

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ**



**ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ  
ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ  
ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ**

**ΜΕΘΟΔΟΙ ΕΠΙΒΛΕΨΗΣ  
ΕΠΙΔΗΜΙΟΛΟΓΙΚΩΝ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ  
ΜΕ ΧΡΗΣΗ ΣΤΑΤΙΣΤΙΚΩΝ ΣΑΡΩΣΗΣ**

Μαρία Π. Αμπατζιάδου

Διπλωματική Εργασία

που υποβλήθηκε στο Τμήμα Στατιστικής και Ασφαλιστικής  
Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των  
απαιτήσεων για την απόκτηση του Μεταπτυχιακού  
Διπλώματος Ειδίκευσης στην Εφαρμοσμένη Στατιστική

Πειραιάς  
Αύγουστος 2014



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ**



**ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ  
ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ**

**ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ  
ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ**

**ΜΕΘΟΔΟΙ ΕΠΙΒΛΕΨΗΣ  
ΕΠΙΔΗΜΙΟΛΟΓΙΚΩΝ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ  
ΜΕ ΧΡΗΣΗ ΣΤΑΤΙΣΤΙΚΩΝ ΣΑΡΩΣΗΣ**

Μαρία Π. Αμπατζιάδου

Διπλωματική Εργασία

που υποβλήθηκε στο Τμήμα Στατιστικής και  
Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς ως  
μέρος των απαιτήσεων για την απόκτηση του  
Μεταπτυχιακού Διπλώματος Ειδίκευσης στην Εφαρμοσμένη  
Στατιστική

Πειραιάς

Αύγουστος 2014

Πανεπιστήμιο Πειραιώς

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίσθηκε από τη ΓΣΕΣ του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ' αριθμ. .... συνεδρίασή του σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών στην Εφαρμοσμένη Στατιστική

Τα μέλη της Επιτροπής ήταν:

- Μπερσίμης Σωτήριος (Επιβλέπων)
- Μπούτσικας Μιχαήλ
- Παναγιωτάκος Δημοσθένης

Η έγκριση της Διπλωματικής Εργασίας από το Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς δεν υποδηλώνει αποδοχή των γνώμων του συγγραφέα.

**UNIVERSITY OF PIRAEUS**



**DEPARTMENT OF STATISTICS  
AND INSURANCE SCIENCE**

**POSTGRADUATE PROGRAM IN  
APPLIED STATISTICS**

**SURVEILLANCE METHODS OF  
EPIDEMIOLOGICAL  
CHARACTERISTICS USING SCAN  
STATISTICS**

By

Maria P. Ampatziadou

MSc Dissertation

submitted to the Department of Statistics and  
Insurance Science of the University of Piraeus in  
partial fulfilment of the requirements for the degree  
of Master of Science in Applied Statistics

Piraeus, Greece  
August 2014

Πανεπιστήμιο Πειραιώς

# Ευχαριστίες

Η παρούσα διπλωματική εργασία εκπονήθηκε από τη φοιτήτρια Αμπατζιάδου Μαρία του μεταπτυχιακού προγράμματος σπουδών Εφαρμοσμένης Στατιστικής του τμήματος Στατιστικής και Αναλογιστικής Επιστήμης του Πανεπιστημίου Πειραιώς.

Μέσω της διπλωματικής αυτής εργασίας μου δόθηκε η δυνατότητα να διευρύνω τις γνώσεις μου στον κλάδο της στατιστικής, και ιδιαίτερα στον τομέα της βιοστατιστικής, εργαζόμενη πάνω σε πραγματικά ιατρικά δεδομένα, με χρήση στατιστικών μεθόδων άγνωστες μέχρι πρότεινος σε εμένα.

Θα ήθελα να ευχαριστήσω τον Δρ. Μπερσίμη Σωτήριο, λέκτορα στο Πανεπιστήμιο Πειραιώς, για την καθοδήγηση, την υποστήριξη και βοήθειά του καθ' όλη τη διάρκεια διεκπαιρέωσης της παρούσας μελέτης. Τον ευχαριστώ θερμά για τις γνώσεις που μου παρείχε, αλλά και για το αμείωτο ενδιαφέρον του και υπομονή του κατά τη συγγραφή και διόρθωση της εργασίας.

Ιδιαίτερες ευχαριστίες θα ήθελα να απευθύνω στον κ. Ιωαννίδη Αναστάσιο, υπεύθυνο ιατρό του μικροβιολογικού εργαστηρίου του Αιγινήτειου νοσοκομείου, για την πολύτιμη βοήθεια που μου προσέφερε όσον αφορά τα ιατρικά δεδομένα, τη συλλογή τους και τη συγγραφή του ιατρικού σκέλους της παρούσας διπλωματικής μελέτης. Τον ευχαριστώ θερμά για τη συνεργασία και το ενδιαφέρον του.

Τέλος, ευχαριστίες οφείλω στην οικογένειά μου και στο φίλο μου Γιώργο Μασούρα, για την ηθική υποστήριξη και βοήθεια όλο το διάστημα των μεταπτυχιακών μου σπουδών, όπως επίσης και για την κατανόηση και ενθάρρυνση κατά τη διάρκεια εκπόνησης της διπλωματικής μου εργασίας.



# Περίληψη

Η πολλαπλή σκλήρυνση, επίσης γνωστή και ως σκλήρυνση κατά πλάκας ή διάχυτη εγκεφαλομυελίτιδα, είναι μια φλεγμονώδης ασθένεια στην οποία τα μονωτικά καλύμματα των νευρικών κυττάρων στον εγκέφαλο και τη σπονδυλική στήλη καταστρέφονται. Αυτή η καταστροφή διαταράσσει την ικανότητα τμημάτων του νευρικού συστήματος να επικοινωνούν, με αποτέλεσμα ένα ευρύ φάσμα ενδείξεων και συμπτωμάτων. Αν και το αίτιο δεν είναι ακόμα σαφές, πιστεύεται ότι ο υποκείμενος μηχανισμός είναι είτε η καταστροφή από το ανοσοποιητικό σύστημα είτε η αποτυχία των κυττάρων που παράγουν μυελίνη. Δεν υπάρχει κάποια γνωστή θεραπεία ίασης για την πολλαπλή σκλήρυνση, ενώ οι ήδη υπάρχουσες επιχειρούν να βελτιώσουν τη λειτουργία του οργανισμού μετά από μια κρίση για την πρόληψη νέων.

Η παρατήρηση και η ανάλυση δεδομένων που βρίσκονται κατανεμημένα στο χώρο ή/και στο χρόνο μπορεί να επιτευχθεί με διάφορες στατιστικές συναρτήσεις σάρωσης (*scan statistics*), τόσο σε μονοδιάστατες περιπτώσεις, όσο και σε διδιάστατες. Αναφέρονται επιγραμματικά έννοιες που αφορούν τη μονοδιάστατη περίπτωση, ενώ εμβάθυνση γίνεται σε γενικεύσεις αυτών, εφαρμόσιμες σε προβλήματα δύο ή παραπάνω διαστάσεων.

Στην παρούσα διπλωματική εργασία, εφαρμόστηκε η στατιστική συνάρτηση σάρωσης (*spatial scan statistics*) που εισήχθη από τον Kuldorff (1997) κάνοντας χρήση ενός μοντέλου Poisson. Τα δεδομένα που συλλέχθηκαν αφορούσαν 179 περιπτώσεις ασθενών οι οποίοι νοσηλεύθηκαν σε διάφορα νοσοκομεία της Ελλάδος και διαγνώστηκαν θετικοί στη νόσο της Σκλήρυνσης Κατά Πλάκας.

Με βάση τη διεύθυνση κατοικίας του εκάστοτε ασθενούς και με χρήση του στατιστικού πακέτου SaTScan, εξετάζεται η υπόθεση τυχαίας κατανομής των γεγονότων εμφάνισης στο χώρο. Η γεωγραφική περιοχή που μελετήθηκε είναι η Στερεά Ελλάδα, η Πελοπόννησος και η Ήπειρος.

Λόγω πολυπλοκότητας του προβλήματος, η παρούσα διπλωματική μελέτη περιορίστηκε μόνο στην εξέταση των γεγονότων στο χώρο, χωρίς να ληφθούν υπόψη άλλοι παράγοντες, όπως φύλο, ηλικία και λοιπά. Ωστόσο, κρίνεται απαραίτητη περαιτέρω μελέτη με εμφανείς τις προοπτικές επέκτασης των αποτελεσμάτων.

# Abstract

Multiple sclerosis, also known as disseminated sclerosis or encephalomyelitis disseminata, is an inflammatory disease in which the insulating covers of nerve cells in the brain and spinal cord are damaged. This damage affects the ability of parts of the nervous system to communicate, resulting in a wide range of signs and symptoms. Despite the fact that the cause is not clear yet, the underlying mechanism is thought to be either destruction by the immune system or the inability of the cells to produce myelin. Until nowadays there is not any known cure for multiple sclerosis. Treatments attempt to improve function after an attack and prevent new attacks.

Data analysis of Observations that are distributed in space or even in time, can be obtained using various scan statistics, either in one or two dimensional cases. This study presents in brief some ideas related to the one dimensional case, while its main aim is the investigation of two dimensional problems.

In particular, spatial scan statistics introduced by Kulldorff (1997) have been employed, using a Poisson model. Data have been derived from 179 cases of patients which are hospitalized in three major geographic departments of Greece: Sterea Ellada, Epirus and Peloponesus.

By considering the home address of each patient and using the statistical program SaTScan, we test the hypothesis of a random distribution of cases appearing in space, versus the hypothesis of spatial clustering which could imply location dependence.

Due to the complexity of the problem, the present study focuses only in the presence of events in space, without taking into consideration factors such as gender, age etc. However, a further research would be judged necessary.

Πανεπιστήμιο Πειραιώς

# Περιεχόμενα

Περίληψη	viii
Abstract	ix
Περιεχόμενα	xii
Κατάλογος Πινάκων	xvii
Κατάλογος Σχημάτων	xviii
<b>1. ΕΙΣΑΓΩΓΗ</b>	<b>1</b>
1.1 Η ΝΟΣΟΣ ΤΗΣ ΣΚΛΗΡΥΝΣΗΣ ΚΑΤΑ ΠΛΑΚΑΣ	1
1.2 ΕΙΣΑΓΩΓΗ ΣΤΙΣ ΣΤΑΤΙΣΤΙΚΕΣ ΣΑΡΩΣΗΣ	4
1.2.1 Η Μονοδιάστατη Περίπτωση	4
1.2.2 Η Διδιάστατη Περίπτωση	6
1.2.3 Εφαρμογές	8
1.2.4 Περιγραφή Ανάλυσης	10
1.3 ΜΕΘΟΔΟΛΟΓΙΑ ΣΤΑΤΙΣΤΙΚΩΝ ΣΑΡΩΣΗΣ	12
1.3.1 Πιθανοθεωρητική Προσέγγιση	13
1.3.2 Χωρική Σάρωση με Χρήση Πιθανοφάνειας	14
1.3.3 Σύγκριση των Δύο Προσεγγίσεων	17
1.4 ΤΟ ΛΟΓΙΣΜΙΚΟ ΠΑΚΕΤΟ SaTScan	18
1.5 ΣΥΝΟΨΗ	21
<b>2. Η ΝΟΣΟΣ ΤΗΣ ΣΚΛΗΡΥΝΣΗΣ ΚΑΤΑ ΠΛΑΚΑΣ</b>	<b>22</b>
2.1 ΕΙΣΑΓΩΓΗ	22
2.2 ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ	23
2.3 ΦΛΕΓΜΟΝΩΔΕΙΣ ΑΠΟΜΥΕΛΙΝΩΤΙΚΕΣ ΝΟΣΟΙ	24
2.3.1 Πολλαπλή Σκλήρυνση (Σκλήρυνση Κατά Πλάκας-Multiple Sclerosis)	24
2.3.1.1 Γενικά Χαρακτηριστικά	25
2.3.1.2 Αιτιοπαθογένεια	26
2.3.1.3 Παθοφυσιολογία	27

2.3.1.4	Παθολογική ανατομική	28
2.3.1.5	Επιδημιολογία	29
2.3.1.6	Κλινική Εικόνα	30
2.3.1.7	Εργαστηριακές Εξετάσεις	32
2.3.1.8	Διάγνωση	34
2.3.1.9	Πρόγνωση	38
2.3.1.10	Θεραπεία	39
2.3.2	Συγκεντρική Σκλήρυνση του Baló	40
2.3.3	Οπτική Νευρομυελίτιδα (Νόσος του Devic)	40
2.3.4	Διάχυτη Σκλήρυνση ή Διάχυτη Μυελινοκλαστική Σκλήρυνση (Νόσος του Schilder)	41
2.3.5	Προϊούσα Πολυεστιακή Λευκοεγκεφαλοπάθεια	41
2.3.6	Οξεία Διάσπαρτη Εγκεφαλομυελίτιδα	42
2.3.7	Οξεία Αιμορραγική Εγκεφαλομυελίτιδα	42
2.4	ΜΗ-ΦΛΕΓΜΟΝΩΔΕΙΣ ΑΠΟΜΥΕΛΙΝΩΤΙΚΕΣ ΝΟΣΟΙ	43
2.4.1	Δυσμυελινωτικές Νόσοι ή Λευκοδυστροφίες	43
2.4.2	Τροφικές-Τοξικές Απομυελινωτικές Νόσοι	43
2.4.2.1	Κεντρική Γεφυρική Μυελινόλυση	44
2.4.2.2	Πρωτοπαθής Εκφύλιση του Μεσολοβίου (Νόσος των Marchiafava-Bignami)	44
2.5	ΣΥΝΟΨΗ	44
<b>3.</b>	<b>ΣΤΑΤΙΣΤΙΚΕΣ ΣΑΡΩΣΗΣ – ΑΝΑΣΚΟΠΗΣΗ ΜΕΘΟΔΩΝ</b>	<b>46</b>
3.1	ΕΙΣΑΓΩΓΗ	46
3.2	Η ΜΟΝΟΔΙΑΣΤΑΤΗ ΠΕΡΙΠΤΩΣΗ	47
3.2.1	Διωνυμική Κατανομή Γεγονότων	48
3.2.2	Poisson Κατανομή Γεγονότων	50
3.3	Η ΔΙΔΙΑΣΤΑΤΗ ΠΕΡΙΠΤΩΣΗ	51
3.3.1	Πιθανοθεωρητική Προσέγγιση	52
3.3.2	Χωρική Στατιστική Συνάρτηση Σάρωσης του Kulldorff	54
3.3.2.1	Εισαγωγή	54
3.3.2.2	Μοντέλα Poisson και Bernoulli	57
3.3.2.3	Έλεγχος Λόγου Πιθανοφάνειας (Likelihood Ratio Test)	59

3.4	ΣΥΝΟΨΗ	64
<b>4.</b>	<b>ΕΦΑΡΜΟΓΕΣ ΣΤΑΤΙΣΤΙΚΩΝ ΣΥΝΑΡΤΗΣΕΩΝ ΣΑΡΩΣΗΣ</b>	<b>65</b>
4.1	ΕΙΣΑΓΩΓΗ	65
4.2	ΕΦΑΡΜΟΓΕΣ ΣΤΗΝ ΕΠΙΔΗΜΙΟΛΟΓΙΑ	66
4.2.1	Εφαρμογές στην Επιδημιολογία του Καρκίνου	67
4.2.2	Εφαρμογές στην Επιδημιολογία Μολυσματικών Νόσων	71
4.3	ΕΦΑΡΜΟΓΕΣ ΣΤΗΝ ΠΑΡΑΣΙΤΟΛΟΓΙΑ	74
4.4	ΑΛΛΕΣ ΙΑΤΡΙΚΕΣ ΕΦΑΡΜΟΓΕΣ	74
4.5	ΕΦΑΡΜΟΓΕΣ ΣΤΗΝ ΚΤΗΝΙΑΤΡΙΚΗ	76
4.6	ΕΦΑΡΜΟΓΕΣ ΣΤΗΝ ΕΝΤΟΜΟΛΟΓΙΑ ΚΑΙ ΣΤΗΝ ΙΧΘΥΟΛΟΓΙΑ	77
4.7	ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΔΑΣΟΛΟΓΙΑ- ΒΟΤΑΝΟΛΟΓΙΑ	78
4.8	ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΓΕΩΛΟΓΙΑ	78
4.9	ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΨΥΧΟΛΟΓΙΑ	79
4.10	ΕΦΑΡΜΟΓΕΣ ΣΤΑ ΑΤΥΧΗΜΑΤΑ ΚΑΙ ΤΙΣ ΑΥΤΟΚΤΟΝΙΕΣ	79
4.11	ΕΦΑΡΜΟΓΕΣ ΣΤΙΣ ΑΝΘΡΩΠΙΣΤΙΚΕΣ ΕΠΙΣΤΗΜΕΣ	80
4.12	ΣΥΝΟΨΗ	81
<b>5.</b>	<b>ΧΩΡΙΚΗ ΑΝΑΛΥΣΗ ΚΡΟΥΣΜΑΤΩΝ ΣΚΠ ΣΤΗΝ ΗΠΕΙΡΟ, ΤΗΝ ΣΤΕΡΕΑ ΕΛΛΑΔΑ ΚΑΙ ΤΗΝ ΠΕΛΟΠΟΝΝΗΣΟ</b>	<b>82</b>
5.1	ΕΙΣΑΓΩΓΗ	82
5.2	ΜΕΘΟΔΟΛΟΓΙΑ ΑΝΑΛΥΣΗΣ	83
5.2.1	Συλλογή Δεδομένων	83
5.2.2	Επεξεργασία Δεδομένων	84
5.2.3	Στατιστική Μεθοδολογία	88
5.3	ΑΠΟΤΕΛΕΣΜΑΤΑ	90
5.3.1	Γεωγραφική Απεικόνιση	97
5.4	ΣΥΜΠΕΡΑΣΜΑΤΑ	99
5.5	ΣΥΝΟΨΗ	100
	<b>Παραρτήματα</b>	<b>102</b>
Π1	Αρχεία Δεδομένων SaTScan	103

Π2	Αποτελέσματα της Μεθόδου Trial-And-Error	125
Π3	GeoDa Ανάλυση	136
<b>Βιβλιογραφία</b>		137

Πανεπιστήμιο Πειραιώς

Πανεπιστήμιο Πειραιώς



# Κατάλογος Πινάκων

Πίνακας 1	Νέα διαγνωστικά κριτήρια για ΣΚΠ (κριτήρια McDonald <i>et al.</i> )	23
Πίνακας 2	Ενδεικτικό Τμήμα του αρχείου cases.txt	60
Πίνακας 3	Ενδεικτικό Τμήμα του αρχείου coordinates.txt	60
Πίνακας 4	Ενδεικτικό Τμήμα του αρχείου population.txt	61
Πίνακας 5	Χωρική Ανάλυση της Σκλήρυνσης κατά Πλάκας με Μέγιστο Μέγεθος Κάλυψης 20%	66
Πίνακας 6	Αποτελέσματα Επαναληπτικής Μεθόδου Εντοπισμού Στατιστικά Σημαντικών Συστάδων	68

# Κατάλογος Εικόνων

Εικόνα 1	Αξονική τομογραφία εγκεφάλου	22
Εικόνα 2	Αρχική οθόνη του πακέτου SaTScan	62
Εικόνα 3	Επιλογή ανάλυσης στο πακέτο SaTScan	64
Εικόνα 4	Επιλογή ποσοστού πληθυσμού εκτεθειμένο στον κίνδυνο	65
Εικόνα 5	Γραφική απεικόνιση αποτελεσμάτων μέσω του GeoDa	70

Πανεπιστήμιο Πειραιώς

Πανεπιστήμιο Πειραιώς

# Κεφάλαιο 1

## ΕΙΣΑΓΩΓΗ

### 1.1 Η ΝΟΣΟΣ ΤΗΣ ΣΚΛΗΡΥΝΣΗΣ ΚΑΤΑ ΠΛΑΚΑΣ

Σε αυτήν την παράγραφο δίνονται συνοπτικά μερικά από τα βασικά στοιχεία της νόσου για την οποία γίνεται λόγος στην παρούσα διπλωματική, αυτή της σκλήρυνσης κατά πλάκας. Πρόκειται για μια αρκετά διαδεδομένη νόσο η οποία προκαλεί ποικιλία συμπτωμάτων, μέχρι και αναπηρία. Θεραπεία δεν έχει βρεθεί, ωστόσο, βάσει νέων στοιχείων και προσεγγίσεων, γίνεται προσπάθεια από γιατρούς για την εύρεση ίασης της.

Η σκλήρυνση κατά πλάκας ή πολλαπλή σκλήρυνση (*multiple sclerosis, MS*) αποτελεί τη συχνότερη εκφυλιστική νευροεγκεφαλοπάθεια των ενηλίκων ηλικίας από 20 έως 40 ετών, προσβάλλοντας περίπου 2,5 εκατομμύρια άτομα παγκοσμίως ([www.wikipedia.org/multiplesclerosis](http://www.wikipedia.org/multiplesclerosis)) .

Έχει παρατηρηθεί ότι παγκοσμίως, η πολλαπλή σκλήρυνση είναι περισσότερο διαδεδομένη σε ανθρώπους που ζουν μακριά από τον

Ισημερινό, με μερικές εξαιρέσεις (Compston and Coles, 2008). Η αιτία αυτού του γεωγραφικού καταμερισμού δεν είναι πλήρως ξεκάθαρη ακόμη, ωστόσο αρκετά άρθρα έχουν συνταχθεί σχετικά με τη μελέτη αυτού του φαινομένου (Milo and Kahana, 2010). Συχνότερη είναι η εμφάνιση της νόσου σε περιοχές με βόρειο-Ευρωπαϊκούς πληθυσμούς με άμεση συνέπεια την ερμηνεία της γεωγραφικής διακύμανσης της νόσου σε παγκόσμιο επίπεδο (Milo and Kahana, 2010). Η μειωμένη έκθεση στον ήλιο οδηγεί σε μείωση της βιταμίνης D γεγονός που έχει προταθεί ως μια λογική εξήγηση της νόσου (Ascherio et al., 2010) Επίσης, η εποχή γέννησης και η εμφάνιση της νόσου ίσως σχετίζονται με βάση κάποιες έρευνες, κατά τις οποίες άτομα που γεννήθηκαν στο βόρειο ημισφαίριο και Νοέμβριο εμφάνισαν τη νόσο σε μικρότερο ποσοστό σε σχέση με άτομα που γεννήθηκαν τον Μάιο (Kulle et al., 2009). Ακόμη, περιβαλλοντικοί παράγοντες ενδεχομένως να παίζουν ρόλο κατά τη διάρκεια της παιδικής ηλικίας ενός ανθρώπου, αφού άνθρωποι που μετανάστευσαν σε μια διαφορετική περιοχή πριν την ηλικία των 15 ετών απέκτησαν το ρίσκο για εμφάνιση της νόσου της νέας περιοχής. Αν η μετανάστευση γινόταν μετά την ηλικία των 15 ετών, το άτομο κατείχε το ποσοστό ρίσκου της χώρας που προερχόταν (Compston and Coles, 2008).

Είναι μια χρόνια νευρολογική, αυτοάνοση ασθένεια, στην οποία το ανοσοποιητικό σύστημα επιτίθεται στο κεντρικό νευρικό σύστημα προκαλώντας απομυελίνωση. Η πολλαπλή σκλήρυνση μπορεί να προκαλέσει μια ποικιλία συμπτωμάτων όπως αλλαγές στην αφή, προβλήματα στην όραση, μυϊκή ατροφία, κατάθλιψη, προβλήματα στον προσανατολισμού και στην ομιλία, εξάντληση, νοητική βλάβη,

προβλήματα με την ισορροπία και πόνο. Σε ακραίες περιπτώσεις μπορεί να προκαλέσει και αναπηρία.

Η σκλήρυνση κατά πλάκας επηρεάζει τους νευρώνες, τα κύτταρα του εγκεφάλου και το νωτιαίο μυελό, τα οποία στέλνουν πληροφορίες, δημιουργούν τη σκέψη και την αντίληψη και επιτρέπουν στον εγκέφαλο να ελέγξει το σώμα. Τα νεύρα του σώματος περιβάλλονται από ένα λεπτό περίβλημα, το οποίο αποτελείται από μυελίνη, μια ουσία που βοηθά στην προστασία των νευρών καθώς και στη μετάδοση των νευρικών ώσεων. Η νόσος προκαλεί σταδιακή καταστροφή της μυελίνης, και όταν η μυελίνη καταστραφεί, οι νευρώνες δεν μπορούν να χειριστούν αποτελεσματικά τα ηλεκτρικά τους σήματα. Το όνομα πολλαπλή σκλήρυνση αναφέρεται στις πολλαπλές ουλές που δημιουργούνται πάνω στη μυελίνη.

Πρόσφατες εξελίξεις στην κατανόηση των μηχανισμών και στην ανάπτυξη νέων θεραπευτικών προσεγγίσεων έδωσαν νέα ελπίδα σε γιατρούς και ασθενείς. Ωστόσο, οι νέες πληροφορίες προκάλεσαν μια κάποια σύγχυση αφού στοιχεία από πολλές πηγές υποδηλώνουν ότι η νόσος είναι παρούσα πολύ πριν από την κλινική της εκδήλωση. Η νόσος φαίνεται πολύ περισσότερο εκτεταμένη και συνεχής, με εκδηλώσεις τόσο στη φαιά, όσο και στη λευκή ουσία του εγκεφάλου (Murray, 2006).

## 1.2 ΕΙΣΑΓΩΓΗ ΣΤΙΣ ΣΤΑΤΙΣΤΙΚΕΣ ΣΑΡΩΣΗΣ

Στην παράγραφο αυτή δίνονται εισαγωγικές έννοιες των στατιστικών συναρτήσεων σάρωσης, τόσο σε επίπεδο μιας διάστασης, αλλά και σε επίπεδο δύο διαστάσεων. Αποτελούν ιδιαίτερα χρήσιμο εργαλείο για πολλές στατιστικές αναλύσεις, τόσο σε προβλήματα χωρικών ή χρονικών δεδομένων, όσο και σε συνδυασμό αυτών.

Σε πολλά επιστημονικά πεδία, ερευνητές λαμβάνουν αποφάσεις δίνοντας ιδιαίτερη σημασία στην εμφάνιση γεγονότων σε χρονικές, χωρικές ή χωροχρονικές συστάδες (*clusters*). Η ανίχνευση συστάδων υπό διάφορες συνθήκες με τη χρήση συναρτήσεων σάρωσης βρίσκει εφαρμογή στις Φυσικές, Κοινωνικές ή Ιατρικές Επιστήμες και η Μοριακή Βιολογία, ο Έλεγχος Ποιότητας ή τα Συστήματα Τηλεπικοινωνιών θα μπορούσαν να είναι μερικά απτά παραδείγματα (Collado, 2003, Witham and Orpenheimer, 2004, Green et al., 2003).

Γενικά, τέτοιες στατιστικές συναρτήσεις χρησιμοποιούνται για το «σκανάρισμα» (*scanning*) του χρόνου ή/και του χώρου, με σκοπό την ανίχνευση περιπτώσεων στις οποίες γεγονότα του προβλήματος που μελετάται εμφανίζονται συγκεντρωμένα σε συστάδες.

### 1.2.1 Η Μονοδιάστατη Περίπτωση

Στη διεθνή βιβλιογραφία, εντοπίζονται πολλές εφαρμογές της θεωρίας των ροών επιτυχιών και των γενικευμένων ροών επιτυχιών που αποτελούν μονοδιάστατες περιπτώσεις των στατιστικών συναρτήσεων

σάρωσης. Οι περισσότερες από αυτές αφορούν την ανίχνευση συστάδων όμοιων αποτελεσμάτων σε μονοδιάστατες ακολουθίες, με τη βοήθεια στατιστικών συναρτήσεων. Οι συναρτήσεις αυτές βασίζονται σε τυχαίες μεταβλητές, οι οποίες συνδέονται με τις κλασικές ροές επιτυχιών ή με τις γενικεύσεις τους, δηλαδή τις στατιστικές συναρτήσεις σάρωσης.

Ο όρος ροή (*run*) σε μια ακολουθία δοκιμών αφορά τη διαδοχή (χωρίς διακοπή) όμοιων αποτελεσμάτων. Στατιστικά, ως ροή ορίζεται η διαδοχή όμοιων αποτελεσμάτων τα οποία ακολουθούνται από διαφορετικά αποτελέσματα σε εκτελέσεις ενός πειράματος τύχης, το οποίο σε κάθε επανάληψή του δίνει δυο δυνατά αποτελέσματα: Επιτυχία(1)-Απότυχία (0).

Τροποποιώντας την έννοια της ροής, μπορούν να οριστούν και να μελετηθούν πολυπλοκότερα σχήματα, όπως αυτό της  $r$ -γενικευμένης ροής επιτυχιών μήκους  $k$ , η οποία χρησιμοποιείται όταν σε μια ακολουθία  $n$  δοκιμών, ενδιαφέρει ο αριθμός των παραθύρων μήκους  $k$  τα οποία περιέχουν τουλάχιστον  $r$  επιτυχίες (Μπερσίμης, 2005). Πιο συγκεκριμένα, δοθέντος  $N$  σημείων τυχαία κατανεμημένων σε μια χρονική περίοδο  $(0, T)$ , ορίζεται η ποσότητα  $S_w$  που απεικονίζει τον μεγαλύτερο αριθμό γεγονότων που εμφανίστηκαν σε ένα παράθυρο μήκους  $w$  του χρόνου. Αυτή η ποσότητα ονομάζεται **στατιστική σάρωσης** (*scan statistic*) από την άποψη ότι σκανάρωντας τη χρονική περίοδο  $(0, T)$  με παράθυρο μεγέθους  $w$ , παρατηρείται ένας μεγάλος αριθμός σημείων. Ως  $W_k$  ορίζεται η μικρότερη χρονική περίοδος, η οποία περιλαμβάνει έναν προκαθορισμένο αριθμό  $k$  γεγονότων. Το διάστημα  $W_{r+1}$  καλείται *r-στατιστική σάρωσης*. Να σημειωθεί ότι οι κατανομές των στατιστικών  $S_w$  και  $W_k$  συνδέονται μεταξύ τους, αφού



$P(W_k > w) = P(S_w < k)$ . Με άλλα λόγια, εάν το μικρότερο παράθυρο που περιέχει  $k$  σημεία, έχει μήκος μεγαλύτερο από  $w$ , τότε δεν υπάρχει παράθυρο μήκους  $w$ , που να περιέχει  $k$  ή περισσότερα σημεία (Glaz, Naus and Wallenstein, 2001).

Τα παραπάνω μπορούν να οριστούν και να εφαρμοστούν το ίδιο και στη διακριτή περίπτωση, δηλαδή σε μια ακολουθία διαδοχικών  $T$  δοκιμών. Συγκεκριμένα, μια ειδική περίπτωση διακριτής συναρτήσεως σάρωσης αποτελεί το μήκος της μεγαλύτερης ροής επιτυχιών σε μια ακολουθία πειραμάτων Bernoulli. Ο μέγιστος αριθμός επιτυχιών που εντοπίζεται ανάμεσα σε  $m$  συνεχόμενες δοκιμές ενός γενικότερου συνόλου  $N$  δοκιμών, καλείται **διακριτή στατιστική σάρωσης** (*discrete scan statistic*) και συμβολίζεται με  $S_m'$ . Η γενική περίπτωση όπου ένα ποσοστό από τουλάχιστον  $k$  επιτυχίες μέσα σε  $m$  συνεχόμενες δοκιμές έχει εντοπιστεί, συμβολίζεται με  $S_m' = k$ .

### 1.2.2 Η Διδιάστατη Περίπτωση

Παρά την ευρεία χρήση των μονοδιάστατων συναρτήσεων σάρωσης σε αρκετές έρευνες, παρατηρήθηκε μεγάλη απώλεια πληροφορίας όταν χρειάστηκε να μελετηθούν περιπτώσεις χωρικών δεδομένων (διδιάστατα δεδομένα) και ο μόνος τρόπος ήταν η απεικόνισή τους σε μια διάσταση.

Προκειμένου να αντιμετωπισθεί αυτή η απώλεια πληροφορίας, εισήχθησαν οι λεγόμενες διδιάστατες στατιστικές συναρτήσεις σάρωσης, οι οποίες κρίθηκαν κατάλληλες για μελέτη προβλημάτων που σχετίζονται με χωρικά δεδομένα. Οι συναρτήσεις αυτές αφορούν την εύρεση συστάδων από όμοια αποτελέσματα σε πλαίσια – περιοχές μεγέθους

$n_1 \times n_2$  και από την εμφάνισή τους το 1965 (Naus, 1965), χρησιμοποιούνται ευρέως για τον εντοπισμό περιοχών με ιδιαίτερα υψηλό (ή χαμηλό) αριθμό γεγονότων ενός συγκεκριμένου τύπου. Πολλοί ερευνητές έχουν ασχοληθεί από τότε με χωρικές συναρτήσεις σάρωσης τόσο σε θεωρητικό επίπεδο όσο και στο επίπεδο της εφαρμογής, ανάμεσά τους, οι Darling and Waterma (1985), Kulldorff (1997), Haiman (2000), Kulldorff (2001), Glaz et al. (2001), Boutsikas and Koutras (2003) και Bersimis et al. (2014).

Όπως ήδη αναφέρθηκε οι μονοδιάστατες στατιστικές σάρωσης χρησιμοποιούν ένα διάστημα ως παράθυρο σάρωσης (*scanning window*) προκειμένου να σαρωθεί ένα συνεχούς χρόνου διάστημα ή μια διακριτή ακολουθία δοκιμών. Στη σάρωση διδιάστατων γεωγραφικών περιοχών μπορεί να εμφανιστούν ορθογώνια παραλληλόγραμμα, κυκλικές περιοχές και γενικότερα περιοχές ακανόνιστου σχήματος. Σύμφωνα με τους Glaz et al. (2001) οι κατανομές των στατιστικών σάρωσης μελετήθηκαν υπό την υπόθεση ότι η εμφάνιση των σημείων (γεγονότων) στο χώρο ακολουθεί διάφορα πρότυπα πιθανοτήτων (την Ομοιόμορφη, την Poisson, την Διωνυμική, Υπεργεωμετρική, κα).

Διάφορες μέθοδοι έχουν προταθεί και χρησιμοποιούνται για ανάλυση και ερμηνεία διδιάστατων δεδομένων και σε πολλές περιπτώσεις αυτές οι μέθοδοι συνδυάζονται, με σκοπό τη εξαγωγή περισσότερης πληροφορίας. Τις τελευταίες δύο δεκαετίες η τοπική-χωρική συσταδοποίηση (*clustering*) απασχολεί πολλούς ερευνητές από διάφορα επιστημονικά πεδία και οι διάφορες μέθοδοι που έχουν προταθεί μπορούν να χωριστούν σε δυο κατηγορίες: α) την Point Based Method (Openshaw et al., 1987) και β) την Area Based Method (Oden, Jacquez and Grimson, 1996). Η

πρώτη κάνει χρήση της ακριβούς τοποθεσίας των γεγονότων που μελετώνται, ενώ η τελευταία χρησιμοποιεί συγκεντρωτικές τιμές για την εμφάνιση γεγονότων σε περιοχές, τιμές που συνήθως προέρχονται από διοικητικές μονάδες.

