test**

			t-test for Equality of Means		
Sex			t	df	Sig. (2-tailed)
Male	P300 amplitudes at Cz	Equal variances assumed	-.701	97	.485
Female	P300 amplitudes at Cz	Equal variances assumed	3.252	70	.002

**Group Statistics**

Group	Sex	N	Mean	Std. Deviation	Std. Error Mean	
Controls	P300 amplitudes at Cz	Male	21	2.6858	4.69377	1.02426
		Female	36	5.4529	4.25684	.70947
Dyslexics	P300 amplitudes at Cz	Male	78	3.4808	4.59204	.51995
		Female	36	2.2207	4.17764	.69627

**Independent Samples Test**

			t-test for Equality of Means		
Group			t	df	Sig. (2-tailed)
Controls	P300 amplitudes at Cz	Equal variances assumed	-2.280	55	.027
Dyslexics	P300 amplitudes at Cz	Equal variances assumed	1.400	112	.164

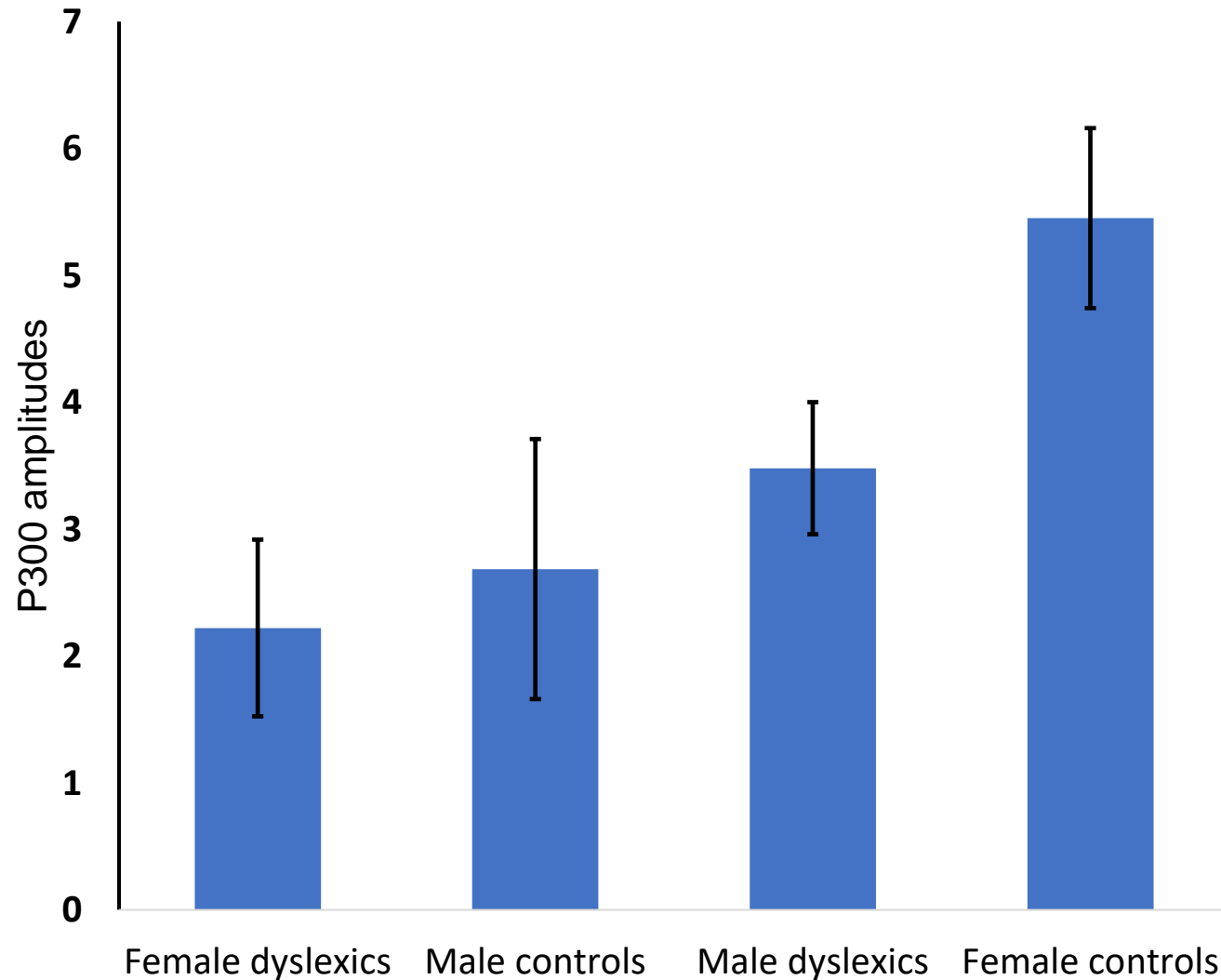
## Methods – Statistical analysis:

