

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

**ΔΠΜΣ: «ΜΑΘΗΜΑΤΙΚΗ ΠΡΟΤΥΠΟΠΟΙΗΣΗ ΣΤΙΣ ΣΥΓΧΡΟΝΕΣ
ΤΕΧΝΟΛΟΓΙΕΣ ΚΑΙ ΤΗΝ ΟΙΚΟΝΟΜΙΑ»**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ



**«Η συμβολή και οι εκτεταμένες χρήσεις του Στατιστικού Ελέγχου
Διεργασιών στον Τομέα της Υγείας».**

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ: κος ΚΟΥΚΟΥΒΙΝΟΣ ΧΡΗΣΤΟΣ

ΟΝΟΜΑΤΕΠΩΝΥΜΟ ΦΟΙΤΗΤΗ: ΤΣΙΛΙΒΗΣ ΒΑΣΙΛΕΙΟΣ

ΑΘΗΝΑ 2012

ΠΕΡΙΛΗΨΗ

Τα τελευταία χρόνια, ο Στατιστικός Έλεγχος Διεργασιών έχει παρουσιάσει ραγδαία εξάπλωση σε όλους τους τομείς των παραγωγικών κοινωνιών, ανά τον κόσμο, από την Οικονομία και το Βιομηχανικό κλάδο, μέχρι τον τομέα της Παραγωγής και της Υγείας. Σκοπός της παρούσας εργασίας, είναι η αναλυτική περιγραφή και η εκτενής παρουσίαση του τρόπου με τον οποίον συντελείται η εφαρμογή του Σ.Ε.Δ. σε ιατρικές και κλινικές πρακτικές. Προφανώς, σε επεμβάσεις που λαμβάνουν χώρα σε ιατρικά κέντρα και νοσοκομειακά ιδρύματα, είναι κομβικής σημασίας να παρακολουθούμε τα αποτελέσματα που σχετίζονται με τις εκάστοτε επεμβάσεις.

Στη μεταμόσχευση οργάνων, είναι απαραίτητη η παρακολούθηση και αξιολόγηση των αποτελεσμάτων των μεταμοσχεύσεων που διεξάγονται στα εν λόγω κέντρα, καθώς και η έγκαιρη προειδοποίηση σε περίπτωση ραγδαίας αύξησης τους ποσοστού αποτυχίας τους. Μία τέτοια λειτουργική και καινοτόμος παρουσία του Σ.Ε.Δ. στον τομέα της Υγείας θα μπορούσε να συντελέσει όχι μόνο στη βελτίωση της αποδοτικότητας των χειρουργικών και νοσοκομειακών μονάδων, αλλά θα μπορούσε ωστόσο να συμβάλει στη θεραπεία χρόνιων νοσημάτων και ασθενειών που θεωρούνταν ανίατες έως τώρα.

Το πρώτο κεφάλαιο, αποτελεί μια εισαγωγή στη γενικότερη θεωρία και φιλοσοφία του Στατιστικού Ελέγχου Διεργασιών, μέσα από την ιστορική αναδρομή της ποιότητας και των βασικών συνιστωσών της, καθώς και των κυριότερων εργαλείων ελέγχου της απόδοσης διεργασιών: των διαγράμματα ελέγχου.

Το δεύτερο κεφάλαιο, πραγματεύεται την παρουσία και χρήση διαφορετικών μεθόδων αξιολόγησης των αποτελεσμάτων, σε εφαρμογές της Υγείας. Η ανάλυση και επεξήγηση των προτερημάτων και μειονεκτημάτων καθεμιάς, μέσα από ένα πλαίσιο σύγκρισης μεταξύ τους, οδηγεί σε ενδιαφέροντα συμπεράσματα σχετικά με τη συνεισφορά τους σε κλινικές και ιατρικές εφαρμογές.

Στο τρίτο κεφάλαιο, αναλύεται ο παράγοντας κινδύνου, ο οποίος χαρακτηρίζει κάθε ασθενή, που πάσχει από μία συγκεκριμένη ασθένεια. Ο κίνδυνος αυτός ο οποίος καθορίζεται από διάφορους παράγοντες, καθορίζει σημαντικά την εξάπλωση και τις επιπτώσεις μιας ασθένειας στη υγεία του εν λόγω ασθενούς. Συγκεκριμένα, η έννοια του κινδύνου αποκτάει ιδιαίτερο νόημα, μέσα από την εφαρμογή της μεθόδου Cusum στην καρδιοχειρουργική επέμβαση.

Τέλος, στο τέταρτο κεφάλαιο, παρουσιάζεται μία εφαρμογή της μεθόδου Cusum, βάσει του μοντέλου Cox, στη μεταμόσχευση νεφρού. Συγκεκριμένα, εφαρμόζονται οι διεργασίες Cusum σε δεδομένα μεταμόσχευσης νεφρού όπως συλλέχθηκαν από 258 κέντρα μεταμόσχευσης των Ηνωμένων Πολιτειών και περιέχονται στη βάση δεδομένων στατιστικής μεταμοσχεύσεων: «*Scientific Registry of Transplant Recipients (SRTR)*».

ABSTRACT

Over the past few years, the Statistical Quality Control has met a huge expansion in the field of productive societies, in all over the world: from the field of economics and industries to production and health applications. The goal is the analytical and detailed presence of the way according to which, the Statistical Quality Control is applied to medical and health applications. Apparently, it is vitally important to record and assess the outcomes of the applications taking place in medical centers and hospital units.

In transplants for example, apart from monitoring and assessing the outcomes, it is equally important the timely warning, in case of an immediate increase in the failure rate. Such practical and innovative presence of SQC in health applications, could contribute not only to the improvement of surgical and medical improvement, but also to the elimination of chronic diseases that considered to be incurable until now.

The first chapter is an introduction to general theory and philosophy related to the Statistical Quality Control, in the frame of quality historical retrospection, the basic parameter estimation techniques, and the main tools which are used in terms of control: the control charts.

The second chapter deals with presence and usage of different assessment methods, in health applications. The analysis and explanation of the advantages and disadvantages of each one, by comparing them each other, leads to interesting conclusions about the contribution of SQC to clinical and medical applications.

In the third chapter, the risk factor, which characterizes any patient that suffers from a specific disease, is discussed extensively. That risk which is defined by several factors, determines the expansion and the effect of the specific disease in the patient's health. Specifically, the meaning of the term "danger" becomes quite clear, in the frame of an application of Cusum method in open-heart surgery.

Finally, in the third chapter, an application of the Cusum method, based on Cox model, is presented, in kidney transplants. Specifically, Cusum procedures are applied in kidney transplant data, which were collected from 250 transplant centers in the United States and are included in the database of SRTR (*Scientific Registry of Transplant Recipients*).

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω τον καθηγητή μου, κ. Κουκουβίνο Χ., για την ανάθεση της εργασίας, την καθοδήγηση, τις αλάνθαστες υποδείξεις και την ευκαιρία μίας εποικοδομητικής συνεργασίας πάνω σε ένα γνωστικό αντικείμενο το οποίο με ενθουσίασε ιδιαίτερος την περίοδο που το διδάχθηκα. Οι επιστημονικές γνώσεις του στον τομέα αυτό, η εμπειρία, ο επαγγελματισμός, η αφοσίωση και η διάθεση του για εποικοδομητική συνεργασία συνέβαλαν καθοριστικά στη θετική μου στάση και ενδιαφέρον ως προς την αποτελεσματική και δημιουργική προετοιμασία του συγκεκριμένου θέματος.

Επίσης, θα ήθελα να ευχαριστήσω θερμά τον υποψήφιο διδάκτορα Ανδρουλάκη Εμμανουήλ, για την καθοδήγηση και πολύτιμη βοήθεια που μου προσέφερε καθ' όλη τη διάρκεια εκπόνησης και προετοιμασίας της συγκεκριμένης εργασίας, όπως επίσης και για τις πολύ εύστοχες παρατηρήσεις και υποδείξεις που συνέβαλαν στην άρτια παρουσίαση της.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένεια μου για την ανεκτίμητη βοήθεια και συμπαράσταση που μου παρείχαν με κάθε δυνατό τρόπο, ηθικά και ψυχολογικά, καθ' όλη τη διάρκεια των σπουδών μου. Για την αγάπη, την αφοσίωση, την κατανόηση, την υπομονή, την υποστήριξη και γενικώς για όλα όσα μου προσέφεραν τους οφείλω την αμέριστη ευγνωμοσύνη μου και όλη μου την αγάπη.

ΠΕΡΙΕΧΟΜΕΝΑ

Σελίδα

ΠΕΡΙΛΗΨΗ.....	1
ABSTRACT.....	3
ΕΥΧΑΡΙΣΤΙΕΣ.....	4

ΚΕΦΑΛΑΙΑ

1. Η ΜΕΘΟΔΟΛΟΓΙΑ ΤΟΥ ΣΤΑΤΙΣΤΙΚΟΥ ΕΛΕΓΧΟΥ ΔΙΕΡΓΑΣΙΩΝ

1.1. Εισαγωγή.....	7
1.1.1. Η έννοια της Ποιότητας και οι Διαστάσεις της.....	8
1.1.2. Ιστορική Αναδρομή της Ποιότητας.....	9
1.1.3. Συνιστώσες του Στατιστικού Ελέγχου Ποιότητας.....	13
1.2. Το πρόβλημα του Στατιστικού Ελέγχου Διεργασιών.....	14
1.2.1. Περιγραφή και χρήση ενός Διαγράμματος Ελέγχου.....	17
1.3. Ο Σ.Ε.Δ στον τομέα της Υγείας: Αξιολόγηση ιατρικών και νοσοκομειακών εφαρμογών.....	22
1.3.1. Η συμβολή της ανάλυσης κλινικών περιστατικών.....	24
1.3.2. Η συνεισφορά του Στατιστικού Ελέγχου Διεργασιών.....	25
1.3.3. Εφαρμογές του Σ.Ε.Δ στις Υπηρεσίες Υγείας.....	28
1.4. Κριτήρια αξιολόγησης διαγραμμάτων ελέγχου: ARL, ANOS, ATS.....	29

2. ΜΕΘΟΔΟΙ ΑΞΙΟΛΟΓΗΣΗΣ ΚΑΙ ΣΥΓΚΡΙΣΗΣ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

2.1. Εφαρμογές διαγραμμάτων ελέγχου και μεθόδων πιθανοφάνειας.....	34
2.1.1. Μέθοδοι CUSUM και EWMA – Πλεονεκτήματα & Μειονεκτήματα.....	35
2.1.2. <i>Resetting Sequential Probability Ratio Test:</i> Εφαρμογές στον τομέα της Υγείας.....	43
2.1.3. <i>Sets Method:</i> Ανίχνευση αύξησης στο ρυθμό σπανίων ασθενειών.....	48

2.2. League Tables, Διαγράμματα σύγκρισης και Funnel Charts.....	49
2.2.1. Ανάλυση και Σύγκριση μεθόδων.....	49
2.2.2. Συμπεράσματα.....	52

3. Η ΠΡΟΣΑΡΜΟΓΗ ΤΟΥ ΚΙΝΔΥΝΟΥ ΣΤΑ ΔΙΑΓΡΑΜΜΑΤΑ ΕΛΕΓΧΟΥ

3.1. Η έννοια του κινδύνου στην εξάπλωση και τις επιπτώσεις μίας ασθένειας.....	55
3.2. Μελέτη των risk-adjusted αποτελεσμάτων στην καρδιοχειρουργική επέμβαση.....	59
3.2.1. Επισκόπηση της ενότητας.....	59
3.2.2. Ασθενείς και Μέθοδοι: Μελέτη, Ανάλυση και Συμπεράσματα	61
3.3. Καταγραφή και παρακολούθηση της χειρουργικής δραστηριότητας με χρήση διαγραμμάτων <i>risk-adjusted</i> σωρευτικών αθροισμάτων.....	67
3.3.1. Standard CUSUM / Risk – adjusted CUSUM: Περιγραφή και σύγκριση των διαγραμματικών μεθόδων.....	70
3.3.2. Εφαρμογή του risk-adjusted CUSUM σε καρδιοχειρουργική επέμβαση.....	75
3.3.3. Χαρακτηριστικά της <i>risk-adjusted</i> διεργασίας CUSUM	80

4. ΔΙΑΓΡΑΜΜΑ CUSUM ΣΕ ΣΥΝΕΧΗ ΧΡΟΝΟ, ΒΑΣΕΙ ΤΟΥ ΜΟΝΤΕΛΟΥ

COX: ΕΦΑΡΜΟΓΗ ΣΤΗ ΜΕΤΑΜΟΣΧΕΥΣΗ ΟΡΓΑΝΩΝ

4.1. Το risk-adjusted CUSUM βάσει του μοντέλου Cox: Εισαγωγή και Υποθέσεις.....	83
4.2. Το διάγραμμα CUSUM σε συνεχή και διακριτό χρόνο.....	87
4.3. Έλεγχος και αξιολόγηση της μεθόδου, με χρήση προσομοιώσεων.....	91
4.3.1. Ελέγχοντας το ποσοστό των εσφαλμένων «συναγερμών»	96
4.4. Εφαρμογή σε επέμβαση μεταμόσχευσης νεφρού.....	100
4.4.1. Προσαρμογή κινδύνου	100
4.4.2. Επιλογή Ορίων Ελέγχου	105

BIBΛΙΟΓΡΑΦΙΑ

107

ΚΕΦΑΛΑΙΟ 1

Η μεθοδολογία του Στατιστικού Ελέγχου Διεργασιών

1.1. Εισαγωγή

Κάθε άνθρωπος που αποφασίζει να αγοράσει ένα προϊόν (ή υπηρεσία) αντιμετωπίζει το πρόβλημα της επιλογής ανάμεσα σε ομοειδή προϊόντα που ικανοποιούν την ίδια ανάγκη. Η επιλογή του προϊόντος από τον καταναλωτή καθορίζεται κυρίως από δύο παράγοντες, την ποιότητά του και την τιμή του.

Η αναγνώριση αυτών των δύο κύριων παραγόντων οι οποίοι καθορίζουν τις αγοραστικές συνήθειες των καταναλωτών οδήγησε τις επιχειρήσεις στην αναζήτηση επιχειρηματικών τακτικών που θα οδηγήσουν στη βέλτιστη δυνατή ποιότητα των παραγόμενων προϊόντων με το ελάχιστο δυνατό κόστος. Η αναζήτηση αυτή οδήγησε σε μια σειρά παρεμβάσεων στο επιχειρηματικό περιβάλλον οι οποίες τελικά διαμόρφωσαν νέα πλαίσια λειτουργίας και διοίκησης των επιχειρήσεων με τελικό στόχο και σκοπό τη διαρκή βελτίωση της ποιότητας. Η στρατηγική / φιλοσοφία που επιζητεί τη συνεχή βελτίωση στην ποιότητα εκτέλεσης όλων των διεργασιών, προϊόντων και υπηρεσιών, σε μια επιχείρηση ονομάζεται Διοίκηση Ολικής Ποιότητας (Total Quality Management). Σύμφωνα με αυτή τη φιλοσοφία “η διοίκηση πρέπει να κατευθύνει και το εργατικό δυναμικό να ενεργεί με στόχο τη βελτίωση της ποιότητας σε κάθε επίπεδο λειτουργίας της επιχείρησης”. Σημαντικό ρόλο στη διαμόρφωση των νέων πλαισίων λειτουργίας και διοίκησης των επιχειρήσεων διαδραματίζει ο Διεθνής Οργανισμός Τυποποίησης (ISO, International Standards Organization) ο οποίος με τα πρότυπα ISO καθορίζει τις διαδικασίες που πρέπει να υιοθετηθούν από μια επιχείρηση με τελικό σκοπό τη συνεχή προαγωγή και διασφάλιση της ποιότητας. Σημαντικό ρόλο στις διαδικασίες αυτές κατέχει ο Στατιστικός Έλεγχος Ποιότητας (Statistical Quality Control).

Ο Στατιστικός Έλεγχος Ποιότητας αποτελεί την παλαιότερη και γνωστότερη μέθοδο ελέγχου παραγωγικών διεργασιών για τη βελτίωση της ποιότητας των παραγόμενων προϊόντων. Ένας από τους βασικούς στόχους του είναι η έγκαιρη ανακάλυψη μη συμμορφωμένων με τις προδιαγραφές παραγόμενων προϊόντων η οποία σηματοδοτεί τη λήψη διορθωτικών ενεργειών για την απομάκρυνση των αιτιών που είναι υπεύθυνες για τις αποκλίσεις, συμβάλλοντας έτσι στη διατήρηση της ποιότητας των προϊόντων. Θα μπορούσαμε να πούμε ότι ο Στατιστικός Έλεγχος Ποιότητας επηρεάζει σημαντικά αποφάσεις σχετικές με τις προδιαγραφές, την παραγωγή και τον έλεγχο των παραγόμενων προϊόντων μιας επιχείρησης.

1.1.1. Η Έννοια της Ποιότητας και οι Διαστάσεις της

Η έννοια της ποιότητας μπορεί να οριστεί με πολλούς τρόπους. Οι περισσότεροι άνθρωποι έχουν μια διαισθητική άποψη για το τι είναι ποιότητα και τι συνδέουν με τα επιθυμητά χαρακτηριστικά / διαστάσεις που πρέπει να έχει ένα προϊόν (ή μια υπηρεσία). Οι διαστάσεις αυτές, βάσει των οποίων ο καταναλωτής αξιολογεί τα προϊόντα, καθορίζουν την τελική ποιότητα ενός προϊόντος και σχετίζονται με την ποιότητα του σχεδιασμού και τη διαδικασία παραγωγής του προϊόντος.

Σύμφωνα με τα όσα αναφέρθηκαν για τις διαστάσεις της ποιότητας μπορούμε να ορίσουμε ότι ποιότητα είναι η προσαρμογή των χαρακτηριστικών ενός προϊόντος (ή υπηρεσίας) στις απαιτήσεις του καταναλωτή / χρήστη, δηλαδή ότι ποιότητα σημαίνει καταλληλότητα προς χρήση (παραδοσιακός ορισμός της ποιότητας). Έτσι η ποιότητα μπορεί να “μετρηθεί” μέσω της ικανοποίησης που προσφέρει το προϊόν στον καταναλωτή. Δυστυχώς ο παραπάνω ορισμός έχει συσχετιστεί περισσότερο με την έννοια της συμμόρφωσης των παραγόμενων προϊόντων με τις προδιαγραφές που πρέπει να πληρούν τα παραγόμενα προϊόντα, και λιγότερο με την έννοια της ποιότητας του σχεδιασμού των. Για παράδειγμα όλα τα αυτοκίνητα αποσκοπούν στην ικανοποίηση της ανάγκης ασφαλούς μεταφοράς των ανθρώπων, ωστόσο διαφέρουν σε μέγεθος, εμφάνιση, απόδοση, εφόδια, κτλ.

Οι διαφορές οφείλονται σε διαφορετικές σχεδιαστικές μεθόδους που εξαρτώνται από τον τύπο των υλικών που χρησιμοποιούνται στην παραγωγή, τις προδιαγραφές των εξαρτημάτων, το σχεδιασμό και την ανάπτυξη νέων κινητήρων, κτλ. Ένας νεότερος ορισμός ορίζει την ποιότητα ως αντιστρόφως ανάλογη της μεταβλητότητας των χαρακτηριστικών της παραγωγικής διαδικασίας που προσδιορίζουν την ποιότητα του προϊόντος.

1.1.2. Ιστορική Αναδρομή της Ποιότητας

Ο ποιοτικός έλεγχος, όπως προαναφέρθηκε είναι μια διαδικασία κατά την οποία καταγράφονται οι αποκλίσεις των προϊόντων, υπηρεσιών από δεδομένες προδιαγραφές. Για να εξασφαλισθεί όμως ότι το τελικό προϊόν, θα είναι σύμφωνο με τις συγκεκριμένες προδιαγραφές, χρησιμοποιείται το σύστημα διασφάλισης ποιότητας. Όταν οι έλεγχοι σε προϊόντα, ακόμα και σε υπηρεσίες, είναι περιοδικοί, τότε μειώνεται η πιθανότητα να εμφανιστούν παρεκκλίσεις ή ελαττώματα από το μέσο επίπεδο της εκάστοτε διεργασίας. Όταν ελέγχονται, για παράδειγμα όλες οι μονάδες στο μηχανισμό παροχής ιατρικών υπηρεσιών μίας Καρδιοχειρουργικής Μονάδας, εκμηδενίζεται ή τουλάχιστον ελαττώνεται σημαντικά η πιθανότητα εμφάνισης σφαλμάτων κατά την εξέλιξη μίας συγκεκριμένης διαδικασίας. Συνεπώς, μειώνεται και το αντίστοιχο κόστος παραγωγής ή προώθησης προϊόντων / υπηρεσιών.