Στην παρούσα διπλωματική εργασία γίνεται αναφορά σε δυο από τις πιο δημοφιλείς μεθοδολογίες της πρώτης κατηγορίας που αναφέρθηκε (Point Based Method): στη μεθοδολογία που προτάθηκε από τους Chen and Glaz (1996) και σε αυτή που προτάθηκε από τον Kuldorff (1997). Και οι δύο μέθοδοι λειτουργούν με παρόμοιο τρόπο, ενώ ο κύριος στόχος της ανάλυσης είναι η απόρριψη της τυχαίας εμφάνισης περιστατικών (γεγονότων) ή ισοδύναμα, ο εντοπισμός μιας ή περισσότερων στατιστικά σημαντικών μεμονωμένων συστάδων περιστατικών. Η συνάρτηση χωρικής σάρωσης (*Spatial Scan Statistic*) που προτείνει ο Kuldorff (1997) θα χρησιμοποιηθεί στην παρούσα διπλωματική.

### **1.2.3 Εφαρμογές**

Οι συναρτήσεις σάρωσης βρίσκουν εφαρμογή σε πολλούς τομείς, καθώς οι ερευνητές δίνουν μεγάλη βάση στη συσταδοποίηση γεγονότων. Στον κλάδο Δημόσιας Υγείας αναζητούνται συνήθεις αιτιογόνοι παράγοντες οι οποίοι εξηγούν συστάδες περιπτώσεων καρκίνου ή γενετικών ανωμαλιών. Οι μοριακοί βιολόγοι ερευνούν παλίνδρομες συστάδες στο DNA για στοιχεία σχετικά με την προέλευση της αντιγραφής των ιών. Επίσης, ειδικοί ποιότητας ελέγχου παρατηρούν και μελετούν συστάδες ελαττωματικών προϊόντων. Οι πιθανότητες

διαφορετικών τύπων συστάδων κάτω από διάφορες προϋποθέσεις αποτελούν απαραίτητα εργαλεία σε φυσικές και κοινωνικές επιστήμες.

Το 1965, πρώτος ο Naus δημοσίευσε ένα άρθρο με τίτλο “*Clustering of random points in two dimensions*”, σχετικό με χωρικές συναρτήσεις σάρωσης και έδωσε την αφορμή για περαιτέρω μελέτη της στατιστικής θεωρίας σε χωρικά δεδομένα, καθώς και εμβάθυνση σε μια πληθώρα εφαρμογών της σε πολλά επιστημονικά πεδία. Μερικά που μπορούν να καταγραφούν είναι τα πεδία της αρχαιολογίας, της αστρονομίας, της εγκληματολογίας, της εγκεφαλικής απεικόνισης, της δημογραφίας, της ψυχολογίας ή της κτηνιατρικής (Costa and Kulldorff, 2009). Στο κεφάλαιο 4 της παρούσας διπλωματικής, ιδιαίτερη αναφορά θα γίνει στις εφαρμογές των χωρικών στατιστικών συναρτήσεων σάρωσης στην ιατρική απεικόνιση και ειδικότερα στους τομείς της επιδημιολογίας και των μολυσματικών νόσων.

Σε γενικές γραμμές, δεδομένα διαφορετικών μορφών απαιτούν διαφορετικές μορφές χωρικών συναρτήσεων σάρωσης αλλά η βασική δομή είναι αυτή που καινοτόμα εισήχθη από τον Naus (1965). Η μέθοδος εφαρμόζεται συχνότερα σε δεδομένα που αφορούν την εμφάνιση διαφόρων μορφών καρκίνων, μολυσματικές ασθένειες και κτηνιατρική επιδημιολογία, αντικείμενα της γενικής επιδημιολογίας και τομείς με αξιοσημείωτο παρελθόν στην εμφάνιση συστάδων ασθένειας.

Ο αριθμός των εφαρμογών των στατιστικών σάρωσης σε μη ιατρικούς τομείς είναι περιορισμένος, όμως με την αυξανόμενη χρήση των γεωγραφικών πληροφοριακών συστημάτων, θα υπάρξει και αύξηση των αντίστοιχων μεθόδων στατιστικής συμπερασματολογίας. Επίσης, τελευταία εμφανίζονται εφαρμογές των χωρικών συναρτήσεων σάρωσης

παρατηρούνται και σε οικονομικούς τομείς, με σκοπό την παρατήρηση κάποιας ομαδοποιημένης δραστηριότητας. Τέλος, χρήση των εν λόγω συναρτήσεων παρατηρείται και από υπηρεσίες υποβολής του νόμου για σκιαγράφηση και απεικόνιση εγκληματικών δραστηριοτήτων (Costa and Kulldorff, 2009).

#### 1.2.4 Περιγραφή Ανάλυσης

Όπως έχει προαναφερθεί, πολλές μέθοδοι έχουν προταθεί για την στατιστική ανάλυση χωρικών δεδομένων, με επικρατέστερες αυτές που εισήχθησαν από τους Chen and Glaz (1996) και από τον Kulldorff (1997), του οποίου η μεθοδολογία θα χρησιμοποιηθεί στην παρούσα διπλωματική.

Στο άρθρο του “*Spatial Disease Clusters: Detection and Inference*” που δημοσιεύθηκε το 1995 σε συνεργασία με τον Nagarwalla, παρουσιάστηκε η νέα μέθοδος ανίχνευσης και εξαγωγής συμπερασμάτων χωρικών συστάδων μιας ασθένειας κάνοντας εκτίμηση με έναν στατιστικό έλεγχο λόγου πιθανοφαινιών. Η μέθοδος βασίστηκε σε υπάρχουσες μεθοδολογίες όπως αυτή της ανίχνευσης επικαλυπτόμενων κυκλικών ζωνών από τετράγωνα που εισήχθη από τον Openshaw *et al.* (1987) και την αντίστοιχη του Turnbull *et al.* (1990) για κυκλικές ζώνες με προκαθορισμένο μέγεθος πληθυσμού.

Η στατιστική συνάρτηση σάρωσης χρησιμοποιείται κατά κύριο λόγο για να ελεγχθεί εάν ένα σημείο ή συμβάν, βρίσκεται τυχαία κατανεμημένο σε μια διάσταση ή εάν υπάρχει σημαντική συστάδα τέτοιων γεγονότων που μπορεί να ανιχνευθεί. Η χωρική συνάρτηση

σάρωσης γενικεύεται για την ανίχνευση κλάσεων στο χώρο (διδιάστατη περίπτωση) ή και σε περισσότερες διαστάσεις. Το παράθυρο με το οποίο γίνεται η σάρωση διαθέτει μεταβλητό μέγεθος και η κύρια μεθοδολογία βασίζεται στο γεγονός ότι τα δεδομένα ακολουθούν κατανομές Poisson ή Bernoulli. Κύριος στόχος είναι η ανίχνευση συστάδων που δεν μπορούν να εξεξηγηθούν από τις αρχικές κατανομές που θεωρήθηκαν.

Παρόλο που στην περιγραφή της μεθόδου γίνεται χρήση κυκλικών ζωνών μεταβλητού μεγέθους, η συνάρτηση ελέγχου μπορεί να τροποποιηθεί κατάλληλα για μια εναλλακτική υπόθεση ζωνών διαφορετικών σχημάτων και το γεγονός αυτό έκανε τη χρήση της, τόσο διαδεδομένη. Ο έλεγχος λόγου πιθανοφανειών βασίζεται σε ανομοιογενή πληθυσμό και επίσης μπορεί να προσαρμοστεί σε ποσοστά εμφάνισης νόσου, βάσει της ηλικίας. Το τελευταίο φαίνεται χρήσιμο σε περιπτώσεις που ο κίνδυνος μιας ασθένειας αυξάνεται με την ηλικία, όπως για παράδειγμα στη νόσο του Alzheimer.

Συγκρίνοντας την ισχύ του ελέγχου που προτάθηκε από τον Kulldorff (1997) με τη μέθοδο του Turnbull et al. (1990), παρατηρήθηκε ότι η τελευταία εξάγει ισχυρότερα συμπεράσματα όταν το πραγματικό μέγεθος της κλάσης που εντοπίστηκε συγκλίνει στην πληθυσμιακή ακτίνα που επιλέχθηκε στη μέθοδο. Από την άλλη πλευρά, ο έλεγχος λόγου πιθανοφανειών είναι ισχυρότερος, όταν το μέγεθος των κλάσεων διαφοροποιείται από την πληθυσμιακή ακτίνα.

Ο σκοπός του ελέγχου ανίχνευσης κλάσεων δεν περιορίζεται μόνο στη γένεση αιτιολογικών μελετών. Σε πολλά σύνολα δεδομένων είναι πιθανό να μη βρεθούν στατιστικά σημαντικές κλάσεις γεγονότων, συμπέρασμα που παραμένει εξίσου σημαντικό. Όπως πολλοί ερευνητές έχουν

παρατηρήσει (Kulldorff and Nagarwalla, 1995), μεγάλο πλήθος ερευνών και πόρων έχουν επενδυθεί στη μελέτη περιπτώσεων που κρίθηκαν ύποπτες, τις περισσότερες φορές άσκοπα, αφού η εμφάνιση πολλών από αυτές χαρακτηρίστηκαν ως τυχαίες κυμάνσεις στα ποσοστά εμφάνισης. Με τη μέθοδο μέγιστης πιθανοφάνειας οι αμφισβητήσεις για πιθανές στατιστικά σημαντικές κλάσεις μειώνονται, και επενδεδυμένοι πόροι μπορούν να χρησιμοποιηθούν σε άλλες μελέτες.

Να σημειωθεί τέλος, ότι το πόσο σημαντική κρίνεται η εμφάνιση μιας συστάδας που παρατηρήθηκε, εξαρτάται από το μέγεθος της περιοχής που μελετάται και είναι ανούσιο να χαρακτηριστεί μια συστάδα στατιστικά σημαντική που χρήζει περαιτέρω ανάλυση, χωρίς να έχει γίνει αναφορά στην περιοχή της μελέτης και τα χαρακτηριστικά της, στην πληθυσμιακή πυκνότητα, στην παρουσία εξωτερικού παράγοντα που μπορεί να επηρεάσει την εξαγωγή συμπερασμάτων (εργοστάσιο στην περιοχή), και άλλα.

## **1.3 ΜΕΘΟΔΟΛΟΓΙΑ ΣΤΑΤΙΣΤΙΚΩΝ ΣΑΡΩΣΗΣ**

Στην τρίτη παράγραφο παρουσιάζονται εκτενέστερα οι δύο προσεγγίσεις που έχουν προταθεί σχετικά με τις στατιστικές συναρτήσεις σάρωσης. Παρατίθεται αρχικά η πιθανοθεωρητική προσέγγιση των Chen and Glaz (1996) και στη συνέχεια η σάρωση και διάγνωση με χρήση

στατιστικής μέγιστης πιθανοφάνειας που προτάθηκε από τον Kulldorff (1997). Επιπλέον, γίνεται και μια σύγκριση των δύο προσεγγίσεων.

### 1.3.1 Πιθανοθεωρητική Προσέγγιση

Σε προβλήματα μιας διάστασης, τα  $N$  σημεία βρίσκονται κατανομημένα σε ένα μοναδιαίο διάστημα (*unit interval*) ή σε έναν μοναδιαίο κύκλο (*unit circle*). Η στατιστική σάρωσης  $S_w$  συμβολίζει τον μεγαλύτερο αριθμό σημείων που εμφανίζονται σε κάποιο υπο-διάστημα  $[0,1)$  μήκους  $w$ . Η στατιστική σάρωσης  $W_k$  από την άλλη, υποδηλώνει το μέγεθος του μικρότερου υπο-διαστήματος  $[0,1)$ , το οποίο περιέχει  $k$ -σημεία. Επίσης, οι κατανομές των δύο στατιστικών συναρτήσεων  $S_w$  και  $W_k$  συνδέονται, αφού  $P(W_k > w) = P(S_w < k)$ .

Σε προβλήματα, τώρα, δύο διαστάσεων, τα  $N$  σημεία βρίσκονται τυχαία κατανομημένα σε μια διδιάστατη περιοχή συγκεκριμένου σχήματος. Σε αντιστοιχία με τα διαστήματα σάρωσης στην μονοδιάστατη περίπτωση, εδώ χρησιμοποιούνται παράθυρα σάρωσης και συγκεκριμένα το μοναδιαίο τετράγωνο (*unit square*) ως γενίκευση του μοναδιαίου διαστήματος ή η επιφάνεια μιας σφαίρας ως γενίκευση του μοναδιαίου κύκλου.

Η στατιστική σάρωσης  $S_w$  μπορεί να γενικευτεί ως ο μεγαλύτερος αριθμός γεγονότων σε κάποιο παράθυρο διαμέτρου  $w$ , ενώ η  $W_k$  αφορά τώρα το μήκος του μικρότερου παραθύρου σάρωσης στο οποίο περιέχονται  $k$  σημεία. Σκανάροντας κανείς μια διδιάστατη περιοχή, ενδιαφέρεται να εντοπίσει τη μεγαλύτερη συστάδα γεγονότων μέσα σε ένα παράθυρο σάρωσης συγκεκριμένου σχήματος. Οι Eggleton και



Kermack (1944) και ο Mack (1950) εξέτασαν το πρόβλημα σάρωσης μοναδιαίου τετραγώνου με ένα μικρότερο ορθογώνιο παραλληλόγραμμο μήκους  $u$  και ύψους  $v$ , παράλληλες στις πλευρές του τετραγώνου που σαρώνεται. Τέτοιας μορφής προβλήματα προϋποθέτουν τη διάκριση σε δύο κατηγορίες, τη δεσμευμένη (*conditional*) και τη μη-δεσμευμένη (*unconditional*) περίπτωση. Στην πρώτη, ο αριθμός των  $N$  γεγονότων που κατανέμονται στο μοναδιαίο τετράγωνο είναι γνωστός εκ των προτέρων, ενώ στη μη-δεσμευμένη περίπτωση του προβλήματος, ο αριθμός των γεγονότων στο μοναδιαίο τετράγωνο αποτελεί τυχαία μεταβλητή που κατανέμεται με βάση την κατανομή Poisson (Glaz et al., 2001).

### 1.3.2 Χωρική Σάρωση με Χρήση Πιθανοφάνειας

Μεγάλος αριθμός επεκτάσεων και τροποποιήσεων της στατιστικής  $S_w$  έχουν προταθεί και εφαρμοστεί από πολλούς ερευνητές (Marshall, 1991, Hills and Alexander, 1989) για μελέτη επιδημιολογικών προβλημάτων. Αυτές οι τεχνικές, μέσω υπολογιστικών προγραμμάτων, σκανάρουν την υπό μελέτη περιοχή χρησιμοποιώντας μια στατιστική συνάρτηση σάρωσης, με διάφορα μεγέθη και σχήματα παραθύρων σάρωσης. Για τέτοιες περιπτώσεις, ευρέως διαδεδομένη είναι η χρήση του λογισμικού SaTScan που εισήχθη από τον Kulldorff (2004), βασίζεται στην εργασία του Kulldorff (1997) και είναι αυτό που θα χρησιμοποιηθεί και στην ανάλυση των δεδομένων της παρούσας διπλωματικής.

Η προσέγγιση του Kulldorff βασίζεται στη χρήση ενός κυκλικού παραθύρου πάνω στο χάρτη της περιοχής που μελετάται του οποίου το κέντρο μετακινείται σε συνεχόμενα διαδοχικά σημεία έτσι ώστε σε διαφορετικές θέσεις του παραθύρου να περιλαμβάνονται διαφορετικά σύνολα γειτονικών εγγραφών. Για κάθε τοποθεσία, τα δεδομένα είναι απαραίτητο να περιέχουν πληροφορίες σχετικά με τον αριθμό των συμβάντων σε εκείνη την τοποθεσία. Εάν το παράθυρο περιλαμβάνει το κεντροειδές (*centroid*) της υπό μελέτη περιοχής, τότε όλη η περιοχή συμπεριλαμβάνεται στο παράθυρο. Επιπλέον, όσον αφορά το κεντροειδές κάθε κύκλου, η ακτίνα κάθε κυκλικού παραθύρου μεταβάλλεται συνεχώς από το μηδέν μέχρι μια μέγιστη ακτίνα, γεγονός που κάνει το παράθυρο σάρωσης ευέλικτο, ως προς το μέγεθος και την τοποθεσία του. Επομένως, η μέθοδος αυτή δημιουργεί έναν άπειρο αριθμό ξεχωριστών γεωγραφικών κύκλων με διαφορετικά γειτονικά σύνολα δεδομένων μέσα σε αυτούς. Ο κάθε κύκλος αποτελεί μια πιθανή υποψήφια κλάση.

Ως εναλλακτική επιλογή του κύκλου, ο ερευνητής μπορεί να ορίσει ένα παράθυρο ελλειπτικού σχήματος και σε αυτή την περίπτωση το σύνολο των ελλείψεων διαφορετικού σχήματος και γωνιών χρησιμοποιείται ως παράθυρο σάρωσης μαζί με τον κύκλο. Με αυτόν τον τρόπο, παρέχεται μεγαλύτερη ισχύς στον έλεγχο εντοπισμού πραγματικών κλάσεων που μπορεί να έχουν μακρύ και στενό σχήμα και δύσκολα θα ανιχνεύονταν από έναν κύκλο, αλλά χαμηλότερη ισχύ στον εντοπισμό κυκλικών ή άλλων πιο συμπαγών σχημάτων.

Συνήθως, το μεγαλύτερο ενδιαφέρον έγκειται στον εντοπισμό συστάδων και στη σύγκρισή τους με άλλες που βρέθηκαν στον ίδιο χάρτη μελέτης και όχι με μια εξωτερική περιοχή ή χρονική περίοδο. Για το λόγο

αυτό προτιμάται η χρήση της δεσμευμένης περίπτωσης. Θεωρώντας γνωστό τον συνολικό αριθμό των γεγονότων  $N$ , ορίστηκε από τον Kulldorff (1997) η στατιστική συνάρτηση σάρωσης ως ο λόγος μέγιστης πιθανοφάνειας, όλων των πιθανών κύκλων  $Z$ . Συγκεκριμένα,

$$S = \frac{\max_Z \{L(Z)\}}{L_0} = \max_Z \left\{ \frac{L(Z)}{L_0} \right\},$$

όπου  $L(Z)$  η μέγιστη πιθανοφάνεια κύκλου  $Z$ , και  $L_0$  η συνάρτηση πιθανοφάνειας υπό τη μηδενική υπόθεση ότι τα γεγονότα που εμφανίζονται κατανέμονται τυχαία στη γεωγραφική περιοχή που μελετάται.

Όταν ο παραπάνω λόγος μεγιστοποιείται για όλους τους υπο μελέτη κύκλους, εντοπίζεται εκείνος που αποτελεί την πιο πιθανή κλάση γεγονότων. Στη συνέχεια υπολογίζεται το  $p$ -value μέσω Monte Carlo, και αποφασίζεται η απόρριψη ή όχι της μηδενικής υπόθεσης.

Στην περίπτωση που ο υποκείμενος ρυθμός εμφάνισης της νόσου ορίζεται στην μηδενική υπόθεση του προβλήματος, αυτό μπορεί να προσεγγιστεί από την μη- δεσμευμένη μορφή του, η οποία φαίνεται να έχει καλύτερα αποτελέσματα σε χωρο-χρονικές συναρτήσεις σάρωσης όπου δεν υφίσταται δέσμευση στο συνολικό αριθμό των περιπτώσεων.

Η χωρική στατιστική συνάρτηση σάρωσης που προτείνει ο Kulldorff θεωρείται η καταλληλότερη για μελέτη χωρικών προβλημάτων αφού μπορεί να προσαρμοστεί τόσο σε ανισομεγέθεις πληθυσμούς, όσο και σε οποιοδήποτε αριθμό συγχυτικών μεταβλητών (*confounding variables*). Επιπλέον, αποφεύγεται η μεροληψία αφού η μέθοδος αναζητά συστάδες γεγονότων ακαθόριστου μεγέθους ή τοποθεσίας και τέλος, ο έλεγχος που βασίζεται στο λόγο μέγιστης πιθανοφάνειας, λαμβάνει υπόψη

αποτελέσματα πολλών διαφορετικών ελέγχων (για κάθε διαφορετικό κύκλο  $Z$ ) και προκύπτει μια μόνο τιμή  $p$ -value για τον έλεγχο της μηδενικής υπόθεσης. Αν η υπόθεση αυτή απορριφθεί, είναι εύκολο πλέον να εντοπιστεί η τοποθεσία της κλάσης που προκάλεσε αυτή την απόρριψη, με άλλα λόγια, να εντοπιστεί στατιστικά σημαντική κλάση συγκεντρωμένων γεγονότων.

Η μεθοδολογία θα αναλυθεί διεξοδικότερα σε επόμενα κεφάλαια, αφού πάνω σε αυτή βασίζεται το πρόγραμμα SaTScan που θα χρησιμοποιηθεί για τη μελέτη και ανάλυση των δεδομένων σκλήρυνσης κατά πλάκας.

### **1.3.3 Σύγκριση των Δύο Προσεγγίσεων**

Συνοψίζοντας την πιθανοθεωρητική προσέγγιση των στατιστικών συναρτήσεων σάρωσης, προκύπτει το συμπέρασμα ότι η μεγαλύτερη πρόκληση αυτών των συναρτήσεων είναι η εύρεση αναλυτικών αποτελεσμάτων που αφορούν τις πιθανότητες να παρατηρηθεί κάποια κλάση συγκεκριμένης έκτασης και πολλές μαθηματικές θεωρίες έχουν αναπτυχθεί με σκοπό την απόκτηση προσεγγίσεων και ορίων για αυτές τις πιθανότητες κάτω από ένα πλήθος υποθέσεων. Κοινά χαρακτηριστικά όλων των μελετών είναι ότι χρησιμοποιείται κάποιο παράθυρο σάρωσης σταθερού μεγέθους και ότι ασχολούνται με μετρήσιμα δεδομένα όπου υπό την μηδενική υπόθεση, ο παρατηρηθής αριθμός περιπτώσεων κατανέμεται ομοιόμορφα, έτσι ώστε ο αναμενόμενος αριθμός τους σε μία περιοχή να είναι ανάλογος του μεγέθους αυτής της περιοχής.

Στην παρακολούθηση ασθενειών, οι παραπάνω υποθέσεις δεν ισχύουν αφού δεν είναι γνωστό εκ των προτέρων το μέγεθος συστάδας και ο πληθυσμός που βρίσκεται σε κίνδυνο είναι γεωγραφικά ανομοιογενής. Κάτω από την μηδενική υπόθεση ίσου κινδύνου ασθένειας εντός και εκτός συστάδας, αναμένεται να βρεθούν περισσότερες εμφανίσεις νόσου σε μια πόλη από ότι σε μια περιοχή ίδιου μεγέθους στην εξοχή, λόγω της υψηλότερης πληθυσμιακής πυκνότητας στην πόλη. Η στατιστική σάρωσης που εισήγαγε ο Kulldorff (1997) και χρησιμοποιείται στο στατιστικό πρόγραμμα SaTScan, δημιουργήθηκε για να επιλυθούν τα παραπάνω προβλήματα όμως λόγω του ότι δεν έχουν βρεθεί ακόμη αναλυτικές μέθοδοι για την εύρεση των πιθανοτήτων για τα πιο πάνω σύνθετα θέματα, χρησιμοποιείται Monte-Carlo για την εύρεση των αντίστοιχων  $p$ -values.

## 1.4 ΤΟ ΛΟΓΙΣΜΙΚΟ ΠΑΚΕΤΟ SaTScan

Στην τελευταία παράγραφο του πρώτου κεφαλαίου παρουσιάζεται συνοπτικά το στατιστικό πακέτο που θα χρησιμοποιηθεί για την ανάλυση που θα ακολουθήσει στο τελευταίο κεφάλαιο της διπλωματικής εργασίας. Πρόκειται για ένα πρόγραμμα που δημιουργήθηκε από τον Martin Kulldorff σε συνεργασία με την Information Management Services Inc. και απώτερο σκοπό την επίλυση χωρικών ή/και χρονικών στατιστικών προβλημάτων. Χρησιμοποιείται μέχρι και σήμερα ενώ η φήμη του ολοένα και μεγαλώνει. Η έκδοση που χρησιμοποιήθηκε είναι η v9.3.

Πρόκειται για ένα ελεύθερο λογισμικό που αναλύει χωρικά, χωρο-χρονικά ή χρονικά δεδομένα χρησιμοποιώντας αντίστοιχες στατιστικές σάρωσης και σχεδιάστηκε για τους εξής σκοπούς:

α) Γεωγραφική επιτήρηση μιας ασθένειας, εντοπισμό συσταδοποιημένων συμβάντων και έλεγχο στατιστικής σημαντικότητας τους,

β) Έλεγχο για το εάν μια ασθένεια είναι τυχαία κατανεμημένη στο χώρο, στο χρόνο ή στο χωρο-χρόνο,

γ) Εκτίμηση στατιστικής σημαντικότητας εμφάνισης συστάδων γεγονότων.

Το SaTScan μπορεί να χρησιμοποιηθεί τόσο για διακριτές όσο και για συνεχείς συναρτήσεις σάρωσης.

Στις διακριτές συναρτήσεις σάρωσης, το SaTScan κάνει χρήση είτε διακριτού μοντέλου Poisson, όπου ο αριθμός των γεγονότων σε μια γεωγραφική περιοχή ακολουθεί Poisson κατανομή με γνωστό τον αριθμό του πληθυσμού σε κίνδυνο, είτε μοντέλου Bernoulli με δεδομένα 0/1 όπως συμβάντα (*cases*) και μη-συμβάντα (*controls*) ή ασθενείς και μάρτυρες. Πληθώρα μοντέλων για τη διακριτή περίπτωση μπορεί επίσης να χρησιμοποιηθεί, όμως δεν θα γίνει περαιτέρω αναφορά στη συγκεκριμένη εργασία. Ενδεικτικά αναφέρονται τα μοντέλα χωρο-χρονικής αντιμετάθεσης, τα πολυωνυμικά μοντέλα για κατηγορικά δεδομένα ή τα εκθετικά. Κοινό τους χαρακτηριστικό είναι ότι οι γεωγραφικές τοποθεσίες όπου έχουν παρατηρηθεί δεδομένα, είναι μη-τυχαία και καθορισμένα από τον ερευνητή.

Για το μοντέλο Bernoulli είναι επίσης απαραίτητο να οριστεί ο αριθμός των υγείων περιπτώσεων (*controls*) σε κάθε τοποθεσία που ορίστηκε ενώ στο διακριτό μοντέλο Poisson, ο ερευνητής πρέπει να ορίσει το μέγεθος του πληθυσμού σε αυτές, που μπορεί να αλλάζει στο χρόνο.

Οι στατιστικές σάρωσης χρησιμοποιούνται για τον εντοπισμό και αξιολόγηση συστάδων συμβάντων σε ένα χρονικό, χωρικό ή χωρο-χρονικό σύνολο δεδομένων. Κάτι τέτοιο μπορεί να επιτευχθεί με σταδιακή σάρωση ενός παραθύρου κατά μήκος του χρόνου ή/και του χώρου και σημειώνοντας τον αριθμό των παρατηρήσεων που βρέθηκαν καθώς και τον αναμενόμενο αριθμό τους μέσα στο παράθυρο σε κάθε τοποθεσία. Το παράθυρο σάρωσης στο SaTScan έχει τη μορφή διαστήματος (για σάρωση στο χρόνο), κύκλου ή έλλειψης (για σάρωση στο χώρο) και κυλίνδρου με κυκλική ή ελλειπτική βάση (για σάρωση στο χωρο-χρόνο). Επίσης, είναι δυνατόν να οριστεί κάποιο μη-Ευκλείδιο σχήμα από τον χρήστη. Στη διαδικασία που περιγράφηκε, χρησιμοποιούνται παράθυρα μεταβαλλόμενου μεγέθους και το παράθυρο με τη μέγιστη πιθανοφάνεια αποτελεί την πιο πιθανή κλάση, δηλαδή, την κλάση που είναι λιγότερο πιθανό να δημιουργήθηκε κατά τύχη. Για αυτή τη συστάδα που εντοπίστηκε, υπολογίζεται μια τιμή  $p$ -value, βάσει της οποίας αποφασίζεται η απόρριψη ή όχι της υπόθεσης της τυχαίας κατανομής των γεγονότων στο παράθυρο. Το  $p$ -value για τον έλεγχο λαμβάνεται συγκρίνοντας το βαθμό της μέγιστη πιθανοφάνειας του πραγματικού συνόλου δεδομένων με αυτόν των μέγιστων πιθανοφανειών των τυχαίων συνόλων που δημιουργήθηκαν. Αν ο βαθμός αυτός συμβολιστεί με  $R$ , τότε

$p = R / (1 + \# \text{επαναλήψεων})$ . Συνήθως, ο αριθμός επαναλήψεων ορίζεται να είναι 999 ή κάποιος αριθμός που να τελειώνει σε 999, για πιο καλαίσθητα αποτελέσματα στην τιμή του p-value.

Οι συναρτήσεις σάρωσης κάνουν χρήση πολλών διαφορετικών μοντέλων πιθανοτήτων, ανάλογα με τη φύση των δεδομένων. Για όλα τα μοντέλα όμως, οι συναρτήσεις προσαρμόζονται στην ανομοιογένεια της πυκνότητας τους πληθυσμού και η ανάλυση «δεσμεύεται» στον συνολικό αριθμό των περιπτώσεων που βρέθηκαν.

## 1.5 ΣΥΝΟΨΗ

Στο Κεφάλαιο 1 έγινε μια εισαγωγή σε βασικές έννοιες και εργαλεία που χαρακτηρίζονται αναγκαία σε μετέπειτα κεφάλαια. Συγκεκριμένα, παρουσιάστηκε αρχικά η νόσος της σκλήρυνσης κατά πλάκας αφού αυτή αποτέλεσε το κίνητρο μελέτης. Δόθηκαν ορισμένες έννοιες και χαρακτηριστικά της, που οδηγούν σε περαιτέρω ανάλυση στο Κεφάλαιο 2 της παρούσας διπλωματικής εργασίας. Στη συνέχεια, αναφέρθηκε η στατιστική μεθοδολογία που θα ακολουθηθεί, όπως επίσης και διάφορες εφαρμογές της. Η μαθηματική παρουσίαση των στατιστικών συναρτήσεων σάρωσης γίνεται στο Κεφάλαιο 3 ενώ οι εφαρμογές παρατίθενται αναλυτικά στο τέταρτο κεφάλαιο. Τέλος, γίνεται συνοπτική αναφορά στο λογισμικό πακέτο του Kulldorff, SaTScan, καθώς αποτέλεσε απαραίτητο εργαλείο για τη διεξαγωγή του πρακτικού τμήματος της εν λόγω εργασίας, και πιο συγκεκριμένα στην μελέτη και ανάλυση των δεδομένων σκλήρυνσης κατά πλάκας που συλλέχθηκαν.



# Κεφάλαιο 2

## Η ΝΟΣΟΣ ΤΗΣ ΣΚΛΗΡΥΝΣΗΣ ΚΑΤΑ ΠΛΑΚΑΣ

### 2.1 ΕΙΣΑΓΩΓΗ

Όλα τα νεύρα στο ανθρώπινο σώμα, έχουν ένα περίβλημα, όπως ακριβώς τα καλώδια του ηλεκτρικού ρεύματος έχουν το μονωτικό πλαστικό. Το περίβλημα αυτό ονομάζεται μυελίνη (*myelin*). Η κατάσταση κατά την οποία το περίβλημα αυτό καταστρέφεται κατά τόπους, ονομάζεται απομυελίνωση (*demyelination*).

Η απομυελίνωση έχει ως αποτέλεσμα να επιβραδύνεται ή να καταργείται η αγωγή των νευρικών ερεθισμάτων κατά μήκος ενός νεύρου. Με πιο απλά λόγια, η απομυελίνωση προκαλεί δυσλειτουργία των νεύρων, με αποτέλεσμα συμπτώματα όπως αδυναμία στα χέρια ή τα πόδια, μούδιασμα, διπλωπία, θόλωση της όρασης κ.ο.κ.

Οι νευρολογικές παθήσεις που σχετίζονται με την απομυελίνωση ονομάζονται απομυελινωτικές (*demyelinating diseases*). Τέτοιες παθήσεις είναι η πολλαπλή σκλήρυνση (η πιο συχνή), η οπτική νευρίτιδα, η εγκάρσια μυελίτιδα, η νόσος Devic και το σύνδρομο Guillain-Barré ([www.neurocenter.gr/sklirinsi.html](http://www.neurocenter.gr/sklirinsi.html)).

## 2.2 ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ

Οι απομυελινωτικές νόσοι διακρίνονται σε *φλεγμονώδεις* και *μη φλεγμονώδεις*. Οι επίκτητες φλεγμονώδεις απομυελινωτικές νόσοι ταξινομούνται σε πέντε κατηγορίες: α) την πολλαπλή σκλήρυνση (Σκλήρυνση κατά πλάκας), β) τις παραλλαγές της πολλαπλής σκλήρυνσης που είναι η κακοήθης οξεία μορφή (τύπος Marburg), η οπτική νευρομυελίτιδα (νόσος Devic) και η συγκεντρωτική σκλήρυνση (νόσος Balot), γ) την οξεία διάσπαρτη εγκεφαλομυελίτιδα (μετά από λοίμωξη ή μετά από εμβολιασμό), δ) την οξεία αιμορραγική λευκοεγκεφαλοπάθεια και ε) την προϊούσα πολυεστιακή λευκοεγκεφαλοπάθεια (Lassman, 2001).

Οι μη φλεγμονώδεις απομυελινωτικές νόσοι διακρίνονται σε: α) *Δυσμυελινωτικές* ή λευκοδυστροφίες και β) *Τροφικές – τοξικές απομυελινωτικές*. Η πρώτη κατηγορία αφορά συνήθως κληρονομικές νόσους με διαταραχές του μεταβολισμού. Εδώ υπάγονται η μεταχρωματική λευκοδιαστροφή, η οξεία νηπιακή λευκοδιαστροφή, η νόσος του Krabe και άλλες. Στη δεύτερη κατηγορία υπάγονται η κεντρική

γεφυρική μυελινόλυση και η πρωτοπαθής εκφύλιση του μεσολοβίου (Hartung, 2002).

## **2.3 ΦΛΕΓΜΟΝΩΔΕΙΣ ΑΠΟΜΥΕΛΙΝΩΤΙΚΕΣ ΝΟΣΟΙ**

Όταν σημειωθεί κάποια φλεγμονή στο περίβλημα της μυελίνης που καλύπτει τους νευροάξονες των περιφερικών νεύρων, τότε γίνεται λόγος για φλεγμονώδη απομυελινωτική νόσο. Αυτό συνεπάγεται μια βαθμιαία εξελισσόμενη και αυτοάνοση μυική αδυναμία των άκρων, καθώς και άλλες σταδιακές παραλύσεις, αφού η καταστροφή της μυελίνης επιβραδύνει ή παρεμποδίζει τη μετάδοση των νευρικών ερεθισμάτων ([www.iatronet.gr/iatriko-lexiko.html](http://www.iatronet.gr/iatriko-lexiko.html)).