The effect of Group, Gender and their two-way interaction on the P300 amplitudes was tested with the General Linear Model, followed by post hoc comparisons of the effect of Group within each Gender and the effect of Gender within each Group. The level of significance was set at 0.05.

## Results:

The General Linear Model showed that both Group and Gender did not have a significant effect on P300 amplitudes, however their significant effect was qualified by the significance of the Group X Gender interaction effect ( $p=0.009$ ). Subsequent pairwise comparisons revealed that Group differences were significant only for female subjects ( $p=0.002$ ), while Gender differences were significant only for the control group ( $p=0.027$ ).

## Εναλλακτικός τρόπος παρουσίασης των αποτελεσμάτων



Δημιουργούμε μια μεταβλητή που περιέχει τις 4 κατηγορίες παιδιών με βάση την ομάδα και το φύλο, όπως φαίνεται στο διάγραμμα αριστερά. Στο διάγραμμα φαίνεται καθαρά ότι η ομάδα που διαφέρει στατιστικά σημαντικά από τις υπόλοιπες τρεις είναι η Female controls.

Σε περίπτωση που έχουμε τρεις παράγοντες (a, b, c), το πολυπαραγοντικό μοντέλο πρέπει να μελετά όλες τις παρακάτω επιδράσεις.

- Main effects of a
- Main effects of b
- Main effects of c
- Effect of two-way interaction a x b
- Effect of two-way interaction a x c
- Effect of two-way interaction b x c
- Effect of three-way interaction a x b x c



# Πολλαπλή γραμμική παλινδρόμηση

- Στην απλή γραμμική παλινδρόμηση έχουμε την εξαρτημένη μεταβλητή  $y$  και την ανεξάρτητη μεταβλητή  $x$  και ψάχνουμε να βρούμε την εξίσωση γραμμικής παλινδρόμησης  $y = a + bx$ , όπου θέλουμε να αποδείξουμε ότι ο συντελεστής  $b$  διαφέρει από το 0. Επίσης αναφέρουμε την τιμή  $R^2$  (συντελεστής αιτιολόγησης)
- Στην πολλαπλή γραμμική παλινδρόμηση έχουμε την εξαρτημένη μεταβλητή  $y$  και δύο ή περισσότερες ανεξάρτητες μεταβλητές (ας υποθέσουμε εδώ τρεις –  $x_1, x_2, x_3$ ) και ψάχνουμε να βρούμε την εξίσωση γραμμικής παλινδρόμησης  $y = a + b_1x_1 + b_2x_2 + b_3x_3$ , όπου θέλουμε να αποδείξουμε ότι ο κάθε ένας συντελεστής  $b_1, b_2, b_3$  διαφέρει από το 0. Επίσης και εδώ αναφέρουμε την τιμή  $R^2$  (συντελεστής πολλαπλής αιτιολόγησης)

## Πολυπαραγοντική ανάλυση

Μια εξαρτημένη

Δύο διωνυμικοί παράγοντες

Διωνυμική μεταβλητή: 28-day mortality (0-Survivor (n=182) 1-Non survivor (n=28))

1. Occurrence of High mobility group box 1 (HMGB1) peak: 0-Early peak 1-Late peak
2. Inflammatory disease: 0-Absence 1-Presence