Αντιθέτως, αν οι ποιοτικοί έλεγχοι είναι ελάχιστοι, ή δε συντελούνται καθόλου, τότε το συνολικό κόστος αυξάνεται από τα ελαττωματικά προϊόντα και όχι από τους επαναλαμβανόμενους ποιοτικούς ελέγχους. Επομένως, ένας κατάλληλος αριθμός ελέγχων ποιότητας, δίνει τα καλύτερα αποτελέσματα.

Ο Στατιστικός Έλεγχος Ποιότητας (Statistical Quality Control) είναι ένα κρίσιμο εργαλείο της επιχείρησης που επηρεάζει αποφάσεις σχετικές με τις προδιαγραφές, την παραγωγική διαδικασία και τον έλεγχο των παραγόμενων προϊόντων. Για την αποτελεσματική χρήση του Στατιστικού Ελέγχου Ποιότητας απαιτείται η ενσωμάτωσή του σε ένα πλαίσιο λειτουργίας και διοίκησης της επιχείρησης που έχει ως στόχο τη διαρκή βελτίωση της ποιότητας σε όλα τα επίπεδα της επιχείρησης, γνωστού ως Διοίκησης Ολικής Ποιότητας (Total Quality Management ή Total Quality Assurance). Θα μπορούσαμε να πούμε ότι ο Στατιστικός Έλεγχος Ποιότητας αποτέλεσε τον προάγγελο της Διοίκησης Ολικής Ποιότητας. Η ιστορία της ποιότητας ξεκινά στις αρχές του 20ου αιώνα με τη δημιουργία των πρώτων εργαστηρίων για πρότυπα στη Μεγάλη Βρετανία και με τη γραμμή παραγωγής στην αυτοκινητοβιομηχανία Ford στο Highland Park στις ΗΠΑ (1905). Τα βήματα στη συνέχεια ήταν αλματώδη και σήμερα η ποιότητα έχει αναδειχθεί σε κύριο μέλημα τόσο της βιομηχανίας όσο και των υπηρεσιών.

Στον πίνακα της επόμενης σελίδας, παρουσιάζονται τα κυριότερα ιστορικά σημεία στην ανάπτυξη της ποιότητας από το 1900 μέχρι σήμερα.

Πίνακας 1.2. Ιστορική αναδρομή της ποιότητας

1901	Τα πρώτα εργαστήρια προτύπων (standards) ιδρύονται στη Μ. Βρετανία
1907	Η AT&T Bell Laboratories αρχίζει τη συστηματική επιθεώρηση και έλεγχο προϊόντων και υλικών
1919	Η Ένωση Τεχνικών Επιθεωρητών ιδρύεται στην Αγγλία η οποία αργότερα μετονομάζεται σε Ινστιτούτο Διασφάλισης της Ποιότητας
1920	Στα εργαστήρια της AT&T Bell Laboratories ιδρύεται τμήμα ποιότητας
1924	Ο W. A. Shewhart εισάγει τα διαγράμματα ελέγχου σε ένα Technical Report στην AT&T Bell
1928	Το δειγματοληπτικό σχέδιο αποδοχής σωρού αναπτύσσεται από τους Dodge και Romig
1931	Ο W.A. Shewhart εκδίδει το περιοδικό Economic Control of Quality of Manufactured Product
1932	Ο W.A. Shewhart δίνει διαλέξεις σε στατιστικές μεθόδους στην παραγωγή και στα διαγράμματα ελέγχου στο Πανεπιστήμιο του Λονδίνου
1938	Ο W.E. Deming προσκαλεί τον Shewhart για σεμινάρια στα διαγράμματα ελέγχου στο U.S. Department of Agriculture Graduate School
1940	Το Υπουργείο Πολέμου των Η.Π.Α. εκδίδει ένα οδηγό για την ανάλυση δεδομένων με χρήση διαγραμμάτων ελέγχου
1946	Ιδρύεται η American Society for Quality Control (ASQC)
1946	Ο W.E. Deming προσκαλείται στην Ιαπωνία για να δώσει σεμινάρια Στατιστικού Ποιοτικού Ελέγχου
1948	Ο Καθηγητής G. Taguchi αρχίζει τη μελέτη των πειραματικών σχεδιασμών
1950	Ο W.E. Deming αρχίζει την εκπαίδευση ανώτατων στελεχών βιομηχανιών της Ιαπωνίας
1950	Ο K. Ishikawa εισάγει το διάγραμμα αιτίου-αποτελέσματος (cause and effect diagram)
1954	Ο J. M. Juran προσκαλείται από την Ιαπωνία για να δώσει διαλέξεις σε θέματα διοίκησης και βελτίωσης της ποιότητας. Ο E.S. Page εισάγει το διάγραμμα ελέγχου CUSUM
1959	Ο S. Roberts εισάγει το διάγραμμα ελέγχου EWMA. Ξεκινά η έκδοση του επιστημονικού περιοδικού Technometrics
1960	Η ιδέα των κύκλων ποιότητας εισάγεται από τον K. Ishikawa
1969	Τα περιοδικά Quality Progress και Journal of Quality Technology αρχίζουν να εκδίδονται
1975	Εμφανίζονται τα πρώτα βιβλία σε σχεδιασμό πειραμάτων
1989	Αρχίζει να εκδίδεται το περιοδικό Quality Engineering. Η Motorola εισάγει την έννοια six-sigma.
1990	Σταδιακή αύξηση της ζήτησης στην βιομηχανία για πιστοποίηση κατά ISO 9000
1997	Η προσέγγιση six-sigma της Motorola υιοθετείται και από άλλες βιομηχανίες

Στην Ιαπωνία, μετά το τέλος του πολέμου, η ανέλιξη νεαρών ατόμων στην ιεραρχία των επιχειρήσεων που ήταν δεκτικά σε νέες ιδέες, οδήγησε διακεκριμένους Αμερικανούς επιστήμονες να επισκεφθούν τότε την Ιαπωνία, αφού οι ιδέες τους ήταν πιθανό να βρουν πρόσφορο έδαφος αποδοχής και εφαρμογής.

Ένας από αυτούς ήταν ο **Edward Deming**, ο οποίος είναι υπεύθυνος για την ποιοτική επανάσταση που επιτεύχθηκε στην Ιαπωνία μετά το Β΄ Παγκόσμιο Πόλεμο & κατέληξε στη διεθνή κυριαρχία της Ιαπωνικής βιομηχανίας. Το κεντρικό σημείο της φιλοσοφίας του *Deming* ήταν η μείωση της μεταβλητότητας στην παραγωγική διαδικασία με ιδιαίτερη έμφαση στη στατιστική επιμέλεια και έλεγχο.

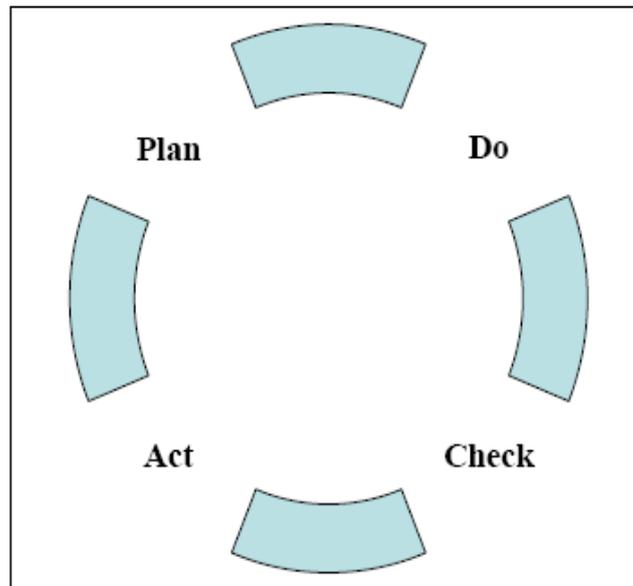
Τα 14 βασικά σημεία της φιλοσοφίας του Deming παρουσιάζονται στον πίνακα:

Πίνακας 1.3. Οι 14 κανόνες του Deming

Κανόνας 1:	Δημιούργησε το κλίμα για ένα σταθερό πρόγραμμα που να αποβλέπει στη συνεχή βελτίωση των προϊόντων και των υπηρεσιών
Κανόνας 2:	Υιοθέτησε τη νέα φιλοσοφία της απόρριψης της κακής δουλειάς, των ελαττωματικών και των μη ικανοποιητικών υπηρεσιών
Κανόνας 3:	Μην βασίζεστε στον τελικό έλεγχο όλης της παραγωγής για την επίτευξη υψηλής ποιότητας
Κανόνας 4:	Σταματήστε τη συνήθεια αγοράς προμηθειών με τη χαμηλότερη τιμή
Κανόνας 5:	Προσπαθήστε να βελτιώνεται χωρίς διακοπή όλα τα συστήματα παραγωγής και υπηρεσιών
Κανόνας 6:	Θεσπίστε ένα πρόγραμμα σύγχρονης εκπαίδευσης και εφαρμόστε το σε όλους τους εργαζόμενους
Κανόνας 7:	Υιοθετήστε σύγχρονους τρόπους επίβλεψης
Κανόνας 8:	Εξαλείψτε το φόβο. Μη διστάζετε να ρωτάτε και να αναφέρετε προβλήματα
Κανόνας 9:	Εξαφανίστε τα εμπόδια μεταξύ των τμημάτων της επιχείρησης. Η από κοινού δουλειά μεταξύ των διαφόρων τμημάτων της επιχείρησης είναι απαραίτητο στοιχείο για την ανάπτυξη της ποιότητας
Κανόνας 10:	Εξαλείψτε τους στόχους και τα συνθήματα για μηδέν ελαττωματικά προϊόντα
Κανόνας 11:	Εξαφανίστε τους αριθμητικούς στόχους
Κανόνας 12:	Εξαλείψτε τα εμπόδια που αφαιρούν από τον εργαζόμενο το δικαίωμα να νιώθει υπερήφανος για την δουλειά του
Κανόνας 13:	Εφαρμόστε ένα πρόγραμμα συνεχούς ενημέρωσης και εκπαίδευσης για όλους τους εργαζομένους
Κανόνας 14:	Δημιουργήστε ένα ανώτατο κλιμάκιο διοίκησης το οποίο θα εργαστεί σθεναρά για την επίτευξη των πρώτων 13 κανόνων

Στόχος της Διοίκησης Ολικής Ποιότητας είναι η συνεχής βελτίωση στην ποιότητα εκτέλεσης όλων των διεργασιών, προϊόντων και υπηρεσιών, σε μια επιχείρηση. Αυτό επιτυγχάνεται με τέσσερα βήματα τα οποία επαναλαμβάνονται το ένα μετά το άλλο και όταν ολοκληρωθεί το τελευταίο βήμα η διαδικασία ξεκινά πάλι από την αρχή.

Τα βήματα αυτά φαίνονται στον τροχό του Deming:



Σχήμα 1.1. Ο τροχός του Deming

Ένας δεύτερος ακαδημαϊκός, εξίσου γνωστός για τη συμβολή του στην παραγωγή προϊόντων υψηλής ποιότητας στην Ιαπωνία, ήταν ο **Joseph Juran**. Όπως και ο *Deming*, υποστήριξε τη μείωση της μεταβλητότητας στις διαδικασίες ως τον πιο ενδεδειγμένο τρόπο για τη βελτίωση της ποιότητας. Κεντρική αρχή της φιλοσοφίας του είναι η έμφαση που δίνει στις δράσεις της διοίκησης για την βελτίωση της ποιότητας. Υποστήριξε ότι τα συστήματα που ελέγχονται από τη διεύθυνση είναι κυρίως υπεύθυνα (80%) για τη βελτίωση της ποιότητας ενώ ένα μικρό μέρος (20%) οφείλεται στην απόδοση του εργατικού δυναμικού.

Ο **Philip Crosby** θεωρείται ο τρίτος μεγάλος θεωρητικός της Διοίκησης Ολικής Ποιότητας αν και η επιρροή του δεν ήταν τόσο μεγάλη όσο των *Deming* και *Juran*. Το κεντρικό σημείο της φιλοσοφίας του ήταν ότι η ποιότητα πρέπει να κτίζεται και όχι να επιθεωρείται. Η εφαρμογή των ιδεών του *Crosby* επαφίεται κυρίως σε αλλαγές στη συμπεριφορά και νοοτροπία του εργατικού

δυναμικού της επιχείρησης παρά στη χρήση του Στατιστικού Ελέγχου Ποιότητας. Αναγνώρισε την τεράστια σπατάλη που παρουσιαζόταν κατά τις διαδικασίες επιθεώρησης και επιδιόρθωσης ελαττωματικών προϊόντων και την ανάγκη για μείωση των δαπανών αυτών. Ο Crosby ήταν υπέρμαχος του προγράμματος **Μηδέν Ελαττωματικά (Zero Defects Program)**, το οποίο έχει δημιουργήσει πολλές παρανοήσεις και διαφωνίες. Ο Crosby πιστεύει ότι η αποδοχή ορισμένου ποσοστού ελαττωματικών στην παραγωγή είναι απαράδεκτη και ότι κάθε εργαζόμενος έχει την ευθύνη να κάνει την οποιαδήποτε δουλειά σωστά από την αρχή, να προλαμβάνει λάθη, σε αντίθεση με τους Deming και Juran που πιστεύουν ότι οποιεσδήποτε ατέλειες οφείλονται στο σχεδιασμό των συστημάτων παραγωγής και όχι στους εργαζομένους.

Αξίζει να σημειωθεί ότι παρά τις διαφορές μεταξύ των τριών (Deming, Juran, Crosby), εντούτοις προκύπτουν και κάποια κοινά σημεία στις θεωρίες τους όπως: (α) η επιθεώρηση δεν αποτελεί την απάντηση στη βελτίωση της ποιότητας, (β) η ανάμειξη και καθοδήγηση από την ανώτερη διεύθυνση είναι ζωτικής σημασίας για τη δημιουργία της απαραίτητης κουλτούρας αφοσιωμένης στην ποιότητα, και (γ) ένα πρόγραμμα για ποιότητα απαιτεί προσπάθειες από ολόκληρο τον οργανισμό και μακροπρόθεσμη δέσμευση καθώς επίσης και την απαραίτητη επένδυση για επιμόρφωση.

1.1.3. Συνιστώσες του Στατιστικού Ελέγχου Ποιότητας

Ο Στατιστικός Έλεγχος Ποιότητας αποτελείται από ένα σύνολο μεθόδων στατιστικής ανάλυσης δεδομένων. Το σύνολο αυτό μπορεί να χωριστεί σε τρία βασικά υποσύνολα που το καθένα περιέχει στατιστικές μεθόδους προσανατολισμένες σε διαφορετικές φάσεις της παραγωγικής διαδικασίας.

Τα τρία υποσύνολα είναι τα ακόλουθα:

- Σχεδιασμός και Ανάλυση Πειραμάτων (Design of Experiments)
- Στατιστικός Έλεγχος Διεργασιών (Statistical Process Control)
- Δειγματοληψία Αποδοχής (Acceptance Sampling)

Ο Σχεδιασμός και η Ανάλυση Πειραμάτων περιέχει όλες εκείνες τις στατιστικές τεχνικές οι οποίες μας βοηθούν στην ανακάλυψη της επίδρασης που έχουν τα διάφορα επίπεδα των

παραγόντων (μεταβλητών) που επηρεάζουν τις ποιοτικές παραμέτρους του τελικού προϊόντος και συνεπώς διαδραματίζει σημαντικό ρόλο στη βέλτιστη σχεδίαση της παραγωγικής διεργασίας. Ο Στατιστικός Έλεγχος Διεργασιών περιέχει στατιστικές τεχνικές που είναι απαραίτητες για τον έλεγχο της παραγωγικής διεργασίας κατά την διάρκεια της παραγωγής των προϊόντων. Η Δειγματοληψία Αποδοχής περιέχει στατιστικές τεχνικές (δειγματοληπτικές) που είναι απαραίτητες για να αποφασίσουμε αν μια συγκεκριμένη παρτίδα (σωρός) προϊόντων θα γίνει δεκτή ή θα απορριφθεί.

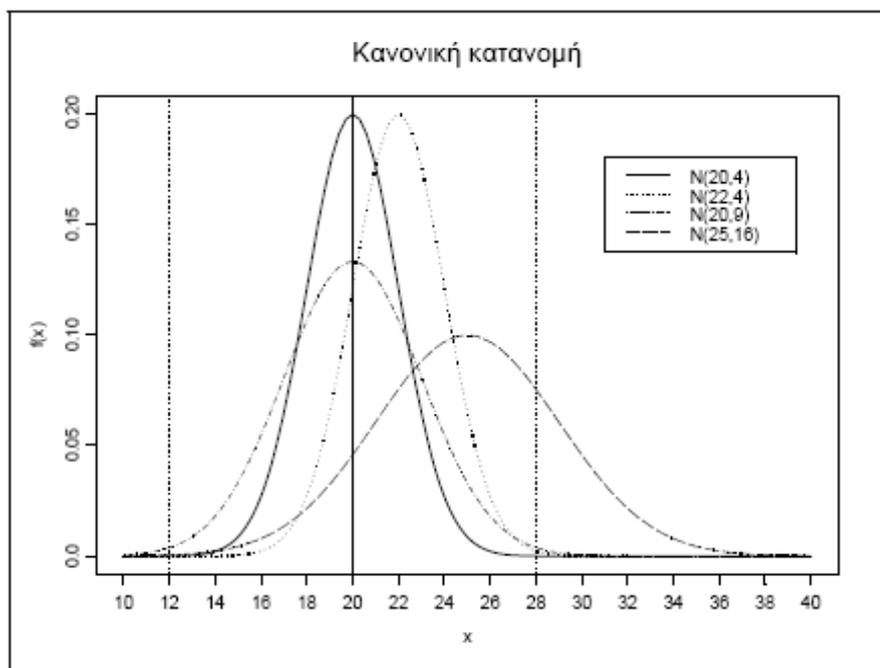
1.2. Το πρόβλημα του Στατιστικού Ελέγχου Διεργασιών

Σε κάθε παραγωγική διεργασία, ανεξάρτητα από το πόσο καλά σχεδιασμένη είναι και το πόσο προσεκτικά επιβλέπεται και συντηρείται, θα υπάρχει πάντα μια μορφή φυσικής μεταβλητότητας που θα τη συνοδεύει. Δηλαδή, όσο καλά ρυθμισμένα και να είναι τα μηχανήματα, όσο ικανοί και να είναι οι χειριστές των μηχανημάτων, όσο ικανοποιητική και να είναι η πρώτη ύλη, ποτέ δύο παραγόμενα προϊόντα δεν θα είναι τα ίδια (θα υπάρχει κάποιο μετρήσιμο μέγεθος του προϊόντος του οποίου η τιμή θα είναι διαφορετική στα δύο προϊόντα). Αυτή η φυσική μεταβλητότητα είναι το αθροιστικό αποτέλεσμα πολλών μικρών αιτιών οι οποίες αναφέρονται ως *κοινές ή συνήθεις ή τυχαίες αιτίες μεταβλητότητας* (common or chance causes of variation). Η φυσική μεταβλητότητα είναι συνήθως μικρή σε μέγεθος και δεν μπορεί να αποδοθεί σε ελέγξιμους παράγοντες. Μια διεργασία (σύστημα) η οποία λειτουργεί μόνο με την παρουσία φυσικής μεταβλητότητας λέμε ότι είναι **εντός (στατιστικού) ελέγχου** διεργασία (in – statistical- control process), ή ότι λειτουργεί σε **ευσταθή κατάσταση** (stable state).

Όμως σε μια διεργασία μπορεί να εμφανίζονται περιστασιακά και άλλες μορφές μεταβλητότητας οι οποίες δεν οφείλονται σε τυχαίες αιτίες αλλά αφορούν τη συστηματική αλλαγή στο επίπεδο κάποιου ή κάποιων παραγόντων που καθορίζουν την ποιότητα του προϊόντος. Αυτές οι μορφές μεταβλητότητας οφείλονται συνήθως στους ακόλουθους λόγους: **(α)** λανθασμένα ρυθμισμένες μηχανές, **(β)** λάθη των χειριστών των μηχανημάτων, και **(γ)** κακής ποιότητας ή ελαττωματική πρώτη ύλη. Η μεταβλητότητα που οφείλεται στους παραπάνω λόγους είναι σε μέγεθος πολύ μεγαλύτερη της φυσικής μεταβλητότητας και η παρουσία της οδηγεί συνήθως σε μη αποδεκτά επίπεδα λειτουργίας της παραγωγικής διεργασίας. Αυτή η μεταβλητοτητα αναφέρεται ως ειδική μεταβλητότητα και οι αιτίες που οδηγούν σε αυτή ονομάζονται ειδικές ή προσδιορισμένες αιτίες μεταβλητότητας (special or assignable causes of variation).