### **2.3.1 Πολλαπλή Σκλήρυνση (Σκλήρυνση Κατά Πλάκας- Multiple Sclerosis)**

Στην πολλαπλή σκλήρυνση, υπάρχει ανοσολογικού τύπου «επίθεση» του οργανισμού στο περίβλημα μυελίνης των νεύρων και αυτό έχει ως αποτέλεσμα να προκαλούνται νευρολογικά συμπτώματα. Η κλινική εικόνα δεν έχει σε όλους την ίδια βαρύτητα ή κατανομή στο χρόνο – άλλοι ασθενείς εμφανίζουν παροδικά και αναστρέψιμα συμπτώματα και σε άλλους εγκαθίσταται και παραμένει πχ η αδυναμία στα χέρια ή ένα

μούδιασμα. Πολλές φορές η έναρξη των συμπτωμάτων είναι όχι τυπική, με αποτέλεσμα η διάγνωση στην αρχή τουλάχιστον να μην είναι βέβαιη ([www.neurocenter.gr/sklirinsi.html](http://www.neurocenter.gr/sklirinsi.html)).

### **2.3.1.1 Γενικά Χαρακτηριστικά**

Είναι μια αυτοάνοση φλεγμονώδης απομυελινοτική νόσος του κεντρικού νευρικού συστήματος (ΚΝΣ) στην οποία οι βλάβες εμφανίζουν διασπορά στο χρόνο και τόπο. Δηλαδή, παρουσιάζονται σε διαφορετικές θέσεις στο ΚΝΣ και επέρχονται σε διαφορετικές χρονικές στιγμές.

Αν και κλασικά περιγράφεται ως νόσος της νεαρής ηλικίας με χαρακτηριστική πορεία εξάρσεων και υφέσεων, η νόσος παρουσιάζει μεγάλη κλινική ανομοιογένεια. Αυτή η ανομοιογένεια αφορά την ηλικία έναρξης, τον τρόπο εμφάνισης των πρώτων κλινικών εκδηλώσεων, τη συχνότητα, τη βαρύτητα και τα υπολείμματα μιας υποτροπής, την ποικιλία και τον διάσπαρτο χαρακτήρα των συμπτωμάτων, την πορεία της νόσου (με εμφάνιση αρκετές φορές προϋούσης επιδείνωσης χωρίς εξάρσεις), τη συσσωρευόμενη αναπηρία με την πάροδο του χρόνου και σε ποσοστό 20% την απουσία οποιασδήποτε αναπηρίας, ακόμα και δέκα έτη μετά την έναρξη της νόσου.

Η μεγάλη ποικιλία της κλινικής εικόνας είναι η αντανάκλαση των πολλαπλών περιοχών με καταστροφή της μυελίνης στο ΚΝΣ (απομυελινοτικές πλάκες), αν και υπάρχει συχνά μεγάλη διαφορά μεταξύ

των κλινικών ευρημάτων και των απομυελινωτικών εστιών, όπως αποκαλύπτονται στη μαγνητική τομογραφία.

### **2.3.1.2 Αιτιοπαθογένεια**

Η αιτιολογία της Σκλήρυνσης Κατά Πλάκας παραμένει άγνωστη. Θεωρείται ότι είναι μια επίκτητη αυτοάνοσος πάθηση, αλλά πιθανώς υπάρχει πολυπαραγοντική αιτιολογία για την εμφάνισή της. Η επικρατέστερη άποψη είναι ότι ένας εξωγενής παράγοντας, πιθανόν ένας ιός ή πολλαπλές ιώσεις, προκαλεί σε γενετικά προδιατεθειμένα άτομα μια παθολογική ανοσολογική αντίδραση που κατευθύνεται εναντίον των πρωτεϊνών της μυελίνης. Η ύπαρξη ενός ιού ως παράγοντα που προκαλεί την έναρξη της παθογένειας της νόσου δεν έχει πάντως επιβεβαιωθεί, αλλά η ύπαρξη ενός ιού (ή άλλου εξωγενούς παράγοντα) υποστηρίζεται από μελέτες (Mathews, 2001, Milonas, Tsonuis and Logothetis, 1990) σε μετανάστες και από τις «επιδημίες» της νόσου στα νησιά Φερόες

Η ύπαρξη γενετικής προδιάθεσης στηρίζεται στη διαφορετική συχνότητα της νόσου στις διάφορες φυλές (λιγότερο συχνή στη μαύρη και πολύ σπανιότερη στην κίτρινη φυλή), στην ύπαρξη οικογενών μορφών, στη μεγαλύτερη συχνότητα εμφάνισης στους μονοωογενείς (26%) από ότι στους διωογενείς (3,5%) διδύμους και στη συσχέτιση με κάποια συγκεκριμένα αντιγόνα. Η νόσος πάντως δεν είναι κληρονομική, παρόλο που η συχνότητά της στους συγγενείς πρώτου βαθμού των πασχόντων είναι 20-30 φορές μεγαλύτερη από ότι στον γενικό πληθυσμό. Έχει παρατηρηθεί ότι μόνο το 4% των πρώτου βαθμού συγγενών του

ασθενούς θα εμφανίσουν τη νόσο. Εκείνο που ουσιαστικά υπάρχει είναι μια γενετική προδιάθεση που χρειάζεται επιπρόσθετα την επίδραση ενός άγνωστου μέχρι σήμερα εξωγενούς παράγοντα για να ξεκινήσει ο παθογενετικός μηχανισμός που προκαλεί τη νόσο.

### **2.3.1.3 Παθοφυσιολογία**

Όταν άτομα που είναι επιρρεπή για να αναπτύξουν σκλήρυνση κατά πλάκας από το γενετικό τους υλικό, εκτεθούν σε έναν ειδικό άγνωστο περιβαλλοντικό παράγοντα, μια ομάδα από βρισκόμενα σε λανθάνουσα κατάσταση T-λεμφοκύτταρα ενεργοποιούνται, οπότε μπορούν να διαπεράσουν τον αιματοεγκεφαλικό φραγμό και εισερχόμενα στο κεντρικό νευρικό σύστημα επιτίθενται στο στόχο-αντιγόνο, που είναι η μυελίνη. Έτσι αρχίζει η φλεγμονώδης διαδικασία της νόσου. Μια τυπική περιοχή με απομυελινωτική φλεγμονή περιβάλλει μια κεντρική φλέβα με διαπερατό εγκεφαλικό φραγμό. Με την καταστροφή της μυελίνης οι νευροάξονες μένουν γυμνοί και έτσι δεν μπορούν να μεταφέρουν αποτελεσματικά τις εντολές, προκαλώντας τα τυπικά συμπτώματα της νόσου. Για κάποια άγνωστη αιτία τα συμπτώματα είναι πιθανό να ελαχιστοποιηθούν. Η φλεγμονή υποχωρεί, ο αιματοεγκεφαλικός φραγμός αποκαθίσταται και συχνά οι νευρικές ίνες επαναμυελίνονται. Αυτό είναι το παθοφυσιολογικό υπόστρωμα μιας απομυελινωτικής πλάκας που προκαλεί τα ανάλογα με τη θέση της στο κεντρικό νευρικό σύστημα συμπτώματα (την ώση, όπως αναφέρεται στη σκλήρυνση κατά πλάκας η εμφάνιση συμπτωματολογίας), που άλλοτε υποχωρούν πλήρως (είναι ο

κανόνας στα αρχικά στάδια της νόσου) και άλλοτε μερικώς (Paty and Ebens, 1998). Υπάρχουν όμως και πολλές κλινικά “σιωπηλές” πλάκες που ανακαλύπτονται νευροακτινολογικά, νευροφυσιολογικά ή στην αυτοψία. Παλαιότερα επικρατούσε η άποψη ότι στη νόσο, ενώ υπάρχει έντονη προσβολή της μυελίνης, η καταστροφή των νευροαξόνων συνέβαινε σε χρόνιες πλάκες, στα προχωρημένα στάδια της νόσου. Πρόσφατες μελέτες όμως έχουν δείξει ότι η καταστροφή των νευροαξόνων συμβαίνει και σε πρώιμες βλάβες. Η καταστροφή των νευροαξόνων συμβάλλει στη δημιουργία μη αναστρέψιμων βλαβών και άρα ανάλογης μη αναστρέψιμης συμπτωματολογίας (Logothetis and Milonas, 2004).

Να τονιστεί ότι οι νευροάξονες με προσβολή της μυελίνης είναι ιδιαίτερα ευαίσθητοι σε έναν αριθμό εξωγενών παραγόντων, για παράδειγμα στην αυξημένη θερμοκρασία του σώματος (πυρετός, πολύ ζεστό μπάνιο, υπερβολική άσκηση). Οι παράγοντες αυτοί μπορεί να προκαλέσουν διακοπή αγωγιμότητας και επιδείνωση υπαρχόντων συμπτωμάτων ή εμφάνιση από σιωπηλές κλινικά πλάκες (φαινόμενο του Uhthoff).

#### **2.3.1.4 Παθολογική ανατομική**

Μακροσκοπική εμφάνιση μικρών φαιόχρωων περιγεγραμμένων περιοχών στη λευκή ουσία του εγκεφάλου, αποτελούν τις απομυελινωτικές πλάκες. Ιδιαίτερα αυτές συσσωρεύονται περικοιλιακά στα ημισφαίρια, γύρω από τον υδραγωγό και την τέταρτη κοιλία, στην

παρεγκεφαλίδα, το νωτιαίο μυελό και τα οπτικά νεύρα. Μικροσκοπικά, στις πλάκες υπάρχει έντονη απομυελίνωση, αλλά και μειωμένη πυκνότητα νευροαξόνων σε σύγκριση με αυτή της λευκής ουσίας που περιβάλλει την πλάκα. Σε πρόσφατες πλάκες υπάρχει οίδημα και ιστολογικά αποκαλύπτεται η ύπαρξη περιαγγειακής διήθησης από λεμφοκύτταρα και μακροφάγα. Στις παλαιότερες πλάκες προέχει αύξηση της νευρογλοίας και δημιουργία ουλώδους ιστού.

### **2.3.1.5 Επιδημιολογία**

Η συχνότητα της νόσου αυξάνεται με την αύξηση του γεωγραφικού πλάτους τόσο προς το βόρειο όσο και προς το νότιο ημισφαίριο. Ο επιπολασμός της για χώρες μακριά από τον Ισημερινό φθάνει το 40 με 80 σε 100.000 πληθυσμό, αλλά σε ορισμένες περιοχές όπως στη Σκωτία, είναι και πολύ μεγαλύτερος. Γενικά, ζώνες με μεγαλύτερη συχνότητα εμφάνισης είναι η βόρεια Ευρώπη, ο Καναδάς, οι Βόρειες πολιτείες των ΗΠΑ και η Αυστραλία. Στις περιοχές της Νότιας Ευρώπης ο επιπολασμός βρίσκεται μεταξύ 10 και 40 σε 100.000 (στη χώρα μας έχει βρεθεί στο 29,5), ενώ κοντά στον Ισημερινό η συχνότητα είναι μικρότερη από 5 σε 100.000 (Paty and Ebens, 1998).

Η νόσος είναι συχνότερη στις γυναίκες από ότι στους άνδρες (1,5-2 προς 1). Η συχνότερη ηλικία έναρξης είναι η νεαρή ηλικία (20-30), αλλά η έναρξη της νόσου δεν είναι σπάνια ακόμη και στην 5η δεκαετία. Σπάνια είναι η έναρξη πριν από το 15ο έτος ηλικίας και μετά το 60ό . Στις μεγαλύτερες ηλικίες η νόσος τείνει να εμφανίζεται με τη χρόνια



προοδευτική μορφή με νωτιαίο εντόπιση, αλλά υπάρχουν και περιπτώσεις με εξάρσεις και υφέσεις. Η ύπαρξη περιβαλλοντικού παθολογικού παράγοντα υποστηρίζεται από το γεγονός ότι άτομα που μετανάστευσαν σε μικρή ηλικία από χώρες με υψηλή συχνότητα της νόσου (π.χ. Μ. Βρετανία) σε χώρες με χαμηλότερη (π.χ. Ν. Αφρική), εμφάνιζαν τη συχνότητα της χώρας προορισμού τους, ενώ αν μετανάστευαν σε ηλικία μεγαλύτερη των 15 ετών, εμφάνιζαν συχνότητα ανάλογη με της χώρας προέλευσής τους (Compston and Coles, 2008). Το γεγονός αυτό είναι ενδεικτικό της δράσης κάποιου περιβαλλοντικού παράγοντα στα πρώτα χρόνια ζωής του πάσχοντα. Τον περιβαλλοντικό παράγοντα υποστηρίζουν και οι λεγόμενες "επιδημίες" της νόσου στις Φερόες νήσους (ακτές της Σκωτίας), όπου ενώ η νόσος ήταν άγνωστη, εμφανίσθηκε σε πολλή μεγάλη συχνότητα μετά την κατάληψη των νήσων από βρετανικά στρατεύματα (Milonas, Tsonuis and Logothetis, 1990).

Σε γενικές γραμμές η παραπάνω υπόθεση αποτελεί σημαντικό κίνητρο για τη διεξαγωγή της παρούσας διπλωματικής εργασίας, αφού είναι ένας παράγοντας με ιδιαίτερα αυξημένο ενδιαφέρον και για τον οποίο διεξάγεται έρευνα σε αρκετά ανεπτυγμένο επίπεδο.

#### **2.3.1.6 Κλινική Εικόνα**

Χαρακτηριστική είναι η πολυεστιακή διάσπαρτη σημειολογία με πορεία εξάρσεων και υφέσεων ή λιγότερο συχνά, με πορεία προϊούσας εξέλιξης. Οι εξάρσεις μπορεί να διαρκούν από τουλάχιστον 24 ώρες μέχρι ημέρες, εβδομάδες ή και περισσότερο και είναι δυνατόν να αφήνουν ή όχι

υπολείμματα που όταν συσσωρεύονται μπορεί να προκαλούν κάποιου βαθμού φυσική αναπηρία. Μεταξύ εξάρσεων οι υφέσεις μπορεί να διαρκούν εβδομάδες, μήνες ή και χρόνια (Logothetis and Milonas, 2004).

Δεν υπάρχει κάποια κλασική κλινική εικόνα της νόσου, ωστόσο, συχνότερα είναι τα συμπτώματα σε ορισμένες περιοχές, όπως τα οπτικά νεύρα, το εγκεφαλικό στέλεχος, την παρεγκεφαλίδα και το νωτιαίο μυελό. Συμπτώματα στο εγκεφαλικό φλοιό και σε βασικά γάγγλια (βιολογική μάζα ιστών και πιο συχνά μάζα φορέων νευρικών κυττάρων είναι σπάνια ([www.news-medical.net/health/Ganglion.aspx](http://www.news-medical.net/health/Ganglion.aspx))).

Συνηθισμένη είναι η έναρξη με παραισθησίες στα άκρα, διπλωπία ή αδυναμία ενός άκρου. Συχνή είναι επίσης, η έναρξη με απώλεια όρασης από το ένα μάτι, λόγω οπισθοβολβικής νευρίτιδας με παρεμβολή συνήθως αρκετού χρόνου μέχρι την εμφάνιση επόμενων συμπτωμάτων. Το σημείο Lhermitte -με αίσθημα διαπέρασης ρεύματος στο μήκος της σπονδυλικής στήλης ή των άκρων σε κάμψη της κεφαλής- είναι αρκετά σταθερό σύμπτωμα με την αρχική εγκατάσταση της νόσου.

Στην πορεία της νόσου οι εξάρσεις διαδέχονται η μια την άλλη σε διαφορετικά χρονικά διαστήματα. Συμπτώματα με σχετική σειρά συχνότητας είναι μυϊκή αδυναμία, παρέσεις (μικρές παραλύσεις μυών), οπτικές διαταραχές (θάμβος, κεντρικά σκοτώματα κ.ά), μουδιάσματα στα άκρα, αστάθεια, διαταραχές κύστεως, δυσαρθρία, ζάλη και άλλα.

Η αντικειμενική σημειολογία είναι γενικά διάσπαρτη χωρίς συστηματοποίηση. Μερικά συνήθη αντικειμενικά ευρήματα είναι: σπαστική πάρεση (βαρύτερη και συχνότερη στα κάτω άκρα), απουσία κοιλιακών αντανακλάσεων, οπτική ατροφία, παρεγκεφαλιδική αταξία,

νυσταγμός, τρόμος, δυσαρθρία, διαπυρηνική οφθαλμοπληγία, ελάττωση της έν τω βάθει αισθητικότητας ή σπαστικοαταξικό βάδισμα. Υπολογίζεται ότι πάνω από το 75% των ασθενών παρουσιάζουν συνδυασμό σπαστικοαταξικών εκδηλώσεων, κυστικών διαταραχών και οπτική ατροφία.

Συχνά οι ασθενείς καταβεβλημένοι από τη σκλήρυνση κατά πλάκας παραπονιούνται για αίσθημα κόπωσης ενώ λιγότερο συχνά είναι τα παροξυντικά συμπτώματα που διαρκούν κάποια δευτερόλεπτα και επαναλαμβάνονται πολλές φορές την ημέρα. Τέτοια συμπτώματα είναι διπλωπία, τονικές συστάσεις άκρων, παροξυντική εμφάνιση παραισθησιών και άλλα. Από την ψυχική οπτική σφαίρα, αναφέρονται καταθλίψεις και λιγότερο συχνά ευφορία και συναισθηματική ακράτεια. Σπάνια είναι η άνια και μόνο στα προχωρημένα στάδια, αλλά η μικρού βαθμού ψυχονοητική έκπτωση είναι συχνότερη.

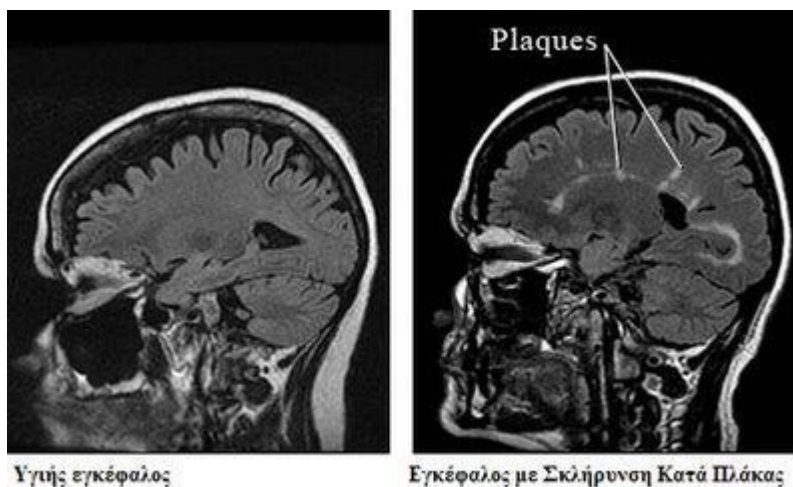
### **2.3.1.7 Εργαστηριακές Εξετάσεις**

Σε περίοδο εξάρσεων το λεύκωμα (θρεπτική ουσία των κυττάρων) στο εγκεφαλονωτιαίο υγρό (ENY) είναι πιθανό να είναι ελαφρά αυξημένο και τα μονοπύρινα κύτταρα να φτάνουν τα 30-50 ανά  $\text{mm}^3$ . Η σχέση  $\gamma$ -σφαιρίνης/λευκωματίνης είναι αυξημένη στο 70% των περιπτώσεων και χαρακτηριστική είναι η διάσπαση της IgG σφαιρίνης σε ολιγοκλωνικές ζώνες στην ανοσοηλεκτροφόρηση των λευκωμάτων του ENY. Οι ολιγοκλωνικές αυτές ζώνες εμφανίζονται σε συχνότητα μεγαλύτερη του 90%. Επίσης, ο δείκτης της IgG ο οποίος συγκρίνει τη σχέση

IgG/λευκωματίνης στο ENY και στον ορό, είναι αυξημένος στο 90% των ασθενών με τη βέβαιη μορφή της νόσου. Στο ENY βρίσκεται επίσης αυξημένη σε περιόδους εξάρσεων και η βασική πρωτεΐνη της μυελίνης (Andreadou et al., 2013).

Η πιο ευαίσθητη εργαστηριακή εξέταση είναι η μαγνητική τομογραφία που αποκαλύπτει την ύπαρξη πολλαπλών απομυελινωτικών εστιών, τόσο παλιών όσο και νέων, στο κεντρικό νευρικό σύστημα (ΚΝΣ) (Εικόνα 1). Εκεί, μετά τη χορήγηση ενδοφλέβιου σκιαστικού, εντοπίζονται οι πιο πρόσφατες εστίες. Τα ευρήματα στη μαγνητική τομογραφία είναι συνήθως πολλά περισσότερα από τη συμπτωματολογία που αναφέρθηκε από τον ασθενή. Να σημειωθεί ότι στην πορεία της νόσου εμφανίζονται νέες εστίες στη μαγνητική τομογραφία ακόμα και χωρίς υποτροπή της. Έχει υπολογιστεί ότι η μαγνητική είναι φυσιολογική σε πάρα πολύ μικρό ποσοστό ασθενών με πολλαπλή σκλήρυνση (ίσως μικρότερο του 3%). Παρόλα αυτά, τα ευρήματα της εξέτασης δεν είναι παθογνωμικά για τη νόσο. Επίσης, ένα ποσοστό μέχρι 4% φυσιολογικών ατόμων, είναι πιθανό να εμφανίσουν μικρές αλλοιώσεις στα εγκεφαλικά ημισφαίρια που μοιάζουν με αλλοιώσεις της πολλαπλής σκλήρυνσης (Milonas, 2003). Οι απομυελινωτικές εστίες σπάνια απεικονίζονται στην αξονική τομογραφία, ουσιαστικά μόνο όταν έχουν πάρει μεγάλη έκταση.

**Εικόνα 1:** Αξονική τομογραφία εγκεφάλου



### 2.3.1.8 Διάγνωση

Η διάγνωση βασίζεται στην έναρξη της νόσου σε μικρή ηλικία, στο χαρακτηριστικό ιστορικό των εξάρσεων και υφέσεων, με την εμφάνιση συμπτωμάτων με διασπορά στο ΚΝΣ και στο χρόνο, στην πολυεστιακή σημειολογία και στις εργαστηριακές εξετάσεις που αναφέρθηκαν πιο πάνω. Στην κλινική πράξη, οι ασθενείς κατατάσσονται σε τρεις κατηγορίες, αναφορικά με τη βεβαιότητα της διάγνωσης (McDonald et al., 2001).

- *Βέβαιη πολλαπλή σκλήρυνση*

Αν εμφανιστεί τουλάχιστον ένας από τους παρακάτω συνδυασμούς:

- α) Ιστορικό 2 προσβολών και κλινική ένδειξη για 2 ξεχωριστές βλάβες.

β) Ιστορικό 2 προσβολών, κλινική ένδειξη μίας βλάβης και παρακλινική ένδειξη (νευροαπεικόνιση) μιας ξεχωριστής βλάβης.

γ) Ιστορικό 2 προσβολών, κλινική ή παρακλινική ένδειξη μίας βλάβης και αύξηση της γ-σφαιρίνης ή και ολιγοκλωνικές ζώνες στο ENY.

δ) Ιστορικό μιας προσβολής, κλινική ή παρακλινική ένδειξη 2 ξεχωριστών βλαβών και αύξηση της γ-σφαιρίνης ή και παρουσία ολιγοκλωνικών ζωνών στο ENY.

• *Δυνατή πολλαπλή σκλήρυνση*

Αν εμφανιστεί ένας από τους παρακάτω συνδυασμούς:

α) Ιστορικό 2 προσβολών και κλινική ένδειξη μίας βλάβης.

β) Ιστορικό μιας προσβολής και κλινική ένδειξη 2 ξεχωριστών βλαβών.

γ) Ιστορικό μίας προσβολής, κλινική ένδειξη μίας βλάβης και παρακλινική ένδειξη μιας ακόμα βλάβης.

δ) Ιστορικό 2 προσβολών και αύξηση της γ-σφαιρίνης ή και ολιγοκλωνικές ζώνες στο ENY.

• *Πιθανή πολλαπλή σκλήρυνση*

Ο ασθενής κατατάσσεται σε αυτήν την διαβάθμιση όταν συνδυάζονται μόνο 2 από τα παραπάνω στοιχεία, δηλαδή από το ιστορικό, την κλινική ή παρακλινική εξέταση.

Να σημειωθεί ότι η αναφορά σε δυο προσβολές αφορά διαφορετικά μέρη του ΚΝΣ και τις χωρίζει περίοδος τουλάχιστον ενός μήνα. Σε κάθε

περίπτωση, για να θεωρηθεί μια προσβολή πλήρης, είναι απαραίτητο να διαρκεί τουλάχιστον 24 ώρες.

Πρόσφατα προτάθηκαν καινούρια κριτήρια για τη διάγνωση της νόσου (McDonald *et al.*(2001)), τα οποία χρησιμοποιούν ευρήματα της μαγνητικής τομογραφίας. Με αυτά τα κριτήρια, αν υπάρξουν νέα ευρήματα 3, 6 ή 12 μήνες μετά την πρώτη εμφάνιση ύποπτου συμπτώματος για τη νόσο και ευρήματα στην τότε μαγνητική τομογραφία, μπορεί να τεθεί η διάγνωση της πολλαπλής σκλήρυνσης ακόμα και χωρίς να υπάρχει κλινική υποτροπή (Πίνακας 1)

Να σημειωθεί επίσης ότι πριν διαγνωστεί η νόσος, είναι απαραίτητο να αποκλειστούν άλλες καταστάσεις που προκαλούν παρόμοια συμπτωματολογία και εμφανίζουν παρόμοια ευρήματα στη μαγνητική τομογραφία. Ο αποκλεισμός γίνεται κυρίως με έναν πλήρη ανοσολογικό έλεγχο προς αποκλεισμό ανοσολογικών παθήσεων ή έλεγχο αντισωμάτων.

**Πίνακας 1:** Νέα διαγνωστικά κριτήρια για ΣΚΠ (κριτήρια McDonald *et al.*)

Όταν πληρούνται όλα τα κριτήρια, η διάγνωση είναι ΣΚΠ, αν δεν πληρούνται όλα η διάγνωση είναι πιθανή, αν δεν πληρούνται κανένα αναζητείται άλλη διαγνωση

Κλινικές Ώσεις	Αντικειμενικά Κλινικά Ευρήματα	Επιπλέον Διαγνωστικά Ευρήματα
$\geq 2$	$\geq 2$	Κανένα
$\geq 2$	1	Διασπορά <b>στον χώρο</b> στην MRI ή δύο ή περισσότερες εστίες στην MRI συμβατές με ΣΚΠ και θετικό ENY ή άλλη ώση από προσβολή άλλης περιοχής του ΚΝΣ
1	$\geq 2$	Διασπορά <b>στο χρόνο</b> στην MRI ή 2 <sup>η</sup> κλινική ώση
1 (με ένα σύμπτωμα)	1	Διασπορά <b>στο χώρο</b> στην MRI ή λιγότερες από 2 εστίες στην MRI και θετικό ENY, <b>και</b> Διασπορά στο χρόνο στην MRI ή 2 <sup>η</sup> κλινική ώση
0 (σταδιακή νευρολογική επιδείνωση από την αρχή)	1	<b>Θετικό ENY και</b> Διασπορά στο χώρο στην MRI με 9 ή παραπάνω εστίες στην T <sub>2</sub> ακολουθία στον εγκέφαλο ή πάνω από 2 εστίες στον NM ή 4-8 εστίες στον εγκέφαλο και 1 στον NM ή 4-8 εστίες στον εγκέφαλο και παθολογικά ΟΠΔ ή λιγότερες από



4 εγκεφαλικές εστίες και 1 εστία στον NM και παθολογικά ΟΠΔ και Διασπορά **στο χρόνο** στην MRI ή συνεχής επιδείνωση για 1 χρόνο

---

\* **MRI:** Μαγνητική Τομογραφία, **ΚΝΣ:** Κεντρικό Νευρικό Σύστημα, **ΕΝΥ:** Εγκεφαλονωτιαίο Υγρό, **NM:** Νωτιαίος Μυελός, **ΟΠΔ:** Οπτικά Προκλητά Δυναμικά

### 2.3.1.9 Πρόγνωση

Πρώιμη εμφάνιση παρεγκεφαλιδικών και πυραμιδικών διαταραχών έχει βαρύτερη πρόγνωση σε αντίθεση με αισθητικές και οπτικές διαταραχές. Καλής πρόγνωσης είναι οι αραιές, ήπιες και βραχείας διάρκειας εξάρσεις και οι μακράς διάρκειας και πλήρεις υφέσεις. Και σε αυτές τις περιπτώσεις όμως, δεν αποκλείεται η εμφάνιση ώσης (διακύμανση της νόσου με περιόδους εξάρσεων και υφέσεων) με σοβαρή συμπτωματολογία που να αφήσει υπολειμματική κατάσταση. Η μορφή με εξάρσεις και υφέσεις έχει καλύτερη πρόγνωση από τη χρόνια προοδευτική μορφή. Σε ταχέως εξελισσόμενες μορφές οι άρρωστοι εμφανίζουν βαριές αναπηρίες και μπορεί να καταλήξουν σε αναπηρική καρέκλα μέσα σε 5 με 10 χρόνια.

Μετά από ένα μεμονωμένο μονοσυμπτωματικό επεισόδιο, ο μελλοντικός κίνδυνος εμφάνισης πολλαπλής σκλήρυνσης είναι μεγάλος, αν υπάρξουν απομυελινωτικές εστίες σε ποσοστό 70-80% σε περίοδο 10 χρόνων, στη μαγνητική τομογραφία. Σε περίπτωση τέτοιου μονοσυμπτωματικού επεισοδίου, ο κίνδυνος μελλοντικής εμφάνισης της νόσου είναι μειωμένος, αν η μαγνητική τομογραφία είναι φυσιολογική με ευρήματα 15% μέσα σε 10 χρόνια.

### 2.3.1.10 Θεραπεία

Θεραπεία για ίαση μέχρι στιγμής δεν υπάρχει, αλλά η θεραπευτική αντιμετώπιση της σκλήρυνσης κατά πλάκας διακρίνεται σε 4 τομείς:

- θεραπεία των οξέων υποτροπών (ώσεων),
- υποστηρικτική θεραπεία (συμβουλευτική, φυσική αποκατάσταση, ψυχολογική στήριξη),
- συμπτωματική θεραπεία (σπαστικότητα, αισθήματος κόπωσης, κυστικών διαταραχών και άλλων) και
- προληπτική θεραπεία.

Τα φάρμακα που προτιμώνται στις εξάρσεις είναι τα στεροειδή, ενώ για προφυλακτική θεραπεία χορηγούνται ανοσοκατασταλτικά ή ανοσοτροποποιητικά. Επίσης, η φυσιοθεραπεία αποτελεί απαραίτητο συμπλήρωμα της θεραπείας βοηθώντας στην αντιμετώπιση των σπαστικών παραλύσεων των μυών (Milonas, 2003).

Συνοπτικά, η αιτία της νόσου, αν και δεν είναι σαφώς καθορισμένη, φαίνεται ότι έχει σχέση με το ανοσοποιητικό σύστημα του οργανισμού, δηλ. πιθανώς ανήκει στα αυτοάνοσα νοσήματα (*autoimmune diseases*). Για το λόγο αυτό, τα φάρμακα που χορηγούνται στην πολλαπλή σκλήρυνση έχουν σχέση με το σύστημα ανοσίας. Έχουν διατυπωθεί θεωρίες που συσχετίζουν την πολλαπλή σκλήρυνση με ιούς, γονιδιακούς ή τοξικούς-περιβαλλοντικούς παράγοντες, αλλά καμιά θεωρία δεν έχει καθολική ισχύ.

### **2.3.2 Συγκεντρική Σκλήρυνση του Balo**

Πρόκειται για μια φλεγμονώδη απομυελινωτική νόσο αγνώστου αιτιολογίας με χαρακτηριστική παθολογοανατομική εικόνα, κατά την οποία, δακτύλιοι με φυσιολογική μυελίνη εναλλάσσονται με δακτυλίους με απομυελίνωση. Η νόσος εμφανίζεται συνήθως σε νέα άτομα και αφήνει σοβαρά νευρολογικά ελλείμματα (Hartung, 2002).

### **2.3.3 Οπτική Νευρομυελίτιδα (Νόσος του Devic)**

Είναι μια απομυελινωτική νόσος με κύρια χαρακτηριστικά την οπτική νευρίτιδα και την εγκάρσια μυελίτιδα, που προσβάλλουν συχνότερα άτομα νεαρής ηλικίας. Η αιτία είναι άγνωστη, πιθανώς όμως υπεισέρχεται λοιμώδης παράγοντας ή μεταλοιμώδης αυτοανοσοβιολογική αντίδραση. Η έναρξη είναι οξεία, με απώλεια της κεντρικής όρασης μέχρι και τέλεια τύφλωση. Ακολουθεί το στάδιο της μυελίτιδας ενώ η αποκατάσταση είναι μερική (Lassman, 2001).

### **2.3.4 Διάχυτη Σκλήρυνση ή Διάχυτη Μυελινοκλαστική Σκλήρυνση (Νόσος του Schilder)**

Πρόκειται για μια σπάνια και αμφισβητούμενη νόσο ως ξεχωριστή κλινική οντότητα που προσβάλλει την παιδική ηλικία. Υπάρχουν διάχυτες ή συρρέουσες περιοχές απομυελίνωσης στη λευκή ουσία των ημισφαιρίων που σταματά απότομα λίγα χιλιοστά από τη φαιά ουσία. Χαρακτηριστική είναι η έναρξη με έκπτωση της όρασης ή με κώφωση, αν η απομυελίνωση αρχίσει στους κροταφικούς λοβούς. Αμφοτερόπλευρη βλάβη των πυραμιδικών ινών οδηγεί σε σπαστική τετραπληγία και ψευδοπρομηκική παράλυση. Ψυχο-διανοητική έκπτωση, ψυχασικές αντιδράσεις, επιληπτικές κρίσεις και τελικό *stupor* (απώλεια γνωστικής λειτουργίας με τον ασθενή να αποκρίνεται μόνο σε βασικά ερεθίσματα, όπως ο πόνος) είναι συνηθισμένα επακόλουθα. Συχνή είναι και η αύξηση της ενδοκρανιακής πίεσης ενώ τα ευρήματα στο ΕΝΥ είναι τα ίδια με της πολλαπλής σκλήρυνσης. Η αξονική τομογραφία είναι ενδεικτική μεγάλων περιοχών απομυελίνωσης στη λευκή ουσία των ημισφαιρίων και το ηλεκτροεγκεφαλογράφημα έχει διάχυτους βραδείς ρυθμούς. Η πρόγνωση είναι βαριά με μέση επιβίωση 3 χρόνων (Lassman, 2001).

### **2.3.5 Προϊούσα Πολυεστιακή Λευκοεγκεφαλοπάθεια**

Είναι μια απομυελινωτική νόσος που εμφανίζεται με την ευκαιρία κακοήθων νόσων ως παρανελασματικό σύνδρομο. Η κατάσταση

θεωρείται ότι οφείλεται σε λοίμωξη από βραδείς δρώντες ιούς (Logothetis and Milonas, 2004).

### **2.3.6 Οξεία Διάσπαρτη Εγκεφαλομυελίτιδα**

Είναι μια σπάνια, οξεία φλεγμονώδης απομυελινωτική νόσος του κεντρικού νευρικού συστήματος προκαλούμενη από μια αυτοάνοσο προσβολή, συνήθως αποτέλεσμα λοίμωξης από ιούς. Η μαγνητική τομογραφία δείχνει εστίες στη λευκή ουσία των εγκεφαλικών ημισφαιρίων, αλλά και στην παρεγκεφαλίδα, εγκεφαλικό στέλεχος και νωτιαίο μυελό (Logothetis and Milonas, 2004).