### LATE PEAKS OF HMGB1 AND SEPSIS OUTCOME: EVIDENCE FOR SYNERGY WITH CHRONIC INFLAMMATORY DISORDERS

Univariate statistics

Crosstab

		28-day outcome		Total	
		Survival	Death		
Peak occurrence	Early peak	Count	132	14	146
		% within Peak occurrence	90.4%	9.6%	100.0%
	Late peak	Count	50	14	64
		% within Peak occurrence	78.1%	21.9%	100.0%
Total		Count	182	28	210
		% within Peak occurrence	86.7%	13.3%	100.0%

Fisher exact test –  $p=0.026$

Logistic regression – OR=2.64 95% CI=1.18-5.93

Crosstab

		28-day outcome		Total	
		Survival	Death		
Presence of at least one inflammatory disease	No	Count	104	12	116
		% within Presence of at least one inflammatory disease	89.7%	10.3%	100.0%
	Yes	Count	78	16	94
		% within Presence of at least one inflammatory disease	83.0%	17.0%	100.0%
Total		Count	182	28	210
		% within Presence of at least one inflammatory disease	86.7%	13.3%	100.0%

Fisher exact test –  $p=0.220$

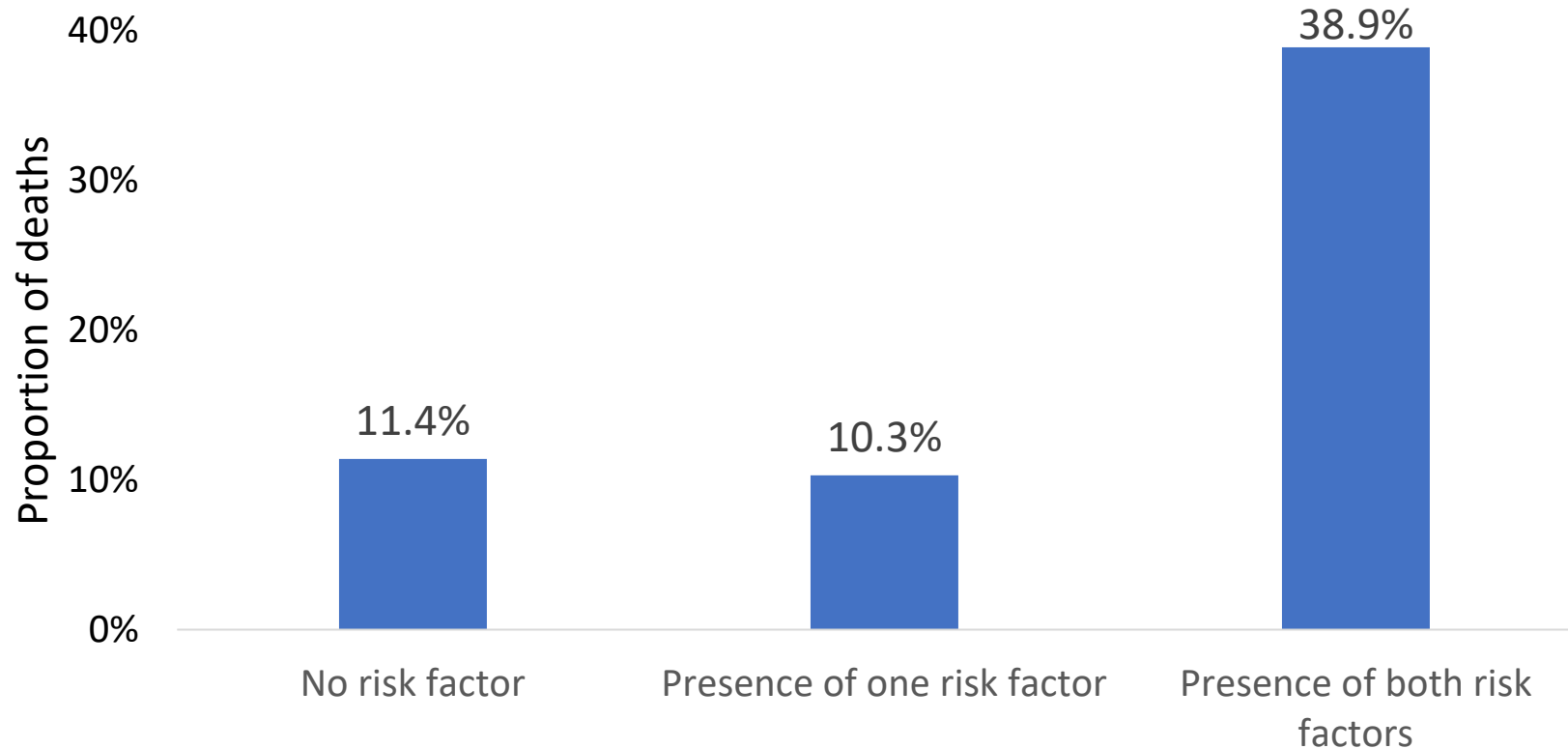
Logistic regression – OR=1.78 95% CI=0.80-3.97