Μια διεργασία (σύστημα) η οποία λειτουργεί με την παρουσία ειδικής μεταβλητότητας λέμε ότι είναι εκτός- στατιστικού- ελέγχου διεργασία (out of (statistical) control process) ή ότι λειτουργεί σε **ασταθή κατάσταση** (unstable state). Άμεσα συνδεδεμένη έννοια με την παραγωγή ενός προϊόντος είναι τα **όρια προδιαγραφών των ποιοτικών χαρακτηριστικών** του προϊόντος τα οποία καθορίζονται στη φάση σχεδιασμού του. Αυτά είναι το **κάτω και το άνω όριο προδιαγραφών** (lower and upper specification limits, *LSL* and *USL*) και εντός αυτών των ορίων πρέπει να βρίσκονται οι τιμές του ποιοτικού χαρακτηριστικού για κάθε παραγόμενο προϊόν προκειμένου να είναι ποιοτικά αποδεκτό.

Επίσης, στη φάση σχεδιασμού του προϊόντος ορίζεται και μια επιθυμητή τιμή για το ποιοτικό χαρακτηριστικό που ονομάζεται **τιμή στόχος T** που είναι συνήθως το μέσο του διαστήματος [*LSL*, *USL*]. Κάτω από συνθήκες φυσικής μεταβλητότητας η συντριπτική πλειοψηφία των τιμών του ποιοτικού χαρακτηριστικού στα παραγόμενα προϊόντα βρίσκεται εντός των ορίων των προδιαγραφών. Όμως κάτω από συνθήκες ειδικής μεταβλητότητας δεν ισχύει το ίδιο. Για παράδειγμα ας υποθέσουμε ότι η ποιότητα ενός προϊόντος εκφράζεται με την τιμή ενός μετρήσιμου ποιοτικού χαρακτηριστικού του για το οποίο $T = 20$, $LSL = 12$, $USL = 28$, και ότι κάτω από συνθήκες φυσικής μεταβλητότητας (εντός ελέγχου διεργασία) οι τιμές του ποιοτικού χαρακτηριστικού στα παραγόμενα προϊόντα περιγράφονται από την κατανομή $N(20, 4)$. Το αποτέλεσμα που επιφέρουν οι κοινές και οι ειδικές αιτίες μεταβλητότητας στην παραγωγική διεργασία περιγράφεται στο ακόλουθο σχήμα:



Υπό συνθήκες φυσικής μεταβλητότητας σχεδόν όλα τα παραγόμενα προϊόντα πληρούν τις προδιαγραφές. Η εμφάνιση ειδικής μεταβλητότητας μπορεί να επιφέρει αλλαγή στην κατανομή των τιμών του ποιοτικού χαρακτηριστικού, είτε αλλάζοντας (μετατοπίζοντας, *shift*) μόνο το μέσο της κατανομής ($N(22,4)$), είτε αλλάζοντας μόνο τη διακύμανση της κατανομής ($N(20,9)$), είτε αλλάζοντας το μέσο και τη διακύμανση της κατανομής ταυτόχρονα ($N(25,16)$). Σε κάθε περίπτωση, η επίδραση της ειδικής μεταβλητότητας μεταφράζεται σε αύξηση των παραγόμενων προϊόντων που έχουν τιμές ποιοτικού χαρακτηριστικού εκτός των ορίων προδιαγραφών.

Το κύριο αντικείμενο του Στατιστικού Ελέγχου Διεργασιών είναι η έγκαιρη ανίχνευση της εμφάνισης ειδικών αιτιών μεταβλητότητας σε μια διεργασία έτσι ώστε να προχωρήσουμε σε έρευνα και να προβούμε στις απαραίτητες διορθωτικές ενέργειες προτού κατασκευαστούν αρκετά προϊόντα μη συμμορφωμένα με τις προδιαγραφές. Τα **διαγράμματα ελέγχου** (*control charts*) είναι μια τεχνική που χρησιμοποιείται ευρέως για την ανίχνευση σε πραγματικό χρόνο της εμφάνισης ειδικών αιτιών μεταβλητότητας σε μια διεργασία (*on-line process-monitoring*).

Ειδικά για το προηγούμενο παράδειγμα θα μας ενδιέφερε η κατασκευή διαγραμμάτων ελέγχου για την παρακολούθηση της μέσης τιμή και της διακύμανσης (ή της τυπικής απόκλισης) του ποιοτικού χαρακτηριστικού μέσω τυχαίων δειγμάτων από τα παραγόμενα προϊόντα. Για να είναι αποτελεσματικός ο Στατιστικός Έλεγχος Διεργασιών θα πρέπει να συνοδεύεται απαραίτητα με ένα **εκτός ελέγχου πρόγραμμα δράσης** (*out-of-control action plan, OCAP*) το οποίο θα πρέπει να ενεργοποιείται κάθε φορά που το διάγραμμα ελέγχου παρέχει ενδείξεις εμφάνισης ειδικών αιτιών μεταβλητότητας στη διεργασία. Το *OCAP* περιγράφει την ακολουθία των ενεργειών που πρέπει να γίνουν με σκοπό την εξάλειψη των ειδικών αιτιών μεταβλητότητας σε μια διεργασία και ο σχεδιασμός του απαιτεί τη συνεργασία ατόμων από διάφορα τμήματα της επιχείρησης.

Οι γενικές αρχές κατασκευής και χρήσης των διαγραμμάτων ελέγχου θα παρουσιαστούν σε επόμενες παραγράφους. Ωστόσο πριν κλείσουμε την παρούσα παράγραφο κρίνεται σκόπιμο να αναφερθούμε στις έννοιες **μη συμμορφωμένο ή ελαττωματικό προϊόν** (*nonconforming or defective product*) και **αριθμός ελαττωμάτων ή ατελειών** (*defects or nonconformities*) ενός προϊόντος προκειμένου να αποφύγουμε τη συσχέτιση των διαγραμμάτων ελέγχου μόνο με την περίπτωση που οι τιμές του ποιοτικού χαρακτηριστικού του προϊόντος περιγράφονται από μια **συνεχή τυχαία μεταβλητή** (*variable*), δηλαδή από μια μεταβλητή που παίρνει τιμές σε μια συνεχή κλίμακα (όπως το παράδειγμα που εξετάσαμε προηγουμένως).

Με τον όρο **μη συμμορφωμένο** ή **ελαττωματικό** προϊόν ονομάζουμε το προϊόν για το οποίο τουλάχιστον ένα ποιοτικό χαρακτηριστικό του έχει τιμή η οποία βρίσκεται εκτός των ορίων προδιαγραφών, δηλαδή παρουσιάζει τουλάχιστον ένα ελάττωμα ή ατέλεια. Όμως σε πολλές περιπτώσεις ανάλογα με τον αριθμό και τη σοβαρότητα των ελαττωμάτων (ατελειών) που παρουσιάζει ένα προϊόν μπορεί να χαρακτηριστεί συμμορφωμένο ή μη ελαττωματικό προϊόν, και να προωθηθεί προς πώληση στην αγορά. Για παράδειγμα, ένα ρούχο μπορεί να χαρακτηριστεί ελαττωματικό εάν το τελικό προϊόν έχει ελαττώματα στις ραφές, ή ελαττώματα στο χρώμα, ή ελαττώματα στο ύφασμα, κτλ. Όμως αρκετές φορές το ρούχο δε θα χαρακτηριστεί (πρακτικά) ελαττωματικό αν παρουσιάζει μόνο ένα από τα παραπάνω ελαττώματα, ή ακόμα και αν παρουσιάζει περισσότερα ελαττώματα, αλλά σε μικρότερο βαθμό.

Σε τέτοιες περιπτώσεις κατασκευάζονται διαγράμματα ελέγχου για τον αριθμό των ελαττωματικών προϊόντων ή ακόμα και για τον αριθμό των ελαττωμάτων ενός προϊόντος που ως ποιοτικά χαρακτηριστικά ενός προϊόντος δεν μετρώνται σε μια συνεχή κλίμακα αλλά παίρνουν αριθμήσιμο πλήθος τιμών και περιγράφονται με **διακριτές τυχαίες μεταβλητές** (*attributes*).

Τέτοια διαγράμματα χρησιμοποιούνται και στην περίπτωση που η προς πώληση μονάδα είναι μια συσκευασία η οποία περιέχει έναν αριθμό προϊόντων και η οποία μπορεί να χαρακτηριστεί ως ελαττωματική ή μη, βάσει του αριθμού των ελαττωματικών προϊόντων που περιέχει. Σύμφωνα με τα όσα προαναφέρθηκαν, στα διαγράμματα ελέγχου μπορούμε να διακρίνουμε δύο βασικές κατηγορίες ανάλογα με το είδος της μεταβλητής που περιγράφει το ποιοτικό χαρακτηριστικό του προϊόντος:

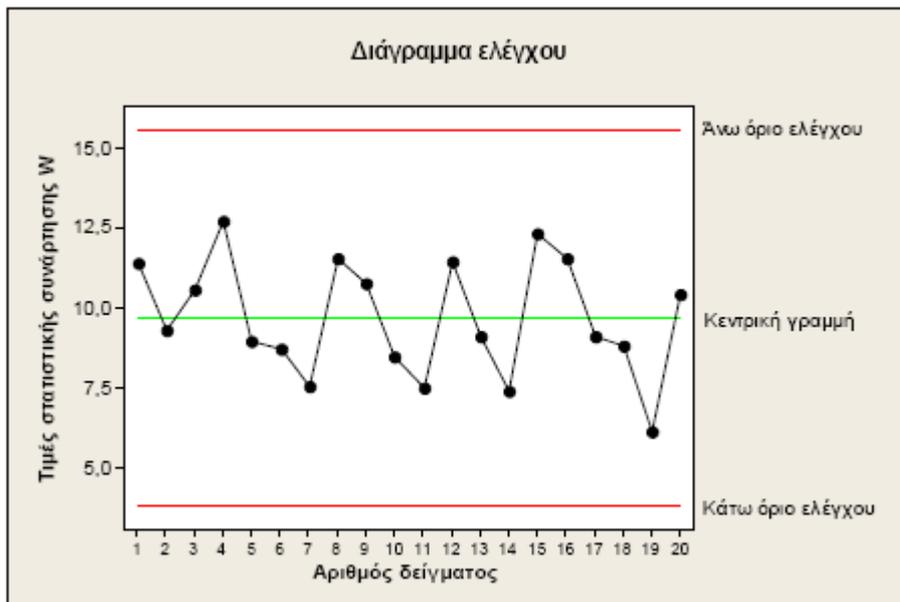
- διαγράμματα ελέγχου για **συνεχή** χαρακτηριστικά – μεταβλητές (*charts for variables*)
- διαγράμματα ελέγχου για **διακριτά** χαρακτηριστικά – ιδιότητες (*charts for attributes*).

1.2.1. Περιγραφή και Χρήση ενός Διαγράμματος Ελέγχου

Στις παραγωγικές διεργασίες μας ενδιαφέρει η παρακολούθηση της συμπεριφορά μιας κρίσιμης ποσότητας ενός (μετρήσιμου) χαρακτηριστικού X (τυχαία μεταβλητή) των προϊόντων που παράγονται (για παράδειγμα το χαρακτηριστικό X μπορεί να είναι μήκος, βάρος, όγκος προϊόντων). Η παρακολούθηση της κρίσιμης ποσότητας βασίζεται σε μετρήσεις του χαρακτηριστικού X (τυχαία μεταβλητή), που προκύπτουν από την επιλογή τυχαίων δειγμάτων προϊόντων από την παραγωγή σε διαφορετικές χρονικές στιγμές, έστω τα X_1, X_2, \dots

Χρησιμοποιώντας τα τυχαία δείγματα $\mathbf{X}_1, \mathbf{X}_2, \dots$ υπολογίζουμε την τιμή $W_i = g(\mathbf{X}_i)$, $i = 1, 2, \dots$, μιας κατάλληλης στατιστικής συνάρτησης (τυχαίας μεταβλητής) που εκτιμά (συνήθως αμερόληπτη εκτιμήτρια) την κρίσιμη ποσότητα που μας ενδιαφέρει (π.χ. μέση τιμή ή διακύμανση της X). Έτσι η (διαχρονική) παρακολούθηση της συμπεριφοράς της κρίσιμης ποσότητας επιτυγχάνεται με την παρακολούθηση των τιμών που λαμβάνει η στατιστική συνάρτηση W στα διάφορα δείγματα.

Για παράδειγμα ας υποθέσουμε ότι ενδιαφερόμαστε να παρακολουθήσουμε τη συμπεριφορά της μέσης τιμής της διαμέτρου X των κυλίνδρων που παράγει μια μηχανή. Για το σκοπό μπορούμε να επιλέξουμε τυχαία δείγματα μεγέθους n ($n \geq 1$) κυλίνδρων από την παραγωγή της μηχανής σε διαφορετικά χρονικά διαστήματα και να χρησιμοποιήσουμε τη στατιστική συνάρτηση $W_i = g(\mathbf{X}_i) = (\mathbf{X}_{i1} + \mathbf{X}_{i2} + \dots + \mathbf{X}_{in}) / n$ (η οποία είναι αμερόληπτη εκτιμήτρια του μέσου της X) για την παρακολούθηση της συμπεριφοράς της μέσης τιμής. Ένα τυπικό διάγραμμα ελέγχου Shewhart (τα διαγράμματα αυτά τα εισήγαγε το 1924 ο W.A. Shewhart) είναι μία γραφική παράσταση με την ακόλουθη μορφή:

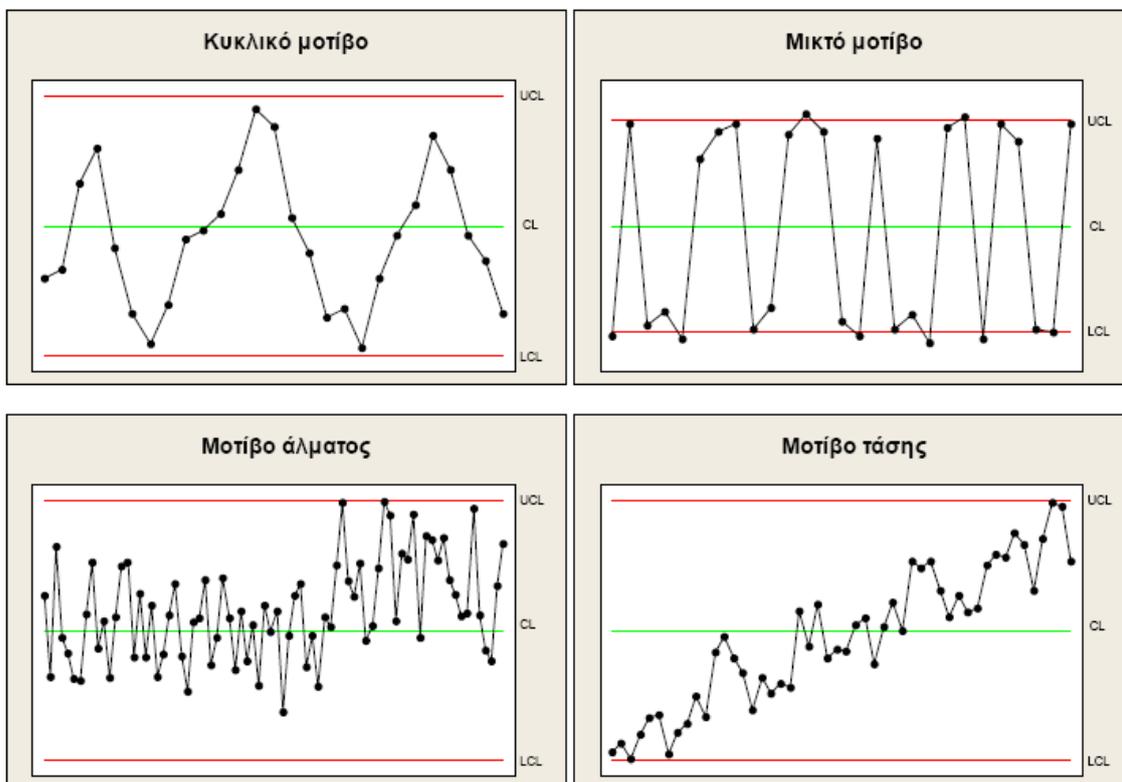


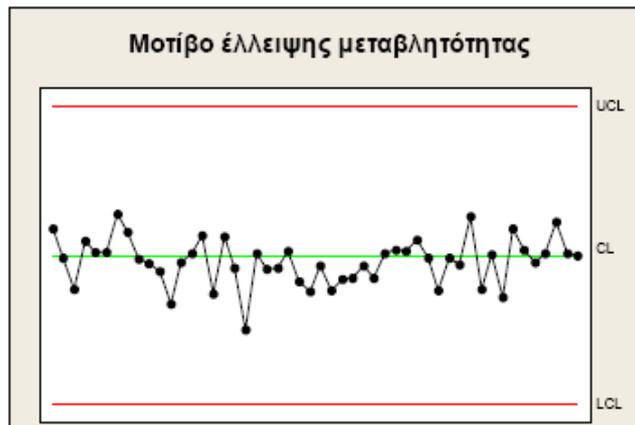
Στο παραπάνω σχήμα, εκτός από τις παρατηρούμενες τιμές της W που έχουν παρασταθεί με σημεία (κουκκίδες) τα οποία έχουν συνδεθεί με μια τεθλασμένη γραμμή, έχουν σχεδιαστεί και άλλες τρεις γραμμές. Η **κεντρική γραμμή** (*center line, CL*) ή **μέσο επίπεδο** της διεργασίας παριστάνει συνήθως τη μέση τιμή (*mean value*) της W όπως αυτή προκύπτει από τη λειτουργία μιας εντός ελέγχου διεργασίας ή την τιμή στόχο του προϊόντος.

Οι δύο ακραίες γραμμές που εμφανίζονται ονομάζονται **άνω και κάτω όρια ελέγχου** (*upper and lower control limits, UCL and LCL*). Όσο οι τιμές (σημεία, δεδομένα) της *W* εμφανίζονται εντός των ορίων ελέγχου και η συμπεριφορά τους είναι «τυχαία» μπορούμε να υποθέσουμε ότι η διεργασία παραμένει εντός ελέγχου (για την ακρίβεια εντός στατιστικού ελέγχου) και δεν χρειάζεται να προβούμε σε κάποια διορθωτική ενέργεια.

Αν όμως κάποιο σημείο βρεθεί εκτός των ορίων ελέγχου λέμε ότι υπάρχει **ένδειξη** ότι η διεργασία είναι εκτός ελέγχου οπότε αντιμετωπίζουμε κατάσταση **συναγερμού** (*alarm*) και πρέπει να προχωρήσουμε σε έρευνα για να ανακαλύψουμε τις ειδικές αιτίες μεταβλητότητας που είναι υπεύθυνες για αυτή τη συμπεριφορά και αν κριθεί απαραίτητο να προβούμε σε διορθωτικές ενέργειες.