### **2.3.7 Οξεία Αιμορραγική Εγκεφαλομυελίτιδα**

Πρόκειται για μια πολύ επιθετική μορφή της οξείας διάσπαρτης εγκεφαλομυελίτιδας. Είναι ακόμη πιο σπάνια και παθολογοανατομικά υπάρχουν επιπρόσθετα εκτεταμένες νεκρωτικές αιμορραγικές βλάβες. Το ENY έχει μεγάλο αριθμό κυττάρων, ενίοτε και πάνω από 1000. Η πρόγνωση είναι κακή και οι περισσότεροι ασθενείς καταλήγουν (Logothetis and Milonas, 2004).

## **2.4 ΜΗ-ΦΛΕΓΜΟΝΩΔΕΙΣ**

### **ΑΠΟΜΥΕΛΙΝΩΤΙΚΕΣ ΝΟΣΟΙ**

Στην κατηγορία αυτήν υπάγονται οι δυσμυελινωτικές νόσοι και οι τροφικές ή/και τοξικές απομυελινωτικές νόσοι.

#### **2.4.1 Δυσμυελινωτικές Νόσοι ή Λευκοδυστροφίες**

Παρά το γεγονός ότι αυτή η διαταραχή επηρεάζει την παραγωγή και ανάπτυξη της μυελίνης, οι λευκοδυστροφίες είναι διαφορετικές από τις απομυελινωτικές διαταραχές, όπως η σκλήρυνση κατά πλάκας, όπου ενώ η μυελίνη έχει συσταθεί κανονικά, χάνεται λόγω ανοσολογικής δυσλειτουργίας ή από άλλα αίτια. Επίσης εδώ, κύριο χαρακτηριστικό είναι η μη φυσιολογική μυελίνωση και η ελαττωματική δομή του ελύτρου της μυελίνης.

#### **2.4.2 Τροφικές-Τοξικές Απομυελινωτικές Νόσοι**

Είναι η δεύτερη κατηγορία των μη φλεγμονώδων απομυελινωτικών νόσων.

#### **2.4.2.1 Κεντρική Γεφυρική Μυελινόλυση**

Είναι μια κατάσταση συμμετρικής καταστροφής της μυελίνης των νευρικών ινών στο βασικό μέρος της γέφυρας. Συναντάται συνήθως σε αλκοολικούς ή γενικά σε βαριές διαταραχές στη θρέψη του οργανισμού. Εκδηλώνεται με γρήγορη εμφάνιση τετραπληγίας και προμηκικής-ψευδοπρομηκικής παράλυσης συχνά με εικόνα εγκλεισμού (*lock in syndrome*), κατάσταση στην οποία η μόνη επικοινωνία γίνεται με κινήσεις των βλεφάρων και των βομβών. Η πρόγνωση είναι βαριά.

#### **2.4.2.2 Πρωτοπαθής Εκφύλιση του Μεσολοβίου (Νόσος των Marchiafava-Bignami)**

Είναι μια κατάσταση οξείας απομυελινοτικής βλάβης του μεσολοβίου, αλλά μερικές φορές και μεγάλου τμήματος της λευκής ουσίας. Παρατηρείται σε αλκοολικούς και χαρακτηρίζεται κλινικά από επιληπτικές κρίσεις, σύγχυση, σπαστική παράλυση, ψυχικές διαταραχές με τάση για βίαιες ενέργειες και άνοια. Στα τελικά στάδια της νόσου ο ασθενής πέφτει σε κώμα και επέρχεται ο θάνατος.

### **2.5 ΣΥΝΟΨΗ**

Στο Κεφάλαιο 2 παρουσιάστηκαν συνοπτικά τα διάφορα είδη απομυελινοτικών νόσων ενώ αναλύθηκε διεξοδικότερα η νόσος της

σκλήρυνσης κατά πλάκας αφού είναι εκείνη που μελετάται στην παρούσα διπλωματική. Τόσο οι φλεγμονώδεις όσο και οι μη φλεγμονώδεις απομυελινωτικές νόσοι απασχολούν την ιατρική κοινότητα δημιουργώντας μια βιβλιογραφία πολλών άρθρων και διατριβών. Η σκλήρυνση κατά πλάκας αποτελεί μια νόσο η οποία εμφανίζεται με ποικίλες μορφές και ολοένα αυξάνονται τα κρούσματά της. Θεραπεία της νόσου δεν έχει βρεθεί ακόμη, ενώ η αντιμετώπισή της διακρίνεται σε κάποιες συγκεκριμένες φάσεις θεραπείας.

Πανεπιστήμιο Πειραιώς



# Κεφάλαιο 3

## ΣΤΑΤΙΣΤΙΚΕΣ ΣΑΡΩΣΗΣ – ΑΝΑΣΚΟΠΗΣΗ ΜΕΘΟΔΩΝ

### 3.1 ΕΙΣΑΓΩΓΗ

Στο Κεφάλαιο 3 της παρούσας διπλωματικής εργασίας γίνεται μια αναλυτική περιγραφή των στατιστικών μεθόδων που υπάρχουν και χρησιμοποιούνται σε προβλήματα σάρωσης χώρου ή/και χρόνου, τόσο σε μονοδιάστατες περιπτώσεις, όσο και σε διδιάστατες. Πιο συγκεκριμένα, στη μονοδιάστατη περίπτωση παρουσιάζονται έννοιες όπως ροή (*run*), συνάρτηση σάρωσης (*scan*) και σχηματισμός (*pattern*). Αναλύονται επίσης κάποιες κατανομές που φαίνεται να προσεγγίζουν καλύτερα προβλήματα κατανομής σημείων στο χώρο ή/και στο χρόνο. Στη διδιάστατη περίπτωση αναλύονται διεξοδικότερα δύο βασικές προσεγγίσεις τέτοιων προβλημάτων, η πιθανοθεωρητική προσέγγιση και η προσέγγιση με λόγο πιθανοφάνειας του Kulldorff (1997).

## 3.2 Η ΜΟΝΟΔΙΑΣΤΑΤΗ ΠΕΡΙΠΤΩΣΗ

Ο όρος ροή σε μια ακολουθία δοκιμών, από μια μη εξειδικευμένη οπτική γωνία, αφορά την αδιάκοπη διαδοχή όμοιων αποτελεσμάτων. Από στατιστικής απόψεως, ως ροή μπορεί να οριστεί η διαδοχή (χωρίς διακοπή) όμοιων αποτελεσμάτων, τα οποία όμως ακολουθούνται από διαφορετικά αποτελέσματα σε  $n > 0$  εκτελέσεις ενός πειράματος τύχης το οποίο σε κάθε επανάληψη του μπορεί να δώσει δύο δυνατά αποτελέσματα: επιτυχία (1) ή αποτυχία (0). Ο αριθμός  $k$  των όμοιων στοιχείων μιας ροής αναφέρεται ως μήκος της ροής.

Έχοντας ως βάση την έννοια της ροής μπορεί να οριστεί μια σειρά τυχαίων μεταβλητών που αναφέρονται σε  $n$  αποτελέσματα ενός πειράματος τύχης με δύο δυνατά αποτελέσματα. Ενδεικτικά αναφέρεται το μήκος  $L_n$  της ροής επιτυχιών που έχει το μέγιστο μήκος. Επίσης, είναι δυνατό να οριστούν τυχαίες μεταβλητές που σχετίζονται με το χρόνο αναμονής μέχρι την εμφάνιση της πρώτης ή γενικότερα της  $r$ -οστής ροής επιτυχιών οποιουδήποτε είδους.

Τροποποιώντας την έννοια της ροής, μπορούν να μελετηθούν πολύπλοκότερα σχήματα, όπως είναι η έννοια της γενικευμένης ροής. Με τον όρο  $r$ -γενικευμένη ροή επιτυχιών μήκους  $k$ , ορίζεται ο αριθμός παραθύρων μήκους  $k$  τα οποία περιέχουν τουλάχιστον  $r$  επιτυχίες, σε μια ακολουθία  $n$  δοκιμών. Τέτοιες συναρτήσεις θα αναλυθούν πιο διεξοδικά στη συνέχεια.

Τέλος, γενικεύοντας κάποιες από τις έννοιες των ροών και των συναρτήσεων σάρωσης μπορούν να διαμορφωθούν και να μελετηθούν ακόμα πιο πολύπλοκοι σχηματισμοί (*patterns*). Τέτοιες περιπτώσεις εμφανίζονται συνήθως όταν μελετώνται ακολουθίες  $n$  εκτελέσεων ενός

πειράματος τύχης το οποίο καταλήγει σε περισσότερα από δυο δυνατά αποτελέσματα. Για παράδειγμα, σε  $n=10$  διαδοχικές επαναλήψεις ενός πειράματος τύχης με τρία δυνατά αποτελέσματα 0, 1 ή 2 προέκυψε η ακολουθία αποτελεσμάτων 1210212120. Παρατηρείται ότι ο σχηματισμός 12 εμφανίζεται 3 φορές, ενώ ο σχηματισμός 121 εμφανίζεται 2 φορές (Μπερσίμης, 2005).

Στο παραπάνω παράδειγμα, ορίστηκε μια ακολουθία γεγονότων σε διακριτό χρόνο. Το ίδιο συχνά μπορεί να οριστεί και μια ακολουθία σε συνεχές διάστημα. Στην πρώτη περίπτωση επομένως, η κατανομή της μέγιστης ροής όμοιων ή επιθυμητών αποτελεσμάτων προσομοιώνεται μέσα από μία ακολουθία δοκιμών Bernoulli, ενώ στη δεύτερη, η εμφάνιση όμοιων ή επιθυμητών αποτελεσμάτων σε συνεχή χρόνο, εύκολα προσομοιώνεται από μια Poisson διαδικασία.

### 3.2.1 Διωνυμική Κατανομή Γεγονότων

Έστω  $X_1, X_2, \dots, X_N$  μια ακολουθία ανεξάρτητων και ισόνομων τυχαίων διακριτών μεταβλητών, όπου  $P(X_i = 1) = p = 1 - P(X_i = 0)$ . Τα  $X$  αναφέρονται ως μια ακολουθία δοκιμών Bernoulli, ενώ η διαδικασία που αποτελούν, ως διαδικασία Bernoulli. Για έναν ακέραιο αριθμό  $m$  και για  $i = 1, 2, \dots, N - m + 1$  ορίζονται οι τυχαίες μεταβλητές (Glaz et al., 2001)

$$Y_i(m) = \sum_{j=i}^{i+m-1} X_j.$$

Οι ποσότητες  $Y_i(m)$  αποτελούν ένα κινητό άθροισμα των  $X_i$ . Η στατιστική συνάρτηση σάρωσης  $S'_m$  ορίζεται ως το μέγιστο των κινητών αθροισμάτων, δηλαδή ως ο μέγιστος αριθμός των μονάδων που εμφανίζονται σε  $m$  συνεχόμενες δοκιμές. Ισχύει δηλαδή

$$S'_m = \max_{1 \leq i \leq N-m+1} \{Y_i(m)\}.$$

Μια άλλη στατιστική συνάρτηση σχετιζόμενη με την παραπάνω, είναι η  $W_k$ , η οποία απεικονίζει το μικρότερο αριθμό συνεχόμενων δοκιμών στις οποίες εμπεριέχονται  $k$  μονάδες. Ισχύει δηλαδή

$$W_k = \min_{k \leq m \leq N} \{m : S'_m = k\}.$$

Δοθείσης μιας διαδικασίας Bernoulli στο  $(1, \infty)$  ορίζεται και η ποσότητα  $T_{k,m}$  ως ο χρόνος αναμονής μέχρι την παρατήρηση για πρώτη φορά τουλάχιστον  $k$  μονάδων σε οποιοδήποτε διάστημα μήκους  $m$ . Με άλλα λόγια,  $T_{k,m}$  είναι το μικρότερο  $i \geq m$  για το οποίο θα ισχύει  $Y_{i-m+1} \geq k$ . Οι τρεις στατιστικές σάρωσης που αναφέρθηκαν, συνδέονται μεταξύ τους ως εξής:

$$P(S'_m \geq k) = P(W_k \leq m) = P(T_{k,m} \leq N)$$

### 3.2.2 Poisson Κατανομή Γεγονότων

Όταν τα γεγονότα προέρχονται από μια διαδικασία Poisson, η κατανομή της στατιστικής σάρωσης διαφοροποιείται. Ο αριθμός των σημείων στο χρόνο δεν αντιμετωπίζεται ως σταθερός εκ των προτέρων γνωστός αριθμός  $N$ , αλλά ως μια τυχαία μεταβλητή με γνωστή συνάρτηση πιθανότητας. Συγκεκριμένα, η κατανομή του συνολικού αριθμού των σημείων είναι Poisson με παράμετρο  $\lambda$ . Δηλαδή κατά μέσο όρο εμφανίζονται  $\lambda$  σημεία στη μονάδα του χρόνου. Η χρησιμότητα της συγκεκριμένης υπόθεσης έγκειται στο γεγονός ότι ο ερευνητής μπορεί να χρησιμοποιήσει τη στατιστική σάρωσης για τη μελέτη μελλοντικών δεδομένων ή για το σχεδιασμό ενός συστήματος στο οποίο στο οποίο θα αποφεύγεται ή θα προλαμβάνεται η υπερφόρτωσή του, αφού θα είναι γνωστός από παρατηρήσεις ο ρυθμός άφιξης μονάδων σε αυτό.

Η διαδικασία Poisson χρησιμοποιείται για τη μοντελοποίηση πολλών φαινομένων που σχετίζονται με την εμφάνιση συμβάντων στο χώρο ή στο χρόνο. Ενδεικτικά αναφέρονται ο σχεδιασμός τηλεφωνικών συστημάτων, οι αφίξεις πελατών και οι τοποθεσίες ή το πλήθος περιπτώσεων διαφόρων ασθενειών.

Όπως προαναφέρθηκε, σε αρκετές εφαρμογές, δεν είναι απαραίτητο να είναι εκ των προτέρων γνωστή η κατανομή της στατιστικής συνάρτησης σάρωσης δεδομένου του συνολικού αριθμού περιστατικών  $N$  στο χρονικό διάστημα  $[0, T)$ . Αντί αυτού, θεωρείται ότι τα περιστατικά συμβαίνουν με τυχαίο τρόπο μέσα στο χρονικό διάστημα  $[0, T)$ , σύμφωνα με κάποια διαδικασία. Εάν  $\lambda$  είναι ο αναμενόμενος μέσος αριθμός περιστατικών στη μονάδα του χρόνου, τότε ο αριθμός των περιστατικών  $Y_i(w)$  σε

οποιοδήποτε διάστημα  $[t, t+w)$  ακολουθεί την Poisson κατανομή με μέσο  $\lambda w$ , δηλαδή

$$P(Y_\tau(w) = k) = e^{-\lambda w} (\lambda w)^k / k!, \quad k = 0, 1, 2, \dots$$

Ο αριθμός των συμβάντων σε οποιαδήποτε μη επικαλυπτόμενα διαστήματα είναι κατανεμημένος ανεξάρτητα. Για τη διαδικασία Poisson, οι χρόνοι αφίξεων  $W_k$  ανάμεσα στα σημεία αποτελούν ανεξάρτητες τυχαίες μεταβλητές από την Εκθετική κατανομή. Θεωρώντας γνωστό τον αριθμό των συνολικών σημείων  $N$  από την Poisson διαδικασία στον διάστημα  $[0, T)$ , αυτά τα  $N$  σημεία κατανέμονται ομοιόμορφα στο διάστημα αυτό.

Δεδομένου ότι τα περιστατικά εμφανίζονται με τυχειότητα στο χρόνο, ορίζεται η ποσότητα  $T_{k,w}$  ως ο χρόνος αναμονής μέχρι να παρατηρηθούν τουλάχιστον  $k$  περιστατικά σε οποιοδήποτε διάστημα μήκους  $w$ . Αναλυτικότερα,  $T_{k,w} = X_{(i+k-1)}$  για το μικρότερο  $i$  για το οποίο θα ισχύει ότι  $X_{(i+k-1)} - X_{(i)} \leq w$ . Οι τρεις στατιστικές σάρωσης  $S_w$ ,  $W_k$  και  $T_{k,w}$  συνδέονται μέσω της σχέσης  $P(S_w \geq k) = P(W_k \leq w) = P(T_{k,w} \leq T)$ .

### 3.3 Η ΔΙΔΙΑΣΤΑΤΗ ΠΕΡΙΠΤΩΣΗ

Πολλές μελέτες έχουν γίνει σε προβλήματα δύο ή και περισσότερων διαστάσεων. Ενδεικτικά αναφέρονται οι Orford (2000) με τις μελέτες του στο χώρο της αστρονομίας και την ανάλυση ασυνήθιστων εκρήξεων με υψηλή ενέργεια στο διάστημα. Οι δύο από τις βασικότερες προσεγγίσεις

τέτοιων χωρικών ή/και χρονικών προβλημάτων παρουσιάζονται παρακάτω.

### 3.3.1 Πιθανοθεωρητική Προσέγγιση

Έστω μια ορθογώνια περιοχή  $[0, T_1] \times [0, T_2]$ . Ορίζεται επίσης η ποσότητα  $h_i = \frac{T_i}{N_i} > 0$ , όπου  $N_i$  θετικοί ακέραιοι αριθμοί,  $T_i$  οι χρονικές

στιγμές και  $i = 1, 2$ . Σε πολλές εφαρμογές η ακριβής τοποθεσία των παρατηρηθέντων γεγονότων στην ορθογώνια περιοχή δεν είναι γνωστή.

Η πληροφορία που είναι συνήθως διαθέσιμη είναι οι μετρήσεις σε μικρές ορθογώνιες υπό-περιοχές. Για  $1 \leq i \leq N_1$  και  $1 \leq j \leq N_2$ , ορίζεται η

ποσότητα  $X_{ij}$  ως ο αριθμός των γεγονότων που παρατηρήθηκαν στην

ορθογώνια υπό-περιοχή  $[(i-1)h_1, ih_1] \times [(j-1)h_2, jh_2]$ . Κύριος στόχος

είναι η ανίχνευση ασυνήθιστων συστάδων αυτών των γεγονότων, υπό την μηδενική υπόθεση ότι οι  $X_{ij}$  είναι ανεξάρτητες και ισόνομες μη

αρνητικές και ακέραιες τυχαίες μεταβλητές από μια συγκεκριμένη

κατανομή. Για  $1 \leq i_1 \leq N_1 - m_1 + 1$  και  $1 \leq i_2 \leq N_2 - m_2 + 1$ , η ποσότητα

$Y_{i_1, i_2} = \sum_{j=i_2}^{i_2+m_2-1} \sum_{j=i_1}^{i_1+m_1-1} X_{ij}$  απεικονίζει τον αριθμό των γεγονότων που

σημειώθηκαν σε μια ορθογώνια περιοχή που αποτελείται από  $m_1 \times m_2$

παρακείμενες ορθογώνιες υπό-περιοχές, με την υπό-περιοχή  $h_1 h_2$  και τη

νοτιοδυτική γωνία να βρίσκονται στο σημείο  $((i_1 - 1)/h_1, (i_2 - 1)/h_2)$ . Αν

η ποσότητα  $Y_{i_1, i_2}$  ξεπεράσει μια προκαθορισμένη τιμή  $k$ , τότε  $k$  γεγονότα αποτελούν πιθανή συστάδα μέσα στην υπό μελέτη περιοχή. Ορίζεται επίσης μια διακριτή στατιστική συνάρτηση σάρωσης δύο διαστάσεων (*two-dimensional discrete scan statistic*), ως ο μέγιστος αριθμός συμβάντων σε οποιεσδήποτε γειτονικές ορθογώνιες υπό-περιοχές μεγέθους  $m_1 h_1 \times m_2 h_2$ . Η συνάρτηση θα είναι

$$S'_{m_1, m_2} = \max \{ Y_{i_1, i_2}; 1 \leq i_1 \leq N_1 - m_1 + 1, 1 \leq i_2 \leq N_2 - m_2 + 1 \}.$$

Η  $S'_{m_1, m_2}$  μπορεί να θεωρηθεί επέκταση της διακριτής στατιστικής συνάρτησης σάρωσης μιας διάστασης  $S'_m$ , που ορίστηκε παραπάνω, και σκοπός είναι η εκτίμηση της ποσότητας  $P(S'_{m_1, m_2} \geq k)$ . Η στατιστική σάρωσης  $S'_{m_1, m_2}$  χρησιμοποιείται για τον έλεγχο της μηδενικής υπόθεσης ότι οι ποσότητες  $X_{ij}$  είναι ανεξάρτητες και ισόνομες τυχαίες μεταβλητές από την κατανομή Bernoulli με παραμέτρους  $n_0$  και  $0 < p_0 < 1$ , ή από την κατανομή Poisson με μέσο  $\lambda_0 > 0$ . Για την εναλλακτική υπόθεση ύπαρξης συσταδοποιημένων παρατηρήσεων, ορίζεται η ορθογώνια υπό-περιοχή

$$R(i_1, i_2) = [(i_1 - 1)h_1, (i_1 + m_1 - 1)h_1] \times [(i_2 - 1)h_2, (i_2 + m_2 - 1)h_2]$$

έτσι ώστε για κάθε  $i_1 \leq i \leq i_1 + m_1 - 1$  και  $i_2 \leq j \leq i_2 + m_2 - 1$ , οι  $X_{ij}$  να ακολουθούν τη Διωνυμική κατανομή με παραμέτρους  $n_0$  και  $p_1 > p_0$ , ή



την Poisson με μέσο  $\lambda_1 > \lambda_0$ . Για οποιοδήποτε σημείο  $(i, j) \notin [i_1, i_1 + m_1 - 1] \times [i_2, i_2 + m_2 - 1]$ , οι  $X_{ij}$  κατανέμονται σύμφωνα με την κατανομή που ορίστηκε στην μηδενική υπόθεση.

Εύκολα μπορεί να συμπεράνει κανείς ότι ένας έλεγχος απορρίπτει την μηδενική υπόθεση, όταν η  $S'_{m_1, m_2}$  ξεπεράσει την τιμή  $k$ , η οποία προσδιορίζεται από συγκεκριμένο επίπεδο σημαντικότητας στη διαδικασία του ελέγχου. Επομένως, εάν η εναλλακτική υπόθεση ισχύει, σε μία τουλάχιστον υπό-περιοχή διαστάσεων  $m_1 h_1 \times m_2 h_2$ , τα γεγονότα θα εμφανίζονται με μεγαλύτερη συχνότητα. Για να εφαρμοστεί η στατιστική συνάρτηση σάρωσης δύο διαστάσεων, είναι απαραίτητες κάποιες προσεγγίσεις για την τιμή  $P(S'_{m_1, m_2} \geq k)$  καθώς δεν υπάρχουν ακριβείς τιμές στη διαθέσιμη βιβλιογραφία.

### **3.3.2 Χωρική Στατιστική Συνάρτηση Σάρωσης του Kulldorff**

#### **3.3.2.1 Εισαγωγή**

Όπως αναφέρεται σε σχετικό του άρθρο ο Kulldorff (1997), η στατιστική σάρωσης χρησιμοποιείται για τον εντοπισμό συστάδων σε κάποιο σύνολο σημείων και αποτελεί χρήσιμο εργαλείο στην επιστήμη της επιδημιολογίας αφού αποσκοπεί στον εντοπισμό και την παρακολούθηση ενδεχόμενων κινδύνων δημόσιας υγείας. Έχει μελετηθεί στη μονοδιάστατη περίπτωση από τον Naus (1965) και πολλούς άλλους (Wallenstein et al., 1993, Kulldorff, 1997). Μια ανασκόπηση μερικών

μεθόδων βρίσκεται στις εργασίες των Marshall (1991) και των Hills and Alexander (1989), ενώ αναφορά σε πιο σύγχρονες μορφές, γίνονται στο άρθρο του Jacques (1993).

Οι επιδημιολόγοι κατά κύριο λόγο, ενδιαφέρονται για τον εντοπισμό περιπτώσεων ασθένειας, εφόσον όμως έχουν προσαρμοστεί οι χωρικές διακυμάνσεις στην πυκνότητα του πληθυσμού που μελετάται. Το παραπάνω κρίνεται απαραίτητο στη μελέτη τέτοιου είδους δεδομένα, διότι είναι πιθανή η εξαγωγή λανθασμένου συμπεράσματος αν μια απλή συγκέντρωση πληθυσμού ερμηνευθεί λανθασμένα ως στατιστικά σημαντική συστάδα.

Σε γενικές γραμμές, για ένα σύνολο σημείων στο διάστημα  $[a, b]$ , ορίζεται ένα παράθυρο  $[t, t+w]$  σταθερού  $w < b-a$  το οποίο κινείται κατά μήκος του διαστήματος. Για όλες τις δυνατές τιμές του  $t$ , σημειώνεται ο μέγιστος αριθμός σημείων στο παράθυρο και έπειτα μέσω της μηδενικής υπόθεσης, ελέγχεται αν τα σημεία προέρχονται από κάποια τυχαία Poisson κατανομή.

Η προσπάθεια του Kulldorff επικεντρώθηκε στην εύρεση μιας στατιστικής σάρωσης που να γενικεύεται όσο το δυνατό περισσότερο, για περισσότερες από μία διαστάσεις, για διωνυμικά δεδομένα, για ανίχνευση πολλαπλών συστάδων, οποιουδήποτε μεγέθους και οποιασδήποτε τοποθεσίας. Μοναδική παραδοχή είναι ότι η ανάλυση θα βασίζεται στο γεγονός ότι ο συνολικός αριθμός των παρατηρηθέντων σημείων θα είναι εκ των προτέρων γνωστός (Kulldorff, 1997). Το παράθυρο σάρωσης μπορεί να πάρει οποιοδήποτε μέγεθος ή σχήμα και αυτό να αλλάζει κατά τη διάρκεια της σάρωσης, γεγονός πολύ χρήσιμο όταν δεν είναι εκ των προτέρων γνωστό το μέγεθος της περιοχής που εμφανίζεται συστάδα

γεγονότων. Επίσης, η μέθοδος που προτείνεται βρίσκει εφαρμογή και σε περιπτώσεις αυθαίρετης αλλά γνωστής, υποκείμενης έντασης που διέπει την κατανομή των σημείων κάτω από την μηδενική υπόθεση. Η μορφή που παίρνει, μπορεί να διαφέρει ανάλογα με την εφαρμογή. Για παράδειγμα, η μέθοδος μοντελοποιείται ως ένα μέτρο  $\mu$  σε κάποιον γεωγραφικό χώρο  $G$ . Όταν ο χώρος αυτός ορίζεται η ευθεία και το μέτρο  $\mu$  να κατανέμεται ομοιόμορφα στο  $[a,b]$ , το πρόβλημα που μελετάται ανάγεται στην κλασική μονοδιάστατη περίπτωση. Με την εισαγωγή του μέτρου Lebesgue, μπορεί να εμφανιστεί και μια χωρική διαδικασία Poisson όπου και απαιτείται περαιτέρω ανάλυση.

Στη μονοδιάστατη περίπτωση, η ακριβής κατανομή του στατιστικού ελέγχου που χρησιμοποιείται είναι γνωστή μόνο σε ορισμένες περιπτώσεις, ενώ στη διεθνή βιβλιογραφία καταγράφονται αρκετές προσεγγίσεις (Loader, 1991, Wallenstein et al., 1993). Στις περισσότερες διαστάσεις η στατιστική θεωρία γίνεται αρκετά σύνθετη και χρησιμοποιείται Monte-Carlo προσομοίωση για τον έλεγχο υποθέσεων. Ενδεικτικές μελέτες είναι αυτές του Naus (1965) ο οποίος εισήγαγε όρια κατανομής για διδιάστατη στατιστική συνάρτηση σάρωσης χρησιμοποιώντας σταθερό τετραγωνικό παράθυρο, ενώ ο Loader (1991) ασχολήθηκε με το ίδιο πρόβλημα αλλά κάνοντας χρήση παραθύρου μεταβλητού μεγέθους.

Όταν το παράθυρο σάρωσης είναι σταθερού μεγέθους, η στατιστική ελέγχου λαμβάνεται πάντα ως το μέγιστο αριθμό σημείων που βρέθηκαν στο παράθυρο, για οποιαδήποτε χρονική στιγμή. Με παράθυρο σάρωσης μεταβλητού μεγέθους, γίνεται χρήση του ελέγχου λόγου πιθανοφάνειας για την εξαγωγή συμπερασμάτων.

### 3.3.2.2 Μοντέλα Poisson και Bernoulli

Έστω  $N$  ένα σύνολο σημείων στο χώρο, με  $N(A)$  να είναι ο τυχαίος αριθμός σημείων στο σύνολο  $A \subset G$  και όπου  $G$  κάποιος γεωγραφικός χώρος. Καθώς το παράθυρο σάρωσης κινείται στην περιοχή μελέτης, δημιουργείται ένα σύνολο  $Z$  από ζώνες  $Z \subset G$ . Εναλλακτικά, θα χρησιμοποιείται ο συμβολισμός  $Z$ , τόσο για κάποιο υποσύνολο του  $G$ , όσο και για το σύνολο των παραμέτρων που ορίζουν μια ζώνη.

Για το μοντέλο Bernoulli, λαμβάνονται υπόψη μόνο μέτρα  $\mu$ , τέτοια ώστε η ποσότητα  $\mu(A)$  να είναι ακέραιος αριθμός για όλα τα υποσύνολα  $A \subset G$ . Κάθε μονάδα αυτού του μέτρου αντιστοιχεί σε ένα άτομο, το οποίο μπορεί να λάβει μία από τις δύο δυνατές καταστάσεις, για παράδειγμα, φορέας ή όχι της νόσου. Τα άτομα σε μία από τις δύο δυνατές καταστάσεις, θεωρούνται ως σημεία (*points*) και η τοποθεσία του καθενός θεωρείται μια διαδικασία (*point process*). Στο μοντέλο, υπάρχει ακριβώς μια ζώνη  $Z \subset G$  τέτοια ώστε, κάθε άτομο της μελέτης που θα βρεθεί μέσα σε αυτή, έχει πιθανότητα  $p$  να νοσήσει (αποτελέσει σημείο), ενώ άτομα εκτός ζώνης, έχουν πιθανότητα  $q$ , ανεξάρτητη από οποιουδήποτε άλλου ατόμου. Η μηδενική υπόθεση που ελέγχεται είναι η

$$H_0 : p = q$$

έναντι της εναλλακτικής

$$H_1 : p > q, Z \in \mathbf{Z}$$

Υπό την  $H_0$ , ισχύει  $N(A) \sim \text{Bin}(\mu(A), p)$  για όλα τα σύνολα  $A$ . Υπό την  $H_1$ , ισχύει  $N(A) \sim \text{Bin}(\mu(A), p)$  για όλα τα σύνολα  $A \subset Z$ , ενώ  $N(A) \sim \text{Bin}(\mu(A), q)$  για όλα τα σύνολα  $A \subset Z^c$ .

Στο μοντέλο Poisson, τα σημεία προέρχονται από μια μη ομογενή διαδικασία Poisson. Και εδώ, υπάρχει ακριβώς μια ζώνη  $Z \subset G$  τέτοια ώστε  $N(A) \sim P(p\mu(A \cap Z) + q\mu(A \cap Z^c)) \forall A$ . Η μηδενική υπόθεση είναι και πάλι

$$H_0 : p = q$$

ενώ η εναλλακτική

$$H_1 : p > q, Z \in Z$$

Υπό την  $H_0$ ,  $N(A) \sim P(p\mu(A)) \forall A$  ενώ η παράμετρος  $Z$  εξαφανίζεται από το μοντέλο.

Η κατάλληλη επιλογή του παραθύρου σάρωσης και κατά συνέπεια η αντίστοιχη συλλογή  $Z$  ζωνών, εξαρτάται πάντοτε από την εφαρμογή που μελετάται με συνηθέστερα τα κυκλικά παράθυρα σταθερού ή μεταβλητού μεγέθους και τα ορθογώνια παραλληλόγραμμα, σταθερού ή μεταβλητού μεγέθους.

Η επιλογή μοντέλου ανάμεσα στο Bernoulli και Poisson, δε φαίνεται να διαδραματίζει σημαντικό ρόλο όταν ο συνολικός αριθμός των σημείων είναι μικρός σε σχέση με τον αριθμό των σημείων σε ολόκληρο των γεωγραφικό χώρο  $G$ , αφού τα δύο μοντέλα προσεγγίζουν αρκετά καλά το ένα το άλλο. Σε άλλες περιπτώσεις, η επιλογή εξαρτάται από το πρόβλημα που μελετάται. Αν οι μετρήσεις είναι διωνυμικής φύσεως,

όπως για παράδειγμα η εμφάνιση ή όχι κάποιας ασθένειας, τότε χρησιμοποιείται το μοντέλο Bernoulli. Διαφορετικά, αν τα δεδομένα σχετίζονται με κάποιο συνεχή παράγοντα κινδύνου, όπως για παράδειγμα με δεδομένα από το μετρητή ραδιενέργειας Geiger ή με μετρήσεις βλαβερών ουσιών στον ανθρώπινο οργανισμό, τότε το μοντέλο που επιλέγεται είναι το Poisson.

### 3.3.2.3 Έλεγχος Λόγου Πιθανοφάνειας (Likelihood Ratio Test)

Έστω  $n_Z$  ο αριθμός των σημείων που έχουν παρατηρηθεί στη ζώνη  $Z$ , και έστω  $n_G$  ο συνολικός αριθμός σημείων που έχουν παρατηρηθεί. Η συνάρτηση πιθανοφάνειας για το μοντέλο Bernoulli, εκφράζεται ως εξής:

$$L(Z, p, q) = p^{n_Z} (1 - p)^{\mu(Z) - n_Z} q^{n_G - n_Z} (1 - q)^{(\mu(G) - \mu(Z)) - (n_G - n_Z)}.$$

Για να ανιχνευθεί η ζώνη η οποία έχει τις περισσότερες πιθανότητες να αποτελέσει συστάδα συμβάντων, εντοπίζεται η ζώνη  $\hat{Z}$  που μεγιστοποιεί τη συνάρτηση πιθανοφάνειας. Με άλλα λόγια, η  $\hat{Z}$  είναι ο εκτιμητής μέγιστης πιθανοφάνειας της παραμέτρου  $Z$ .

Η διαδικασία γίνεται σε δύο βήματα. Καταρχήν απαιτείται η μεγιστοποίηση της συνάρτησης πιθανοφάνειας (στο  $Z$ ).

$$L(Z) \stackrel{def}{=} \sup_{p>q} L(Z, p, q) = \left( \frac{n_Z}{\mu(Z)} \right)^{n_Z} \left( 1 - \frac{n_Z}{\mu(Z)} \right)^{\mu(Z)-n_Z} \\ \times \left( \frac{n_G - n_Z}{\mu(G) - \mu(Z)} \right)^{n_G - n_Z} \left( 1 - \frac{n_G - n_Z}{\mu(G) - \mu(Z)} \right)^{(\mu(G) - \mu(Z)) - (n_G - n_Z)}, \text{ \acute{o}ταν} \\ \frac{n_Z}{\mu(Z)} > \frac{(n_G - n_Z)}{(\mu(G) - \mu(Z))}$$

$$\text{και } L(Z) = \left( \frac{n_G}{\mu(G)} \right)^{n_G} \left( \frac{\mu(G) - n_G}{\mu(G)} \right)^{\mu(G) - n_G}, \text{ διαφορετικά.} \\ (3.1)$$

Στη συνέχεια, προκύπτει η λύση  $\hat{Z} = \{Z : L(Z) \geq L(Z') \forall Z' \in \mathbf{Z}\}$ .