# Multivariate model (Stepwise logistic regression)

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
								Lower	Upper
Step 1 <sup>a</sup>	Peak occurrence by Presence of at least one inflammatory disease	1.562	.463	11.377	1	.001	4.766	1.924	11.809
	Constant	-2.203	.248	78.697	1	.000	.110		

a. Variable(s) entered on step 1: Peak occurrence \* Presence of at least one inflammatory disease .



Το ίδιο πρέπει να γίνεται (και δεν γίνεται) και στο multivariate Cox regression

## Activate: Randomized Clinical Trial of BCG Vaccination against Infection in the Elderly

**Table 3. Univariate and Multivariate Analysis of the Effects of Covariates on the Incidence of at Least One Infection until Month 12**

Covariates	No. Infection (n = 99)	At Least One New Infection (n = 51)	Univariate Analysis			Multivariate Analysis		
			HR	95% CI	p value	HR	95% CI	p value
BCG vaccination, no. (%) <sup>a</sup>	54 (54.4)	18 (35.3)	0.55	0.31–0.97	0.035	0.56	0.32–0.99	0.048
Male gender, no. (%) <sup>b</sup>	40 (40.4)	27 (52.9)	0.66	0.38–1.15	0.145			
CCI > 4, no. (%) <sup>c</sup>	61 (61.6)	40 (78.4)	2.11	1.08–4.12	0.028			
Type 2 diabetes mellitus, no. (%) <sup>b</sup>	31 (31.6)	21 (41.2)	1.34	0.76–2.33	0.310			
Chronic heart failure, no. (%) <sup>b</sup>	26 (26.5)	17 (33.3)	1.35	0.75–2.42	0.313			
Chronic renal disease, no. (%) <sup>a</sup>	13 (13.3)	13 (25.2)	1.96	1.05–3.69	0.036	1.95	1.04–3.66	0.038
COPD, no. (%) <sup>a</sup>	10 (10.2)	13 (25.5)	2.14	1.14–4.02	0.018	2.12	1.13–3.99	0.019

## Αθροιστική και πολλαπλασιαστική επίδραση παραγόντων κινδύνου (Additive and multiplicative effect of risk factors)

HLA DQA1X0501 and DRB1\*0301 antigens do not independently convey susceptibility to Graves' disease  
March 2001 Journal of endocrinological investigation 24(2):88-91

- HLA DQA1X0501 has an effect with  $OR=4.2$
- DRB1\*0301 has an effect with  $OR=4.5$
- If the two risk factors act additively and independently then the overall effect would be  $OR=4.2+4.5-1=7.7$
- If the two risk factors act in a multiplicative manner, i.e. synergistically in an interaction fashion then the overall effect would be  $OR=4.2 \times 4.5 = 18.9$
- In fact the joint effect was found to be  **$OR=27$  !!!**
- The presence of both antigens practically determines the development of Graves' disease.
- This signifies the presence of **synergy**.