Θα πρέπει να σημειώσουμε ότι ακόμη και στην περίπτωση που όλα τα σημεία βρίσκονται εντός των ορίων ελέγχου αλλά συμπεριφέρονται με ένα **συστηματικό ή μη τυχαίο τρόπο** τότε και αυτό αποτελεί ένδειξη ότι η διεργασία είναι εκτός ελέγχου. Οι βασικότερες μορφές **προτύπων ή μοτίβων** (*patterns*) μη τυχαίου τρόπου συμπεριφοράς μιας χρονοσειράς σημείων είναι τα πρότυπα κύκλων (*cycles*), τα μικτά πρότυπα (*mixtures*), τα πρότυπα αλμάτων (*shifts in the process level*), τα πρότυπα τάσεων (*trends*), και τα πρότυπα έλλειψης μεταβλητότητας (*stratification*) τα οποία περιγράφονται οπτικά στο ακόλουθο σχήμα:





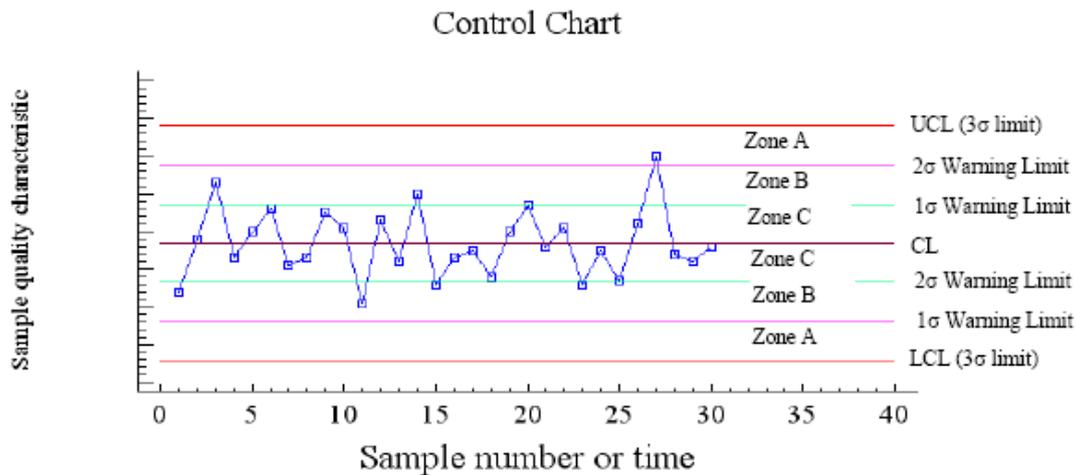
Σχήμα: Τα κυριότερα μη τυχαία μοτίβα

Στα κυκλικά ή περιοδικά πρότυπα (*cycles*) εμφανίζεται μια περιοδικότητα στη διάταξη των σημείων με αποτέλεσμα να σχηματίζονται επαναλαμβανόμενοι κύκλοι. Πιθανές αιτίες είναι οι περιοδικές διακυμάνσεις των περιβαλλοντικών συνθηκών και συνθηκών παραγωγής όπως η θερμοκρασία, η ηλεκτρική τάση, εναλλαγή εργατοτεχνιτών, εναλλαγή μηχανών, κτλ. Τα μικτά πρότυπα (*mixture*) φαίνεται να προκύπτουν από δύο διαφορετικές κατανομές του χαρακτηριστικού X μεταξύ των οποίων παλινδρομεί η διεργασία.

Τα **πρότυπα αλμάτων** (*shift in the process level*) οφείλονται συνήθως σε νέους εργαζόμενους, νέες μεθόδους παραγωγής, νέες μηχανές, διαδοχική επεξεργασία παρτίδων πρώτων υλών διαφορετικής ποιότητας, κτλ. Τα **πρότυπα τάσεων** (*trend*) οφείλονται συνήθως σε παράγοντες όπως η σταδιακή φθορά εργαλείων, η κόπωση των εργαζομένων, η παρουσία επιθεωρητή, κτλ.

Στα **πρότυπα έλλειψης μεταβλητότητας** (*stratification*) τα σημεία του διαγράμματος κινούνται με τεχνητό τρόπο γύρω από την κεντρική γραμμή χωρίς να υπάρχει κάποια αξιοσημείωτη μεταβλητότητα. Οφείλονται συνήθως σε εσφαλμένη σχεδίαση των ορίων ελέγχου.

Σε πολλές περιπτώσεις για να κάνουμε περισσότερο ευαίσθητο ένα διάγραμμα ελέγχου ως προς την ικανότητά του να ανιχνεύει πιο γρήγορα εκτός ελέγχου διεργασίες, εκτός από τη σχεδίαση των ορίων ελέγχου, σχεδιάζουμε επίσης και **προειδοποιητικά όρια** (*warning limits*) εσωτερικά των ορίων ελέγχου όπως δείχνει το παρακάτω σχήμα (η σχεδίαση των προειδοποιητικών ορίων προϋποθέτει την ανάπτυξη διαγραμμάτων ελέγχου χρησιμοποιώντας το μοντέλο των **ορίων τριών σίγμα (3σ)**).

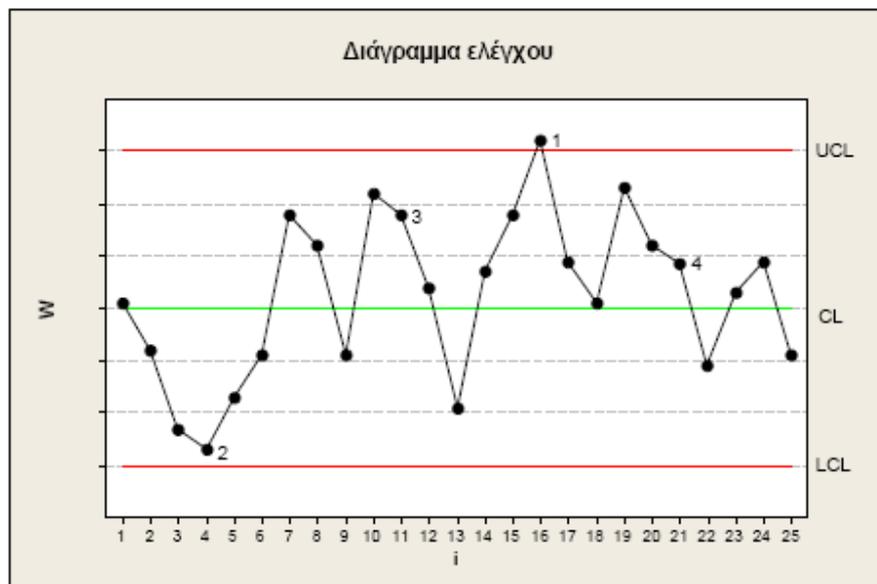


Σχήμα: Ζώνες A , B και C σε ένα διάγραμμα ελέγχου

Τα προειδοποιητικά όρια χρησιμοποιούνται μαζί με κάποιους “κανόνες” που περιγράφουν την εμφάνιση ειδικών μοτίβων σε ένα διάγραμμα ελέγχου. Στην περίπτωση που εμφανιστεί το μοτίβο που περιγράφει ο κανόνας τότε θεωρούμε ότι η διεργασία είναι εκτός ελέγχου χωρίς απαραίτητα να έχουμε κάποιο σημείο του διαγράμματος εκτός των ορίων ελέγχου (*UCL* και *LCL*). Οι σημαντικότεροι κανόνες που χρησιμοποιούνται για την ευαισθητοποίηση ενός διαγράμματος ελέγχου είναι οι ακόλουθοι:

1. Ένα ή περισσότερα σημεία εκτός των ορίων ελέγχου.
2. Δύο από τρία συνεχόμενα σημεία στην Ζώνη A (σε μια από τις δύο ζώνες A).
3. Τέσσερα από πέντε συνεχόμενα σημεία πέραν της Ζώνης C (σε μια από τις δύο περιοχές).
4. Οκτώ συνεχόμενα σημεία στην ίδια μεριά (επάνω ή κάτω) της κεντρικής γραμμής.
5. Έξι συνεχόμενα σημεία σε αύξουσα ή φθίνουσα διάταξη.
6. Δεκαπέντε συνεχόμενα σημεία στην ολική Ζώνη C.
7. Δεκατέσσερα συνεχόμενα σημεία σε εναλλασσόμενη μορφή «πάνω-κάτω».
8. Οκτώ συνεχόμενα σημεία εκτός της ολικής Ζώνης C.
9. Οποιαδήποτε ασυνήθιστη ή μη τυχαία ακολουθία σημείων.
10. Ένα ή περισσότερα σημεία κοντά στα προειδοποιητικά όρια ή τα όρια ελέγχου.

Οι πρώτοι τέσσερις κανόνες είναι γνωστοί ως *Western Electric Rules*. Στο ακόλουθο διάγραμμα ελέγχου έχουν σημειωθεί τα σημεία που «χτυπούν» για πρώτη φορά ο καθένας από τους *Western Electric Rules*.



Έτσι, στο σημείο 4 «χτύπησε» ο Κανόνας 2, στο σημείο 11 «χτύπησε» ο Κανόνας 3, στο σημείο 16 «χτύπησε» ο Κανόνας 1 και στο σημείο 21 «χτύπησε» ο Κανόνας 4.

1.3. Ο Σ.Ε.Δ στον τομέα της Υγείας:

Αξιολόγηση ιατρικών και νοσοκομειακών εφαρμογών.

Η ποιότητα των παραγόμενων υπηρεσιών ή των παρερχόμενων υπηρεσιών καθορίζουν τη μελλοντική πορεία και εξέλιξη ενός οργανισμού. Όπως αναφέρθηκε και στην προηγούμενη ενότητα, ο Στατιστικός Έλεγχος Διεργασιών (SPC) είναι ένα πολύτιμο εργαλείο που χρησιμεύει στον έλεγχο και την παρακολούθηση της μεταβλητότητας των κρίσιμων παραμέτρων μίας διεργασίας. Αν και αρχικά ο SPC εφαρμόστηκε σε βιομηχανικές και παραγωγικές διεργασίες, τις τελευταίες δεκαετίες η χρήση του έχει εξαπλωθεί και στον τομέα των υπηρεσιών. Ιδιαίτερο ενδιαφέρον παρουσιάζει η χρήση του στις υπηρεσίες Υγείας, όπου η ποιότητα είναι συνώνυμη με την ασφάλεια, την ποιότητα και την αποτελεσματικότητα. Στην παρούσα ενότητα, παρουσιάζουμε και δίνουμε έμφαση στις εφαρμογές του SPC και

ειδικότερα των διαγραμμάτων ελέγχου, στην αξιολόγηση των ιατρικών και νοσοκομειακών εφαρμογών.

Παρατηρείται ότι με την πάροδο των χρόνων ολοένα και περισσότερες εφαρμογές καταγράφονται στο συγκεκριμένο τομέα. Πιο αναλυτικά, την περίοδο 1951 – 1988 δεν καταγράφηκε κάποια αναφορά, την περίοδο 1989 – 1991 υπήρξαν δύο αναφορές, την περίοδο 1992 – 1995 βρέθηκαν 26 και τέλος την περίοδο 1996 – 2004 υπήρξαν 71 αναφορές. Επίσης, έχουν εκδοθεί αρκετά βιβλία σχετικά με το SPC στη υγεία, καθώς και οργανισμοί, όπως:

Joint Commission on Accreditation of Healthcare Organizations στις ΗΠΑ και *National Health Service Modernization Agency* στο Ηνωμένο Βασίλειο, έχουν επισήμως υιοθετήσει το SPC.

Βέβαια πρέπει να γίνει αντιληπτό ότι λόγω της ποικιλίας των παραγόντων, της φύσης των προβλημάτων και των οργανωτικών θεμάτων που εμπλέκονται στο σύστημα υγείας, είναι απαραίτητο να γίνει συστηματική αξιολόγηση της διεργασίας, εφαρμόζοντας απλά και κατανοητά κριτήρια. Μόνο με αυτόν τον τρόπο θα μπορέσουν να τεθούν προτεραιότητες στην παρακολούθηση και βελτίωση της δημόσιας υγείας, που θα επιφέρουν ορατό και άμεσο αποτέλεσμα στη βελτίωση της ποιότητας παροχής των υπηρεσιών.

Για παράδειγμα, υπάρχουν ορισμένα κριτήρια για την επιλογή της ασθένειας που είναι προτιμότερο να μελετηθεί ως προς συγκεκριμένα χαρακτηριστικά (π.χ. ο ρυθμός κρουσμάτων ή εξάπλωσης της σε έναν πληθυσμό), σε σχέση με άλλες. Τα κριτήρια αυτά μπορεί να είναι το υψηλό κόστος που προκύπτει κατά την περίθαλψη των ασθενών, ο αριθμός των εισαγωγών που αφορούν στην ασθένεια, το μέγεθος του πληθυσμού που προσβάλλεται, αναγνωρισμένα προβλήματα σχετικά με την ποιότητα φροντίδας που παρέχεται στους ασθενείς. Τέλος, η ύπαρξη καλά καθορισμένης θεραπείας για κάποια ασθένεια αποτελεί σημαντικό κριτήριο επιλογής, αφού κατ' αυτόν τον τρόπο διευκολύνεται η μελέτη, κατά συνέπεια και η ερευνητική διαδικασία.

Εφόσον επιλεγεί η ασθένεια που θα μελετηθεί, συγκεντρώνεται η σχετική βιβλιογραφία, καθώς και η γνώμη των ειδικών για τη διάγνωση και τη θεραπεία της. Δημιουργείται έτσι ένας οδηγός κλινικής πρακτικής (*CPG – clinical practice guideline*). Τα τελευταία χρόνια έχουν αναπτυχθεί οδηγοί κλινικής πρακτικής σχεδόν για κάθε ασθένεια. Ιδιαίτερη προσοχή έχει δοθεί σε ασθένειες, όπως ο διαβήτης, το άσθμα και η υπέρταση. Συνήθως στις περισσότερες περιπτώσεις, οι ειδικοί αποφασίζουν ποιο CGP κρίνεται κατάλληλο προς χρήση στην εκάστοτε τοπική κοινότητα. Αν κάποιο από τα διαθέσιμα CGP δεν είναι κατάλληλο, θεωρείται απαραίτητη η κατάρτιση ενός νέου με βάση τα βιβλιογραφικά δεδομένα.

Όταν έχει πλέον δημιουργηθεί το κατάλληλο CPG, μπορεί να αρχίσει η διαχείριση των δεδομένων που συλλέγονται κατά την εξάσκηση της κλινικής πρακτικής, ώστε να είναι αποτελεσματική η λήψη των αποφάσεων. Αν και η κατηγοριοποίηση των ασθενειών που πρέπει να μελετώνται είναι σημαντική, πρέπει να δίνεται έμφαση στην πρόληψη που αποτελεί την κατάλληλη θεραπεία.

1.3.1. Η συμβολή της ανάλυσης κλινικών περιστατικών.

Κατά τη διάρκεια των ιατρικών εφαρμογών, παράγονται πολύ ευχάριστα αποτελέσματα, όπως υγιή νεογέννητα, ασθενείς θεραπευμένοι από χρόνιες, ή μεταδοτικές ασθένειες, καθώς και η αποκατάσταση καρδιακής λειτουργίας ασθενών. Δυστυχώς όμως, συχνά προκύπτουν και ανεπιθύμητα αποτελέσματα, όπως λοιμώξεις, ανεπιτυχείς επεμβάσεις, κ.λπ. Πολλοί επαγγελματίες στον χώρο της Υγείας, θεωρούν τη δεύτερη κατηγορία ως απλώς ατυχή περιστατικά που συνέβησαν χωρίς λόγο, με αποτέλεσμα αυτά να επανεμφανίζονται τακτικά. Αντιθέτως, όσο ανεπιθύμητα κι αν είναι αυτά τα αποτελέσματα, πρέπει να διερευνηθούν οι αιτίες πρόκλησής τους, ώστε να αποτραπεί η επανεμφάνισή τους.

Στις Η.Π.Α, τη Βρετανία και την Αυστραλία έχουν δημιουργηθεί ειδικά προγράμματα για την καταγραφή των περιστατικών στον τομέα της Υγείας. Ωστόσο, δεν αρκεί η απλή καταγραφή, καθώς χρειάζεται βαθύτερη κατανόηση που θα οδηγήσει στον εντοπισμό της αιτίας ή των αιτιών του προβλήματος. Συνήθως υπάρχει μία αλυσίδα γεγονότων και μία ποικιλία παραγόντων που οδηγούν σε ένα συμβάν. Πρέπει να διερευνηθεί τι ακριβώς συνέβη και προέκυψε το πρόβλημα, καθώς και να μελετηθούν σε βάθος οι αιτίες για να αποκαλυφθούν τα «κενά» και οι ανεπάρκειες στο σύστημα της Υγείας. Επομένως, το περιστατικό λειτουργεί σαν ένα «παράθυρο» μέσω του οποίου διαφαίνεται η γενικότερη κατάσταση του συστήματος Υγείας. Άλλωστε το κλειδί για τη συνεχή βελτίωση της ποιότητας είναι η μέτρηση, ανάλυση και ερμηνεία της μεταβλητότητας.

Τα δύο βασικά χαρακτηριστικά που πρέπει να διέπουν τη φροντίδα και την παρακολούθηση της δημόσιας υγείας, είναι η ασφάλεια και η αποτελεσματικότητα. Για να εξασφαλίζονται τα παραπάνω, θα πρέπει οι υπηρεσίες στην Υγεία να χαρακτηρίζονται από υψηλό επίπεδο ποιότητας. Η ποιότητα στις υπηρεσίες της Υγείας εστιάζεται: **α)** στον έλεγχο και τη αξιολόγηση μίας διεργασίας ιατρικής, όπως ο έλεγχος άσθματος ενός ασθενούς, **β)** στην απόδοση μίας μονάδας, όπως η πνευμονολογική ή η καρδιοχειρουργική μονάδα ενός νοσοκομείου, **γ)** στον έλεγχο ικανοποίησης ενός ασθενούς από τη φροντίδα σε μία κλινική ή

ένα νοσοκομείο. Ωστόσο, η μελέτη των παραμέτρων στην κλινική πράξη δεν είναι εύκολη υπόθεση.

Για παράδειγμα, για την εκτίμηση της αποτελεσματικότητας μίας χειρουργικής επέμβασης γίνεται χρονομέτρηση του κάθε σταδίου, δηλαδή της παραλαβής του κάθε ασθενούς, της προετοιμασίας, της αναισθησίας, της επαναφοράς του ασθενούς μετά τη χειρουργική επέμβαση, κ.λ.π. Δε διαφαίνεται όμως με αυτές τις μετρήσεις το αποτέλεσμα της διεργασίας, δηλαδή η βελτίωση της υγείας του ασθενούς. Μετρήσεις, όπως χρόνος παραμονής στο νοσοκομείο, επίπεδα μολύνσεων, ικανοποίηση ασθενούς χρησιμοποιούνται γιατί μπορούν να εκτιμηθούν σε σύντομο χρονικό διάστημα. Σε ασθένειες, όπως η υπέρταση που ανήκει σε χρόνια νοσήματα, μπορεί να γίνει παρακολούθηση της εξέλιξης με την ανταπόκριση του ασθενούς στην αντίστοιχη εφαρμογή, π.χ. στη χορήγηση φαρμάκων. Μετρήσεις, όπως ποσοστά θανάτων μετά από χειρουργική επέμβαση, χρησιμοποιούνται για εκτίμηση της λειτουργίας του υγειονομικού συστήματος και σύγκρισης μεταξύ διαφορετικών νοσηλευτικών ιδρυμάτων.

1.3.2. Η συνεισφορά του Στατιστικού Ελέγχου Διεργασιών

Αν και σε αρκετές περιπτώσεις υπάρχουν δεδομένα για τις διεργασίες στο χώρο των νοσοκομείων, των ερευνητικών ιδρυμάτων και όχι μόνο, το σύνηθες πρόβλημα είναι η δύσκολη επεξεργασία τους με αποτέλεσμα αυτά να παραμένουν ανεκμετάλλευτα. Χρησιμοποιούνται διάφορα εργαλεία στατιστικού ελέγχου, όπως τα ιστογράμματα, τα διαγράμματα ροής και τα διαγράμματα ελέγχου. Το ιστόγραμμα είναι σαν ένα στιγμιότυπο που αποτυπώνει την εικόνα των δεδομένων σε μία δεδομένη χρονική στιγμή, ενώ τα διαγράμματα ελέγχου προσομοιάζουν σε ένα φιλμ που αποτυπώνει την εικόνα των δεδομένων με την πάροδο του χρόνου.

Η μεγάλη λοιπόν συνεισφορά του SPC έγκειται στο γεγονός ότι συνδυάζει μία σειρά αναλυτικών μεθόδων με τη γραφική παρουσίαση των δεδομένων, προσφέροντας μία πιο διεισδυτική «ματιά» στα δεδομένα, όπως επίσης και σαφώς καλύτερη κατανόηση της διεργασίας. Συνεπώς, ο Στατιστικός Έλεγχος Διεργασιών (SPC) και το βασικό του εργαλείο, το διάγραμμα ελέγχου, προσφέρουν στους ερευνητές έναν πιο αποτελεσματικό τρόπο για τη βελτίωση της ποιότητας της.