Υπολογίζεται επίσης η ποσότητα

$$L_0 \stackrel{def}{=} \sup_{p=q} L(Z, p, q) = \left( \frac{n_G}{\mu(G)} \right)^{n_G} \left( \frac{\mu(G) - n_G}{\mu(G)} \right)^{\mu(G) - n_G}$$

η οποία συμβολίζει τη συνάρτηση πιθανοφάνειας υπό τη μηδενική υπόθεση και συγκεκριμένα, για την περίπτωση όπου  $p=q$ .

Ο λόγος πιθανοφάνειας  $\lambda$ , μπορεί να γραφεί ως:

$$\lambda = \frac{\sup_{Z \in \mathcal{Z}, p>q} L(Z, p, q)}{\sup_{p=q} L(Z, p, q)} = \frac{L(\hat{Z})}{L_0} \\ (3.2)$$

Παρατηρείται ότι ο παρονομαστής εξαρτάται μόνο από τον συνολικό αριθμό σημείων  $n_G$  και όχι από τη χωρική κατανομή των σημείων. Ο λόγος  $\lambda$  είναι αυτός που χρησιμοποιείται ως στατιστική ελέγχου.

Να σημειωθεί επίσης ότι, οι παραπάνω ισότητες βρίσκουν εφαρμογή στην περίπτωση που ο σκοπός της μελέτης είναι ο εντοπισμός ασυνήθιστα μεγάλου αριθμού γεγονότων σε μία περιοχή. Εάν αντιθέτως ζητείται ο εντοπισμός ασυνήθιστα μικρού αριθμού συγκεντρωμένων σημείων σε μία περιοχή, η μόνη αλλαγή γίνεται στην αρχική εξίσωση (3.1) και συγκεκριμένα, στη φορά των δύο ανισοτήτων.

Η συνάρτηση πιθανοφάνειας για το μοντέλο Poisson είναι

$$\frac{e^{-p\mu(Z)-q(\mu(G)-\mu(Z))} [p\mu(Z) + q(\mu(G) - \mu(Z))]^{n_G}}{n_G!}.$$

Η συνάρτηση πυκνότητας  $f(x)$  ενός συγκεκριμένου σημείου που παρατηρήθηκε στην τοποθεσία  $x$ , θα είναι

$$\begin{cases} \frac{p\mu(x)}{p\mu(Z) + q(\mu(G) - \mu(Z))}, x \in Z \\ \frac{q\mu(x)}{p\mu(Z) + q(\mu(G) - \mu(Z))}, x \notin Z \end{cases}$$

Η συνάρτηση πιθανοφάνειας μπορεί να γραφεί ως



$$\begin{aligned}
L(Z, p, q) &= \frac{e^{-p\mu(Z)-q(\mu(G)-\mu(Z))} [p\mu(Z) + q(\mu(G) - \mu(Z))]^{n_G}}{n_G!} \times \\
&\times \prod_{x_i \in Z} \frac{p\mu(x_i)}{p\mu(Z) + q(\mu(G) - \mu(Z))} \prod_{x_i \notin Z} \frac{q\mu(x_i)}{p\mu(Z) + q(\mu(G) - \mu(Z))} = \\
&= \frac{e^{-p\mu(Z)-q(\mu(G)-\mu(Z))}}{n_G!} p^{n_Z} q^{(n_G-n_Z)} \prod_{x_i} \mu(x_i).
\end{aligned}
\tag{3.3}$$

Όμοια με την εξίσωση (3.2) ορίζεται ο λόγος πιθανοφάνειας με

$$L_0 = \sup_p \frac{e^{-p\mu(G)} p^{n_G}}{n_G!} \prod_{x_i} \mu(x_i) = \frac{e^{-n_G}}{n_G!} \left( \frac{n_G}{\mu(G)} \right)^{n_G} \prod_{x_i} \mu(x_i)$$

αφού  $p\mu(G) = n_G$  και  $p = \frac{n_G}{\mu(G)}$ .

Για τον αριθμητή, υπολογίζεται πρώτα το *supremum* για όλα τα  $p$  και  $q$  με  $Z$  σταθερό. Η εξίσωση πιθανοφάνειας (3.3) λαμβάνει μέγιστο όταν  $p = n_Z / \mu(Z)$  και  $q = (n_G - n_Z) / (\mu(G) - \mu(Z))$ , και επομένως:

$$L(Z) = \begin{cases} \frac{e^{-n_G} \left( \frac{n_Z}{\mu(Z)} \right)^{n_Z} \left( \frac{n_G - n_Z}{\mu(G) - \mu(Z)} \right)^{n_G - n_Z} \prod_{x_i} \mu(x_i), & \text{if } \frac{n_Z}{\mu(Z)} > \frac{n_G - n_Z}{\mu(G) - \mu(Z)} \\ \frac{e^{-n_G} \left( \frac{n_G}{\mu(G)} \right)^{n_G} \prod_{x_i} \mu(x_i), & \text{διαφορετικά.} \end{cases}$$

Η στατιστική συνάρτηση ελέγχου  $\lambda$  του λόγου πιθανοφάνειας μπορεί τώρα να γραφεί ως

$$\begin{aligned} \lambda &= \frac{\sup_{Z \in \mathfrak{Z}} L(Z)}{\frac{e^{-n_G} \left( \frac{n_G}{\mu(G)} \right)^{n_G} \prod_{x_i} \mu(x_i)}{\left( \frac{n_G}{\mu(G)} \right)^{n_G}}} = \sup_{Z \in \mathfrak{Z}} \frac{\left( \frac{n_Z}{\mu(Z)} \right)^{n_Z} \left( \frac{n_G - n_Z}{\mu(G) - \mu(Z)} \right)^{n_G - n_Z}}{\left( \frac{n_G}{\mu(G)} \right)^{n_G}} \mathbb{I} \left( \frac{n_Z}{\mu(Z)} > \frac{n_G - n_Z}{\mu(G) - \mu(Z)} \right) \end{aligned}$$

αν υπάρχει τουλάχιστον μια ζώνη  $Z$  τέτοια ώστε  $\frac{n_Z}{\mu(Z)} > \frac{n_G - n_Z}{\mu(G) - \mu(Z)}$  ή

ως  $\lambda = 1$ , διαφορετικά.

Πράγματι,

$$\lambda = \frac{\sup_{Z \in \mathfrak{Z}} L(Z)}{\frac{e^{-n_G} \left( \frac{n_G}{\mu(G)} \right)^{n_G} \prod_{x_i} \mu(x_i)}{\left( \frac{n_G}{\mu(G)} \right)^{n_G}}} = \frac{\frac{e^{-n_G} \left( \frac{n_G}{\mu(G)} \right)^{n_G} \prod_{x_i} \mu(x_i)}{\left( \frac{n_G}{\mu(G)} \right)^{n_G}}}{\frac{e^{-n_G} \left( \frac{n_G}{\mu(G)} \right)^{n_G} \prod_{x_i} \mu(x_i)}{\left( \frac{n_G}{\mu(G)} \right)^{n_G}}} = 1$$

Η κατανομή της  $\lambda$  εξαρτάται από την κατανομή του υποκείμενου πληθυσμού και γενικά δεν έχει κάποια απλή αναλυτική μορφή. Αν ο συνολικός πληθυσμός είναι πολύ μικρός, τότε είναι πιθανό να βρεθεί η ακριβής κατανομή, απαριθμώντας όλα τα δυνατά αποτελέσματα στα άτομα του πληθυσμού και υπολογίζοντας τη τιμή του ελέγχου για κάθε αποτέλεσμα.

Για μεγάλα σύνολα δεδομένων, είναι πρακτικά αδύνατο να γίνει ένας τέτοιος υπολογισμός και για το λόγο αυτό χρησιμοποιείται η μέθοδος Monte Carlo για την προσομοίωση της ακριβούς κατανομής του  $\lambda$ . Η ιδέα του ελέγχου σημαντικότητας που να βασίζεται στην τυχαιοποιημένη κατανομή ενός ελέγχου σημαντικότητας, οφείλεται στον Fisher (1935). Η χρήση της Monte Carlo μεθόδου για δειγματοληψία από μια τυχαιοποιημένη κατανομή με σκοπό τη διεξαγωγή ενός ελέγχου σημαντικότητας, προτάθηκε από τον Dwass (1957). Για πρώτη φορά εφαρμόστηκε σε μελέτες χωρικών σημείων από τους Besag και Diggle (1977).

### **3.4 ΣΥΝΟΨΗ**

Στο Κεφάλαιο 3 παρουσιάστηκαν δύο από τα βασικότερα εργαλεία που χρησιμοποιούνται για την ανίχνευση χωρικών συστάδων τόσο σε προβλήματα μίας διάστασης, αλλά και σε διδιάστατα. Στην πρώτη μέθοδο που αναφέρθηκε χρησιμοποιούνται πιθανοθεωρητικά μοντέλα στοχεύοντας σε μια πιο θεωρητική προσέγγιση, ενώ στην τελευταία παρουσιάζεται μια πιο πρακτική προσέγγιση της μεθόδου ικανή να χρησιμοποιηθεί για αναλύσεις χωρικών ή/και χρονικών δεδομένων από διάφορους ερευνητές.

# Κεφάλαιο 4

## ΕΦΑΡΜΟΓΕΣ ΣΤΑΤΙΣΤΙΚΩΝ ΣΥΝΑΡΤΗΣΕΩΝ ΣΑΡΩΣΗΣ

### 4.1 ΕΙΣΑΓΩΓΗ

Έχοντας αναλύσει σε προηγούμενο κεφάλαιο τη δομή και τα χαρακτηριστικά των δύο προσεγγίσεων, δηλαδή της πιθανοθεωρητικής και της χωρικής συνάρτησης σάρωσης, κρίνεται αναγκαία η αναφορά στους τομείς που τέτοιες συναρτήσεις μπορούν και χρειάζεται να εφαρμοστούν, με σκοπό τη διεξαγωγή έγκυρων συμπερασμάτων. Όπως προαναφέρθηκε, κύριος σκοπός μιας στατιστικής συνάρτησης σάρωσης είναι ο εντοπισμός και η εκτίμηση της στατιστικής σημαντικότητας μιας χωρικής συστάδας δεδομένων που έχει βρεθεί και δεν μπορεί να εξηγηθεί από το μοντέλο πιθανότητας το οποίο ορίζεται από την μηδενική υπόθεση χωρικής τυχειότητας.

Οι στατιστικές συναρτήσεις σάρωσης δύο και περισσότερων διαστάσεων όπως εισήχθησαν από τον Naus (1965), έχουν εφαρμοστεί σε διάφορα πεδία, όπως η καρδιολογία, η επιδημιολογία, η ρευματολογία, οι

νευρολογικές ασθένειες, οι αυτοάνοσες ασθένειες, οι ασθένειες ήπατος, ο διαβήτης, οι καταχρήσεις, η παρασιτολογία, η κτηνιατρική, τα ατυχήματα και οι αυτοκτονίες, η ψυχολογία, καθώς επίσης και στην εγκληματολογία, στη γεωλογία και δασολογία, στην αστρονομία, στη δημογραφία και τις ανθρωπιστικές επιστήμες.

Τόσο η προσέγγιση του Kulldorff (1997) σύμφωνα με την οποία γίνεται χρήση ενός παραθύρου σάρωσης συγκεκριμένου μεγέθους και σχήματος, όσο και η πιθανοθεωρητική προσέγγιση του Glaz (1996) εφαρμόζονται αποτελεσματικά στα παραπάνω όπως επίσης και σε μια πληθώρα κλάδων ανάλογα με το πρόβλημα που μελετάται.

Να αναφερθεί επίσης ότι, η χωρική συνάρτηση σάρωσης μπορεί να χαρακτηριστεί ως ένα από τα σημαντικότερα εργαλεία για την ανάλυση γεωγραφικών και χωρικών δεδομένων. Άλλες σημαντικές μέθοδοι περιλαμβάνουν περιγραφική στατιστική αναλογιών και ποσοστών, χωρικές μεθόδους εξομάλυνσης, παλινδρόμηση για χωρικός συσχετισμένα δεδομένα, τεχνικές οπτικοποίησης (*visualization techniques*) και άλλα χαρακτηριστικά (Costa and Kulldorff, 2009).

Η χωρική μέθοδος για δύο διαστάσεις, όπως και η χώρο-χρονική (*space-time scan statistics*) για τρεις διαστάσεις, μπορεί να πάρει διάφορες μορφές, ανάλογα με το πρόβλημα που εξετάζεται, γεγονός που καθιστά κάθε παραλλαγή της μεθόδου πολύ χρήσιμη για διαφορετικού τύπου εφαρμογές (Ντζούφρας, 2006).

## **4.2 ΕΦΑΡΜΟΓΕΣ ΣΤΗΝ ΕΠΙΔΗΜΙΟΛΟΓΙΑ**

Ως Επιδημιολογία ορίζεται η επιστήμη που μελετάει την κατανομή και την εξέλιξη διαφόρων νοσημάτων ή χαρακτηριστικών στον ανθρώπινο

πληθυσμό και των παραγόντων που τον διαμορφώνουν ή τον επηρεάζουν. Επίσης εξετάζει χαρακτηριστικά που συνδέονται τόσο με τον περιβαλλοντικό όσο και με το γενετικό έλεγχο. Αντικείμενο της Επιδημιολογίας είναι και η περιγραφική κατανομή των παραγόντων που επηρεάζουν τον πληθυσμό όπως για παράδειγμα ηλικία, φύλο, φυλή, γεωγραφική περιοχή και επάγγελμα.

Στην Επιδημιολογία κάθε νόσος περιγράφεται κυρίως από τη συχνότητα εμφάνισης, το χρόνο διάρκειας και την εξέλιξή της. Μέσα σε μελέτες που αφορούν επιδημιολογικές νόσους, ιδιαίτερα διευρυμένη είναι η χρήση στατιστικών συναρτήσεων σάρωσης. Αναλυτικά περιγράφονται παρακάτω εφαρμογές τους, τόσο σε μολυσματικές ασθένειες, όσο και στην επιδημιολογία του καρκίνου.

#### **4.2.1 Εφαρμογές στην Επιδημιολογία του Καρκίνου**

Ο δείκτης θνησιμότητας, το ποσοστό επιπολασμού ή η συχνότητα εμφάνισης του καρκίνου, μπορεί να διαφέρουν ανά γεωγραφική περιοχή για ποικίλους λόγους, όπως χωρική διαφοροποίηση περιβαλλοντικών παραγόντων κινδύνου ή γενετική σύνθεση του πληθυσμού. Οι συστάδες εμφάνισης καρκίνου εμφανίζονται τυχαία σε διάφορα σημεία του χάρτη και για το λόγο αυτό οι χωρικές συναντήσεις σάρωσης χρησιμοποιούνται για τον εντοπισμό και την εκτίμηση τέτοιων στατιστικά σημαντικών συστάδων. Η λευχαιμία ήταν η πρώτη μορφή καρκίνου που μελετήθηκε με αυτή τη μέθοδο, από τον Turnbull et al.(1990) στην άνω Νέα Υόρκη, και τους Hjalmarsson et al.(1996) στη Σουηδία και συγκεκριμένα

παρατηρήθηκε η παιδική λευχαιμία στην περιοχή. Η μελέτη των Hjalmar et al. δεν έδωσε ένδειξη στατιστικά σημαντικής συστάδας στα δεδομένα τους, αν και υπήρξε ένα προειδοποιητικό σήμα εμφάνισης μιας συστάδας λευχαιμίας μερικά χρόνια πριν. Παρόλο που ανιχνεύθηκε η συστάδα, δεν κρίθηκε στατιστικά σημαντική και δεν συγκαταλέχθηκε ούτε στις τρεις κυριότερες συστάδες. Από την άλλη πλευρά, οι Viel et al. (2000) εντόπισαν στατιστικά σημαντική συστάδα σαρκώματος (*soft-tissue sarcoma*) και λεμφώματος non-Hodgkins (*non-Hodgkin's lymphoma*) γύρω από έναν αποτεφρωτήρα στερεών αποβλήτων, μια περιοχή με ιδιαίτερα αυξημένα ποσοστά εκπομπής διοξινών, στη Γαλλία.

Υπό την μηδενική υπόθεση, ο δείκτης θνησιμότητας λόγω καρκίνου, θεωρείται ότι ακολουθεί την κατανομή Poisson, με τον αναμενόμενο αριθμό συμβάντων σε μια συγκεκριμένη τοποθεσία να είναι ανάλογος με τον προσαρμοσμένο σε συμμεταβλητές πληθυσμό, της συγκεκριμένης τοποθεσίας. Συνήθως η ηλικία και άλλες συμμεταβλητές προσαρμόζονται χρησιμοποιώντας κανονικοποίηση (*standardization*). Έστω  $b(i,k)$ , ο πληθυσμός στην ηλικιακή ομάδα  $k$  στην περιοχή  $i$ . Έστω επίσης  $B(k)$  ο συνολικός πληθυσμός στην ηλικιακή ομάδα  $k$  και  $C(k)$  ο συνολικός αριθμός των συμβάντων στην ηλικιακή ομάδα  $k$ . Ο αναμενόμενος αριθμός συμβάντων, προσαρμοσμένος στην ηλικία στην τοποθεσία  $i$ , είναι

$$\mu_i = \sum_k b_{i,k} \times C_k / B_k .$$

Η ηλικία είναι απαραίτητο να λαμβάνεται πάντοτε υπόψη σε μελέτες που σχετίζονται με την καρκινική θνησιμότητα ή το δείκτη εμφάνισής του διότι διαφορετικά, θα εντοπίζονται στατιστικά σημαντικές συστάδες σε περιοχές που κατοικούνται από άτομα προχωρημένης ηλικίας και αυτό γιατί, οι γηραιότεροι είναι εκτεθειμένοι σε υψηλότερο κίνδυνο να διαγνωστούν με τη νόσο. Ιδιαίτερο ενδιαφέρον έχει επίσης η προσαρμογή του μοντέλου σε άλλους παράγοντες κινδύνου, όπως η εθνικότητα, το μορφωτικό επίπεδο ή οι βιολογικές μεταβλητές όπως το χρώμα του δέρματος, για μελέτες σχετικά με τον καρκίνο του δέρματος ή ο αριθμός των γεννήσεων που έχει κάνει μια γυναίκα, για μελέτες που σχετίζονται με την εμφάνιση καρκίνου του μαστού (Kulldorff *et al.*, 1997, Hsu *et al.*, 2004).

Με τη χωρική συνάρτηση σάρωσης ανιχνεύεται και εκτιμάται η στατιστική σημαντικότητα των μεμονωμένων συστάδων, αλλά δεν παρέχεται κάποια εκτίμηση του ποσοστού θνησιμότητας για όλη την περιοχή μελέτης. Για το λόγο αυτό, έχουν προταθεί κάποιες συμπληρωματικές στατιστικές μέθοδοι από τους Thomas and Carlin (2003) και Buntix *et al.* (2003).

Στις περισσότερες περιπτώσεις καρκίνου, ενδέχεται να περάσει αρκετό διάστημα από τη στιγμή της προσβολής μέχρι τη διάγνωση και ακόμα περισσότερο από τη στιγμή προσβολής μέχρι το θάνατο. Οι Han *et al.* (2004) παρουσίασαν μια ενδιαφέρουσα μελέτη για τον καρκίνο του μαστού και την ύπαρξη συστάδων, λαμβάνοντας υπόψη τον τόπο κατοικίας της γυναίκας (ως χωρικές συντεταγμένες) τη στιγμή της γεννήσεώς της, τη στιγμή ενάρξεως της εμμήνου ρήσεώς της και τη στιγμή γεννήσεως του πρώτου της παιδιού. Με αυτό τον τρόπο, η



μέθοδος παρείχε τη δυνατότητα να εξεταστεί η γεωγραφική συσταδοποίηση του καρκίνου του μαστού σε διάφορες χρονικές περιόδους της ζωής της γυναίκας. Να σημειωθεί ότι υπήρχαν πολύ λιγότερες συστάδες που εντοπίστηκαν όταν τα δεδομένα αναλύθηκαν χρησιμοποιώντας χωρικές συντεταγμένες από τον τόπο κατοικίας μόνο τη στιγμή της διάγνωσης της νόσου.

Οι στατιστικές συναρτήσεις σάρωσης δύο διαστάσεων, μπορούν επίσης να χρησιμοποιηθούν για τη μελέτη της γεωγραφικής κατανομής ενός συγκεκριμένου τύπου καρκίνου, ούτως ώστε να προσδιοριστεί αν υπάρχουν γεωγραφικές συστάδες σε προχωρημένο στάδιο καρκίνου ή καρκίνου συγκεκριμένου τύπου ή βαθμού. Η ανίχνευση, για παράδειγμα, γεωγραφικής συστάδας με υψηλό ποσοστό εμφάνισης καρκίνου του μαστού προχωρημένου σταδίου, ίσως παρέχει ένδειξη ότι είναι απαραίτητη η βελτίωση της διαδικασίας της μαστογραφίας σε εκείνη την περιοχή. Ως πληθυσμός ορίζεται ο συνολικός αριθμός των περιπτώσεων που διαγνώστηκαν με καρκίνο, ενώ ως περιπτώσεις ορίζονται αυτές που εμφάνισαν καρκίνο συγκεκριμένου τύπου.

Ενδιαφέρον επίσης μπορεί να αποτελέσει η μελέτη της γεωγραφικής διασποράς του χρόνου επιβίωσης από τη στιγμή της διάγνωσης και μετά, για να προσδιοριστεί εάν υπάρχει κάποια γεωγραφική περιοχή με χαμηλό ποσοστό επιβίωσης. Τέτοιες αναλύσεις απαιτούν χειρισμό λογοκρινμένων (*censored*) δεδομένων και συνεχείς τροποποιήσεις βάσει διαφορών στους προγνωστικούς παράγοντες, όπως η ηλικία του ασθενή, το στάδιο του καρκίνου και άλλων.

## 4.2.2 Εφαρμογές στην Επιδημιολογία Μολυσματικών Νόσων

Στην παρακολούθηση μολυσματικών ασθενειών, οι συναρτήσεις σάρωσης δύο και τριών διαστάσεων χρησιμοποιούνται για δύο διαφορετικούς σκοπούς. Οι πρώτες λειτουργούν αναδρομικά, αφού μελετώνται ιστορικά δεδομένα για την ανίχνευση γεωγραφικών περιοχών με αυξημένο τον αριθμό συμβάντων εμφάνισης της νόσου. Τέτοιες συστάδες είναι δυνατό να είναι είτε προσωρινές, λόγω κάποιας διαταραχής, είτε μακροχρόνιες αν η περιοχή ή ο πληθυσμός εκεί, είναι ιδιαίτερα ευάλωτος σε μολύνσεις.

Οι Cousens *et al.* (2001) παρατήρησαν χωρικά 84 περιπτώσεις που εμφάνιζαν τη νόσο variant Creutzfeldt-Jacobs (vCJD) μια σπάνια και θανατηφόρα ασθένεια από τον ίδιο μεταδοτικό παράγοντα που σχετίζεται με τη σπογγώδη εγκεφαλοπάθεια των βοοειδών (Νόσος των Τρελών Αγελάδων) και επομένως υποθετικά συσχετίστηκε με την κατανάλωση βοδινών προϊόντων. Με τη χωρική συνάρτηση σάρωσης εντοπίστηκε μια στατιστικά σημαντική συστάδα με πέντε περιστατικά. Περαιτέρω έρευνα οδήγησε στην αποκάλυψη ενός τοπικού κρεοπωλείου ως την πιθανότερη αιτία μόλυνσης.

Οι Fevre *et al.* (2001) χρησιμοποίησαν χωρική συνάρτηση σάρωσης για τη μελέτη της ασθένειας του ύπνου (*sleeping sickness*) στην Ουγκάντα. Η ασθένεια του ύπνου προκαλείται από ένα παράσιτο που μεταδίδεται στον άνθρωπο από τη μύγα τσε-τσε, η οποία λαμβάνει τη μόλυνση από εγχώριες κτηνοτροφικές μονάδες. Εφαρμόστηκε μια χωρική ανάλυση χρησιμοποιώντας τον αριθμό των περιστατικών που είχαν

διαγνωστεί μέσα σε 32 ημέρες, από τη στιγμή της πρώτης καταγραφής γεγονότος, μέχρι τη στιγμή που τέθηκαν σε ισχύ μέτρα για τον έλεγχο των φορέων και είχε τη μορφή μελέτης ασθενών-μαρτύρων (*case-control study*). Κάθε ασθενής αντιστοιχίστηκε με έναν υγιή βάσει της ηλικίας, του φύλου και του μήνα καταμέτρησης και επομένως η ανάλυση έγινε υποθέτοντας ότι οι δύο κατηγορίες δεδομένων ακολουθούσαν κατανομή Bernoulli. Εντοπίστηκε μια στατιστικά σημαντική συστάδα, γύρω από μια περιφερειακή αγορά βοοειδών.

Η λιστερίωση (*listeriosis*) είναι ένα βακτηριακό τροφιμογενές παθογόνο, το οποίο μπορεί να εμφανιστεί στο 1 με 5% των τροφίμων που είναι έτοιμα για κατανάλωση και μπορεί να προκαλέσει μια σπάνια και σοβαρή επεμβατική εκδήλωση ασθένειας, καθώς και το θάνατο σε μερικούς ανθρώπους. Πιο συγκεκριμένα, αφού η εξάπλωση του βακτηρίου, είναι άμεσα συνδεδεμένη με τα μολυσμένα τρόφιμα, η πηγή έκθεσης και μόλυνσης πιθανώς να προέρχεται είτε από τη παγκόσμια διανομή φαγητού, είτε από τοπικές πηγές. Συνεπώς, μια χωροχρονική συσταδοποίηση ίσως να εντοπίζει μικρές ή μεγάλες συστάδες. Η ανάλυση συστάδων διενεργήθηκε χρησιμοποιώντας διαφορετικές υποκατηγορίες γενετικού υλικού (*ribotyping*) από μη-μολυσμένες τοποθεσίες. Συστάδες με την ίδια μοριακή υποκατηγορία, είναι πιθανό να αποτελούν συστάδες με κοινή πηγή έκθεσης, αυξάνοντας την ικανότητα ανίχνευσης εστιών μόλυνσης (Sauders and Fortes, 2003).

Οι Wylie, Carbal και Jolly (2005) μελέτησαν σεξουαλικά μεταδιδόμενα νοσήματα χρησιμοποιώντας τη χωρική στατιστική συνάρτηση σάρωσης, διαχωρίζοντας τα περιστατικά βάσει του γονότυπου. Σύμφωνα με τους συγγραφείς, η υπόθεση που εξετάζεται

είναι ότι δυο άτομα που μολύνθηκαν από το ίδιο στέλεχος του μολυσματικού παράγοντα είναι πιο πιθανό να εμφανίσουν έναν επιδημιολογικό σύνδεσμο μεταξύ τους από ότι δύο άτομα που μολύνθηκαν από διαφορετικό στέλεχος.

Ο δεύτερος λόγος χρήσης στατιστικών σάρωσης για μολυσματικές νόσους έχει προοδευτική (*prospective*) μορφή, αφού συλλέγονται συνεχώς δεδομένα κι αναλύονται σε πραγματικό χρόνο με σκοπό τη γρήγορη ανίχνευση μιας επικίνδυνης ασθένειας.

Οι Mostashari et al. (2003) πρότειναν ένα σύστημα επίβλεψης για τη νόσο West Nile, σύμφωνα με το οποίο καταγράφονταν σε καθημερινή βάση ο αριθμός νεκρών πτηνών που εντοπιζόνταν. Παρατηρήθηκε ότι η πυκνότητα και ο αριθμός των κορακιών, σχετίζονταν άμεσα με τα επίπεδα της νόσου το 2000, γεγονός που οδήγησε στο συμπέρασμα ότι η επίβλεψη τέτοιων συμβάντων θα μπορούσε εύκολα να ανιχνεύσει τα επόμενα κρούσματα. Πολλαπλές αναφορές νεκρών πτηνών στην ίδια περιοχή και την ίδια ημέρα καταγράφονταν μια φορά, ενώ τα αποτελέσματα έδειξαν ότι στις περισσότερες περιπτώσεις οι συστάδες των νεκρών πτηνών που εντοπίστηκαν, όχι μόνο απέτρεψαν τη σπατάλη χρόνου στη συλλογή κουνουπιών και πτηνών που βρέθηκαν θετικά στον ιό West Nile, αλλά και την αναφορά μόλυνσης ανθρώπων κοντά στην περιοχή της συστάδας.

Οι χωροχρονικές συναντήσεις σάρωσης χρησιμοποιούνται επίσης σε περιπτώσεις συνδρομικής επιτήρησης (*syndromic surveillance*), κατά την οποία συλλέγονται καθημερινές καταγραφές ενός αυτοματοποιημένου συστήματος υγείας, με σκοπό την έγκαιρη ανίχνευση εξάρσεων μολυσματικών ασθενειών (Kulldorff et al., 2005).

### 4.3 ΕΦΑΡΜΟΓΕΣ ΣΤΗΝ ΠΑΡΑΣΙΤΟΛΟΓΙΑ

Πολλές έρευνες έχουν γίνει και στον τομέα της παρασιτολογίας, χρησιμοποιώντας χωρικές στατιστικές συναρτήσεις. Οι Enemark et al. (2002) παρουσίασαν μια ανάλυση συστάδων για το *Cryptosporidium parvum*, ένα παράσιτο το οποίο μολύνει το γαστρεντερικό σωλήνα και αναφέρεται ως η κύρια αιτία της διάρροιας. Οι Odoi et al. (2004) εφάρμοσαν τη χωρική συνάρτηση σάρωσης στη μελέτη της γιαρδίασης (*giardiasis*) στον Καναδά. Πρόκειται για μια παρασιτική ασθένεια που προκαλείται από το πρωτόζωο *Giardia intestinalis*, το οποίο προσβάλλει το πεπτικό σύστημα οικόσιτων και άγριων ζώων, καθώς και των ανθρώπων. Αποτελεί την κυριότερη αιτία γαστρεντερίτιδας στον άνθρωπο, μολύνοντας περίπου 200 εκατομμύρια παγκοσμίως. Τέλος, οι Reperant και Deplazes (2005) χρησιμοποίησαν τη στατιστική σάρωσης για τη μελέτη του νηματοειδούς πρωτόζωου *Capillaria hepatica* που ευθύνεται για την λοίμωξη *Hepatic Capellariasis* πολλών θηλαστικών, στη Σουηδία.

### 4.4 ΑΛΛΕΣ ΙΑΤΡΙΚΕΣ ΕΦΑΡΜΟΓΕΣ

Η υποπλαστική αριστερή δυσπλασία στην καρδιά (*hypoplastic left heart malformation*) είναι μια καρδιαγγειακή δυσπλασία που εμφανίζεται εκ γενετής. Η έκθεση του γονέα σε διάφορους διαλύτες, είναι άμεσα συσχετισμένη με την εμφάνιση της νόσου σε νεογέννητα. Οι Kuehl και Loffredo (2006) χρησιμοποίησαν στατιστικές συναρτήσεις σάρωσης για την εύρεση συστάδων ασθενών και αποδεικτικά στοιχεία που ίσως

συσχέτιζαν τη νόσο με την απελευθέρωση διαλυτών από τη βιομηχανική περιοχή στην Μέρυλαντ, στις Ηνωμένες Πολιτείες. Μετά τον εντοπισμό γεωγραφικών συστάδων τα αποτελέσματα προσαρμόστηκαν σε διάφορα μοντέλα λογιστικής παλινδρόμησης στρωματοποιημένων κατά τον τόπο διαμονής των γονέων μέσα ή έξω από την συστάδα, τη στιγμή της σύλληψης.

Πολλές μελέτες έχουν γίνει επίσης χρησιμοποιώντας στατιστικές συναρτήσεις σάρωσης σε παιδιατρικά δεδομένα. Οι Sankoh *et al.* (2001) ανέλυσαν την παιδική θνησιμότητα στη βορειοδυτική περιοχή της Μπουργκίνα Φάσο στην Αφρική κατά την περίοδο 1993-1998. Εφαρμόστηκε χωρική ανάλυση για κάθε ένα έτος και έτσι συγκεντρώθηκαν χρονικά ανεξάρτητες μελέτες που ανίχνευαν συστάδες για συγκεκριμένα χρόνια. Τα αποτελέσματα έδειξαν ότι ένα συγκεκριμένο χωριό αποτέλεσε στατιστικά σημαντική συστάδα τόσο χωρικά όσο και χωρο-χρονικά. Απομονώνοντας αυτό το χωριό από το δείγμα, διενεργήθηκε νέα ανάλυση, η οποία αναγνώρισε τη δεύτερη πιο σημαντική συστάδα, ως την πιο πιθανή. Η εξαίρεση δεδομένων, είναι μια τεχνική ιδιαίτερα χρήσιμη αν οι μελετητές επιθυμούν να συγκεντρωθούν σε χωρικές συστάδες, πέραν της πιο στατιστικά σημαντικής.

Οι Sabel *et al.* (2003) ανίχνευσαν και εκτίμησαν γεωγραφικές συστάδες αμυοτροφικής πλευρικής σκλήρυνσης (*amyotrophic biliary cirrhosis*) στη Φινλανδία. Διεξήχθησαν επιπλέον ξεχωριστές αναλύσεις, χρησιμοποιώντας ως συντεταγμένες τον τόπο γεννήσεως και τον τόπο θανάτου και η συστάδα που βρέθηκε με τη χρήση του τόπου γεννήσεως επικαλύφθηκε από την αντίστοιχη στατιστικά σημαντική συστάδα που βρέθηκε κάνοντας χρήση του τόπου θανάτου.

Οι Ala et al. (2006) χρησιμοποίησαν επίσης στατιστική σάρωσης, με σκοπό να δείξουν ότι ο επιπολασμός των ασθενών που διαγνώστηκαν με πρωτοπαθή χολική κίρρωση και αναμένουν μόσχευμα, ήταν υψηλότερος κοντά σε μια χωματερή τοξικών αποβλήτων στη Νέα Υόρκη.

Άλλες εφαρμογές των χωρικών συναρτήσεων σάρωσης συναντώνται μεταξύ άλλων, σε μελέτες συστημικής σκλήρυνσης (*systemic sclerosis*) στις Ηνωμένες Πολιτείες (Walsh and DeChello, 2001), διαβήτη στον Καναδά (Green et al., 2003), πολλαπλής σκλήρυνσης στη Σκωτία (Donnan et al., 2005) και άσθματος στις Ηνωμένες Πολιτείες (Cook, Gold and Li, 2007). Χρησιμοποιήθηκαν επίσης από τους Hanson και Wiczorek (2002) για τη μελέτη της γεωγραφικής κατανομής των καταχρήσεων αλκοόλ και ναρκωτικών, καθώς και της έκθεσης σε φυτοφάρμακα από τους Sudakin, Horowitz και Griffin (2002).