Στο σημείο αυτό θα πρέπει να αναφερθεί ότι τα δεδομένα που συλλέγονται στον τομέα της Υγείας, είναι κυρίως διακριτά και ότι στις περισσότερες περιπτώσεις κρίνεται σκόπιμη η προσαρμογή των δεδομένων με βάση παράγοντες, που επηρεάζουν το αποτέλεσμα της διεργασίας (**risk – adjustment**). Για παράδειγμα, το ποσοστό επιβίωσης ενός ασθενούς από μία καρδιοχειρουργική επέμβαση, μπορεί να διαφέρει ανάλογα με την ηλικία του ασθενούς, τη γενικότερη κατάσταση της υγείας του κ.λ.π. Επίσης, πολύ συχνά στον τομέα της Υγείας χρησιμοποιείται το 100% των παρατηρήσεων, όπως συμβαίνει σε περιπτώσεις μελέτης των ποσοστών θνησιμότητας, εμφάνισης κρουσμάτων μίας μεταδοτικής ασθένειας σε ένα πληθυσμό, μεταλλάξεις κ.λ.π. και όχι κάποιο δείγμα του πληθυσμού.

Τέλος, αξίζει να σημειωθεί ότι στο τομέα αυτό η επαναφορά μίας διεργασίας εντός ελέγχου, εφόσον σε αυτήν έχει εμφανιστεί κάποια συστηματική αιτία μεταβλητότητας, είναι πιο δύσκολη και απαιτεί διεξοδικό έλεγχο και μελέτη.

Το βασικό πλεονέκτημα της συνεισφοράς του Σ.Ε.Δ στον τομέα της Υγείας, σε σχέση με τις κλασικές στατιστικές μεθόδους που εφαρμόζονται στον τομέα της Βιομηχανίας και της παραγωγικής διαδικασίας, είναι ο έγκαιρος εντοπισμός της μεταβλητότητας στο δείγμα που μελετούμε. Δηλαδή μέσω του διαγράμματος ελέγχου, καθίσταται δυνατή η ανίχνευση μετατοπίσεων στη διεργασία και των εμφανιζόμενων τάσεων στα αρχικά στάδια. Επομένως, είναι δυνατή η ανάλυση και επεξεργασία των δεδομένων που συλλέγονται σε καθημερινή βάση. Επιπλέον με τη χρήση αυτών των στατιστικών εργαλείων συντελείται μείωση του κόστους, αφού περιορίζονται οι αποτυχημένες επενδύσεις σε αλλαγές που δε θα βελτίωναν ουσιαστικά τη διαδικασία.

Η διευκόλυνση από τη χρήση των διαγραμμάτων ελέγχου βασίζεται στο γεγονός ότι δεν απαιτούν τόσο μεγάλο όγκο δεδομένων, όπως συμβαίνει με τις κλασικές στατιστικές μεθόδους. Ο ακριβής αριθμός των σημείων και ο τύπος του διαγράμματος που επιλέγονται για την ορθή παρακολούθηση της διεργασίας, εξαρτώνται από:

- τον τρόπο που οργανώνονται και συλλέγονται τα δεδομένα.
- τα χαρακτηριστικά των δεδομένων, όσον αφορά την κατανομή τους.
- τη σημασία της γρήγορης ανίχνευσης των αλλαγών στη διαδικασία.

Το πλεονέκτημα τους είναι ότι για τον ίδιο αριθμό δεδομένων, παρέχουν πολύ περισσότερη πληροφορία, ενώ χαρακτηρίζουν τη διεργασία στη ροή του χρόνου. Επίσης, είναι πολύ χρήσιμα για την ανίχνευση των συστηματικών αιτιών.

Όπως έχει ήδη αναφερθεί, εκτός από τα διαγράμματα Shewhart, υπάρχει μία εναλλακτική ομάδα διαγραμμάτων, που ανιχνεύουν έγκαιρα μικρότερες μεταβολές στη διεργασία. Με τον τρόπο αυτόν μειώνονται οι λάθος συναγερμοί σε ένα αποδεκτό επίπεδο, ενώ ταυτόχρονα γίνεται πιο έγκαιρος εντοπισμός της αιτίας ενός προβλήματος. Υπενθυμίζεται ότι τέτοιου είδους διαγράμματα είναι τα CUSUM (Cumulative Sum) και τα EWMA (Geometric Exponentially Weighted Moving Average) τα οποία είναι ιδιαίτερα ευαίσθητα, καθώς ανταποκρίνονται σε μικρές μόνο αλλαγές του μέσου (κυρίως σε μικρές, συγκριτικά με τα Shewhart που δίνουν σήμα κυρίως για μεγάλες, αν και δεν ισχύει πάντα). Με τα διαγράμματα αυτά μπορούν να εντοπιστούν διεργασίες που δεν έχουν σχεδιαστεί σωστά και οι οποίες, ενώ είναι εντός ελέγχου, ανταποκρίνονται σε χαμηλά επίπεδα ποιότητας, δηλαδή εμφανίζουν πολύ μεγάλη ενδογενή μεταβλητότητα. Αναλυτικότερα, θα συζητήσουμε για τα διαγράμματα αυτά και τις αντίστοιχες εφαρμογές τους στον Τομέα της Υγείας, σε επόμενα κεφάλαια.

Έτσι, οι μέθοδοι αυτές είναι κατάλληλες για ανάλυση δεδομένων από νοσοκομεία, όπου η μη ικανοποιητική απόδοση οφείλεται συνήθως σε μικρές ή μέτριες αλλαγές στις μέσες τιμές (*mean values*) των υπό παρακολούθηση χαρακτηριστικών ενός ασθενούς ή μίας συγκεκριμένης ασθένειας.

Ένα ακόμη σημαντικό πλεονέκτημα των EWMA είναι ότι είναι ιδιαίτερα ανθεκτικά στις αποκλίσεις της κατανομής του ποιοτικού χαρακτηριστικού που μελετάται, από την κανονική, γεγονός ιδιαίτερα σημαντικό, καθώς τα δεδομένα από το χώρο των επιστημών Υγείας συνήθως δεν είναι κανονικά κατανομημένα.

1.3.3. Εφαρμογές του Σ.Ε.Δ στις Υπηρεσίες Υγείας.

Σήμερα, αρκετές εφαρμογές του Στατιστικού Ελέγχου Διεργασιών και συγκεκριμένα των διαγραμμάτων ελέγχου λαμβάνουν χώρα στον Τομέα της Υγείας και σε αρκετά νοσοκομεία και ερευνητικά κέντρα, τόσο στην Ελλάδα, όσο και στο εξωτερικό. Τμήματα του Σ.Ε.Δ που εφαρμόστηκαν στους προαναφερθέντες τομείς, συντελέστηκαν ανεξάρτητα από τις κλασικές μεθόδους του Ποιοτικού Ελέγχου, όπως αυτές εφαρμόζονται στο βιομηχανικό και τον παραγωγικό κλάδο. Τα διαγράμματα ελέγχου (*control charts*), προτάθηκαν κυρίως, με στόχο να χρησιμοποιηθούν στην παρακολούθηση και την εξέλιξη της νοσοκομειακής λειτουργίας / απόδοσης. Παραδείγματος χάριν, ένα διάγραμμα ελέγχου θα μπορούσε να καταγράψει με απόλυτη ακρίβεια τον ρυθμό με τον οποίον μεταδίδεται μία ασθένεια σε μία ομάδα ή έναν πληθυσμό, ή γενικότερα το ρυθμό θανάτων των ασθενών.

Η χρήση των διαγραμμάτων ελέγχου σε σχετικές με την Υγεία εφαρμογές, όπως σημειώθηκε προηγουμένως, διαφέρει σημαντικά από την αντίστοιχη χρήση τους στον τομέα της βιομηχανίας και της παραγωγικής διαδικασίας.

Καταρχήν, η χρήση των δεδομένων κυριαρχεί σε μεγαλύτερο βαθμό σε ιατρικές και νοσοκομειακές εφαρμογές, όπως για παράδειγμα η καταγραφή των κρουσμάτων μίας ασθένειας σε μία πόλη, μία συγκεκριμένη χρονική περίοδο, παρά σε βιομηχανικές πρακτικές. Ασφαλώς, το πεδίο των νοσοκομειακών και ιατρικών εφαρμογών καθιστά αναγκαία τη συνεχή χρήση των διαγραμμάτων ελέγχου, καθώς σημαντικό ρόλο διαδραματίζει ο παράγοντας του κινδύνου και ο χρόνος που μεσολαβεί μεταξύ διαδοχικών αποτυχιών μίας εφαρμογής (π.χ. η αποτυχία προσαρμογής ενός μοσχεύματος σε έναν οργανισμό). Συγκεκριμένα σε αυτήν την περίπτωση, τα διαγράμματα ελέγχου εφαρμόζονται με ένα υφιστάμενο γεωμετρικό ή εκθετικό μοντέλο.

Ο Woodall (Woodall *et al.*, 1989) παρείχε μάλιστα μία λίστα με δημοσιεύσεις, αναφορικά με αυτό το θέμα, ωστόσο πραγματοποιήθηκε και σχετικά πρόσφατη έρευνα από τον Yang, τον Xie και πολλούς ακόμη. Ωστόσο, οι μέθοδοι και οι εφαρμογές των γεωμετρικών διαγραμμάτων ελέγχου δεν έχουν ακόμη χρησιμοποιηθεί σε δημοφιλή συγγράμματα, ή σε λογισμικό.

Επιπλέον, διαγράμματα ελέγχου έχουν χρησιμοποιηθεί για να διερευνηθεί η αποτελεσματικότητα των μέτρων για τη βελτίωση της ποιότητας των υπηρεσιών ενός νοσοκομείου, η αποτελεσματικότητα μεθόδων απεικόνισης οργάνων, η αποτελεσματικότητα ενός νέου ιατρικού μηχανήματος, η αποτελεσματικότητα αναισθησίας κατά τη διάρκεια μίας επέμβασης και η αποτελεσματικότητα εμβολίου *Pasteur* για τη λύσσα.

Επίσης, διαγράμματα ελέγχου έχουν χρησιμοποιηθεί κατά την παρακολούθηση ασθενών με πνευμονία και νεφρική ανεπάρκεια, για τον έλεγχο θνησιμότητας μετά από εγχείρηση καρδιάς, θνησιμότητα από βακτηριδιακή σηψαιμία και θνησιμότητα νεογνών. Άλλες περιπτώσεις όπου μπορούν να χρησιμοποιηθούν διαγράμματα ελέγχου είναι για τον έλεγχο χρόνου παραμονής σε μονάδα εντατικής θεραπείας, σε λίστες αναμονής ασθενών και στον έλεγχο ποσοστού θανάτων που οφείλονται σε τραυματισμούς. Αν για παράδειγμα, χρησιμοποιηθεί διάγραμμα ελέγχου για να μελετηθεί η αύξηση του ρυθμού εμφάνισης μίας ασθένειας, τότε σε περίπτωση που η διεργασία βρεθεί εκτός ελέγχου, θα πρέπει να διερευνηθούν οι αιτίες αύξησης του πλήθους των κρουσμάτων και να ληφθούν κατάλληλα μέτρα αφενός ιατροφαρμακευτικά για την εξάλειψη των αιτιών και αφετέρου στατιστικά για να προκύψει βελτίωση της διεργασίας (παράλληλα με τη χρήση διαγραμμάτων ελέγχου απαιτείται και η χρήση διαγραμμάτων αιτίας – αποτελέσματος και διαγραμμάτων Pareto).

1.4. Κριτήρια αξιολόγησης διαγραμμάτων ελέγχου: ARL, ATS, ANOS

Μέσο Μήκος Ροής: ARL

Μια σημαντική έννοια που σχετίζεται με τα διαγράμματα ελέγχου είναι το μέσο μήκο ροής (ή μέσο μήκος διαδρομής) του διαγράμματος (average run length, ARL) που ορίζεται με τη σχέση:

$$ARL = \frac{1}{p}$$

όπου p συμβολίζει την πιθανότητα να βρεθεί ένα σημείο του διαγράμματος ελέγχου εκτός των ορίων ελέγχου. Είναι προφανές ότι το ARL δηλώνει τον αναμενόμενο αριθμό σημείων (δειγμάτων) που πρέπει να σχεδιαστούν σε ένα διάγραμμα ελέγχου, για να εμφανιστεί ένα σημείο εκτός των ορίων ελέγχου αφού το μήκος ροής ή διαδρομής (*run length*) ακολουθεί την κατανομή $G(p)$. Αν η τυχαία μεταβλητή X ακολουθεί τη γεωμετρική κατανομή με πιθανότητα επιτυχίας p ($X \sim G(p)$), δηλαδή: $P(X = x) = p(1 - p)^{x-1}$, $x = 1, 2, 3, \dots, n$, τότε:

$$E(X) = \frac{1}{p} \quad \text{και} \quad V(x) = \frac{1}{p^2}$$

Για μια διεργασία που βρίσκεται εντός ελέγχου και στην περίπτωση που χρησιμοποιούμε 3σ όρια ελέγχου με κανονική κατανομή για την απεικονιζόμενη στατιστική συνάρτηση W , έχουμε ότι το εντός ελέγχου μέσο μήκος ροής ARL_{in} (in-control average run length) είναι ίσο με:

$$ARL_{in} = \frac{1}{\alpha} = \frac{1}{0,0027} \approx 370,4$$

όπου α είναι η πιθανότητα να βρεθεί ένα σημείο εκτός των ορίων ελέγχου ενώ η διεργασία βρίσκεται εντός ελέγχου και το χαρακτηριστικό X του προϊόντος ακολουθεί κανονική κατανομή με - εντός ελέγχου - μέση τιμή μ και διασπορά σ^2 .

Για μια διεργασία που βρίσκεται εκτός ελέγχου λόγω μετατόπισης του μέσου επιπέδου της διεργασίας, το εκτός ελέγχου μέσο μήκος ροής ARL_{out} (out-of-control average run length) είναι ίσο με:

$$ARL_{out} = \frac{1}{1 - \beta}$$

αφού η πιθανότητα να βρεθεί σε αυτήν την περίπτωση ένα σημείο του διαγράμματος ελέγχου, εντός των ορίων ελέγχου, είναι ίσο με β , οπότε βρίσκεται εκτός των ορίων ελέγχου, με πιθανότητα $1 - \beta$. Στην πράξη είναι επιθυμητή μία μικρή τιμή του ARL_{out} , αφού το ARL_{out} δηλώνει τον αναμενόμενο αριθμό δειγμάτων που πρέπει να ληφθούν, για να εντοπιστεί η αλλαγή στο μέσο επίπεδο της διεργασίας, από την στιγμή που θα συμβεί αυτή η μετατόπιση.

Σημαντικό είναι να αναφερθεί πως αυξάνοντας το μέγεθος του δείγματος n , η πιθανότητα β μειώνεται και συνεπώς βελτιώνεται η ικανότητα του διαγράμματος ελέγχου να ανιχνεύει εγκαίρως τις μετατοπίσεις του μέσου επιπέδου της διεργασίας. Με βάση τη μετατόπιση που θέλουμε να ανιχνεύσουμε, είναι δυνατό να επιλέξουμε το κατάλληλο μέγεθος δείγματος, έτσι ώστε να έχουμε συγκεκριμένη τιμή για το ARL_{out} . Αν αντί των 3σ ορίων είχαμε επιλέξει L όρια, τότε θα βλέπαμε πως καθώς αυξάνει η τιμή του L , μειώνεται η πιθανότητα εσφαλμένης ένδειξης εκτός ελέγχου διεργασίας. Ενώ αντίθετα αυξάνεται η πιθανότητα λανθασμένης ένδειξης ότι η διεργασία βρίσκεται εντός ελέγχου ενώ στην πραγματικότητα βρίσκεται εκτός ελέγχου λόγω μετατόπισης του μέσου της διεργασίας.

Τα τελευταία χρόνια ωστόσο η χρήση του ARL ως μέτρου για την περιγραφή της απόδοσης μιας διεργασίας έχει υποστεί κριτική, αφού το ARL που παρατηρείται στην πράξη

διαφέρει συνήθως αρκετά από το θεωρητικό ARL λόγω του ότι η κατανομή του μήκους ροής είναι μη-συμμετρική (παρουσιάζει δεξιά ουρά) και συνεπώς η μέση τιμή δεν μπορεί να χρησιμοποιηθεί ως αντιπροσωπευτικό μέτρο κεντρικής τάσης της κατανομής (ιδιαίτερα για μικρές τιμές του p).

Μια συνήθης πρακτική είναι να συνοδεύεται η τιμή του ARL με την τυπική απόκλιση της κατανομής του μήκους ροής **SDRL** (standard deviation run length). Στην περίπτωση που η κατανομή του μήκους ροής του διαγράμματος είναι γεωμετρική η τιμή του SDRL ταυτίζεται με την τιμή του ARL. Επιπλέον η γνώση της κατανομής του μήκους ροής, η οποία με βάση τα όσα έχουμε δει έως τώρα είναι γεωμετρική κατανομή, μας επιτρέπει τον υπολογισμό και τη χρήση ποσοστιαίων σημείων προκειμένου να έχουμε περισσότερη πληροφορία σχετικά με την απόδοση του διαγράμματος.

Θεωρώντας ένα μοντέλο ορίων 3σ που η κατανομή της W είναι κανονική με γνωστές παραμέτρους έχουμε τον ακόλουθο πίνακα για μια εντός ελέγχου διεργασία.

Πίνακας 1.4: Πίνακας στατιστικών μέτρων για την κατανομή του μήκους ροής

α	ARL	SDRL	Q_1	M	Q_3
0,0027	370,4	370,4	107	257	513

Από τον πίνακα προκύπτει ότι κατά μέσο όρο αναμένονται περίπου 370 δείγματα, μέχρι να βρεθεί σημείο εκτός των ορίων ελέγχου όταν η διεργασία είναι εντός ελέγχου. Από την τιμή της διαμέσου M ο διαχειριστής του διαγράμματος γνωρίζει ότι η πιθανότητα να περιμένω περίπου 257 δείγματα μέχρι να βρεθεί ένα σημείο εκτός των ορίων ελέγχου είναι τουλάχιστον 50%. Επιπλέον από την τιμή Q_3 έχουμε ότι η πιθανότητα να μη δώσει σημείο το διάγραμμα ένδειξη εκτός ελέγχου διεργασίας μετά το 513 απεικονιζόμενο σημείο είναι το πολύ 25%. Έτσι λοιπόν σε πρακτικές εφαρμογές η χρήση των ποσοστιαίων σημείων της κατανομής του μήκους ροής μας παρέχει περισσότερες πληροφορίες σχετικά με την απόδοση του διαγράμματος ελέγχου.

Μέσο Χρόνος Σήματος (ATS) και Μέσο Πλήθος Παρατηρήσεων (ANOS)

Στην προηγούμενη υποενότητα αναφερθήκαμε στο βασικότερο μέτρο απόδοσης ενός διαγράμματος ελέγχου και είδαμε τα κύρια χαρακτηριστικά του. Στην παρούσα υποενότητα, θα αναφερθούμε σε εναλλακτικά μέτρα απόδοσης και αξιολόγησης ενός διαγράμματος ελέγχου.

Ένα μέτρο που χρησιμοποιείται συχνά είναι ο μέσος χρόνος σήματος **ATS** (average time to signal) που ορίζεται από τη σχέση:

$$\mathbf{ATS = ARL * \tau}$$

όπου τ συμβολίζει το χρόνο που μεσολαβεί για τη λήψη δυο διαδοχικών δειγμάτων.

Έτσι η ποσότητα **ATS** δηλώνει το μέσο χρόνο που απαιτείται για να δώσει το διάγραμμα ελέγχου ένδειξη εκτός ελέγχου διεργασίας. Αν ο ρυθμός δειγματοληψίας είναι μεταβλητός τότε ορίζεται μια νέα κατηγορία διαγραμμάτων ελέγχου, τα διαγράμματα με μεταβλητό ρυθμό δειγματοληψίας (*adaptive sampling rate control charts*).

Σε αυτήν την περίπτωση το τ δεν είναι σταθερό, αλλά τυχαία μεταβλητή και ο μέσος χρόνος σήματος δίνεται από τη σχέση:

$$\mathbf{ATS = ARL * E(\tau)}$$

Είναι σημαντικό να αναφέρουμε ότι τα **ARL**, **ATS** που έχουμε αναφέρει έως τώρα υπολογίζουν την απόδοση του διαγράμματος θεωρώντας ότι η μεταβολή στην εντός ελέγχου τιμή της παραμέτρου συμβαίνει ακριβώς τη χρονική στιγμή που αρχίζει η παρακολούθηση της διεργασίας. Σε αυτή την περίπτωση μιλάμε για μέτρα **ARL** και **ATS** αρχικής κατάστασης.

Σε πολλές περιπτώσεις όμως η διεργασία ξεκινάει ευρισκόμενη εντός ελέγχου και η μεταβολή στην τιμή της υπό παρακολούθηση παραμέτρου συμβαίνει σε κάποιο τυχαίο χρονικό σημείο μεταξύ της λήψης δυο διαδοχικών δειγμάτων. Σε αυτή την περίπτωση για τον υπολογισμό της απόδοσης ενός διαγράμματος με μεταβλητό ρυθμό δειγματοληψίας χρησιμοποιείται το μέτρο **ATS** σταθερής κατάστασης το οποίο στην βιβλιογραφία αναφέρεται ως προσαρμοσμένος μέσος χρόνος σήματος (*adjusted average time to signal AATS*). Αναλόγως στην περίπτωση των διαγραμμάτων με σταθερό ρυθμό δειγματοληψίας χρησιμοποιείται το μέτρο **ARL** σταθερής κατάστασης .

Γενικότερα συγκρίσεις μεταξύ των διαφορετικών διαγραμμάτων ελέγχου χωρίς μεταβλητό ρυθμό δειγματοληψίας έχει επικρατήσει να γίνονται χρησιμοποιώντας το ARL αρχικής κατάστασης, ενώ για τα διαγράμματα με μεταβλητό ρυθμό δειγματοληψίας χρησιμοποιείται το AATS. Ωστόσο, το ATS είναι χρήσιμο, όταν ο χρόνος μεταξύ των δειγμάτων, ή μεταξύ των σημείων ενός διαγράμματος ελέγχου, ποικίλει, όπως για παράδειγμα συμβαίνει όταν έχουμε εκθετικά δεδομένα.

Μία άλλη χρήσιμη μεταβλητή που χρησιμοποιείται ως μέτρο καταλληλότητας ενός διαγράμματος ελέγχου, είναι το αναμενόμενο πλήθος των παρατηρήσεων (*average number of observations to signal – ANOS*) που καταγράφονται σε ένα πείραμα, μέχρι να σημάνει «συναγερμός» για ένδειξη εκτός ελέγχου διεργασίας. Δυστυχώς, αυτή η διάκριση μεταξύ των μέτρων απόδοσης, δεν είναι τόσο σαφής, όσον αφορά τη χρήση των διαγραμμάτων ελέγχου στον τομέα τη Υγείας.

ΚΕΦΑΛΑΙΟ 2

Μέθοδοι αξιολόγησης & σύγκρισης αποτελεσμάτων

2.1. Εφαρμογές διαγραμμάτων ελέγχου και μεθόδων πιθανοφάνειας

Όπως αναφέρθηκε στην προηγούμενη ενότητα, το ARL είναι ένα από τα σημαντικότερα κριτήρια για την αξιολόγηση της αποδοτικότητας ενός διαγράμματος ελέγχου, τόσο στον τομέα της βιομηχανίας, όσο και στον Τομέα της Υγείας. Ωστόσο, οι *Sonesson* και *Bock* (2003), υποστήριξαν τη χρήση της πιθανότητας να σημάνει εσφαλμένος συναγερμός, έναντι της χρήσης του εντός ελέγχου ARL. Ο *Frisen* (1992) πρότεινε την προβλεπόμενη τιμή ενός συναγερμού, ώστε να αποφασιστεί αν ο συναγερμός είναι εσφαλμένος, ή μη απόλυτα βασισμένος στον ακριβή χρόνο του συναγερμού. Ωστόσο, για να εφαρμοστούν οι παραπάνω μέθοδοι, θα πρέπει πρώτα να καταστεί συγκεκριμένο το μέγεθος της μετατόπισης που θα συντελεστεί, καθώς επίσης και η κατανομή του χρόνου της μετατόπισης.

Ο *Aylin* (*Aylin et al.*, 2003) υπέδειξε ότι οι μέθοδοι παρακολούθησης της δημόσιας υγείας, δεν έχουν απολύτως προσαρμοστεί με την καταγραφή πολλαπλών μονάδων (νοσοκομεία) ή με την υπερ-διασπορά των δεδομένων σχετικών με την υγεία. Η ταυτόχρονη χρήση ενός μεγάλου πλήθους διαγραμμάτων ελέγχου, λαμβάνει επίσης χώρα και στον τομέα της παραγωγικής διαδικασίας και της βιομηχανίας. Στους τομείς αυτούς φαίνεται ότι το πρόβλημα δεν επιδέχεται συγκεκριμένη λύση, πέρα από την αύξηση του πλάτους των ορίων ελέγχου, προκειμένου να μειωθεί το αναμενόμενο πλήθος των εσφαλμένων συναγερμών. Μία προσέγγιση του *Aylin* (2003) απαιτεί σε πρώτο στάδιο να συγκεκριμενοποιηθεί ο αναμενόμενος αριθμός των διαδικασιών που εντοπίζονται στιγμιαία εκτός ελέγχου.

Στη συνέχεια, οι παραπάνω συγγραφείς υπέθεσαν ότι όλες οι εκτός ελέγχου διαδικασίες μετατοπίζονται κατά το ίδιο ποσό. Υπό αυτές τις συνθήκες, αξιολόγησαν το ρυθμό όλων των εσφαλμένων συναγερμών, το ρυθμό εσφαλμένης ανίχνευσης (**FDR**), καθώς και το ρυθμό των μη εσφαλμένων, όπως επίσης και το ρυθμό επιτυχούς ανίχνευσης (**SDR**).

Αν και οι συγκεκριμένες ιδέες δείχνουν να είναι πρωτοπόρες και σε μεγάλο βαθμό αρεστές, παρόλα αυτά οι υποθέσεις πάνω στις οποίες βασίζονται είναι περιορισμένου πεδίου. Πιθανά αποτελέσματα υπερβολικής διασποράς των αποτελεσμάτων, όταν η διακύμανση της αποκρινόμενης μεταβλητής ξεπερνάει τα αναμενόμενα αποτελέσματα (π.χ. στο μοντέλο *Poisson*), θα μπορούσαν να προκαλέσουν μία ραγδαία και καθόλου αμελητέα αύξηση στο πλήθος των εσφαλμένων συναγερωμών.

Η υπερβολική διασπορά των χαρακτηριστικών δεδομένων, έχει ληφθεί υπ' όψιν στον τομέα της βιομηχανίας και της παραγωγικής διαδικασίας, με τον **Woodall** (1997) να παρέχει μία λίστα δημοσιεύσεων αναφορικά με το θέμα αυτό. Πρόσφατη εργασία πάνω στο συγκεκριμένο αντικείμενο, αναφέρθηκε από τους **Fang** (2003) και **Christensen** (2003). Οι **Hawkins** και **Olwell** πρότειναν την αντικατάσταση του μοντέλου *Poisson* σε συγκεκριμένες περιπτώσεις, από το μοντέλο της διωνυμικής κατανομής, σε περίπτωση εμφάνισης υπερ-διασποράς των αποτελεσμάτων μίας έρευνας. Ωστόσο, αρκετά ζητήματα σχετικά με τη βέλτιστη, κατά το δυνατόν, ρύθμιση της αποδοτικότητας των διαγραμμάτων ελέγχου, εξακολουθούν να παραμένουν άλυτα.

2.1.1. Μέθοδοι CUSUM και EWMA – Πλεονεκτήματα και Μειονεκτήματα

Όπως γίνεται λοιπόν αντιληπτό, από τα προηγούμενα, στη σύγχρονη ιατρική πρακτική, ο ρόλος του κλινικού εργαστηρίου και των Νοσοκομειακών εφαρμογών γίνονται ολοένα και πιο σημαντικοί για τη διάγνωση, την παρακολούθηση, την πρόγνωση και τη θεραπευτική αγωγή των ασθενών. Οι εργαστηριακές εξετάσεις αποτελούν ένα πολύ σημαντικό εργαλείο στην υπηρεσία του κλινικού ιατρού ο οποίος είναι υποχρεωμένος να παίρνει αποφάσεις σε συνθήκες σχετικής αβεβαιότητας, δεδομένου ότι η Ιατρική είναι η επιστήμη της αβεβαιότητας και των πιθανοτήτων, όπως επίσης και ότι κάθε προσδιορισμός στο χώρο του εργαστηρίου, του νοσοκομείου ή της κλινικής εμπεριέχει τον κίνδυνο του τυχαίου ή του συστηματικού σφάλματος. Η απαίτηση, συνεπώς, για διαγνώσεις και αποτελέσματα ακριβή και αξιόπιστα είναι προφανής και αναγκαία. Ιστορικά, η ανάγκη για έλεγχο και διασφάλιση της ποιότητας των ερευνών και των αποτελεσμάτων τους, δεν είναι σύγχρονη ανακάλυψη, αλλά βρίσκεται καταγεγραμμένη στα ιατρικά ιστορικά κείμενα.

Οι δεοντολογικοί και ηθικοί κανόνες του Χαμουραμί, της Παλαιάς Διαθήκης, των αιγυπτιακών παπύρων κ.λ.π έδιναν ιδιαίτερη έμφαση στη διασφάλιση της ποιότητας των παρεχόμενων, προς τον ασθενή, υπηρεσιών, επιβάλλοντας μάλιστα και ανάλογες ποινές σε περίπτωση αστοχίας της ιατρικής παρέμβασης. Όπως λοιπόν τονίστηκε ήδη, η ποιότητα των νοσοκομειακών και εργαστηριακών εφαρμογών είναι βασική απαίτηση για τη σωστή ερμηνεία των αντίστοιχων αποτελεσμάτων, και η αξιοπιστία των αντίστοιχων αναλύσεων και διαγνώσεων διασφαλίζεται με τη διερεύνηση, την ποσοτικοποίηση και τον έλεγχο όλων των πηγών αβεβαιότητας που επηρεάζουν τις συγκεκριμένες δοκιμές και έρευνες.

Η εγκυρότητα της ιατρικής και εργαστηριακής πληροφορίας είναι δυνατόν να περιορίζεται είτε από εγγενείς μεθοδολογικές ατέλειες, κατά τη διάρκεια των αναλύσεων και των εξετάσεων, είτε από συστηματικά ή τυχαία σφάλματα που υπεισέρχονται σε τεχνικό, βιολογικό ή νοσολογικό επίπεδο. Είναι συνεπώς αναγκαίο να υπάρχουν προγράμματα και διαδικασίες που θα εντοπίζουν, θα ελέγχουν και θα διορθώνουν τα σφάλματα που εμφανίζονται, προς όφελος των ασθενών.

Αθροιστικά Διαγράμματα ελέγχου (Cumulative Sum Charts – Cusum)

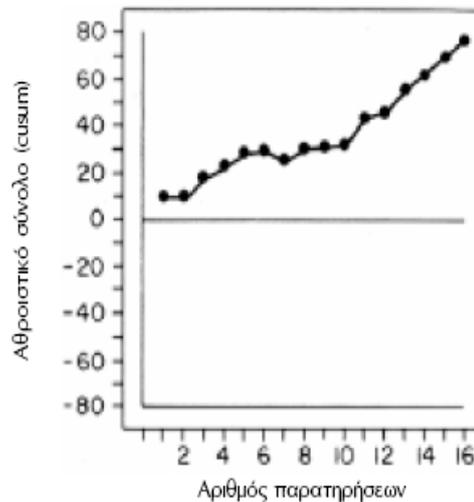
Μία εκ των δύο βασικότερων μεθόδων που εφαρμόζονται ευρέως στον τομέα της Υγείας και του Νοσοκομείου είναι το Διάγραμμα του **σωρευτικού αθροίσματος** (*Cumulative sum chart*) (εικ.2). Τα αθροιστικά διαγράμματα ελέγχου εισήχθησαν το 1954 από το Βρετανό χημικό Page (Page E., 1961) και είναι γνωστά ως διαγράμματα Cusum.

Το σωρευτικό άθροισμα είναι η τρέχουσα ολική τιμή που προκύπτει από το αλγεβρικό άθροισμα των τιμών που λαμβάνονται από τη διαφορά των καθημερινών τιμών προσδιορισμού του παρασκευάσματος ελέγχου (controls) από τη μέση τιμή του (Πίνακας 1).

Πίνακας 1. Τιμές μετρήσεων και αλγεβρικό άθροισμα των τιμών που λαμβάνονται από τη διαφορά των καθημερινών τιμών προσδιορισμού του παρασκευάσματος ελέγχου (control) από τη μέση τιμή του, για την κατασκευή του διαγράμματος CUSUM (εικ. 2).

Αύξοντας αριθμός δείγματος ελέγχου	Τιμή δείγματος ελέγχου	Διαφορές από \bar{x} , d_i	Αθροιστικό σύνολο διαφορών CS_1
1	110	+10	+10
2	100	0	+10
3	108	+8	+18
4	105	+5	+23
5	105	+5	+28
6	101	+1	+29
7	96	-4	+25
8	105	+5	+30
9	101	+1	+31
10	101	+1	+32
11	111	+11	+43
12	102	+2	+45
13	110	+10	+55
14	107	+7	+62
15	107	+7	+69
16	107	+7	+76

Αποτελεί μία περισσότερο ευαίσθητη μέθοδο ανίχνευσης συστηματικών σφαλμάτων (μετατόπιση του μέσου της διαδικασίας), συγκριτικά με διαγράμματα άλλων τύπων, όπως το διάγραμμα **Levey – Jennings**, ο σχεδιασμός του οποίου βασίζεται κυρίως στην υπόθεση ότι οι τιμές υπόκεινται σε μία τυχαία διακύμανση και ότι η διακύμανση αυτή ακολουθεί την κατανομή Gauss. Είναι ιδιαίτερα ευαίσθητο για την ανίχνευση αποκλειστικά των συστηματικών σφαλμάτων



Εικόνα 2. Διάγραμμα αθροιστικών συνόλων (CUSUM), με βάση τις τιμές του πίνακα 1.

(ακόμη και μικρών), σε αντίθεση με το διάγραμμα *Levey –Jennings* που ανιχνεύει τυχαία και συστηματικά σφάλματα.

Στο διάγραμμα αυτό, η απόκλιση από τον στόχο απεικονίζεται γραφικά με τέτοιον τρόπο, ώστε κάθε σημείο του διαγράμματος να αντιστοιχεί στο σύνολο των αποκλίσεων από την τιμή – στόχο. Στον οριζόντιο άξονα καταγράφονται οι μήνες, ενώ στον κάθετο άξονα το αθροιστικό ποσό της απόκλισης, π.χ. αν η μηνιαία καταγραφή κρουσμάτων μίας ασθένειας σε ένα χωριό της Φιλανδίας, είναι: X_1, X_2, X_3 και \bar{X} η τιμή-στόχος, τότε:

$$C_1 = X_1 - \bar{X}, C_2 = C_1 + (X_2 - \bar{X}), C_3 = C_2 + (X_3 - \bar{X}).$$

Με το διάγραμμα CUSUM μεγιστοποιούνται οι τάσεις που δημιουργούνται από τα αποτελέσματα που λαμβάνονται σε κάθε έρευνα και σε κάθε εξέταση γίνονται πιο εμφανείς οι αποκλίσεις από την τιμή-στόχο.

- **Διάγραμμα Tabular Cusum**

Στο διάγραμμα Tabular Cusum είναι απαραίτητος ο υπολογισμός δύο συσσωρευτικών αθροισμάτων για κάθε τιμή ελέγχου. Οι θετικές αποκλίσεις από τον στόχο συναθροίζονται με το άνω συσσωρευμένο άθροισμα C_i^+ (*one – sided upper cusum*), ενώ οι αρνητικές αποκλίσεις από τον στόχο συναθροίζονται με το κάτω συσσωρευμένο άθροισμα C_i^- (*one – sided lower cusum*), σύμφωνα με τις σχέσεις στην επόμενη σελίδα:

$$C_i^+ = \max[0, X_i - (\mu_0 + K) + C_{i-1}^+]$$

$$C_i^- = \max[0, (\mu_0 - K) - X_i + C_{i-1}^-]$$

με $1 \leq i \leq m$ και αρχικές τιμές $C_0^+ = C_0^- = 0$.

Τα αθροίσματα C_i^+ και C_i^- υπολογίζονται από τις διαφορές των τιμών X_i από τη μέση τιμή μ_0 , εφόσον αυτές είναι μεγαλύτερες από την τιμή αναφοράς K . Κάθε φορά που οι διαφορές γίνονται αρνητικές, το άθροισμα (C_i^+ ή C_i^-) μηδενίζεται για να αρχίσει πάλι να αυξάνεται, όταν οι διαφορές γίνουν ξανά μεγαλύτερες του μηδενός. Στο διάγραμμα Tabular Cusum τα αθροίσματα C_i^+ , C_i^- σχεδιάζονται ως δύο διαφορετικές στήλες πάνω και κάτω από τη μέση τιμή. Η τιμή αναφοράς K δεν σχεδιάζεται στο διάγραμμα Tabular Cusum. Το όριο ελέγχου που σχεδιάζεται στο διάγραμμα είναι το διάστημα απόφασης H (*decision interval*). Το ανώτερο H^+ και H^- σχεδιάζονται με δύο ευθείες γραμμές παράλληλες προς το μέσο μ_0 . Η τιμή του διαστήματος απόφασης υποδεικνύει τα ανώτατα επιτρεπτά όρια των αθροισμάτων C_i^+ , C_i^- και επιλέγεται να είναι $H = 5\sigma$.

Θέτουμε $H = h\sigma$, $K = k\sigma = \frac{\delta}{2}\sigma$, $\delta = \frac{|\mu_1 - \mu_0|}{\sigma}$, όπου k το μέγεθος της μετατόπισης που

θέλουμε να ανιχνευθεί, μ_0 η τιμή στόχος (μέση τιμή), μ_1 η εκτός ελέγχου τιμή του μέσου (η ανώτατη επιτρεπτή τιμή των δειγμάτων ελέγχου). Η επιλογή της παραμέτρου k εξαρτάται από το μέγεθος της μετατόπισης που θέλουμε να ανιχνευθεί.

• Τυποποιημένο διάγραμμα Cusum

Σε μερικές περιπτώσεις, είναι προτιμότερο να τυποποιείται η μεταβλητή X_i πριν τους υπολογισμούς των συσσωρευμένων αθροισμάτων. Ορίζεται η μεταβλητή:

$Y_i = \frac{X_i - \mu_0}{\sigma}$ η οποία ακολουθεί την κατανομή $N(0,1)$ και αποτελεί την τυποποιημένη τιμή της X_i .

Τα άνω και κάτω συσσωρευτικά αθροίσματα μετασχηματίζονται ως εξής:

$$C_i^+ = \max[0, Y_i - k + C_{i-1}^+]$$

$$C_i^- = \max[0, -k - Y_i + C_{i-1}^-]$$

με $1 \leq i \leq m$ και αρχικές τιμές $C_0^+ = C_0^- = 0$.

Τα πλεονεκτήματα των τυποποιημένων διαγραμμάτων Cusum, είναι τα εξής:

- i) Υπάρχει η δυνατότητα ύπαρξης πολλών διαγραμμάτων Cusum με τις ίδιες τιμές των k και h , καθώς οι επιλογές των παραμέτρων δεν αποτελούν ακολουθία εξαρτημένων τιμών. Οι παράμετροι δηλαδή δεν εξαρτώνται από την τυπική απόκλιση σ της κάθε διεργασίας, διότι στην τυποποιημένη κανονική κατανομή, $\sigma = 1$.
- ii) Με τη χρήση της τυποποιημένης μεταβλητής Y_i , δημιουργούνται εύκολα τα διαγράμματα για τον έλεγχο της μεταβλητότητας μίας διαδικασίας.

• Διάγραμμα Scale Cusum

Τα διαγράμματα Scale Cusum χρησιμοποιούνται για την παρακολούθηση της μεταβλητότητας που ενέχει μία διαδικασία. Έστω, όπως παραπάνω, η κανονική μεταβλητή X_i , με μέση τιμή ή τιμή στόχο μ_0 και τυπική απόκλιση σ .

Η τυποποιημένη τιμή της X_i είναι η $Y_i = \frac{X_i - \mu_0}{\sigma}$. Με τη νέα μεταβλητή v_i (Hawkins, 1991),

$$v_i = \frac{\sqrt{|Y_i|} - 0.822}{0.349}, 1 \leq i \leq m. \text{ Οι ποσότητες } v_i \text{ είναι ευαίσθητες σε αλλαγές της διασποράς μίας}$$

διαδικασίας και ακολουθούν την τυποποιημένη κανονική κατανομή $N(0, 1)$.

Τα δύο one – sided Scale Cusums είναι τα εξής:

$$S_i^+ = \max[0, v_i - k + S_{i-1}^+]$$

$$S_i^- = \max[0, -k - v_i + S_{i-1}^-]$$

με $1 \leq i \leq m$ και αρχικές τιμές $S_0^+ = S_0^- = 0$.