## 4.5 ΕΦΑΡΜΟΓΕΣ ΣΤΗΝ ΚΤΗΝΙΑΤΡΙΚΗ

Στην κτηνιατρική επιστήμη, οι χωρικές συναρτήσεις σάρωσης βρίσκουν εφαρμογή τόσο στη μελέτη οικόσιτων ζώων, όσο και άγριων. Πολλές μελέτες που βασίζονται στην παρακολούθηση οικόσιτων ζώων έχουν διενεργηθεί και ενδεικτικά αναφέρονται μελέτες βοοειδών (Nordstrom, Pfeiffer and Jarp, 2000), προβάτων (Ward, 2001), πουλερικών (Guerin et al., 2005) και σκύλων (Ward, 2002). Ιδιαίτερα διαδεδομένη είναι η χρήση των εν λόγω συναρτήσεων και σε επιδημιολογικές έρευνες της σπογγώδους εγκεφαλοπάθειας των βοδινών (Νόσος των Τρελών Αγελάδων) στη Γαλλία (Abrial et al., 2003), στην Ιρλανδία (Sheridan et al., 2005) και την Ολλανδία (Heres et al., 2008).

Όσον αφορά δεδομένα που προέρχονται από παρατήρηση άγριων ζώων, οι συναρτήσεις σάρωσης χρησιμοποιούνται για τη μελέτη διάφορων ασθενειών που εντοπίστηκαν σε αλεπούδες στη Γερμανία (Berke et al., 2002), θαλάσσιες ενυδρίδες στην Καλιφόρνια (Miller et al., 2002) ή ασβούς στην Ιρλανδία (Olea-Poppelka et al., 2003). Ένα σύνηθες πρόβλημα που συναντάται στη μελέτη τέτοιων περιπτώσεων, είναι η μη σταθερή διαμονή των ζώων που παρακολουθούνται και η ικανότητά τους να διανύουν πολλά χιλιόμετρα πριν συλληθούν (Hoar et al., 2003). Οι Miller et al. (2002) προσπάθησαν να εντοπίσουν χωρικές συστάδες παρασίτων στις θαλάσσιες ενυδρίδες αλλά πιθανόν λόγω της υψηλής κινητικότητας του είδους, η ανάλυση δεν ανίχνευσε κάποια στατιστικά σημαντική συστάδα. Μια ενδεχόμενη εναλλακτική για την αντιμετώπιση του προβλήματος, είναι η συλλογή στατικών πηγών μελέτης, όπως πτώματα ζώων.

## **4.6 ΕΦΑΡΜΟΓΕΣ ΣΤΗΝ ΕΝΤΟΜΟΛΟΓΙΑ ΚΑΙ ΣΤΗΝ ΙΧΘΥΟΛΟΓΙΑ**

Χρήσεις των συναρτήσεων σάρωσης έχουν αναφερθεί και σε μελέτες εντομολογίας και ιχθυολογίας. Οι Porcasi et al. (2006) μελέτησαν την προσβολή των αγροτικών κατοικιών από το έντομο *Triatona infestans*, τον κυριότερο φορέα του παράσιτου *Trypanosoma cruzi*, σε μια περιοχή στη νότια Αργεντινή. Στη μελέτη τους εντόπισαν στατιστικά σημαντικές συστάδες υψηλής προσβολής, στο νοτιοανατολικό τμήμα της περιοχής και γενικότερα, μέσω της επιλογής από την εφαρμογή πολλαπλών μοντέλων, κατέληξαν ότι συστάδες εμφανίζονται σε περιοχές με χαμηλό



υψόμετρο, λίγα αγροκτήματα, υψηλή μέση θερμοκρασία της επιφάνειας της γης και υψηλό ποσοστό άγονων εκτάσεων.

Τέλος, οι Spindler et al. (2007) χρησιμοποίησαν συναρτήσεις σάρωσης για τη χωρική ανάλυση ενός είδους ψαριού, του ωχρού οξύρυγχου *Scaphirhynchus albus*, στον ποταμό Μισούρι, στη Βόρεια Ντακότα, στις Ηνωμένες Πολιτείες Αμερικής.

## **4.7 ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΔΑΣΟΛΟΓΙΑ- ΒΟΤΑΝΟΛΟΓΙΑ**

Οι Tuia et al. (2008) χρησιμοποίησαν τη στατιστική συνάρτηση σάρωσης για να εντοπίσουν και να εκτιμήσουν χωροχρονικές συστάδες πυρκαγιών στα δάση. Συμπέραναν ότι «η εκτίμηση της παρουσίας χωρικών και χρονικών σχημάτων σε περιπτώσεις πυρκαγιών, καθώς επίσης και η στατιστική τους σημαντικότητα, μπορεί να αποτελέσει σημαντικό αντίκτυπο σε προσεχείς μελέτες για την πρόβλεψη δασικών πυρκαγιών».

Οι Bayon et al. (2007) ασχολήθηκαν με τη γενετική δομή και τη χωρική κατανομή του μυκοπαράσιτου *Sphaerellopsis filum*, το οποίο ευθύνεται για την εμφάνιση σκουριάς και συγκεκριμένα της *Melampsora Larici-epitea* στο φύλλωμα των ιτιών.

## **4.8 ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΓΕΩΛΟΓΙΑ**

Οι Conover, Bement και Iman (1979) εφάρμοσαν τη χωρική συνάρτηση σάρωσης σε γεωλογικά δεδομένα, με σκοπό την εύρεση κοιτασμάτων ουρανίου, χρησιμοποιώντας μετρήσεις ακτινοβολίας που

είχαν ληφθεί από ένα αεροπλάνο. Δεδομένου ότι οι μετρήσεις περιείχαν ικανοποιητική ποσότητα θορύβου, ο σκοπός ήταν η ανίχνευση συστάδων με υψηλές καταγραφές ακτινοβολίας.

## **4.9 ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΨΥΧΟΛΟΓΙΑ**

Οι Margai και Henry (2003) χρησιμοποίησαν σε μελέτη τους τη χωρική στατιστική συνάρτηση σάρωσης, με σκοπό να ανιχνεύσουν γεωγραφικές συστάδες με υψηλό επιπολασμό μαθησιακών δυσκολιών σε παιδιά στο Binghamton στη Νέα Υόρκη, όπου και εντόπισαν μια στατιστικά σημαντική συστάδα στο βορειοανατολικό τμήμα της πόλης. Συμπληρωματικά στη χωρική συνάρτηση σάρωσης, χρησιμοποίησαν το δείκτη χωρικής αυτοσυσχέτισης I (*Moran's I*) για να εκτιμήσουν την ύπαρξη καθολικής συσταδοποίησης σε όλη την πόλη. Μελέτησαν επίσης και ένα σύνολο κοινωνικο-οικονομικών μεταβλητών, που σχετίζονταν με τη χωρική εμφάνιση ατόμων με μαθησιακές δυσκολίες. Τα συμπεράσματα εξάχθηκαν μέσω ελέγχων *t*-test, συγκρίνοντας τις μέσες τιμές αυτών των μεταβλητών, μέσα και έξω από την συστάδα που ανιχνεύθηκε.

## **4.10 ΕΦΑΡΜΟΓΕΣ ΣΤΑ ΑΤΥΧΗΜΑΤΑ ΚΑΙ ΤΙΣ ΑΥΤΟΚΤΟΝΙΕΣ**

Οι Nkhoma et al. (2004) εφάρμοσαν χωρικές συναρτήσεις σάρωσης σε δεδομένα που σχετίζονταν με θάνατο λόγω δηλητηρίασης. Τα περιστατικά διαιρέθηκαν σύμφωνα με συγκεκριμένους τοξικούς παράγοντες και τόσο οι χωρικές όσο και οι χωροχρονικές συναρτήσεις

σάρωσης χρησιμοποιήθηκαν για την εκτίμηση των δεδομένων με παρουσία μιας χρονικής τάσης, και χωρίς. Οι Yiannakoulis *et al.* (2003) έκαναν χρήση χωρικών συναρτήσεων σάρωσης για τη μελέτη της γεωγραφικής κατανομής των τραυματισμών από πτώση ηλικιωμένων ανθρώπων.

Το 2007, οι Exter και Boyle μελέτησαν το φαινόμενο των αυτοκτονιών νέων ενηλίκων στη Σκωτία, σε τρεις διαφορετικές χρονικές περιόδους, το 1980-1982, 1990-1992 και 1999-2001. Εντόπισαν στατιστικά σημαντικές συστάδες στην ανατολική Γλασκώβη και τις τρεις περιόδους που μελέτησαν, περιλαμβάνοντας 92, 159 και 245 περιστατικά, αντίστοιχα. Σε σύγκριση με την υπόλοιπη χώρα, οι περισσότεροι θάνατοι σε αυτές τις συστάδες συνέβησαν από δηλητηρίαση, καθ' όλη τη διάρκεια της μελέτης, εκτός από την τελευταία περίοδο. Και οι τρεις συστάδες που βρέθηκαν, μπορούν να εξηγηθούν από τη συγκέντρωση πολλών κοινωνικο-οικονομικών στερήσεων σε εκείνο το μέρος της Σκωτίας.

## **4.11 ΕΦΑΡΜΟΓΕΣ ΣΤΙΣ ΑΝΘΡΩΠΙΣΤΙΚΕΣ ΕΠΙΣΤΗΜΕΣ**

Η χρήση των στατιστικών συναρτήσεων που μελετώνται δεν είναι ιδιαίτερα διαδεδομένη στον κλάδο των ανθρωπιστικών επιστημών. Ωστόσο, υπάρχουν ορισμένα παραδείγματα από την ανθρωπολογία, την αρχαιολογία και την ιστορία. Σε μια πολύ ενδιαφέρουσα ιστορική μελέτη, οι Witham και Orpenheimer (2004) χρησιμοποίησαν τις εν λόγω συναρτήσεις για να μελετήσουν τους επιπλέον θανάτους στην Αγγλία που οφείλονταν στην έκρηξη ενός ηφαιστείου στην Ισλανδία, το 1783.

Συμπέραναν ότι το ανατολικό τμήμα της Αγγλίας ήταν αυτό που επηρεάστηκε περισσότερο. Στην ανθρωπολογία, οι Usher και Allen (2005), εφάρμοσαν χωρική γενετική ανάλυση για την εκτίμηση ύπαρξης συστάδων συγγενειών στα νεκροταφεία. Ο Waller (2006) έκανε χρήση των χωρικών συναρτήσεων σάρωσης καθώς και πολλών ακόμα χωρικών στατιστικών τεχνικών, για να συγκρίνει τη γεωγραφική κατανομή των αρχαιολογικών μνημείων του πολιτισμού των Ανασάζι στην Αριζόνα, της πρώιμης περιόδου, έναντι της μεταγενέστερης.

## 4.12 ΣΥΝΟΨΗ

Στο κεφάλαιο αυτό παρουσιάστηκε ένα δείγμα των εφαρμογών στις οποίες μπορεί να γίνει χρήση των στατιστικών συναρτήσεων σάρωσης. Διαφορετικού τύπου δεδομένα απαιτούν διαφορετικές μορφές στατιστικών συναρτήσεων σάρωσης, όμως, η κύρια και βασική γραμμή που ακολουθείται είναι η ίδια που εισήχθησε από τον Naus (1965). Η μέθοδος χρησιμοποιείται συχνότερα σε μελέτες καρκίνου, μολυσματικών νόσων και κτηνιατρικής επιδημιολογίας και γενικότερα σε κλάδους επιδημιολογίας με μεγάλο ενδιαφέρον. Σημειώνεται επίσης ότι, η χρήση τέτοιων συναρτήσεων σε μελέτες νόσων πέραν των επιδημιολογικών, συνεχώς και αυξάνεται, καθώς επίσης και οι εφαρμογές τους σε μη-ιατρικές μελέτες, όπου παρόλο που είναι περιορισμένες προς το παρόν, εκφράζονται ελπίδες εξέλιξης στο μέλλον.

# Κεφάλαιο 5

## ΧΩΡΙΚΗ ΑΝΑΛΥΣΗ ΚΡΟΥΣΜΑΤΩΝ ΣΚΠ ΣΤΗΝ ΗΠΕΙΡΟ, ΤΗΝ ΣΤΕΡΕΑ ΕΛΛΑΔΑ ΚΑΙ ΤΗΝ ΠΕΛΟΠΟΝΝΗΣΟ

### 5.1 ΕΙΣΑΓΩΓΗ

Κύριος σκοπός του κεφαλαίου είναι η εφαρμογή της στατιστικής συνάρτησης σάρωσης για την ανάλυση δεδομένων της νόσου της σκλήρυνσης κατά πλάκας (ΣΚΠ). Αναλυτικότερα, στην παρούσα εργασία διερευνάται η εμφάνιση της ασθένειας στην Ελλάδα, και συγκεκριμένα στις περιοχές της Ηπείρου, Στερεάς Ελλάδας και Πελοποννήσου. Κίνητρο της έρευνας αποτέλεσε το γεγονός ότι έχουν παρατηρηθεί πολλά συμβάντα στις προαναφερθείσες περιοχές και με τη βοήθεια στατιστικών συναρτήσεων σάρωσης θα γίνει μια προσπάθεια εντοπισμού τυχόν συστάδων, όπως επίσης και να δοθεί απάντηση στο ερώτημα αν αυτές οι συστάδες εμφανίζονται με τυχαίο τρόπο ή οφείλονται σε κάποιους εξωγενείς παράγοντες.

## 5.2 ΜΕΘΟΔΟΛΟΓΙΑ ΑΝΑΛΥΣΗΣ

Ως πρώτο βήμα για τη μελέτη, απαιτείται η απόρριψη της τυχαίας γεωγραφικής κατανομής των συμβάντων ΣΚΠ στις υπό μελέτη περιοχές ή με άλλα λόγια ο εντοπισμός μίας ή και παραπάνω μεμονωμένης στατιστικά σημαντικής συστάδας. Για το λόγο αυτό, η μεθοδολογία που προτάθηκε από τον Kulldorff κρίνεται καταλληλότερη.

Το μοντέλο που θα χρησιμοποιηθεί είναι το μοντέλο Poisson αφού θεωρείται καταλληλότερο στη μοντελοποίηση ενός αριθμού γεγονότων σε συνεχή χρόνο. Παρόλο που σε χωρικές αναλύσεις προκύπτει το πρόβλημα της υπερδιασποράς (*overdispersion*), όταν μελετάται μόνο το ερώτημα ύπαρξης τυχαιότητας, το μοντέλο Poisson είναι επαρκές αφού η απόρριψη της τυχαιότητας αρκεί και δεν είναι απαραίτητη επιπλέον ανάλυση (δεν γίνεται μοντελοποίηση).

### 5.2.1 Συλλογή Δεδομένων

Απαραίτητο συστατικό για τη διεξαγωγή μιας έρευνας αποτελούν πάντοτε τα δεδομένα, και στην προκειμένη περίπτωση, τα δεδομένα ασθενών που εμφάνισαν τη νόσο και νοσηλεύθηκαν σε όλα τα νοσοκομεία της Ελλάδας. Εφόσον η σκλήρυνση κατά πλάκας είναι μια ασθένεια που μελετάται ακόμη, δεν υπήρχε κάποια βάση έτοιμων δεδομένων. Για το λόγο αυτό, στο μικροβιολογικό εργαστήριο του Αιγινήτειου νοσοκομείου, καταγράφονται δεδομένα εδώ και ένα χρόνο, μέχρι και σήμερα. Καθημερινά, μέχρι και τις ημέρες που έγινε η συλλογή

δεδομένων κατεύθαναν στο εργαστήριο του νοσοκομείου καρτέλες ασθενών οι οποίοι είχαν υποβληθεί σε σχετικές εξετάσεις και δέχτηκαν να λάβουν μέρος στη μελέτη που γίνεται. Από εκεί, ακολουθούσε τηλεφωνική επικοινωνία με το νοσοκομείο νοσηλείας του κάθε ασθενούς, για να λάβουμε τα στοιχεία διεύθυνσής του, το φύλο του και την ημερομηνία γέννησής του. Χαρακτηριστική ήταν η άρνηση αρκετών γραφείων κίνησης, και ιδιαίτερα στρατιωτικών νοσοκομείων, να δώσουν στοιχεία ασθενών, χωρίς κάποια υπεύθυνη δήλωση ή εξουσιοδότηση.

Οι ασθενείς νοσηλεύονταν σε διάφορα νοσοκομεία της Ελλάδας όπου και συλλέγονταν εγκεφαλονωτιαίο υγρό και όρος αίματος από κάθε έναν. Στις ουσίες αυτές μετρήθηκε η περιεκτικότητα σε ανοσοσφαιρίνη Γ (IgG), είδος αντισώματος και σε η αλβουμίνη, είδος πρωτεΐνης. Από αυτές προέκυψαν κάποιοι δείκτες, ο δείκτης *Αλβουμίνης* ( $\frac{AlbENY}{AlbOPOY} * 1000$ ), και ο δείκτης *IgG* ( $\frac{IgGENY}{IgGOPOY} / \frac{AlbENY}{AlbOPOY}$ ). Όταν το εύρημα στο ασθενή για τον δείκτη IgG ήταν πάνω από 0,65 και για τον δείκτη αλβουμίνης πάνω από 10, τότε το αποτέλεσμα ήταν θετικό, και ο ασθενής έφερε τη νόσο της σκλήρυνσης κατά πλάκας.

Η διάγνωση βασίστηκε στη μέθοδο ολιγοκλωνικών ταινιών ανοσοσφαιρινών ή διαφορετικά στη μέθοδο ισοηλεκτρικής εστίασης και Western Blot (Χατζηπαναγιώτου κ. ά, 2010).

## 5.2.2 Επεξεργασία Δεδομένων

Επόμενο βήμα ήταν η δημιουργία μιας βάσης δεδομένων η οποία θα περιείχε όλα τα στοιχεία των ασθενών που βρέθηκαν θετικοί στην

εξέταση και τα οποία θα μπορούν αργότερα να επεξεργαστούν με σκοπό την εξαγωγή στατιστικών συμπερασμάτων. Σε ένα λογιστικό φύλλο του MS Excel καταχωρήθηκαν τα δεδομένα που συλλέξαμε και συγκεκριμένα, το φύλο του ασθενή, η διεύθυνσή του τη στιγμή της εισαγωγής του στο τμήμα νοσηλείας, ο καποδιστριακός δήμος στον οποίο ανήκει, η δημοτική ενότητα, το νοσοκομείο που νοσηλεύθηκε και η ημερομηνία εισαγωγής του. Συνολικά καταγράφηκαν 178 ασθενείς οι οποίοι έφεραν τη νόσο, για τα έτη 2012 - 2013.

Τα δεδομένα αυτά θα έπρεπε με κάποιον τρόπο να εισαχθούν στο στατιστικό πρόγραμμα SaTScan που χρησιμοποιήσαμε και μάλιστα με συγκεκριμένη μορφοποίηση. Έτσι, δημιουργήσαμε εκ νέου τρία καινούρια αρχεία κειμένου (txt files) για να είναι εφικτό να χρησιμοποιηθούν στην ανάλυση. Το αρχείο cases.txt περιείχε κάποιες πληροφορίες για τους ασθενείς. Ενδεικτικά ένα μέρος του απεικονίζεται παρακάτω:

**Πίνακας 2:** Ενδεικτικό Τμήμα του αρχείου cases.txt

tzoumerkwn	0	2012
dodonis	0	2012
zagoriou	0	2013
zitsas	0	2013
ioanniton	1	2013
ioanniton	1	2013
ioanniton	1	2013
konitsas	0	2012
metsovou	1	2012
pagoniou	0	2012
artaiwn	1	2013



Στην πρώτη στήλη αναγράφεται το όνομα του δήμου που ανήκει ο πρώτος ασθενής, στην δεύτερη το φύλο του (0: άνδρας, 1: γυναίκα) και στην τρίτη στήλη το έτος εισαγωγής του. Το ίδιο ισχύει και για τις υπόλοιπες εγγραφές.

Το αρχείο coordinates.txt περιείχε τις συντεταγμένες κάθε δήμου των τριών γεωγραφικών διαμερισμάτων που λήφθηκαν υπόψη στη μελέτη (Στερεάς Ελλάδας, Ηπείρου και Πελοποννήσου). Ενδεικτικά ένα μέρος του απεικονίζεται παρακάτω:

**Πίνακας 3:** Ενδεικτικό Τμήμα του αρχείου coordinates.txt

tzoumerkwn	39.550747	21.000152
dodonis	39.57513	20.748196
zagoriou	39.896568	20.856857
zitsas	39.755026	20.649533
ioanniton	39.666698	20.853767
konitsas	40.049103	20.752916
metsovou	39.771306	21.182842
pagoniou	39.966069	20.476456
artaiwn	39.248207	21.18576

Στην πρώτη στήλη αναγράφεται ο καποδιστριακός δήμος, στη δεύτερη η γεωγραφική τεταγμένη του, ενώ στην τρίτη στήλη η γεωγραφική τετμημένη του κέντρου κάθε δήμου (συνήθως η διεύθυνση του δημαρχείου). Σύμφωνα με το νέο σχέδιο Καλλικράτη που δημοσιεύθηκε στο ΦΕΚ 87Α' τον Ιούνιο του 2010 (N3852/2010), κάθε καποδιστριακός

δήμος εμπεριέχει περισσότερες απο μία υπο-περιοχές. Επομένως, πολλές διευθύνσεις έχουν συγχωνευθεί κάτω από έναν δήμο, και κατά συνέπεια, κάτω από ένα ζευγάρι συντεταγμένων. Για παράδειγμα, ταχυδρομικοί κώδικες Αττικής που υποδήλωναν κοντινές περιοχές, καταγράφηκαν ως μια περιοχή, λαμβάνοντας τις ίδιες συντεταγμένες. Τα στοιχεία βρέθηκαν βάσει της διεύθυνσης που δήλωσε ο κάθε ασθενής και με τη βοήθεια του Google Maps.

Το αρχείο population.txt, απαραίτητο εφόσον θα χρησιμοποιηθεί μοντέλο Poisson, περιείχε πληθυσμιακά στοιχεία για κάθε έναν από αυτούς τους δήμους. Ενδεικτικά, ένα μέρος του αρχείου απεικονίζεται παρακάτω:

**Πίνακας 4:** Ενδεικτικό Τμήμα του αρχείου population.txt

tzoumerkwn	2012	5714
dodonis	2012	9693
zagoriou	2013	3724
zitsas	2013	14766
ioanniton	2013	112486
konitsas	2012	6362
metsovou	2012	6196
pagoniou	2012	6196
artaiwn	2013	4316

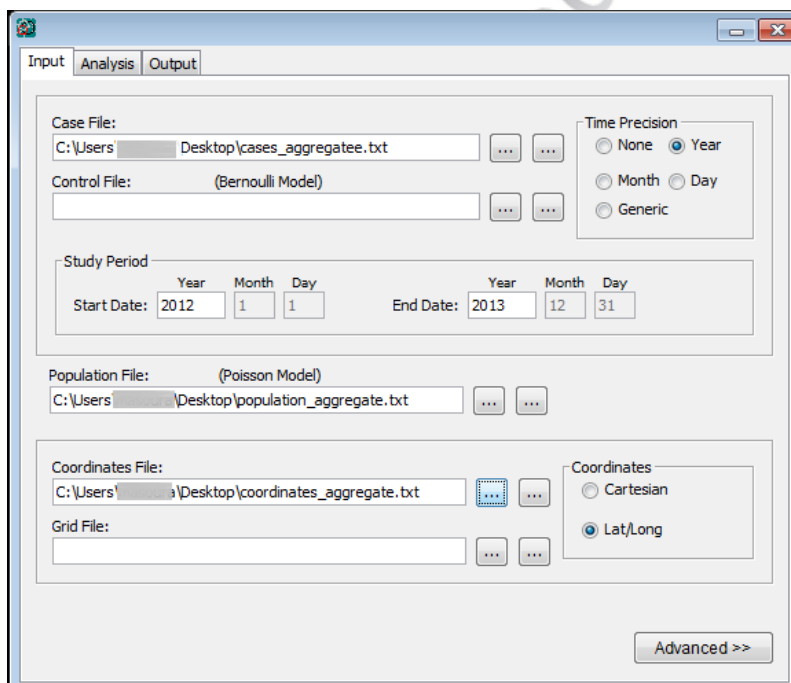
Και πάλι, στην πρώτη στήλη αναγράφεται ο καποδιστριακός δήμος, στην δεύτερη η χρονολογία καταγραφής του ασθενούς και στην τρίτη ο πληθυσμός του δήμου. Σημειώνεται ότι ο συνολικός πληθυσμός για κάθε

δήμο λήφθηκε από τα στοιχεία της ελληνικής απογραφής του έτους 2011 και από την Ελληνική Στατιστική Υπηρεσία ([www.esye.gr](http://www.esye.gr)).

Αφού δημιουργήθηκαν τα τρία αρχεία που εμπεριείχαν τα δεδομένα και σε συμβατή μορφή, στο επόμενο στάδιο εισήχθησαν στο στατιστικό πρόγραμμα με σκοπό την ανάλυσή τους.

Στην παρακάτω εικόνα φαίνεται το παράθυρο του προγράμματος και η θέση που τοποθετείται το κάθε αρχείο.

Εικόνα 2: Αρχική οθόνη του πακέτου SaTScan



### 5.2.3 Στατιστική Μεθοδολογία

Ο αλγόριθμος του Kulldorff, ο οποίος χρησιμοποιείται στο στατιστικό πακέτο SaTScan, κάνει χρήση μιας ακολουθίας κυκλικών παραθύρων,

πάνω στην γεωγραφική περιοχή που μελετάται. Οι κύκλοι αυτοί δημιουργήθηκαν έχοντας ως κέντρο τους, το κέντρο κάθε δήμου. Αν το παράθυρο περιέχει το κέντρο της περιοχής που απογράφεται, τότε και όλη η περιοχή εμπερικλείεται σε αυτό. Για κάθε κέντρο, η ακτίνα του κυκλικού παραθύρου αυξάνεται συνεχόμενα καλύπτοντας όλο και περισσότερους γειτονικούς δήμους, μέχρι να καλυφθεί ένα συγκεκριμένο ποσοστό του συνολικού πληθυσμού της μελέτης το οποίο καθορίζεται από τον χρήστη. Η μηδενική υπόθεση που εξετάζεται είναι ότι υπάρχει πλήρης χωρική τυχαιότητα στο πώς κατανέμονται τα περιστατικά σκλήρυνσης κατά πλάκας στις περιοχές της Ελλάδας. Η πιθανότητα λανθασμένης ανίχνευσης συστάδας σε κάποια περιοχή, ισούται με το 5% του επιπέδου σημαντικότητας, ενώ η στατιστική συνάρτηση σάρωσης ορίζεται ως ένας έλεγχος λόγου πιθανοφάνειας, δεσμευμένης στον συνολικό αριθμό περιπτώσεων  $N$  που παρατηρήθηκε, για όλους τους κύκλους  $Z$ . Συγκεκριμένα,

$$S = \frac{\max_Z \{L(Z)\}}{L_0} = \max_Z \left\{ \frac{L(Z)}{L_0} \right\},$$

όπου  $L(Z)$  η μέγιστη πιθανοφάνεια του κύκλου  $Z$  και  $L_0$  η συνάρτηση πιθανοφάνειας υπό την μηδενική υπόθεση. Όσο αυτός ο λόγος πιθανοφάνειας μεγιστοποιείται για όλους τους κύκλους, εντοπίζει αυτόν που ενδεχομένως αποτελεί την πιο πιθανή συστάδα (Bersimis et al., 2014).

Η κατανομή του στατιστικού ελέγχου προέκυψε με Monte Carlo προσομοίωση, και συγκεκριμένα μετά από 99999 επαναλήψεις, σε καθεμία από τις οποίες υπολογίστηκε η στατιστική συνάρτηση. Όποια συστάδα ανιχνευθεί, κρίνεται στατιστικά σημαντική, σε 1% επίπεδο

σημαντικότητας, εάν η τιμή της στατιστικής συνάρτησης που προέκυψε από το πραγματικό πληθυσμιακό δείγμα (σύνολο δεδομένων), βρεθεί ανάμεσα στο υψηλότερο 1% των 100000 τιμών (99999 οι επαναλήψεις της προσομοίωσης και 1 η πραγματική τιμή).

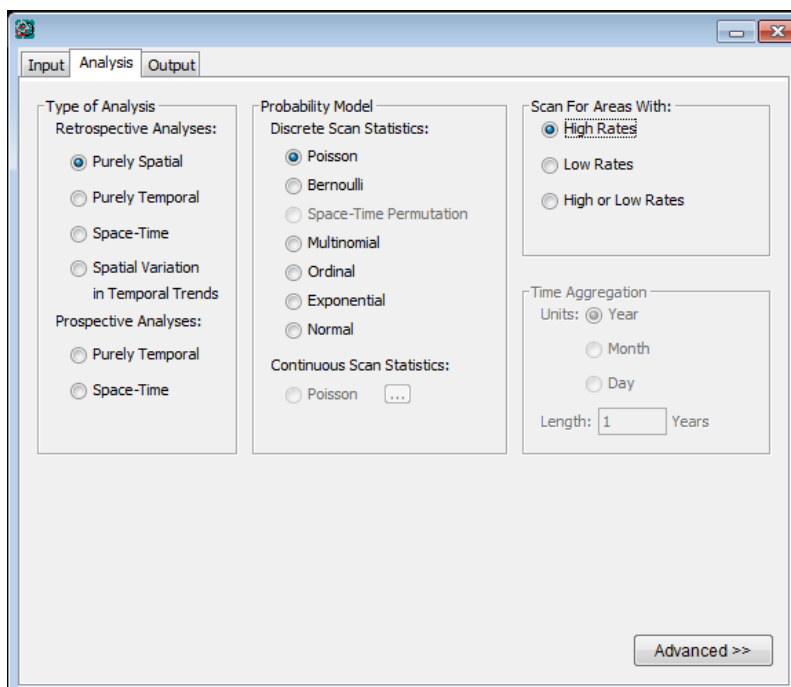
Αναζητούμε “*high-rates*” συστάδες, δηλαδή συστάδες με υψηλή συγκέντρωση συμβάντων μέσα σε αυτές. Ο λόγος αυτής της επιλογής είναι διότι η υπό μελέτη περιοχή περιλαμβάνει μέσα υπο-περιοχές με καθόλου δραστηριότητα.

Από τη στιγμή που ανιχνευθούν οι στατιστικά σημαντικές συστάδες, το μέγεθος του μεγαλύτερου παραθύρου που δημιουργήθηκε, αρχίζει να μειώνεται σταδιακά, ώστε να εκτιμήσει εάν υπάρχουν υπό-συστάδες που να είναι αρκετά ισχυρές για να απορρίψουν την μηδενική υπόθεση.

### 5.3 ΑΠΟΤΕΛΕΣΜΑΤΑ

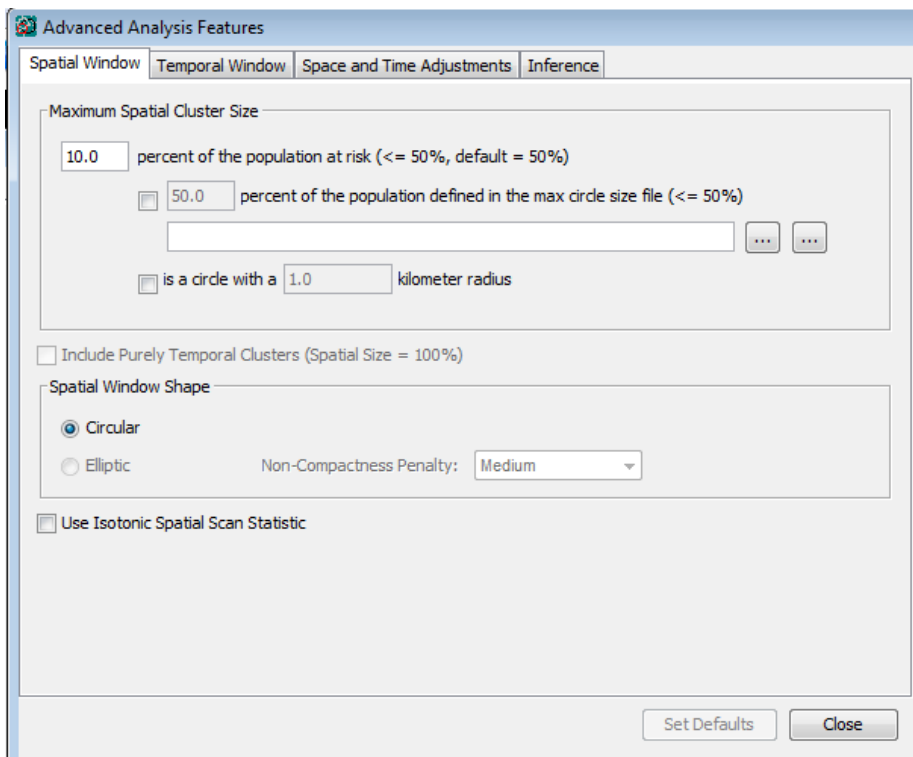
Μέσα από μια ακολουθία δοκιμών και επαναλήψεων (μέθοδος *trial-and-error*) εξετάστηκαν οι συστάδες που προκύπτουν για διάφορα ποσοστά του πληθυσμού εκτεθειμένα στον κίνδυνο. Αρχική υπόθεση ήταν ότι αναζητούμε περιοχές με υψηλή συγκέντρωση γεγονότων. Το μοντέλο που θα χρησιμοποιήσουμε είναι μοντέλο Poisson και πρόκειται για μια ανάλυση αμιγώς χωρική, αφού δεν ψάχνουμε για τροποποιήσεις των συμβάντων σε πέρασ χρόνου.

Εικόνα 3: Επιλογή ανάλυσης στο πακέτο SaTScan



Εφόσον επιλέχθηκε μια μέθοδος δοκιμών για την εξαγωγή εγκυρότερου αποτελέσματος, θεωρήσαμε ότι ως πρώτο σταδιο, το μέγεθος του μέγιστου παραθύρου θα είναι τέτοιο ώστε να καλύπτει το 10% του πληθυσμού που είναι εκτεθειμένο στη νόσο.

**Εικόνα 4:** Επιλογή ποσοστού πληθυσμού εκτεθειμένου στον κίνδυνο



Πιο αναλυτικά, βάσει των συντεταγμένων όλων των δήμων που εισήχθησαν στο πρόγραμμα μέσω του αρχείου population.txt, δημιουργήθηκε ένα πλέγμα, το οποίο απεικονίζει τους δήμους των τριών γεωγραφικών διαμερισμάτων. Πάνω σε αυτό το πλέγμα δημιουργούνται παράθυρα - κυκλικά στην περίπτωση που εξετάζουμε, τα οποία σαρώνουν αυτήν την επιφάνεια.

Αφού προέκυψαν κάποια αποτελέσματα όταν το μέγιστο μέγεθος ήταν 10%, το αυξήσαμε σε 15% και σταδιακά μέχρι 30% ακολουθώντας την ίδια διαδικασία. Παρατηρήθηκε ότι κάποιες συστάδες παρέμεναν

στατιστικά σημαντικές όσο και αν μεταβαλλόταν το μέγεθος του παραθύρου.