Αν η τυπική απόκλιση της διαδικασίας αυξάνεται, τότε θα αυξάνονται και οι τιμές των αθροισμάτων S_i^+ ξεπερνώντας κάποια στιγμή το διάστημα απόφασης H , ενώ αν η τυπική απόκλιση μειώνεται, θα μειώνονται και οι τιμές των S_i^+ μέχρι τελικά να ξεπεράσουν την

τιμή H . Οι τιμές των παραμέτρων k και h επιλέγονται όπως στα διαγράμματα του μέσου. Αν υπάρξει ένδειξη εκτός ελέγχου στο διάγραμμα Scale Cusum, τότε υπάρχει υποψία αλλαγής στην τιμή της διασποράς της διαδικασίας, ενώ αν υπάρξει ένδειξη εκτός ελέγχου και στα δύο διαγράμματα (διαγράμματα μέσου και μεταβλητότητας), τότε υπάρχει υποψία μετατόπισης στο μέσο.

Διαγράμματα ελέγχου με κινητούς μέσους και εκθετικά βάρη (Exponentially Weighted Moving Average Control Charts – EWMA)

Το διάγραμμα EWMA είναι ένα εναλλακτικό διάγραμμα των διαγραμμάτων Shewhart και παρουσιάστηκε το 1959 από τον Roberts (Roberts S.W., 1959). Η χρήση του συνίσταται κυρίως στη περίπτωση που θέλουμε να εντοπίσουμε μικρές μεταβολές στο μέσο μίας διαδικασίας και χρησιμοποιείται επίσης για μεμονωμένες παρατηρήσεις. Οι τιμές προς έλεγχο τοποθετούνται πάνω στο διάγραμμα ως z_i , τα οποία υπολογίζονται βάσει του τύπου: $z_i = \lambda x_i + (1 - \lambda)z_{i-1}$, όπου x_i είναι οι παρατηρήσεις, λ είναι μία παράμετρος που ονομάζεται συντελεστής βαρύτητας (*weighted factor*), με $\lambda \in (0,1]$. Η παράμετρος αυτή καθορίζει τον βαθμό κατά τον οποίο παλαιότερα δεδομένα εισάγονται στον υπολογισμό του EWMA στατιστικού. Όσο πιο κοντά στη μονάδα είναι η τιμή του λ , τόσο μικρότερο βάρος δίδεται στα προγενέστερα δεδομένα. Η πρώτη τιμή $z_0 = \mu_0$ είναι η μέση τιμή – τιμή στόχος (*target value*). Ως αρχική τιμή μπορούμε να χρησιμοποιήσουμε το μέσο από τα προηγούμενα δεδομένα (historical data) ή το \bar{X} (δειγματικός μέσος).

Ο EWMA είναι τελικά ένας σταθμισμένος μέσος όλων των προηγούμενων παρατηρήσεων, ο οποίος περιγράφεται ως εξής:

$$z_i = \lambda \cdot \sum_{j=0}^{i-1} (1-\lambda)^j x_{i-j} + (1-\lambda)^i z_0$$

Θεωρώντας ότι τα x_i είναι ανεξάρτητες τυχαίες μεταβλητές με διασπορά σ^2 , η διασπορά των z_i είναι: $\sigma_{z_i}^2 = \sigma^2 \left(\frac{\lambda}{2-\lambda} \right) (1 - (1-\lambda)^{2i})$. Το διάγραμμα EWMA μπορεί να κατασκευαστεί κάνοντας το γράφημα των z_i , ως προς το δειγματικό αριθμό i .

Τα όρια ελέγχου (control limits) είναι:

$$UCL = \mu_0 + L\sigma\sqrt{\left(\frac{\lambda}{2-\lambda}\right)\left(1-(1-\lambda)^{2i}\right)}, \quad CL = \mu_0, \quad LCL = \mu_0 - L\sigma\sqrt{\left(\frac{\lambda}{2-\lambda}\right)\left(1-(1-\lambda)^{2i}\right)}$$

όπου L το εύρος των ορίων και μ_0 η τιμή στόχος.

Επειδή, $1-(1-\lambda)^{2i} \rightarrow 1$, καθώς αυξάνεται το I, τα όρια μετατρέπονται:

$$UCL = \mu_0 + L\sigma\sqrt{\left(\frac{\lambda}{2-\lambda}\right)}, \quad CL = \mu_0, \quad LCL = \mu_0 - L\sigma\sqrt{\left(\frac{\lambda}{2-\lambda}\right)}$$

Τα άνω και κάτω όρια ελέγχου αποκτούν μία σταθερή τιμή και στο διάγραμμα ελέγχου απεικονίζονται ως δύο ευθείες γραμμές παράλληλες μεταξύ τους. Τα παραπάνω ισχύουν και στην περίπτωση όπου οι παρατηρήσεις είναι μεμονωμένες.

Όπως είχε επισημανθεί και στην *ενότητα 1.4*, το ARL (Average Run Length – Μέσο μήκος ροής) ενός διαγράμματος ελέγχου ορίζεται ως ο αναμενόμενος αριθμός σημείων που πρέπει να απεικονιστούν στο διάγραμμα, μέχρις ότου να υπάρξει ένδειξη εκτός ελέγχου. Διακρίνεται σε εντός ελέγχου (in control) ARL_0 και εκτός ελέγχου (out-of-control) ARL_1 . Το ARL_0 είναι προτιμότερο να είναι μεγάλο, ώστε σε μία εντός ελέγχου διαδικασία να χρειαστεί μεγάλος αριθμός δειγμάτων, μέχρι την εσφαλμένη ένδειξη ότι η διαδικασία είναι εκτός ελέγχου. Αντιθέτως, το ARL_1 είναι προτιμότερο να είναι μικρό, διότι σε μία εκτός ελέγχου διαδικασία πρέπει από όσο γίνεται μικρό αριθμό δειγμάτων να προκύψει ένδειξη ότι η διαδικασία είναι εκτός ελέγχου.

Το πρόβλημα στο EWMA έγκειται στη σωστή επιλογή των παραμέτρων λ και L, ώστε να έχει η ARL ως απόδοση την επιθυμητή τιμή. Όσον αφορά το λ , συνίσταται η χρήση μικρών τιμών του για την ανίχνευση μικρών αλλαγών στη διαδικασία, ενώ για το L χρησιμοποιείται συνήθως η τιμή $L = 3$. Αυξάνοντας το πλάτος των ορίων, μειώνεται το σφάλμα **Τύπου I**, η πιθανότητα δηλαδή να προκύψει σημείο εκτός ελέγχου, ενώ στην πραγματικότητα είναι εντός. Αυξάνεται όμως η πιθανότητα σφάλματος **Τύπου II**, δηλαδή σημεία που φαίνονται να είναι εντός ελέγχου, στην πραγματικότητα είναι εκτός. Έχουν γίνει αρκετές έρευνες σχετικά με την ARL συμπεριφορά του EWMA διαγράμματος, σε σχέση με τις τιμές των λ και L.

Κατά τους Lucas και Saccucci (1990) (Lucas J.M. and M.S. Saccucci), το βέλτιστο διάγραμμα EWMA με κριτήριο την τιμή του ARL, επιτυγχάνεται όταν:

- Έχουμε καθορίσει το επιθυμητό, εντός ελέγχου, ARL.
- Έχουμε καθορίσει το μέγεθος της αλλαγής στο μέσο της διαδικασίας που θέλουμε να εντοπίσουμε.
- Από τα διάφορα ζεύγη λ και L , επιλέγουμε εκείνο που δίνει τη μικρότερη τιμή ARL για τη συγκεκριμένη αλλαγή που μας ενδιαφέρει.

- **Μειονεκτήματα της διαγραμματικής μεθόδου EWMA**

Στην περίπτωση που έχουμε δώσει μικρή τιμή στο λ , παρουσιάζεται το ακόλουθο πρόβλημα. Αν η τιμή του EWMA στατιστικού είναι από τη μία πλευρά της κεντρικής γραμμής (CL) και προκύψει μεταβολή στο μέσο στην αντίθετη κατεύθυνση, αυτό θα οδηγήσει στην ανάγκη χρονικών περιόδων, ώστε να αντιδράσει το EWMA στη μεταβολή αυτή για να το εντοπίσει. Αυτό συμβαίνει, διότι η μικρή τιμή του λ δε δίνει μεγάλο βάρος στις νέες παρατηρήσεις. Αυτό το φαινόμενο είναι γνωστό ως φαινόμενο αδράνειας (*inertia effect*). Επίσης, το EWMA δεν αντιδρά γρήγορα σε μεγάλες αλλαγές στο μέσο, συγκριτικά με τα διαγράμματα Shewhart. Ένας τρόπος να αυξήσουμε την ευαισθησία του διαγράμματος στην ανίχνευση τους, χωρίς να «θυσιάσουμε» την ικανότητα του στην ανίχνευση μικρών αλλαγών, είναι να γίνει ένας συνδυασμός ενός Shewhart και ενός EWMA διαγράμματος. Το διάγραμμα που θα προκύψει από αυτόν το συνδυασμό, θα είναι εξίσου αποτελεσματικό τόσο στις μικρές όσο και στις μεγάλες μεταβολές στο μέσο, ιδίως στην περίπτωση όπου θα χρησιμοποιηθούν ελαφρώς μεγαλύτερα όρια στο Shewhart διάγραμμα.

2.1.2. Resetting Sequential Probability Ratio Test: Εφαρμογές στον τομέα της Υγείας.

- Εισαγωγή

Η μέθοδος «*Resetting Sequential Probability Ratio Test*», (SPRT), βασίζεται σε επαναλήψεις διαδοχικών ελέγχων υποθέσεων και αναπτύχθηκε από τον **Abraham Wald** (Wald, 1989). Το λήμμα που διατύπωσαν το 1933 οι **Neyman** και **Pearson** ενέπνευσαν τον Wald να το αναδιατυπώσει ως ένα πρόβλημα διαδοχικής ανάλυσης.

Ειδικότερα, οι Jerzy Neyman και Egon Pearson *et al.* (1933), έδειξαν με το Λήμμα που διατύπωσαν, ότι όταν διενεργείται ένας έλεγχος υποθέσεων, μεταξύ της μηδενικής υπόθεσης $H_0: \mathbf{p} = \mathbf{p}_0$ και της εναλλακτικής $H_1: \mathbf{p} = \mathbf{p}_1$, τότε το στατιστικό ελέγχου πιθανοφάνειας (likelihood ratio test) που απορρίπτει τη μηδενική υπόθεση, όταν:

$$\Lambda(\mathbf{x}) = \frac{L(\mathbf{p}_0 : \mathbf{x})}{L(\mathbf{p}_1 : \mathbf{x})} \leq \mathbf{n}, \text{ όπου } P(\Lambda(\mathbf{x}) \leq \mathbf{n} : H_0) = \alpha$$

είναι ο πιο ισχυρός έλεγχος για ένα σύνολο \mathbf{n} – κανόνων.

Πλέον, η μέθοδος SPRT, όπως αναδιατυπώθηκε από τον Wald, βασίζεται στην εξής θεωρία:

Έστω η μηδενική υπόθεση $H_0: \mathbf{p} = \mathbf{p}_0$ και η εναλλακτική $H_1: \mathbf{p} = \mathbf{p}_1$, αντίστοιχα.

Στη συνέχεια, υπολογίζεται το σωρευτικό άθροισμα του **log-likelihood ratio**, $\log \Lambda_i$:

$$S_i = S_{i-1} + \log \Lambda_i$$

Ο κανόνας τερματισμού, βασίζεται στο εξής σχήμα:

$$\mathbf{a} < S_i < \mathbf{b}, \text{ ο έλεγχος συνεχίζεται}$$

$$S_i \geq \mathbf{b}, \text{ αποδοχή της } H_1$$

$$S_i \leq \mathbf{a}, \text{ αποδοχή της } H_0$$

όπου \mathbf{a} και \mathbf{b} ($0 < \mathbf{a} < \mathbf{b} < \infty$) εξαρτώνται από τα επιθυμητά σφάλματα Τύπου I και II.

Τα \mathbf{a} , \mathbf{b} επιλέγονται ως εξής: $\mathbf{a} \approx \log \frac{\beta}{1-\alpha}$ και $\mathbf{b} \approx \log \frac{1-\beta}{\alpha}$

Με άλλα λόγια, τα **a** και **b** πρέπει να επιλέγονται εκ των προτέρων, ώστε οι προηγούμενοι κανόνες να τεθούν καταλλήλως. Ο λόγος που χρησιμοποιούνται προσεγγίσεις είναι διότι στη διακριτή περίπτωση, το σήμα μπορεί να ξεπεράσει το όριο ελέγχου μεταξύ των δειγμάτων.

Ενώ αρχικά η μέθοδος SPRT αναπτύχθηκε για χρήση ελέγχου της ποιότητας στο χώρο της βιομηχανίας και της παραγωγής, πλέον έχει διαμορφωθεί και η χρήση της έχει επεκταθεί και στο χώρο της Υγείας, για παράδειγμα στην ανίχνευση αύξησης του ρυθμού κρουσμάτων για το σύνδρομο Down.

Το **σύνδρομο Down** (ή αλλιώς Τρισωμία 21) αποτελεί τη συχνότερη γενετική αιτία της νοητικής υστέρησης. Πιο συγκεκριμένα, το σύνδρομο αυτό περιλαμβάνει ένα σύνολο συμπτωμάτων που παρουσιάζονται, εκ γενετής, σε ένα άτομο και ευθύνονται για ένα βαθμό μαθησιακής δυσκολίας ή νοητικής υστέρησης, καθώς και για σωματικά και εξωτερικά χαρακτηριστικά γνωρίσματα. Η πρώτη περιγραφή του συνδρόμου πραγματοποιήθηκε το 1860 από το γιατρό *John Langdon Down*, ο οποίος παρατήρησε, σε μια ομάδα ιδρυματικών ατόμων, που δε σχετίζονταν μεταξύ τους, παρεμφερή φυσιολογικά χαρακτηριστικά.

Οι άνθρωποι με σύνδρομο Down, παρουσιάζουν μεγάλη ποικιλία ως προς τις αντιληπτικές, κοινωνικές και εκφραστικές ικανότητες και την ανάπτυξη των δικών τους μοναδικών δεξιοτήτων και ταλέντων.

Σε μια προσπάθεια να αποφευχθεί μια, σε βάθος, επιστημονική περιγραφή και ανάλυση, παραθέτουμε, κωδικοποιημένα, τα βασικότερα κοινά χαρακτηριστικά που εμφανίζονται στα παιδιά εξαιτίας του συνδρόμου Down:

- Τα παιδιά με σύνδρομο Down έχουν αρκετά στρογγυλεμένο πρόσωπο, σχιστά μάτια και μικρή μύτη.
- Τα μωρά αμέσως μετά τη γέννησή τους, κατά κανόνα δεν κλαίει.
- Δε διαθέτουν μυϊκή δύναμη.
- Έχουν αδύνατα και κοντά δάχτυλα χεριών.
- Το ανοσοποιητικό τους σύστημα είναι επιρρεπές σε ασθένειες.
- Παρουσιάζουν μαθησιακές δυσκολίες και μειωμένη νοητική ανάπτυξη.
- Παρατηρείται μεγάλη όρεξη για φαγητό και παχυσαρκία.
- Είναι, συνήθως, άτομα που εκδηλώνουν τα συναισθήματά τους.

- **Εφαρμογές της SPRT**

Οι *Morton* και *Lindstein* λοιπόν, το 1976 πρότειναν τη χρήση της μεθόδου SPRT για να ανιχνεύσουν τυχούσα αύξηση στο ρυθμό κρουσμάτων του συνδρόμου Down. Οι μετρήσεις που παρατηρήθηκαν με την πάροδο του χρόνου, υποτέθηκε ότι βασίζονται σε κατανομή Poisson, όπου οι –εντός ελέγχου- τιμές της παραμέτρου Poisson ίσως να παρουσιάζουν μεγάλη διασπορά στο πέρασμα του χρόνου. Συγκεκριμένα, προτίμησαν την εφαρμογή της μεθόδου στο διάγραμμα Cusum, για τα Poisson δεδομένα, καθώς το διάγραμμα Cusum δε θα επέτρεπε για την ώρα διασπορά στις – εντός ελέγχου- τιμές της παραμέτρου. Στήριξαν μάλιστα τη θέση ότι οι πιθανότητες σφαλμάτων Τύπου **I** και **II**, δηλαδή **α** και **β** αντίστοιχα, ήταν ευκολότερο να ερμηνευθούν συγκριτικά με τα ARL's.

Έχει συζητηθεί και έχει υποστηριχθεί το γεγονός ότι η χρήση των πιθανοτήτων σφάλματος με τη συγκεκριμένη μέθοδο, είναι ακατάλληλη και πως η χρήση του διαγράμματος Cusum είναι η καλύτερη επιλογή. Όπως, αναφέρθηκε και στην αρχή αυτής της ενότητας, η SPRT που αναδιατυπώθηκε από τον **Wald**, είναι ένας διαδοχικός έλεγχος υποθέσεων, για τον οποίον η δοκιμή σταματάει μόλις ένα στατιστικό ελέγχου (test statistic) που συνήθως βασίζεται σε ανεξάρτητα και ομοιόμορφα κατανομημένες παρατηρήσεις, ξεπεράσει είτε ένα ανώτατο όριο απόρριψης, είτε ένα κατώτατο όριο αποδοχής.

Η διαδοχική χρήση της SPRT, όπου τα όρια καθορίστηκαν με τις προσεγγίσεις του Wald, προτάθηκε να εφαρμοστεί στο Στατιστικό Έλεγχο Διεργασίας, για διασφάλισης της ποιότητας και της εγκυρότητας της διάγνωσης στο χώρο της Υγείας. Παρά το γεγονός ότι μέθοδοι αυτού του είδους επιδοκιμάζονται ιδιαίτερα από τους παθολόγους, οι οποίοι κατέχουν το αντικείμενο του ελέγχου υποθέσεων, ωστόσο οι πιθανότητες σφαλμάτων δεν είναι ερμηνεύσιμες με επαναλαμβανόμενη χρήση διαδοχικών ελέγχων (sequential tests), όπως υπέδειξαν προσφάτως οι *Grigg* (2003) και *Spiegelhalter* (2004).

Με επαναλαμβανόμενους ελέγχους η πιθανότητα να απορριφθεί τελικά η αληθής μηδενική υπόθεση, οδηγώντας έτσι σε εσφαλμένο συναγερμό, ισούται με 1. Επίσης, κάθε σταθερή μετατόπιση της διαδικασίας μπορεί να ανιχνευθεί, το θέμα είναι όμως πόσος καιρός θα περάσει για να γίνει αυτό.

Ο *Rogers* (2003) στήριξε τη θέση ότι οι πιθανότητες σφάλματος είναι απαραίτητες για να προσδιορίσουμε την ARL απόδοση των διαγραμμάτων Cusum, χωρίς όμως αυτό να είναι μεγίστης σημασίας ζήτημα. Ο *Grunkemeier* (2003) επίσης χρησιμοποίησε πιθανότητες σφαλμάτων στο σχεδιασμό διαγράμματος Cusum.

Στο παρακάτω διάγραμμα (**Figure 1**), απεικονίζεται ένα single SPRT που εφαρμόστηκε από τον Spiegelhalter *et al.* (2003), χρησιμοποιώντας δεδομένα από τη πρακτική του Dr. Harold Shipman, ο οποίος καταδικάστηκε το 2000, για τη δολοφονία 15 ασθενών του και ενεπλάκη στην πρόκληση θανάτου σε ακόμη 200 -300 άτομα. Τα περισσότερα από τα θύματα του ήταν γυναίκες, άνω των 65 ετών. Η συγκεκριμένη SPRT σχεδιάστηκε για να ελέγξει τη μηδενική υπόθεση της σταθερής επίδοσης, ενάντια στις διπλάσιες πιθανότητες θανάτου, σε σχέση με τους τοπικούς γιατρούς.