Καταλήξαμε στην αποδοχή των αποτελεσμάτων που προέκυψαν όταν το μέγιστο ποσοστό πληθυσμού που κάλυπτε το παράθυρο ήταν 20%, αφού εκεί ήταν περισσότερο ορατά τα αποτελέσματα. Εντοπίζεται μια στατιστικά σημαντική συστάδα, ενώ επόμενες απορρίπτονται. Αναλυτικότερα,

**Πίνακας 5:** Χωρική Ανάλυση της Σκλήρυνσης κατά Πλάκας με Μέγιστο Μέγεθος Κάλυψης 20%

ID	Centroid Coordinates	Radius	Number of locations included	Population	Cases			Relative Risk	Log-Likelihood ratio	P-value
					Observed	Expected	Ratio			
1	23.689442, 38.01374	5.81 km	10	1252722	60	33.57	1.79	2.18	10.97061	0.0015
2	24.0077, 37.97492	0 km	1	33821	6	0.91	6.62	6.81	6.31994	0.11475

υπήρξαν συνολικά 179 περιπτώσεις ασθενών που διαγνώστηκαν με Σκλήρυνση Κατά Πλάκας στις περιοχές της Ηπείρου, Στερεάς Ελλάδας και Πελοποννήσου, τα έτη 2012-2013. Εξετάστηκαν συνολικά 178 περιοχές οι οποίες κάλυπταν συνολικό πληθυσμό 6.678.849 ατόμων. Αναλογικά, στους 100.000 ανθρώπους οι 2,7 διαγνώσκονται με Σκλήρυνση Κατά Πλάκας κάθε χρόνο.

Η ανάλυση χωρικής σάρωσης ανίχνευσε μία στατιστικά σημαντική συστάδα, και συγκεκριμένα στην περιοχή με συντεταγμένες (23,689442, 38,01374). Αργότερα με τη βοήθεια περισσότερων στοιχείων και χάρτη,



εντοπίζουμε ποιοί δέκα δήμοι εμπεριέχονται σε αυτήν την περιοχή. Εκεί, παρατηρήθηκαν 60 περιπτώσεις ασθενών που εμφάνισαν τη νόσο της, ενώ αναμενόταν μόνο 33,57 βάσει της κατανομής Poisson που υποθέσαμε. Πιο αναλυτικά, σε ακτίνα 5,81 χιλιομέτρων και σε συνολικό πληθυσμό 1.252.722 ατόμων ο λόγος παρατηρηθέντων συμβάντων προς αναμενόμενο αριθμό συμβάντων βρέθηκε ίσος με 1,79. Αναλογικά, 4,8 (δηλαδή 5) άτομα στα 100.000 θα εμφανίζουν τη νόσο κάθε χρόνο σε αυτές τις περιοχές. Επιπλέον, ο σχετικός κίνδυνος (RR) που ισούται με 2,18 υποδεικνύει ότι αν κάποιος ζει στις συγκεκριμένες περιοχές εμφανίζει 2,18 φορές μεγαλύτερο κίνδυνο να νοσήσει, από κάποιον που δεν ζει εκεί. Η υπόθεση της τυχαιότητας απορρίπτεται, σε 99% επίπεδο εμπιστοσύνης, με  $p$ -value 0,0015.

Αναγνωρίστηκε επίσης μια δεύτερη συστάδα γεγονότων με συντεταγμένες (24,0077, 37,97492) οι οποίες αντιστοιχούν σε έναν μόνο δήμο. Εκεί ανιχνεύθηκαν 6 περιστατικά ενώ αναμενόταν 0,91 και πιο αναλυτικά, σε συνολικό πληθυσμό 33.821 ατόμων ο λόγος παρατηρηθέντων συμβάντων προς αναμενόμενο αριθμό συμβάντων βρέθηκε ίσος με 6,62. Ο σχετικός κίνδυνος (RR) ισούται με 6,81, δηλαδή κάποιος που ζει σε εκείνη την περιοχή εμφανίζει 6,81 φορές μεγαλύτερο κίνδυνο να προσβληθεί από τη νόσο της σκλήρυνσης, σε σχέση με κάποιον που ζει αλλού. Το  $p$ -value του στατιστικού ελέγχου προέκυψε ίσο με 0,11475, γεγονός που σημαίνει ότι σε 99% επίπεδο εμπιστοσύνης δεν απορρίπτεται η υπόθεση τυχαιότητας. Αν ο έλεγχος διεξαγόταν σε «πιο χαλαρά» όρια, θα απορρίπτοταν οριακά.

Αναφέρονται επίσης και άλλες δύο συστάδες οι οποίες όμως δεν κρίθηκαν στατιστικά σημαντικές στα επίπεδα εμπιστοσύνης που ορίστηκαν.

Να σημειωθεί τέλος ότι, οι δήμοι που εμπεριέχονται στην πρώτη συστάδα, η οποία και κρίθηκε στατιστικά σημαντική, αφορά περιοχές της δυτικής Αττικής ενώ η δεύτερη συστάδα αναφέρεται στο δήμο Σπατών, στην ανατολική Αττική.

Σε γενικές γραμμές, υπήρξε ένδειξη απόρριψης της μηδενικής υπόθεσης, και μάλιστα σε οποιοδήποτε διάστημα και αν περιορίζαμε το μέγιστο μέγεθος που θέλαμε να καλύψει το παράθυρο σάρωσης (10%, 15%, 20% και 30%). Η ένδειξη μη-τυχαίας κατανομής συμβάντων στο χώρο επιβεβαιώνεται αργότερα και με άλλες μεθόδους απεικόνισης σημαντικών συστάδων σε μια γεωγραφική περιοχή.

Παρακάτω απεικονίζονται συνοπτικά οι συστάδες που εντοπίστηκαν (στατιστικά σημαντικές και μη-στατιστικά σημαντικές), για κάθε ποσοστό του πληθυσμού που επιλέξαμε να καλύψει το παράθυρο σάρωσης.

Πίνακας 6: Αποτελέσματα Επαναληπτικής Μεθόδου Εντοπισμού Στατιστικά Σημαντικών Συστάδων

	LocID	Population	#cases	Expected cases	RR	LLR	P-value	
10% pop-at-risk	1	Aigalew, peristeri, chaidari, varvara, koridallos	346819	22	9.3	2.56	6.737212	0.06789
	2	Spatwn	33821	6	0.91	6.81	6.314	0.10055
	3	Athina	664046	29	17.80	1.75	3.355261	0.81203
	4	pentelis, vrilissiwn, pallinis, dionisou, paraskeuis, kifisias, amarousiou	362920	17	9.73	1.83	2.377	0.98770
	5	levadewn, orxomenou, distomou	51124	4	1.37	2.96	1.675	0.999
15% pop-at-risk	1	aigalew, peristeriou, xaidariou, varvaras, koridallou	346819	22	9.30	2.56	6.737212	0.07437
	2	spatwn	33821	6	0.91	6.81	6.314	0.10914
	3	galatsiou, filotheis, filadelfias, ionias, anargirwn, zografou, athina	986604	42	26.44	1.77	4.6979	0.39314
	4	pentelis, vrilissiwn, pallinis, dionisou, paraskeuis, kifisias, amarousiou	362920	17	9.73	1.83	2.37706	0.98926
	5	levadewn, orxomenou, distomou	51124	4	1.37	2.96	1.67516	0.99999
20% pop-at-risk	1	peristeriou, aigalew, iliou, xaidariou, petroupolis, anargirwn, varvaras, athinaiwn, koridallou, filadelfias	1252722	60	33.57	2.18	10.9706	0.00150
	2	spatwn	33821	6	0.91	6.81	6.314	0.11475
	3	pentelis, vrilissiwn, pallinis, dionisou, paraskeuis, kifisias, amarousiou	362920	17	9.73	1.83	2.37706	0.99080
	4	levadewn, orxomenou, distomou	51124	4	1.37	2.96	1.67516	0.99999
30% pop-at-risk	1	petroupolis, iliou, peristeriou, xaidariou, anargirwn, aigalew, filadelfias, varvaras, filis, koridallou, ionias, athinaiwn	1365821	64	36.61	2.16	11.18394	0.00147
	2	spatwn	33821	6	0.91	6.81	6.314	0.12319
	3	pentelis, vrilissiwn, pallinis, dionisou, paraskeuis, kifisias, amarousiou	362920	17	9.73	1.83	2.37706	0.99224

### 5.3.1 Γεωγραφική Απεικόνιση

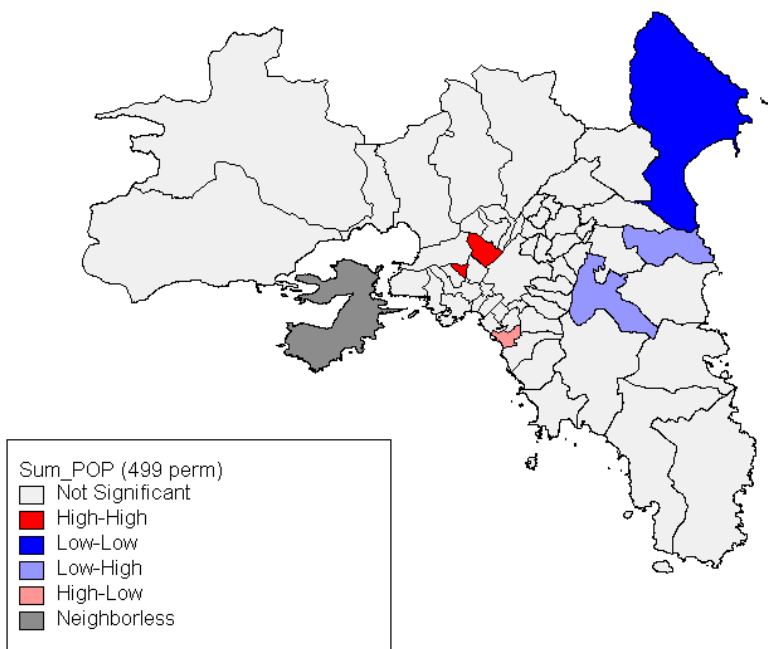
Σε συνέχεια της στατιστικής ανάλυσης έρχεται η γεωγραφική απεικόνιση του χάρτη της περιοχής που μελετήθηκε. Μέσω του λογισμικού GeoDa (Anselin et al., 2006) και αφού έγινε η χαρτογράφηση, επιπλέον εφαρμόστηκε η μέθοδος της τοπικής αυτοσυσχέτισης LISA (Local Indicator of Spatial Autocorrelation). Το πακέτο GeoDa έχει σχεδιαστεί για ειδικότερη ανάλυση χωρικών δεδομένων, τοποθετημένα σε πλέγμα, παρέχοντας μια γραφική απεικόνιση των στατιστικών μεθόδων που έχουν χρησιμοποιηθεί.

Η μέθοδος αυτοσυσχέτισης LISA βασίζεται στις τιμές του στατιστικού μεγέθους  $I$  του Moran (Moran's  $I$ ) λόγω του ότι εστιάζουμε κυρίως σε χωρική εξάρτηση ή ανεξαρτησία. Συγκεκριμένα, θετικό  $I_i$  (θετική αυτοσυσχέτιση), αντιστοιχεί είτε σε περιοχές υψηλής συγκέντρωσης περιστατικών που συνορεύουν με περιοχές υψηλής συγκέντρωσης περιστατικών (*high-high cluster*) είτε με περιοχές χαμηλής συγκέντρωσης περιστατικών (*low-low cluster*). Αρνητική τιμή του  $I_i$  (αρνητική αυτοσυσχέτιση) αντιστοιχεί είτε σε περιοχές που ενώ έχουν χαμηλή συγκέντρωση περιστατικών συνορεύουν με περιοχές υψηλής συγκέντρωσης (*low-high cluster*), είτε σε περιοχές με υψηλή συγκέντρωση που συνορεύουν με περιοχές χαμηλής συγκέντρωσης περιστατικών (*high-low cluster*). Βάσει την συνεισφοράς κάθε περιοχής στο στατιστικό  $I$  του Moran για όλη την περιοχή που μελετάται, δημιουργείται η εκτίμηση για τη δημιουργία συστάδων που σχετίζονται (Anselin, 1995). Η μηδενική υπόθεση είναι

και εδώ ότι δεν υπάρχει χωρική αυτοσυσχέτιση μεταξύ των περιοχών του γραφικού πλέγματος.

Παρακάτω απεικονίζονται γραφικά τα αποτελέσματα μόνο για το γεωγραφικό διαμέρισμα της Αττικής και έχει ιδιαίτερο ενδιαφέρον ότι τα αποτελέσματα της εν μέρη περιγραφικής αυτής ανάλυσης ταυτίζονται με αυτά της εφαρμογής του SatScan.

Εικόνα 5: Γραφική απεικόνιση αποτελεσμάτων μέσω του GeoDa



Στον χάρτη απεικονίζονται οι καποδιστριακοί δήμοι του νομού Αττικής, ενώ χρωματισμένοι είναι αυτοί που προέκυψαν με κάποιο ιδιαίτερο ενδιαφέρον. Πιο αναλυτικά, με κόκκινο χρώμα σημειώνονται οι

δήμοι με υψηλές τιμές συγκέντρωσης περιστατικών που συνορεύουν και με δήμους υψηλής συγκέντρωσης (*high-high clusters*) και αυτές αντιτοιχούν σε δήμους της δυτικής Αττικής. Με άλλα λόγια ολόκληρη η περιοχή της δυτικής Αττικής περιέχει δήμους με υψηλά ποσοστά συγκέντρωσης, όπως αποδεικνύεται και με τη στατιστική ανάλυση. Με μπλέ χρώμα σημειώνονται οι περιοχές χαμηλής συγκέντρωσης περιστατικών που συνορεύουν με περιοχές χαμηλής συγκέντρωσης (*low-low clusters*) και στο χάρτη καταλαμβάνουν μεγάλο τμήμα της βορειοανατολικής Αττικής και περιοχές όπως τον Μαραθώνα, τον Αγ. Στέφανο και γειτονικές τους, όπως Πεντέλη ή Ραφήνα. Με κόκκινο χρώμα φαίνονται οι περιοχές χαμηλής συγκέντρωσης περιστατικών που όμως συνορεύουν με περιοχές υψηλής συγκέντρωσης (*low-high clusters*) ενώ με ροζ οι περιοχές υψηλής συγκέντρωσης που συνορεύουν με περιοχές χαμηλότερης περιεκτικότητας σε συμβάντα (*high-low clusters*). Οι δύο τελευταίες κατηγορίες δεν θα μας απασχολήσουν στη συγκεκριμένη περιοχή.

Συμπερασματικά, εύκολα διακρίνει κανείς ότι τα αποτελέσματα που προέκυψαν εφαρμόζοντας τη μέθοδο LISA με 99999 επαναλήψεις (Monte Carlo προσομοίωση) και σε 95% επίπεδο σημαντικότητας ταυτίζονται με αυτά της στατιστικής μεθόδου και ανάλυσης που προηγήθηκε με 20% του πληθυσμού να βρίσκεται εκτεθειμένο, γεγονός που προμηνύει μια περαιτέρω μελέτη.

## 5.4 ΣΥΜΠΕΡΑΣΜΑΤΑ

Η στατιστική συνάρτηση σάρωσης που χρησιμοποιήθηκε, αποτελεί μια ισχυρή μέθοδο αναγνώρισης συστάδων κάποιας νόσου, καθώς και

επίδειξης σημαντικών χωρικών ή και χώρο-χρονικών αθροίσεων συμβάντων σκλήρυνσης κατά πλάκας στις περιοχές της Ηπείρου, Στερεάς Ελλάδας και Πελοποννήσου. Αυτή η προσέγγιση εδώ εντόπισε μια στατιστικά σημαντική χωρική συστάδα και μια οριακά σημαντική σε 99% επίπεδο σημαντικότητας.

Ειδικότερα, καταγράφηκε υψηλή χωρική συγκέντρωση περιπτώσεων εμφάνισης σκλήρυνσης κατά πλάκας στα δυτικά προάστια της Αττικής, κρίνοντας τη στατιστικά σημαντική λαμβάνοντας υπόψη και την τοπική πληθυσμιακή πυκνότητα. Η συστάδα περιλαμβάνει τις περιοχές του Αιγάλεω, του Περιστερίου, του Ιλίου, του Χαϊδαρίου, της Πετρούπολης, των Αγ. Αναργύρων, της Αγ. Βαρβάρας, του Κορυδαλλού, της Νέας Φιλαδέλφειας και των Αθηναίων. Η επόμενη συστάδα που οριακά δεν απορρίπτεται η μηδενική υπόθεση αντιστοιχεί στο δήμο των Σπατών, στα ανατολικά προάστια της Αττικής.

Είναι εμφανές ότι η ανάλυση που προηγήθηκε επιβεβαιώνει την αρχική υπόθεση μη-τυχαίας κατανομής των περιπτώσεων εμφάνισης της νόσου. Αντιθέτως, διάφοροι παράγοντες, οι οποίοι χρήζουν περαιτέρω ιατρική πλέον μελέτη και έρευνα, φαίνεται να επηρεάζουν την ανάπτυξή της.

## 5.5 ΣΥΝΟΨΗ

Στη διεθνή βιβλιογραφία μέχρι σήμερα έχει χρησιμοποιηθεί μια πληθώρα χωρικών στατιστικών συναρτήσεων σάρωσης με σκοπό την ανάλυση δεδομένων σε διάφορα θέματα που προέκυπταν και σε διάφορους επιστημονικούς τομείς, συμπεριλαμβανομένων των

περιβαλλοντικών μελετών, των μελετών εγκλημάτων και άλλων. Μάλιστα, σε πολλές απο αυτές, έχουν συνδιαστεί διαφορετικές μορφές συναρτήσεων σάρωσης, με αποτέλεσμα τη διεξαγωγή περισσότερης και εγκυρότερης πληροφορίας.

Όσον αφορά τη μελέτη της νόσου της σκλήρυνσης κατά πλάκας, αυτή αποφασίστηκε να διεξαχθεί διότι παρατηρήθηκε αρκετά μεγάλος επιπολασμός της στα γεωγραφικά διαμερίσματα της Ηπείρου, της Στερεάς Ελλάδας και της Πελοποννήσου. Η εμφάνισή της δεν έχει συσχετιστεί ακόμη με κάποιον παράγοντα, όμως συνεχίζονται οι μελέτες για την ανεύρεση κάποιων στοιχείων, που ενδεχομένως να επηρεάζουν την ανάπτυξη της νόσου.

Στην παρούσα εργασία έγινε χρήση στατιστικής συνάρτησης σάρωσης όπως εισήχθη από τον Kulldorff η οποία αναπτύχθηκε για την ανίχνευση επιθετικών μορφών λευχαιμίας και καρκίνο του μαστού (Donnan et al., 2005). Χρησιμοποιεί ένα κινούμενο κυκλικό παράθυρο μεταβλητού μεγέθους το οποίο σκανάρει την περιοχή μελέτης και εντοπίζει σημαντικές συστάδες. Επιπλέον παρέχει έναν έγκυρο στατιστικό έλεγχο για κάθε μια από τις συστάδες που εντοπίστηκε.



# Παραρτήματα

**Π1 Αρχεία Δεδομένων SaTScan**

**Π2 Αποτελέσματα της Μεθόδου Trial-And-Error**

**Π3 GeoDa Ανάλυση**

Πανεπιστήμιο Πειραιώς

## Π1 Αρχεία Δεδομένων SaTScan

### 1. Αρχείο cases.txt

tzoumerkwn	0	2012
dodonis	0	2012
zagoriou	0	2013
zitsas	0	2013
ioanniton	1	2013
ioanniton	1	2013
ioanniton	1	2013
konitsas	0	2012
metsovou	1	2012
pagoniou	0	2012
artaiwn	1	2013
karaiskaki	0	2013
kentrtzoumerkwn0	2013	
skoufa	1	2013
igoumenitsas0	2013	
souliou	0	2013
filiatwn	0	2013
zirou	0	2013
pargas	0	2013
prevezas	1	2013
agias	0	2013
elassonas	1	2013
kileler	0	2013
larisaiwn	0	2013

tempwn	0	2013
tirnavou	0	2013
farsalwn	0	2013
argitheas	0	2013
karditsas	0	2013
plastira	0	2013
mouzakiou	0	2013
palama	0	2013
sofadwn	0	2013
almirou	0	2013
volou	1	2013
volou	1	2013
zagoras	0	2013
phliou	0	2013
feraiou	0	2013
alonnisou	0	2013
skiathou	0	2013
skopelou	0	2013
kalampakas	0	2013
pilis	0	2013
trikkaiwn	1	2013
farkadonas	0	2013
amfikleias	0	2013
domokou	0	2013
lamiewn	1	2013
lokrwn	0	2013
makrakomis	0	2013
molou	0	2013
stilidos	0	2013
aliartou	0	2013

distomou	1	2013
thivaiwn	0	2013
levadewn	1	2013
levadewn	1	2013
orxomenou	1	2013
tanagras	0	2013
dirfiwn	0	2013
eretrias	1	2013
istiaias	0	2013
karistou	0	2013
kimis	0	2013
mantoudiou	0	2013
skirou	0	2013
xalkidewn	1	2012
xalkidewn	1	2012
xalkidewn	1	2013
agrafwn	0	2013
karpenisiou	0	2013
delfwn	1	2013
doridos	0	2013
aigialias	0	2013
achaias	0	2013
erimanthou	0	2013
kalavritwn	0	2013
patrwn	0	2013
agriniou	0	2013
aktiou	0	2013
amfilochias	0	2013
thermou	0	2013
mesologgiou	0	2013

naupaktias	0	2013
ksiromerou	0	2013
killinis	0	2013
krestenwn	0	2013
olimpias	0	2013
zacharws	0	2013
ilidas	0	2013
piniou	0	2013
pirgou	1	2013
pirgou	1	2013
kinourias	0	2013
gortinias	0	2013
megalopolis	0	2013
notkinourias	0	2013
tripolis	1	2013
tripolis	1	2013
argous	1	2013
argous	1	2013
epidaourou	0	2013
ermionidas	0	2013
naupliewn	1	2013
naupliewn	1	2013
velou	0	2013
korinthiwn	1	2013
korinthiwn	1	2013
loutrakiou	1	2013
nemeas	0	2013
ksilokastrou	0	2013
sikiwniwn	0	2013
manis	0	2013





vrilissiwn	0	2013
irakleiou	1	2013
irakleiou	1	2013
kifisias	1	2013
kifisias	1	2013
kifisias	1	2013
kifisias	1	2013
likovrisis	0	2013
metamorfosis1		2013
ionias	1	2013
ionias	1	2013
ionias	1	2013
papagou	1	2013
pentelis	1	2013
filotheis	1	2013
filotheis	1	2013
xalandriou	1	2013
varvaras	0	2013
anargirwn	1	2013
anargirwn	1	2013
anargirwn	1	2013
aigalew	1	2013
aigalew	1	2012
aigalew	1	2013
aigalew	1	2013
iliou	1	2013
iliou	1	2013
iliou	1	2013
peristeriou	1	2013
peristeriou	1	2013



peristeriou 1	2013
peristeriou 1	2013
peristeriou 1	2013
peristeriou 1	2013
peristeriou 1	2013
peristeriou 1	2013
peristeriou 1	2013
petroupolis 1	2013
xaidariou 1	2013
xaidariou 1	2013
xaidariou 1	2013
xaidariou 1	2013
xaidariou 1	2013
xaidariou 1	2013
xaidariou 1	2013
dimitriou 1	2013
alimou 1	2013
alimou 1	2013
glifadas 1	2013
glifadas 1	2013
glifadas 1	2013
glifadas 1	2013
ellinikou 1	2013
ellinikou 1	2013
kalitheas 1	2013
moschatou 1	2013
smirnis 1	2013
smirnis 1	2013
smirnis 1	2013
falirou 1	2013
falirou 1	2013
acharnwn 1	2013

acharnwn	1	2013
varis	1	2013
varis	1	2013
varis	1	2013
dionisou	0	2013
kropias	0	2013
laureotikis	0	2013
marathonos	1	2013
markopoulou	1	2013
paianias	0	2013
pallinis	1	2013
pallinis	1	2013
pallinis	1	2013
rafinas	0	2013
saronikou	0	2013
spatwn	1	2012
spatwn	1	2012
spatwn	1	2012
spatwn	1	2012
spatwn	1	2013
spatwn	1	2013
oropou	0	2013
aspropirgou	1	2013
aspropirgou	1	2013
eleusinas	1	2013
eleusinas	1	2013
mandras	1	2013
megarewn	1	2013
filis	1	2013
keratsiniou	1	2013

koridallou	1	2012
koridallou	1	2012
koridallou	1	2012
koridallou	1	2013
nikaias	1	2013
nikaias	1	2013
nikaias	1	2013
peiraiws	1	2013
peiraiws	1	2013
peiraiws	1	2013
peramatos	0	2013
agkistriou	0	2013
aiginas	0	2013
kithirwn	1	2013
porou	0	2013
salaminos	0	2013
spetswn	0	2013
trizinas	0	2013
idras	0	2013

## 2. Αρχείο coordinates.txt

tzoumerkwn	39.550747	21.000152
dodonis	39.57513	20.748196
zagoriou	39.896568	20.856857
zitsas	39.755026	20.649533
ioanniton	39.666698	20.853767
konitsas	40.049103	20.752916
metsovou	39.771306	21.182842
pagoniou	39.966069	20.476456
artaiwn	39.248207	21.18576

karaiskaki	39.169997	21.218033
kentrtrzkoumerkwn	39.44361721	21.145935
skoufa	39.079441	21.023197
igoumenitsas	39.542441	20.260162
souliou	39.471185	20.572586
filiatwn	39.601092	20.307713
zirou	39.284888	20.831451
pargas	39.257246	20.546494
prevezas	38.989303	20.76416
agias	39.719451	22.758737
elassonas	39.895843	22.189164
kileler	39.510398	22.480774
larisaiwn	39.64866	22.41989
tempwn	39.857574	22.549438
tirnavou	39.740062	22.288299
farsalwn	39.296414	22.384644
argitheas	39.356865	21.53904
karditsas	39.366023	21.921673
plastira	39.309863	21.739197
mouzakiou	39.431685	21.662893
palama	39.470854	22.082348
sofadwn	39.33682	22.091789
almirou	39.052118	22.952778
volou	39.362448	22.942168
zagoras	39.428238	23.122101
phliou	39.245017	23.216171
feraiou	39.468005	22.787018
alonnisou	39.217625	23.916636
skiathou	39.167701	23.46684
skopelou	39.117542	23.716779

kalampakas	39.711413	21.623497
pilis	39.460385	21.62251
trikkaiwn	39.55753	21.767693
farkadonas	39.593817	22.066898
amfikleias	38.64007	22.59304
domokou	39.131325	22.298212
lamiewn	38.899183	22.435455
lokrwn	38.625454	23.027344
makrakomis	38.943322	22.116165
molou	38.809751	22.646298
stilidos	38.913743	22.616301
aliartou	38.375846	23.106222
distomou	38.430631	22.667027
thivaiwn	38.315801	23.05481
levadewn	38.440615	22.875853
orxomenou	38.494175	22.98018
tanagras	38.246809	23.524475
dirfiwn	38.583533	23.83338
eretrias	38.431539	23.884964
istiaias	38.953702	23.152871
karistou	38.014761	24.420075
kimis	38.636919	24.102759
mantoudiou	38.825801	23.411865
skirou	38.907465	24.565859
xalkidewn	38.46945	23.606186
agrafwn	39.137883	21.649525
karpensisiou	38.918819	21.793613
delfwn	38.506266	22.502747
doridos	38.553535	22.104492
aigialias	38.11295	22.292633

achaias	38.158317	21.956177
erimanthou	38.032949	21.788635
kalavritwn	38.034132	22.110372
patrwn	38.250044	21.73542
agriniou	38.62599	21.409607
aktiou	38.91982	20.888915
amfilochias	38.864272	21.166749
thermou	38.571924	22.667185
mesologgiou	38.370934	21.430635
naupaktias	38.527756	21.813354
ksiromerou	38.655488	21.150055
killinis	37.935296	21.14512
krestenwn	37.593322	21.62075
olimpias	37.645466	21.625557
zacharws	37.484871	21.648173
ilidas	37.84341	21.450119
piniou	37.869975	21.335449
pirgou	37.673835	21.44351
kinourias	37.190955	22.862549
gortinias	37.692514	22.005615
megalopolis	37.401256	22.075653
notkinourias	37.134045	22.766418
tripolis	37.579413	22.339325
argous	37.609336	22.683334
epidauroy	37.612634	23.035927
ermionidas	37.387822	23.246942
naupliewn	37.565764	22.801566
velou	37.903845	22.738609
korinthiwn	37.940475	22.932844
loutrakiou	38.011853	23.025284

nemeas	37.822057	22.66119
ksilokastrou	38.079176	22.633295
sikiwniwn	37.929034	22.502747
manis	36.968547	22.287827
elafonisou	36.491283	22.959709
eurota	37.011874	22.532959
monemvasias	36.686592	22.979279
spartis	37.075279	22.43022
dytmanis	36.967449	22.286453
kalamatas	37.087501	22.095566
ichalias	37.304099	21.931458
pilou	36.91425	21.696625
trifilias	37.225408	21.689758
athinaiwn	37.985746	23.72858
vironos	37.960914	23.751583
galatsiou	38.014491	23.753128
dafnis	37.951304	23.741369
zografou	37.977966	23.761797
ilioupolis	37.932013	23.748236
kaisarianis	37.969846	23.757076
filadelfias	38.035856	23.737679
paraskeuis	38.01165	23.825312
amarousiou	38.058229	23.807545
vrilissiwn	38.034842	23.832006
irakleiou	38.053228	23.772354
kifisias	38.080257	23.815098
likovrisis	38.069379	23.780594
metamorfosis	38.064581	23.761024
ionias	38.04559	23.750124
papagou	37.993728	23.805399

pentelis	38.049105	23.869858
filotheis	38.01395	23.772526
xalandriou	38.028352	23.800163
varvaras	37.991394	23.658242
anargirwn	38.028825	23.722315
aigalew	37.99738	23.68103
iliou	38.032543	23.70369
peristeriou	38.013747	23.689442
petroupolis	38.043833	23.684292
xaidariou	38.008878	23.663692
dimitriou	37.937022	23.729525
alimou	37.918337	23.718538
glifadas	37.882712	23.769436
ellinikou	37.91143	23.746262
kalitheas	37.952624	23.700728
moschatou	37.949274	23.678026
smirnis	37.94589	23.712187
falirou	37.927257	23.698411
acharnwn	38.092147	23.737679
varis	37.846935	23.760595
dionisou	38.100591	23.87887
kropias	37.876208	23.835526
laureotikis	37.714398	24.057773
marathonos	38.153584	23.961954
markopoulou	37.884067	23.933287
paianias	37.957395	23.852348
pallinis	38.006578	23.884106
rafinas	38.027676	24.006672
saronikou	37.784283	23.931656
spatwn	37.974921	24.007702



oropou	38.304352	23.755016
aspropirgou	38.062757	23.590908
eleusinas	38.04728	23.541985
mandras	38.075967	23.501
megarewn	38.056742	23.269043
filis	38.104001	23.669229
keratsiniou	37.963147	23.618631
koridallou	37.982769	23.650131
nikaias	37.973162	23.642921
peiraiws	37.946059	23.650088
peramatos	37.968561	23.568077
agkistriou	37.699849	23.3535
aiginas	37.747372	23.500786
kithirwn	36.274725	22.979279
porou	37.522525	23.467827
salaminos	37.92091	23.489456
spetswn	37.266403	23.133087
trizinas	37.583766	23.216858
idras	37.347234	23.467484

### 3. Αρχείο population.txt

tzoumerkwn	2013	5714
dodonis	2013	9693
zagoriou	2013	3724
zitsas	2013	14766
ioanniton	2013	112486

konitsas	2013	6362
metsovou	2013	6196
pagoniou	2013	8960

artaiwn 2013 43166  
karaiskaki 2013 5780  
kentrtzoumerkwn20136178

skoufa 2013 12753  
igoumenitsas2013 25814  
souliou 2013 10063  
filiatwn 2013 7710  
zirou 2013 13892  
pargas 2013 11866  
prevezas 2013 31733  
agias 2013 11470  
elassonas 2013 32121  
kileler 2013 20854  
larisaiwn 2013 162591

tempwn 2013 13712  
tirnavou 2013 25032  
farsalwn 2013 18545  
argitheas 2013 3450  
karditsas 2013 56747  
plastira 2013 4635  
mouzakiou 2013 13122  
palama 2013 16726  
sofadwn 2013 18864  
almirou 2013 18614  
volou 2013 144449

zagoras 2013 5809  
phliou 2013 10216

feraiou	2013	10922
alonnisou	2013	2750
skiathou	2013	6088
skopelou	2013	4960
kalampakas	2013	21991
pilis	2013	14343
trikkaiwn	2013	81355
farkadonas	2013	13396
amfikleias	2013	10922
domokou	2013	11495
lamiewn	2013	75315
lokrwn	2013	19623
makrakomis	2013	16036
molou	2013	12090
stilidos	2013	12750
aliartou	2013	10887
distomou	2013	8188
thivaiwn	2013	36477
levadewn	2013	31315
orxomenou	2013	11621
tanagras	2013	19432
dirfiwn	2013	18800
eretrias	2013	13053
istiaias	2013	21083
karistou	2013	12180
kimis	2013	28437
mantoudiou	2013	12045
skirou	2013	2994
xalkidewn	2013	102223

agrafwn	2013	6976
karpenisiou2013		13105
delfwn	2013	26716
doridos	2013	13627
aigialias	2013	49872
achaias	2013	25916
erimanthou	2013	8877
kalavritwn	2013	11045
patrwn	2013	213984
agriniou	2013	94181
aktiou	2013	17370
amfilochias2013		17056
thermou	2013	8242
mesologgiou2013		34416
naupaktias	2013	27800
ksiromerou	2013	11737
killinis	2013	21581
krestenwn	2013	14109
olimpias	2013	13409
zacharws	2013	8953
ilidas	2013	32219
piniou	2013	21034
pirgou	2013	47995
kinourias	2013	10341
gortinias	2013	10109
megalopolis2013		10687
notkinourias2013		8294
tripolis	2013	47254
argous	2013	42022

epidaurou	2013	8115
ermionidas	2013	13551
naupliewn	2013	33356
velou	2013	19027
korinthiwn	2013	58192
loutrakiou	2013	21221
nemeas	2013	6483
ksilokastrou2013		17365
sikiwniwn	2013	22794
manis	2013	13005
elafonisou	2013	1041
eurota	2013	17891
monemvasias2013		21942
spartis	2013	35259
dytmanis	2013	6945
kalamatas	2013	69849
ichalias	2013	11228
pilou	2013	21077
trifilias	2013	27373
athinaiwn	2013	664046
vironos	2013	61308
galatsiou	2013	59345
dafnis	2013	33628
zografou	2013	71026
ilioupolis	2013	78153
kaisarianis2013		26458
filadelfias2013		35556
paraskeuis	2013	59704
amarousiou	2013	72333

vrilissiwn	2013	30741
irakleiou	2013	49642
kifisias	2013	70600
likovrisis	2013	31002
metamorfosis	2013	29891
ionias	2013	67134
papagou	2013	44539
pentelis	2013	34934
filotheis	2013	26968
xalandriou	2013	74192
varvaras	2013	26550
anargirwn	2013	62529
aigalew	2013	69946
ilou	2013	84793
peristeriou	2013	139981
petroupolis	2013	58979
xaidariou	2013	46897
dimitriou	2013	71294
alimou	2013	41720
glifadas	2013	87305
ellinikou	2013	51356
kalitheas	2013	100641
moschatou	2013	40413
smirnis	2013	73076
falirou	2013	64021
acharnwn	2013	106943
varis	2013	48399

dionisou	2013	40193
kropias	2013	30307
laureotikis2013		25102
marathonos	2013	33423
markopoulou2013		20040
paianias	2013	26668
pallinis	2013	54415
rafinas	2013	20266
saronikou	2013	29002
spatwn	2013	33821
oropou	2013	33769
aspropirgou2013		30251
eleusinas	2013	29902
mandras	2013	17885
megarewn	2013	36924
filis	2013	45965
keratsiniou2013		91045
koridallou	2013	63445
niakaias	2013	105430
peiraiws	2013	163688
peramatos	2013	25389
agkistriou	2013	1142
aiginas	2013	13056
kithirwn	2013	4041
porou	2013	3993
salaminos	2013	39283
spetswn	2013	4027
trizinas	2013	7143

## Π2 Αποτελέσματα της Μεθόδου Trial-And-Error

### 1. Results 10%

#### CLUSTERS DETECTED

1.Location IDs included.: aigalew, peristeriou, xaidariou, varvaras, koridallou

Overlap with clusters.: No Overlap

Coordinates / radius..: (23.681030 N, 37.997380 E) / 3.74 km

Gini Cluster.....: No

Population.....: 346819

Number of cases.....: 22

Expected cases.....: 9.30

Annual cases / 100000.: 3.2

Observed / expected...: 2.37

Relative risk.....: 2.56

Log likelihood ratio..: 6.737212

P-value.....: 0.06789

2.Location IDs included.: spatwn

Overlap with clusters.: No Overlap

Coordinates / radius..: (24.007702 N, 37.974921 E) / 0 km

Gini Cluster.....: No



Population.....: 33821  
Number of cases.....: 6  
Expected cases.....: 0.91  
Annual cases / 100000.: 8.9  
Observed / expected...: 6.62  
Relative risk.....: 6.81  
Log likelihood ratio...: 6.319937  
P-value.....: 0.10055

3.Location IDs included.: athinaiwn  
Overlap with clusters.: No Overlap  
Coordinates / radius...: (23.728580 N, 37.985746 E) / 0 km  
Gini Cluster.....: No  
Population.....: 664046  
Number of cases.....: 29  
Expected cases.....: 17.80  
Annual cases / 100000.: 2.2  
Observed / expected...: 1.63  
Relative risk.....: 1.75  
Log likelihood ratio...: 3.355261  
P-value.....: 0.81203

4.Location IDs included.: pentelis, vrilissiwn, pallinis,  
dionisou, paraskeuis, kifisias,  
amarousiou  
Overlap with clusters.: No Overlap  
Coordinates / radius...: (23.869858 N, 38.049105 E) / 6.99  
km  
Gini Cluster.....: No  
Population.....: 362920  
Number of cases.....: 17  
Expected cases.....: 9.73

Annual cases / 100000.: 2.3  
Observed / expected...: 1.75  
Relative risk.....: 1.83  
Log likelihood ratio...: 2.377060  
P-value.....: 0.98770

5.Location IDs included.: levadewn, orxomenou, distomou  
Overlap with clusters.: No Overlap  
Coordinates / radius...: (22.875853 N, 38.440615 E) /  
23.23 km

Gini Cluster.....: No  
Population.....: 51124  
Number of cases.....: 4  
Expected cases.....: 1.37  
Annual cases / 100000.: 3.9  
Observed / expected...: 2.92  
Relative risk.....: 2.96  
Log likelihood ratio...: 1.675162  
P-value.....: 0.99999

6.Location IDs included.: varis, glifadas, ellinikou,  
alimou  
Overlap with clusters.: No Overlap  
Coordinates / radius...: (23.760595 N, 37.846935 E) / 8.64  
km

Gini Cluster.....: No  
Population.....: 228780  
Number of cases.....: 11  
Expected cases.....: 6.13  
Annual cases / 100000.: 2.4  
Observed / expected...: 1.79  
Relative risk.....: 1.85

Log likelihood ratio..: 1.629691

P-value.....: 0.99999

7.Location IDs included.: mandras, eleusinas, aspropirgou

Overlap with clusters.: No Overlap

Coordinates / radius..: (23.501000 N, 38.075967 E) /

10.08 km

Gini Cluster.....: No

Population.....: 78038

Number of cases.....: 5

Expected cases.....: 2.09

Annual cases / 100000.: 3.2

Observed / expected...: 2.39

Relative risk.....: 2.43

Log likelihood ratio..: 1.473325

P-value.....: 0.99999

## **2. Results 15%**

### SUMMARY OF DATA

Study period.....: 2013/1/1 to 2013/12/31

Number of locations.....: 178

Total population.....: 6678849

Total number of cases.....: 179

Annual cases / 100000.....: 2.7

---

### CLUSTERS DETECTED

1.Location IDs included.: aigalew, peristeriou, xaidariou, varvaras, koridallou

Overlap with clusters.: No Overlap

Coordinates / radius...: (23.681030 N, 37.997380 E) / 3.74 km

Gini Cluster.....: No

Population.....: 346819

Number of cases.....: 22

Expected cases.....: 9.30

Annual cases / 100000.: 6.3

Observed / expected...: 2.37

Relative risk.....: 2.56

Log likelihood ratio...: 6.737212

P-value.....: 0.07437

2.Location IDs included.: spatwn

Overlap with clusters.: No Overlap

Coordinates / radius...: (24.007702 N, 37.974921 E) / 0 km

Gini Cluster.....: No

Population.....: 33821

Number of cases.....: 6

Expected cases.....: 0.91

Annual cases / 100000.: 17.8

Observed / expected...: 6.62

Relative risk.....: 6.81

Log likelihood ratio...: 6.319937

P-value.....: 0.10914

3.Location IDs included.: galatsiou, filotheis, filadelfias, ionias, anargirwn, zografou, athinaiwn

Overlap with clusters.: No Overlap

Coordinates / radius...: (23.753128 N, 38.014491 E) / 4.00  
km

Gini Cluster.....: No  
Population.....: 986604  
Number of cases.....: 42  
Expected cases.....: 26.44  
Annual cases / 100000.: 4.3  
Observed / expected...: 1.59  
Relative risk.....: 1.77  
Log likelihood ratio..: 4.697814  
P-value.....: 0.39314

4.Location IDs included.: pentelis, vrilissiwn, pallinis,  
dionisou, paraskeuis, kifisias,  
amarousiou

Overlap with clusters.: No Overlap  
Coordinates / radius...: (23.869858 N, 38.049105 E) / 6.99

km

Gini Cluster.....: No  
Population.....: 362920  
Number of cases.....: 17  
Expected cases.....: 9.73  
Annual cases / 100000.: 4.7  
Observed / expected...: 1.75  
Relative risk.....: 1.83  
Log likelihood ratio..: 2.377060  
P-value.....: 0.98926

5.Location IDs included.: levadewn, orxomenou, distomou

Overlap with clusters.: No Overlap

Coordinates / radius...: (22.875853 N, 38.440615 E) /

23.23 km

Gini Cluster.....: No  
Population.....: 51124  
Number of cases.....: 4  
Expected cases.....: 1.37  
Annual cases / 100000.: 7.8  
Observed / expected...: 2.92  
Relative risk.....: 2.96  
Log likelihood ratio..: 1.675162  
P-value.....: 0.99999

6.Location IDs included.: varis, glifadas, ellinikou,  
alimou

Overlap with clusters.: No Overlap

Coordinates / radius..: (23.760595 N, 37.846935 E) / 8.64  
km

Gini Cluster.....: No  
Population.....: 228780  
Number of cases.....: 11  
Expected cases.....: 6.13  
Annual cases / 100000.: 4.8  
Observed / expected...: 1.79  
Relative risk.....: 1.85  
Log likelihood ratio..: 1.629691  
P-value.....: 0.99999

### **3. Results 20%**

#### SUMMARY OF DATA

Study period.....: 2013/1/1 to 2013/12/31  
Number of locations.....: 178

Total population.....: 6678849  
Total number of cases.....: 179  
Annual cases / 100000.....: 2.7

---

---

CLUSTERS DETECTED

1.Location IDs included.: peristeriou, aigalew, iliou,  
xaidariou, petroupolis, anargirwn,  
varvaras, athinaiwn, koridallou,  
filadelfias  
Overlap with clusters.: No Overlap  
Coordinates / radius...: (23.689442 N, 38.013747 E) / 5.81  
km

Gini Cluster.....: Yes  
Population.....: 1252722  
Number of cases.....: 60  
Expected cases.....: 33.57  
Annual cases / 100000.: 4.8  
Observed / expected...: 1.79  
Relative risk.....: 2.18  
Log likelihood ratio..: 10.970609  
P-value.....: 0.00150

2.Location IDs included.: spatwn  
Overlap with clusters.: No Overlap  
Coordinates / radius...: (24.007702 N, 37.974921 E) / 0 km  
Gini Cluster.....: No  
Population.....: 33821  
Number of cases.....: 6  
Expected cases.....: 0.91

Annual cases / 100000.: 17.8  
Observed / expected...: 6.62  
Relative risk.....: 6.81  
Log likelihood ratio...: 6.319937  
P-value.....: 0.11475

3.Location IDs included.: pentelis, vrilissiwn, pallinis,  
dionisou, paraskeuis, kifisias,  
amarousiou

Overlap with clusters.: No Overlap

Coordinates / radius...: (23.869858 N, 38.049105 E) / 6.99

km

Gini Cluster.....: No  
Population.....: 362920  
Number of cases.....: 17  
Expected cases.....: 9.73  
Annual cases / 100000.: 4.7  
Observed / expected...: 1.75  
Relative risk.....: 1.83  
Log likelihood ratio...: 2.377060  
P-value.....: 0.99080

4.Location IDs included.: levadewn, orxomenou, distomou

Overlap with clusters.: No Overlap

Coordinates / radius...: (22.875853 N, 38.440615 E) /

23.23 km

Gini Cluster.....: No  
Population.....: 51124  
Number of cases.....: 4  
Expected cases.....: 1.37  
Annual cases / 100000.: 7.8  
Observed / expected...: 2.92



Relative risk.....: 2.96  
Log likelihood ratio...: 1.675162  
P-value.....: 0.99999

5.Location IDs included.: varis, glifadas, ellinikou,  
alimou  
Overlap with clusters.: No Overlap  
Coordinates / radius...: (23.760595 N, 37.846935 E) / 8.64  
km  
Gini Cluster.....: No  
Population.....: 228780  
Number of cases.....: 11  
Expected cases.....: 6.13  
Annual cases / 100000.: 4.8  
Observed / expected...: 1.79  
Relative risk.....: 1.85  
Log likelihood ratio...: 1.629691  
P-value.....: 0.99999

#### **4. Results 30%**

##### SUMMARY OF DATA

Study period.....: 2013/1/1 to 2013/12/31  
Number of locations.....: 178  
Total population.....: 6678849  
Total number of cases.....: 179  
Annual cases / 100000.....: 2.7

---

---

CLUSTERS DETECTED

1.Location IDs included.: petroupolis, iliou, peristeriou, xaidariou, anargirwn, aigalew, filadelfias, varvaras, filis, koridallou, ionias, athinaiwn

Overlap with clusters.: No Overlap

Coordinates / radius...: (23.684292 N, 38.043833 E) / 7.69 km

Gini Cluster.....: Yes

Population.....: 1365821

Number of cases.....: 64

Expected cases.....: 36.61

Annual cases / 100000.: 4.7

Observed / expected...: 1.75

Relative risk.....: 2.16

Log likelihood ratio..: 11.183941

P-value.....: 0.00147

2.Location IDs included.: spatwn

Overlap with clusters.: No Overlap

Coordinates / radius...: (24.007702 N, 37.974921 E) / 0 km

Gini Cluster.....: No

Population.....: 33821

Number of cases.....: 6

Expected cases.....: 0.91

Annual cases / 100000.: 17.8

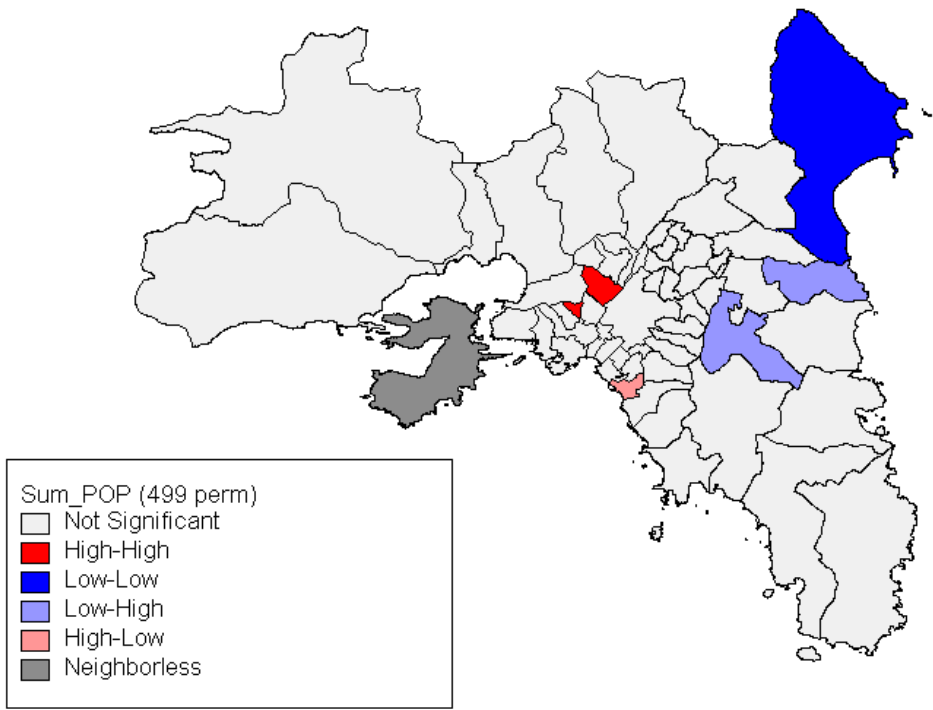
Observed / expected...: 6.62

Relative risk.....: 6.81

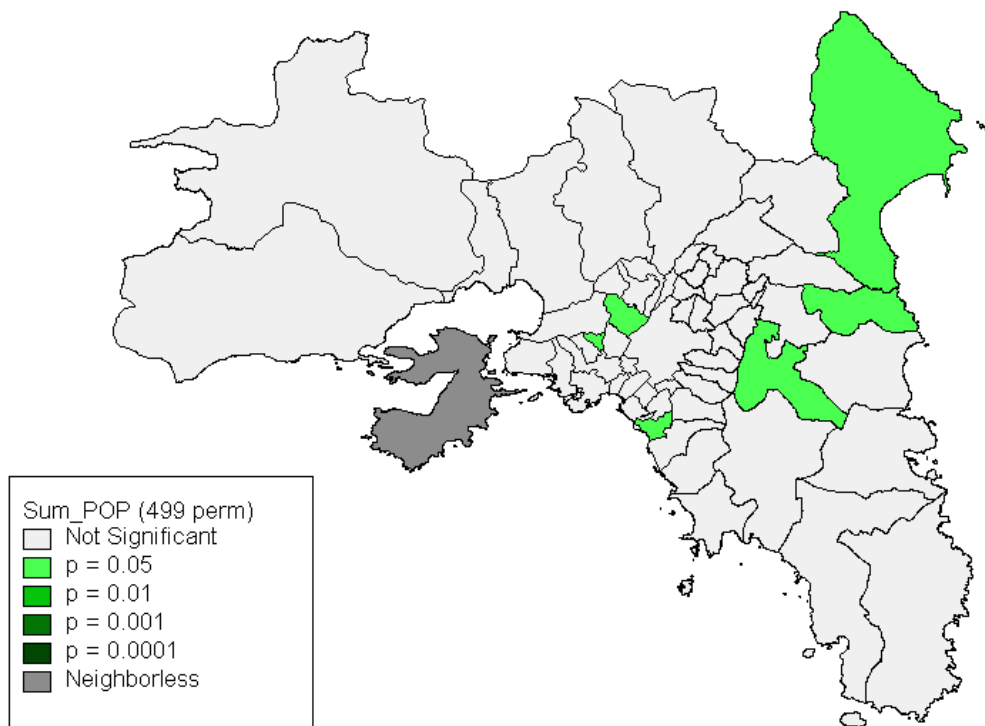
Log likelihood ratio..: 6.319937

P-value.....: 0.12319





Πανεπιστήμιο



Πανεπιστήμιο

# Βιβλιογραφία

## ΕΛΛΗΝΙΚΗ

- Λογοθέτης, Ι. και Μυλωνάς, Ι. (2004). *Νευρολογία Λογοθέτη*, University Studio Press, Θεσσαλονίκη.
- Μπερσίμης, Σ. (2005). *Θεωρία Ροών Επιτυχιών και Εφαρμογές*, Διδακτορική Διατριβή, Παν/μιο Πειραιώς, Πειραιάς.
- Ντζούφρας, Ι. (2006). *Σημειώσεις Βιοστατιστικής II- Επιδημιολογικές Μελέτες*, Παν/μιο Αθηνών, Αθήνα.
- Χατζηπαναγιώτου, Σ., Ανδρεάδου, Ε., Μπερσίμης, Σ., Κυριαζή, Κ., Τσάτση, Α., Μιχαλοπούλου, Μ., Μπουφίδου, Φ. και Νικολάου, Χ. (2010). *Αξιολόγηση του προσδιορισμού ολιγοκλωνικών κλασμάτων IgG στο εγκεφαλονωτιαίο υγρό (ENY), ασθενών με απομυελινωτικές διαταραχές σε σχέση με το δείκτη αυτόχθονης παραγωγής IgG κατά τη διάρκεια τεσσάρων ετών (2006-2009)*, Πανελλήνιο Συνέδριο Ανοσολογίας, Αθήνα.

## ΞΕΝΗ

- Abrial, D., Calavas, D., Lauvergne, N., Morignat, E. and Durcot, C. (2003). Descriptive spatial analysis of BSE in western France, *Veterinary Research*, **34**, 749-760.
- Ala, A., Staanca, C. M., Bu-Ghanim, M., Ahmado, I. et al. (1997). Increased prevalence of primary biliary cirrhosis near superfund toxic waste sites, *Hepatology*, **43**, 525-531.
- Anselin, L. (1995). Local Indicators of Spatial Association LISA, *Geographical Analysis*, **27**, 93-115.
- Anselin, L., Syabri, I. and Kho, Y. (2006). GeoDa: An introduction to spatial data analysis, *Geographical Analysis*, **38**:1, 5-22.
- Ascherio, A., Munger, K. L. and Simon, K. C. (2010). Vitamin D and Multiple Sclerosis, *Lancet*, **9**(6), 599-612.
- Bayon, C., Pei, M. H., Ruiz, C. and Hunter, T. (2007). Genetic structure and spatial distribution of the mycoparasite *Sphaerellopsis filum* on *Melampsora larice-epitea* in a short rotation coppice willow planting, *Plant Pathology*, **56**, 616-623.
- Berke, O., von Keyserlingk, M., Broll, S. and Kreienbrock, L. (2002). On the distribution of *Echinococcus multilocularis* in red foxes in Lower Saxony: identification of a high risk area by spatial epidemiological cluster analysis, *Berliner und Munchener Tierarztliche Wochenschrift*, **115**, 428-434.
- Bersimis, S., Chalkias, C. and Anthopoulou, T. (2014). Detecting and Interpreting Clusters of Economic Activity in Rural areas using Scan Statistic and LISA under a Unified Framework, *Applied Stochastic*

*Models in Business and Industry*, DOI: 10.1002/asmb.2003.

- Besag, J. and Diggle, P. J. (1977). Simple Monte-Carlo tests for spatial pattern, *Applied Statistics*, **26**, 327-333.
- Buntix, F., Geys, H. et al. (2003). Geographical differences in cancer incidence in Belgian province of Limburg, *European Journal of Cancer*, **39**, 2058-2072.
- Compston, A. and Coles, A. (2008). Multiple Sclerosis, *Lancet*, **372**(9648), 1502-17.
- Conover, W. J., Bement, T. R. and Iman, R. L. (1979). On a method for detecting clusters of possible uranium deposits, *Technometrics*, **21**, 277-282.
- Cook, A. J., Gold, D. R. and Li, Y. (2007). Spatial cluster detection for censored outcome data, *Biometrics*, **63**, 540-549.
- Costa, M. A. and Kulldorff, M. (2009). Scan Statistics: Methods and Applications in *Statistics for Industry and Technology*, edited by Glaz, J. et al.
- Cousens, S., Smith, P. G. et al. (2001). Geographical distribution of variant Creutzfeldt-Jacob disease in Great-Britain, *The Lancet*, **357**, 1002-1007.
- Donnan, P. T., Parratt, J. D. E., Wilson, S. V., Forbes, R. B., O'Riordan, J. I. and Swingle, R. J. (2005). Multiple sclerosis in Tayside, Scotland: detection of clusters using a spatial scan statistics, *Multiple Sclerosis*, **11**, 403-408.
- Dwass, M. (1957). Modified randomization tests for nonparametric



- hypotheses, *Annals of Mathematical Statistics*, **28**, 181-187.
- Enemark, H. L., Ahrens, P. et al. (2002). Molecular characterization of Danish *Cryptosporidium parvum* isolates, *Parasitology*, **125**, 331-341.
- Exter, D. J. and Boyle, P. J. (2007). Does young adult suicide cluster geographically in Scotland? *Journal of Epidemiology and Community Health*, **61**, 731-736.
- Fevre, E. M., Coleman, P. G. et al. (2001). The origins of a new *Trypanosoma brucei rhodesiense* sleeping sickness outbreak in eastern Uganda, *The Lancet*, **358**, 625-628.
- Fisher, R. A. (1935). *The design of Experiments*, Oliver and Boyd, Edinburgh.
- Glaz, J., Naus, J. and Wallenstein, S. (2001). *Scan Statistics*, Springer, New York.
- Green, C., Hoppa, R. D., Young, T. K. and Blanchard, J. F. (2003). Geographic analysis of diabetes prevalence in an urban area, *Social Science and Medicine*, **57**, 551-560.
- Guerin, M. T., Martin, S. W., Darlington, G. A. and Rajic, A. (2005). A temporal study of *Salmonella* serovars in animals in Alberta between 1990 and 2001, *Canadian Journal of Veterinary Research*, **69**, 88-89.
- Han, D. W., Rogerson, P. A. et al. (2004). Geographic clustering of residence in early life and subsequent risk of breast cancer (United States), *Cancer Causes and Control*, **15**, 921-929.
- Hanson, C. E. and Wieczorek, W. F. (2002). Alcohol mortality: a

- comparison of spatial clustering methods, *Social Science and Medicine*, **55**, 791-802.
- Hartung, H. I. (2002). Demyelinating diseases, *Current Opinion in Neurology*, **15**:3, 227-283.
- Heres, L., Brus, D. J. and Hagenaars, T. J. (2008). Spatial analysis of BSE cases in the Netherlands, *BMC Veterinary Research*, **4**:21.
- Hills, M. and Alexander, F. (1989). Statistical Methods used in assessing the risk of disease near a source of possible environmental pollution: a review, *Journal of the Royal Statistical Society, Series A*, **152**, 307-325.
- Hjalmars, U., Kulldorff, M., Grustafsson, G. and Nagarwalla, N. (1996). Childhood leukemia in Sweden: using GIS and a spatial scan statistic for cluster detection, *Statistics in Medicine*, **15**, 707-725.
- Hoar, B. R., Chomel, B. B, Rofle, D. L., Chang, C. C., Fritz, C. L., Sacks, B. N. and Carpenter, T. E. (2003). Spatial analysis of *Yersinia pestis* and *Bartonella vinsonii* subsp. *berkhoffi* seroprevalence in California coyotes (*Canis latrans*), *Preventive Veterinary Medicine*, **56**, 299-311.
- Hsu, C. E., Jacobson, H. E. and Soto Mas, F. (2004). Evaluating the disparity of female breast cancer mortality among racial groups – a spatiotemporal analysis, *International Journal of Health Geographics*, **3**:4.
- Jacquez, G. M. (ed.) (1993). Proceedings of the Workshop on Statistics and Computing in Disease Clustering, Port Washington, New York, 1992, *Statistics in Medicine*, **12**, 1751-1968.
- Kuehl, K. S. and Loffredo, C. A. (2006). A cluster of hypoplastic left heart malformation in Baltimore, Maryland, *Pediatric Cardiology*, **27**, 25-31.

- Kulldorff, M. (1997). A Spatial Scan Statistic, *Commun. Statist.-Theory Meth.*, **26**(6), 1481-1496.
- Kulldorff, M. and Nagarwalla, N. (1995). Spatial Disease Clusters: Detection and Inference, *Statistics in Medicine*, **14**, 799-810.
- Kulldorff, M., Feuer, E. J., Miller, B. A. and Freedman, L. S. (1997). Breast cancer in northeastern US: a geographical analysis, *American Journal of Epidemiology*, **146**, 161-170.
- Kulldorff, M., Hefferman, R., Hartman, J. et al. (2005). A space-time permutation scan statistic for the early detection of diseases outbreaks, *PLoS Magazine*, **2**, 216-224.
- Kulle, T., Groff, A., Redmer, J., Hounshell, J. and Schrage, S. (2009). Vitamin D: an evidence-based review, *J Am Board Fam Fed*, **22**(6), 698-706.
- Lassman, H. (2001). Demyelinating diseases, *Current Opinion in Neurology*, edited by Lippincott, Williams, Wilkins, 14:3.
- Loader, C. R. (1961). Large-deviation approximations to the distribution of scan-statistics, *Advances in Applied Probability*, **23**, 751-771.
- Margai, F. and Henry, N. (2003). A community- based assessment of learning disabilities using environmental and contextual risk factors, *Social Science and Medicine*, **56**, 1073-1085.
- Marshall, R. C. (1991). A review of the statistical analysis of spatial patterns of disease, *Journal of the Royal Statistical Society, Series A*, **154**, 421-441.

- Mathews, W. B. (2001). *Multiple Sclerosis: the facts*, 4th ed., Oxford University Press.
- McDonald, W. I., Compston, A., Edan, G. et al. (2001). Recommended diagnostic criteria in multiple sclerosis, *Ann Neurolog*, **50**, 121-127.
- Miller, M. A., Gardner, I. A., Kreuder, C., Paradies, D. M., Worcester, K. R., Jessup, D. A., Dodd, E., Harris, M. D., Ames, J. A., Packham, A. E. and Conrad, P. A. (2002). Coastal freshwater runoff is a risk factor for *Toxoplasma gondii* infection of southern sea otters (*Enhydra lutris nereis*), *International Journal for Parasitology*, **32**, 997-1006.
- Milo, R. and Kahana, E. (2010). Multiple Sclerosis: geoepidemiology, genetics and the environment, *Autoimmun Rev*, **9**(5), A387-94.
- Milonas, I. (2003). *A critique of therapeutic trials in MS ENS*, 13th meeting, Teaching course book.
- Milonas, I., Tsonuis, S. and Logothetis, I. (1990). Epidemiology of multiple sclerosis in northern Greece, *Acta Neurolog Scandin*, **81**, 43-47.
- Mostashari, F., Kulldroff, M. et al. (2003). Dead bird clustering: a potential early warning system for West Nile virus activity, *Emerging Infectious Diseases*, **9**, 641-646.
- Murray, T. J. (2006). Diagnosis and Treatment of Multiple Sclerosis, *British Medical Journal*, **332**(7540), 525-527.
- Naus, J. I. (1965a). The distribution of the size of the maximum cluster of points on the line, *Journal of the American Statistical Association*, **60**, 532-538.

- Naus, J. I. (1965b). Clustering of random points in two dimensions, *Biometrika*, **52**, 263-267.
- Nkhoma, E. T., Hsu, C. E., Hunt, V. I. and Harris, A. M. (2004). Detecting spatiotemporal clusters of accidental poisoning mortality among Texas counties, U.S., 1980-2001, *International Journal of Health Geographics*, **3**:25.
- Norstrom, M., Pfeiffer, D. U. and Jarp, J. (2000). A space-time cluster investigation of an outbreak of acute respiratory disease in Norwegian cattle herds, *Preventive Veterinary Medicine*, **47**, 107-119.
- Oden, N., Jacquez, G. and Grimson, R. (1996). Realistic power simulations compare point and area-based disease cluster tests, *Statistics in Medicine*, **15**, 783-806.
- Odoi, A., Martin, S. W. et al. (2004). Investigation of clusters of giardiasis using GIS and a spatial scan statistic, *International Journal of Health Geographics*, **3**:11.
- Olea-Popelka, F. J., Griffin, J. M., Collins, J. D., McGrath, G. and Martin, S. W. (2003). Bovine tuberculosis in badgers in four areas in Ireland: does tuberculosis cluster? *Preventive Veterinary Medicine*, **59**, 103-111.
- Openshaw, S., Charlton, M., Wymer, C. and Craft, A. (1987). A Mark I geographical analysis machine for the automated analysis of point datasets, *International Journal of Geographical Information Systems*, **1**(4), 3353-58.
- Orford, K. J. (2000). The analysis of cosmic ray data, *Journal of Physics G-Nuclear and Particle Physics*, **26**(4), R1-R26.

- Paty, D. W., Ebers, G. C. (1998). *Multiple Sclerosis*, Philadelphia: Davis.
- Porcasi, X., Catala, S. S., Hrellac, H., Scavuzzo, M. C., Gorla, D. E. (2006). Infestation of rural houses by *Triatoma Infestans* (Hemiptera: *Rediriidae*) in Southern area of Gran Chaco in Argentina, *Journal of Medical Entomology*, **43**, 1060-1067.
- Reperant, L. A. and Deplazes, P. (2005). Cluster of *Capillaria hepatica* infections in non-commensal rodents from the canton of Geneva, Switzerland, *Parasitology Research*, **96**, 340-342.
- Rothman, K. J. (1990). A sobering start for the cluster busters' conference, *American Journal of Epidemiology*, **132**, S6-S13.
- Sabel, C. E., Boyle, P. J., Lythen, M., Gartell, A. C., Jokelainen, M., Flowerdew, R. and Maasilta, P. (2003). Spatial clustering of amyotrophic lateral sclerosis in Finland at place of birth and place of death, *American Journal of Epidemiology*, **157**, 898-905.
- Sankoh, O. A., Ye, Y., Sauerborn, R., Muller, O. and Becher, H. (2001). Clustering of childhood mortality in rural Burkina Faso, *International Journal of Epidemiology*, **30**, 485-492.
- Sauders, B. D. and Fortes, E. R. (2003). Molecular subtyping to detect human listeriosis clusters, *Emerging Infectious Diseases*, **9**, 672-680.
- Sheridan, H. A., McGrath, G., White, P., Fallon, R., Shoukri, M. M. and Martin, S. W. (2005). A temporal-spatial analysis of bovine spongiform encephalopathy in Irish cattle herds, from 1996 to 2000, *Canadian Journal of Veterinary Research*, **69**, 19-25.
- Smith, K. L., DeVos, V., Bryden, H., Price, L. B., Hugh-Jones, M. E. and

- Keim, P. (2000). *Bacillus anthracis* diversity in Kruger National Park, *Journal of Clinical Microbiology*, **38**, 3780-3784.
- Spindler, B. D., Chipps, S. R., Klumb, R. A. and Wimberly, M. C. (2009). Spatial analysis of pallid sturgeon *Scaphirhynchus albus* distribution in the Missouri river, S. Dakota, *Journal of Applied Ichthyology*, **25**, 8-13.
- Sudakin, D. L., Horowitz, Z. and Griffin, S. (2002). Regional variation in the incidence of symptomatic pesticide exposures: applications of geographic information systems, *Journal of Toxicology – Clinical Toxicology*, **40**, 767-773.
- Thomas, A. J. and Carlin, B. P. (2003). Late detection of breast and colorectal cancer in Minnesota counties: an application of spatial smoothing and clustering, *Statistics in Medicine*, **22**, 113-127.
- Tuia, D., Ratle, F., Lasaponara, R., Telesca, L. and Kanevski, M. (2008). Scan statistics analysis of forest fire clusters, *Communications in Nonlinear Science and Numerical Simulation*, **13**, 1689-1694.
- Turnbull, B. W., Iwano, E. J., Burnett, W. S., Howe, H. L. and Clark, L. C. (1990). Monitoring for clusters of disease: applications to leukemia incidence in upstate New York, *American Journal of Epidemiology*, **132**, S136-S143.
- Turnbull, B., Iwano, E. J., Burnett, W. S., Howe, H. L. and Clark, L. C. (1990). Monitoring for clusters of disease: application to leukemia incidence in upstate New York, *American Journal of Epidemiology*, **132**, 136-143.
- Usher, B. M. and Allen, K. L. (2005). Identifying kinship clusters: SaTScan

- for genetic spatial analysis, *American Journal of Physical Anthropology, Supplement*, **126**, S40, 210.
- Viel, J. F., Arreux, P., Baverel, L. and Cahn, J. Y. (2000). Soft-tissue sarcoma and non-Hodgkins lymphoma clusters around a municipal solid waste incinerator with high dioxin emission levels, *American Journal of Epidemiology*, **152**, 13-19.
- Wallenstein, S., Naus, J. and Glaz, J. (1993). Power of the scan statistic for detection of clustering, *Statistics in Medicine*, **12**, 1829-1843.
- Waller, L. A., (2006). *Detection of Clustering in Spatial Data*, Emory University, Biostatistics dept, Technical Report **06-12**.
- Walsh, S. J. and Dechello, J. R. (2001). Geographical variation in mortality from systemic lupus erythematosus in the United States, *Lupus*, **10**, 637-646.
- Ward, M. P. (2001). Blowfly strike in sheep flocks as an example of the use of a space-time scan statistic to control confounding, *Preventive Veterinary Medicine*, **49**, 61-69.
- Ward, M. P. (2002). Clustering of reported cases of leptospirosis among dogs in the United States and Canada, *Preventive Veterinary Medicine*, **56**, 215-226.
- Witham, C. S. and Oppenheimer, C. (2004). Mortality in England during the 1783-4 Laki Craters eruption, *Bulletin of Volcanology*, **67**, 15-25.
- Wylie, J. L., Cabral, T. and Jolly, A. M. (2005). Identification of networks of sexually transmitted infection: a molecular, geographic and social network analysis, *Journal of Infectious Diseases*, **191**, 899-906.



Yiannakoulis, N., Rowe, B. H., Svenson, L. W., Schopflocher, D. P., Kelly, K. and Voaklander, D. C. (2003). Zones of prevention: the geography of injuries in the elderly, *Social Science and Medicine*, **57**, 2065-2073.

## ΑΛΛΕΣ ΠΗΓΕΣ

[www.doctorswithoutborders.org](http://www.doctorswithoutborders.org)

[www.iatronet.gr/iatriko-lexiko.html](http://www.iatronet.gr/iatriko-lexiko.html)

[www.neurocenter.gr/sklirinsi.html#t1](http://www.neurocenter.gr/sklirinsi.html#t1)

[www.news-medical.net/health/Ganglion.aspx](http://www.news-medical.net/health/Ganglion.aspx)

[www.wikipedia.org/wiki/Giardiasis](http://www.wikipedia.org/wiki/Giardiasis)

[www.wikipedia.org/wiki/Capillaria\\_hepatica](http://www.wikipedia.org/wiki/Capillaria_hepatica)

[www.wikipedia.gr/Jacob\\_disease](http://www.wikipedia.gr/Jacob_disease)

SaTScan User's Guide v.9.23