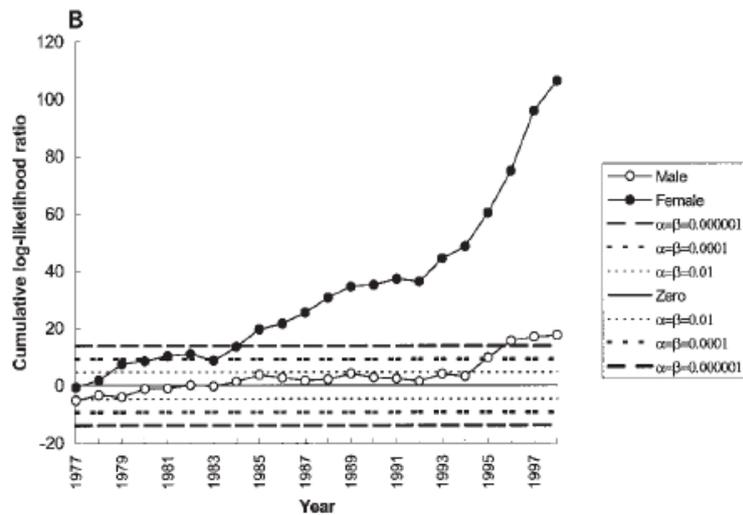


FIGURE 1. Sequential Probability Ratio Test (SPRT) for Detection of a Doubling in Mortality Risk: Age >64 Years and Death in Home/Practice for Dr. Harold Shipman. (Figure 2(B) of Spiegelhalter *et al.* (2003)). Reproduced by permission of the Oxford University Press.

Στο διάγραμμα **Figure 2**, δείχνει ένα, αντίστοιχο με το προηγούμενο, διάγραμμα, μόνο που στη συγκεκριμένη περίπτωση η σύγκριση γίνεται με τους ιατρούς στην Αγγλία και την

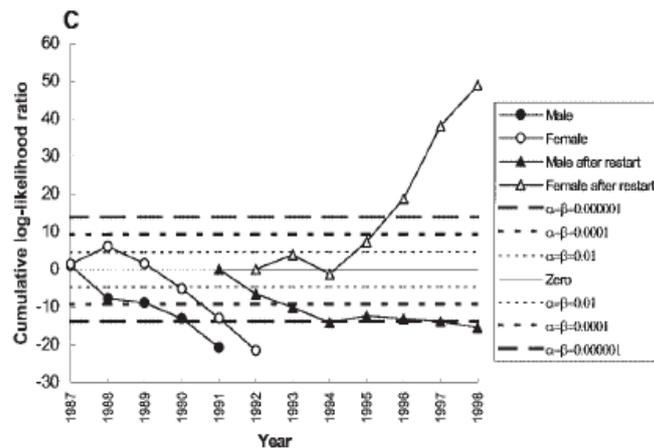


FIGURE 2. Sequential Probability Ratio Test for Detection of a Doubling in Mortality Risk Allowing for Restarts: Age >64 Years. α , False-Positive Error Rate; β , False-Negative Error Rate. (Figure 3(C) of Spiegelhalter *et al.* (2003)). Reproduced by permission of the Oxford University Press.

Ουαλία. Και τα δύο διαγράμματα δείχνουν περισσότερους θανάτους γηραιών γυναικών από ό,τι θα αναμενόταν υπό σταθερή επίδοση.

Το σύνηθες **one-sided** διάγραμμα Cusum μπορεί να θεωρηθεί ως SPRT με κατώτερο όριο αποδοχής στο μηδέν. Η μέθοδος SPRT όπως χρησιμοποιείται συνήθως σε εφαρμογές σχετικές με τον τομέα της Υγείας και της Πρόληψης, παρουσιάζει ένα μειονέκτημα: το στατιστικό ελέγχου της μπορεί να προσεγγίσει το μηδέν, μόνο όταν εντοπιστεί αύξηση π.χ. στο ρυθμό μετάδοσης μίας ασθένειας. Σύμφωνα με την ορολογία που χρησιμοποιείται στο χώρο της Υγείας, οι συγγραφείς χρησιμοποιούν τον όρο «*credit*», όταν αναφέρονται σε αυτό το πρόβλημα, ενώ στο χώρο της Βιομηχανίας και της Παραγωγής, χρησιμοποιείται ο όρος «*inertia*» (αδράνεια).

Για διαγράμματα που σχεδιάστηκαν να ανιχνεύουν μία μετατόπιση στο μέσο μίας κανονικά κατανομημένης διαδικασίας, οι **Woodall** και **Mahmoud** (2007) όρισαν την «αντίσταση ενός διαγράμματος στο σήμα» (*signal resistance of a chart*), ως τη μεγαλύτερη τυποποιημένη μέση τιμή δείγματος, χωρίς να οδηγεί κατ' ανάγκην σε σήμα εκτός ελέγχου. Ωστόσο είναι αναγκαίος ένας άλλος προσδιορισμός αυτής της έννοιας, για τα διαγράμματα, για την παρακολούθηση των παραμέτρων της γεωμετρικής και της εκθετικής κατανομής, όπως προκύπτουν στον τομέα της δημόσιας υγείας. Σε αυτές τις περιπτώσεις, η αύξηση του στατιστικού Cusum, οριοθετείται. Έτσι είναι πιο ουσιαστικό να ορίσουμε την «αντίσταση στο σήμα», αναφορικά με το μέγιστο αριθμό των διαδοχικών παρατηρήσεων, όπου δεν είναι δυνατόν να εντοπιστεί σήμα εκτός ελέγχου.

Τα διαγράμματα ελέγχου τα οποία βασίζονται στην επαναλαμβανόμενη χρήση της SPRT, και προτάθηκαν από τους **Reynolds** και **Stoumbos** για την παρακολούθηση ενός ποσοστού, μπορούν να εφαρμοστούν υπό ακριβή επιθεώρηση. Οι ANOS συγκρίσεις τους, εφαρμόζονται στην προηγούμενη περίπτωση (100% επιθεώρηση). Παρόλα αυτά, οι **Stoumbos** και **Reynolds** έδειξαν ότι καθώς έχουν σχεδιαστεί κατά το βέλτιστο για στατιστικές επιδόσεις, τα SPRTs έχουν ένα όριο αποδοχής πολύ πιο κοντά στο μηδέν, από το αντίστοιχα που παράχθησαν με τις προσεγγίσεις του Wald. Σε μερικές περιπτώσεις, το «καλύτερο» κατώτερο όριο αποδοχής μπορεί να είναι ακόμη και θετικό. Η στατιστική απόδοση των συγκεκριμένων διαγραμμάτων είναι πολύ πιο κοντά στην αντίστοιχη του διαγράμματος Cusum, αλλά στη περίπτωση ενός θετικού κατώτερου ορίου, τα διαγράμματα θα παρουσιάζουν καλύτερες ιδιότητες ως προς την αδράνεια. Γενικά, οι **Reynolds** και **Stoumbos** (1999, 2000) έδειξαν ότι η επαναλαμβανόμενη χρήση της μεθόδου SPRT, στο διαγραμματικό έλεγχο, είναι αρκετά χρήσιμη στη μείωση του κόστους δειγματοληψίας, όταν τα ποσοστά δειγματοληψίας είναι μεταβλητά.

2.1.3. Sets Method: Ανίχνευση αύξησης στο ρυθμό σπανίων ασθενειών

Ο Chen (1988) πρότεινε αυτό που σήμερα είναι γνωστό ως «Sets Method», με σκοπό την ανίχνευση στο ρυθμό κρουσμάτων ενός σπανίου γεγονότος, όπως για παράδειγμα είναι η εμφάνιση γενετήσιων δυσμορφιών. Η χρήση της μεθόδου αυτής, αναλύεται στην παρούσα ενότητα, καθώς και μερικές από τις τροποποιήσεις της.

Μία βασική παραδοχή είναι ότι τα αποτελέσματα όλων των γεννήσεων συλλέγονται, υπό ακριβή επιθεώρηση, κατά χρονολογική σειρά. Υπολογίζεται έπειτα το σύνολο των γεννήσεων (*counts of births*) που προκύπτουν ανάμεσα σε διαδοχικές δυσμορφίες ενός συγκεκριμένου τύπου. Ο τρόπος λειτουργίας της sets method είναι ο εξής:

Αν οι μετρήσεις αυτές των γεννήσεων, μεταξύ n πλήθους, διαδοχικών δυσμορφιών, είναι όλες μικρότερες από μία καθορισμένη σταθερά, έστω k , τότε σηματοδοτείται μία αύξηση στο ρυθμό κρουσμάτων αυτής της δυσμορφίας.

Γενικά, αρκετοί συγγραφείς μελέτησαν το υπόβαθρό και τη λειτουργία αυτής της μεθόδου και έκαναν αρκετές τροποποιήσεις. Οι **Walter** (1987) και **Radaelli** (1992) πρότειναν μεθόδους τύπου *Cuscore*. Στις συγκεκριμένες μεθόδους, μία μεταβλητή ορίζεται να παίρνει την τιμή **1**, αν το πλήθος των γεννήσεων - μεταξύ διαδοχικών δυσμορφιών - προσεγγίζει ένα οριακό σημείο, αλλιώς παίρνει την τιμή **-1**. Στη συνέχεια, αυτές οι μεταβλητές χρησιμοποιούνται σε ένα διάγραμμα σωρευτικού αθροίσματος (*cumulative sum chart*).

Ο **Lie** (1991) παρείχε μία έκθεση της Sets Method και άλλων αντιμαχόμενων προσεγγίσεων της. Σε ίδιο μήκος κύματος, ο **Sego** (2005) ανέλυσε σε σχετική αναφορά του το ίδιο θέμα, συγκρίνοντας ωστόσο και την στατιστική απόδοση όλων των σχετικών μεθόδων. Συγκεκριμένα οι **Lie** και **Sego** σύγκριναν τη sets method με το *Bernoulli* διάγραμμα Cusum (είναι βασισμένο σε δεδομένα τύπου Bernoulli), καθώς πρότερες συγκρίσεις ως προς την απόδοση, περιορίστηκαν κυρίως στο – λιγότερο αποδοτικό – *Poisson* διάγραμμα Cusum.

Ο **Sego** (2005) διαπίστωσε ότι η sets method του **Chen** (1978) είναι μία ειδική περίπτωση των «run rules» που διατύπωσε ο **Page** (1995). Οι **Champ** και **Woodall** (1987) έδειξαν ότι για ένα μεγάλο εύρος κανόνων **run rules**, το αποτέλεσμα των κανόνων αυτών ήταν να αυξηθεί η ευαισθησία του διαγράμματος Shewhart στις μικρές μεταβολές του μέσου της Κανονικής Κατανομής, αλλά η ευαισθησία του διαγράμματος Cusum σε αυτές τις μεταβολές ήταν καλύτερη. Ο **Sego** (2005) έδειξε ότι ένα βέλτιστα σχεδιασμένο, Bernoulli διάγραμμα Cusum, έχει καλύτερη απόδοση από ό,τι οι μέθοδοι τύπου sets methods. Οι sets methods δεν χρησιμοποιούν όλη την πληροφορία των δεδομένων και η απόδοση των διαγραμμάτων

Cusum παρουσιάζει ιδιότητες βελτίστου, όπως συζητήθηκε από τον *Moustakides* (1986) και τους *Hawkins & Olwell* (1998).

Τέλος, ο *Sitter* (1990) πρότεινε μία μέθοδο η οποία εκπέμπει σήμα, όταν δύο σήματα της sets method διαχωρίζονται από λιγότερα του ενός συγκεκριμένου αριθμού περιστατικών. Ωστόσο, ο *Sego* (2005) έδειξε ότι η ανάλυση μέσου μήκους ροής (ARL) για την απόδοση αυτής της μεθόδου, δεν έλαβε υπ' όψιν την επίδραση της σταθεράς **headstart**. Μία ανάλυση σταθερού μήκους ροής (steady-state run length), έδειξε ότι αυτή η μέθοδος δεν ήταν καλύτερη από τη sets method, ως προς την ανίχνευση αργών μετατοπίσεων της διαδικασίας. Ο *Sego* (2005) επισήμανε ότι η **sets method** λόγω της ελαφριάς επίδρασης του headstart, καθιστά αναγκαίες τις συγκρίσεις απόδοσης - βάσει του σταθερού μήκους ροής, σε σχέση με τις αντίστοιχες «zero state».

2.2. League Tables, Διαγράμματα σύγκρισης και Funnel Charts

2.2.1. Ανάλυση και Σύγκριση μεθόδων

Στην παρούσα ενότητα, συζητούνται μερικές από τις γραφικές μεθόδους που χρησιμοποιούνται για να συγκριθεί η απόδοση των νοσοκομειακών ιδρυμάτων. Κατά τις συγκρίσεις των ποσοστών θνησιμότητας, ενός μεγάλου πλήθους νοσοκομείων ή ιατρών, είναι σύνηθες να παρουσιάζονται τα δεδομένα σε έναν πίνακα κατάταξης (**league table**), όπου τα στοιχεία κατατάσσονται από το καλύτερο προς το χειρότερο. Ο πίνακας αυτός ορισμένες φορές αναφέρεται ως ένα *διάγραμμα σύγκρισης* (**comparison chart**). Ένα παράδειγμα αυτού του διαγράμματος δίδεται στο διάγραμμα **Figure 3**. Οι League Tables έχουν λάβει έντονη κριτική από τον *Adab* (2002) και πολλούς ακόμη, κυρίως λόγω του ότι η στατιστική σημαντικότητα της κατάταξης των στοιχείων δεν οδηγεί σε ένα σαφές συμπέρασμα. Με τους *league tables* και τα *διαγράμματα σύγκρισης* (comparison charts), όπως περιγράφεται από τους *Lee* και *McGreevy* (2002), δίδεται ένα διάστημα εμπιστοσύνης για κάθε υπό μελέτη στοιχείο (*unit*) που αποκλίνει από το μέσο όρο, μόνο εάν ο μέσος όρος δεν περιλαμβάνεται στο αντίστοιχο διάστημα.

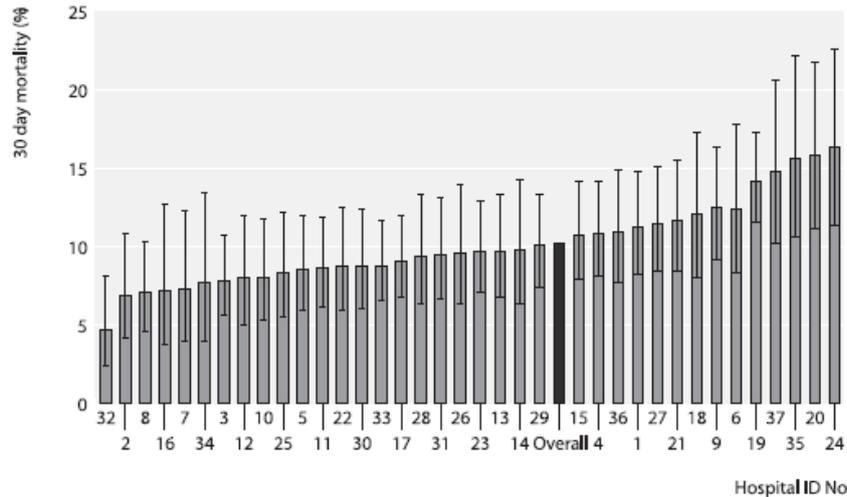


FIGURE 3 . Example of a League Table from Adab et al. (2002). Reproduced with permission from the BMJ Publishing Group.

Στο παραπάνω διάγραμμα λοιπόν, το Νοσοκομείο με τον αύξοντα αριθμό **#32**, θεωρείται ότι παρουσιάζει την καλύτερη, ως προς το μέσο όρο, απόδοση, καθώς σε περίοδο μελέτης 30 ημερών το συγκεκριμένο νοσοκομείο παρουσίασε το μικρότερο ποσοστό (%) θνησιμότητας ασθενών. Αντιθέτως, όπως παρατηρούμε και στο διάγραμμα, τα νοσοκομεία **#19, 20, 24** και **35** έχουν τη χειρότερη, ως προς τη μέσο όρο, απόδοση.

Οι *Adab* (2002) και *Mohammed* (2001) πρότειναν ένα νέο διάγραμμα ελέγχου, όπου τα όρια ελέγχου είναι δεδομένα. Η συγκεκριμένη προσέγγιση είναι ιδιαίτερα χρήσιμη, ωστόσο δεν μπορεί να θεωρηθεί ως ένα τυπικό διάγραμμα ελέγχου, αφού τα δεδομένα δεν είναι σε χρονική σειρά, συνεπώς ο όρος «διάγραμμα ελέγχου» δείχνει ακατάλληλος.

Για να κατανοήσουμε λοιπόν τη φιλοσοφία και τον τρόπο εφαρμογής του διαγράμματος που προτάθηκε από τους Adab και Mohammed, διενεργήθηκε μία έρευνα προκειμένου να αξιολογηθεί η αποδοτικότητα των νοσοκομειακών μονάδων, ανάλογα με το πλήθος των ασθενών που εισάγονται στο κάθε νοσοκομείο μέσα σε μία χρονική περίοδο 30 ημερών και το πλήθος των θανάτων που παρατηρούνται στο νοσοκομείο αυτό, για την ίδια ακριβώς περίοδο.

Τα αποτελέσματα λοιπόν της έρευνας παρουσιάζονται στο διάγραμμα της επόμενης σελίδας.

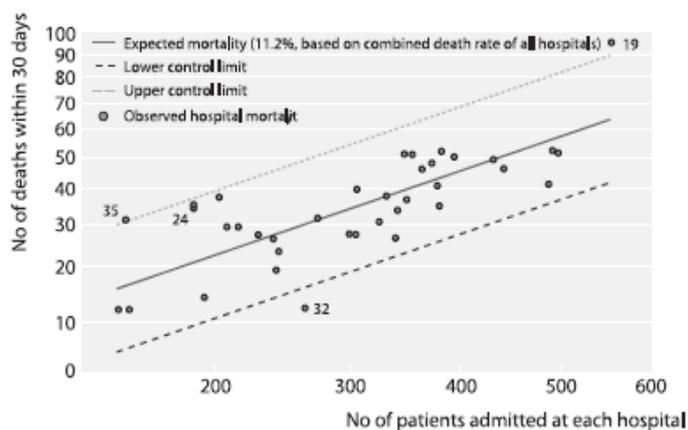


FIGURE 1 Example of Proposed Control Chart by Adab et al. (2002). Reproduced with permission from the BMJ Publishing Group.

Η συνεχής γραμμή εκφράζει το αναμενόμενο ποσοστό θανάτων, της τάξης του **11,2 %**, βάσει του συνδυασμένου ποσοστού θνησιμότητας όλων των νοσοκομείων, ενώ οι διακεκομμένες γραμμές εκφράζουν το κατώτερο και το ανώτερο όριο ελέγχου, αντίστοιχα, για το συγκεκριμένο διάγραμμα. Όπως παρατηρούμε, το νοσοκομείο **#32** παρουσιάζει καλύτερη απόδοση, σε σχέση με το μέσο όρο, καθώς επιδεικνύει το λιγότερο ποσοστό θνησιμότητας, για το σύνολο των ασθενών που εισήχθησαν στο νοσοκομείο το συγκεκριμένο μήνα που διενεργήθηκε η έρευνα. Αντιθέτως, τα νοσοκομεία **#19** και **#35** παρουσιάζουν τη χειρότερη, ως προς τη μέση τιμή, απόδοση για την ίδια χρονική περίοδο.

Το διάγραμμα, γνωστό ως **funnel plot**, προτάθηκε από τον *Spiegelhalter* (2002) και το αντίστοιχο διάγραμμα ελέγχου του ποσοστού θανάτων από τον *Tekkis* (2003). Στο funnel plot, το πλήθος των ασθενών ανά νοσοκομείο/ ιατρό παρουσιάζεται στον οριζόντιο άξονα και το ποσοστό του χαρακτηριστικού που μελετούμε στον κατακόρυφο άξονα. Οι *decision lines* βασίζονται σε ένα πολλαπλάσιο του τυπικού σφάλματος σχετικά με τη συνολική μέση τιμή. Συνεπώς, η στατιστική σημαντικότητα της απόκλισης από τη μέση τιμή αξιολογείται παρόμοια με το διάγραμμα σύγκρισης (*comparison chart*) και τον *league table*, αλλά τα υπό μελέτη στοιχεία δεν κατατάσσονται από το καλύτερο στο χειρότερο, ή κατά αλφαβητική σειρά, όπως συμβαίνει με ένα τυπικό διάγραμμα ελέγχου.

Αντιπροσωπευτικό παράδειγμα της εφαρμογής διαγράμματος τύπου **funnel plot**, παρουσιάζεται στο διάγραμμα της επόμενης σελίδας όπου μελετάται το ποσοστό των ασθενών που εισήχθησαν ξανά στο νοσοκομείο, λόγω ενός σφοδρού ξεσπάσματος μιας ασθένειας. Ειδικότερα, μελετάται η απόδοση συγκεκριμένων νοσοκομείων στην Αγγλία και την Ουαλία αντίστοιχα, τη χρονική περίοδο 2000 – 2001.

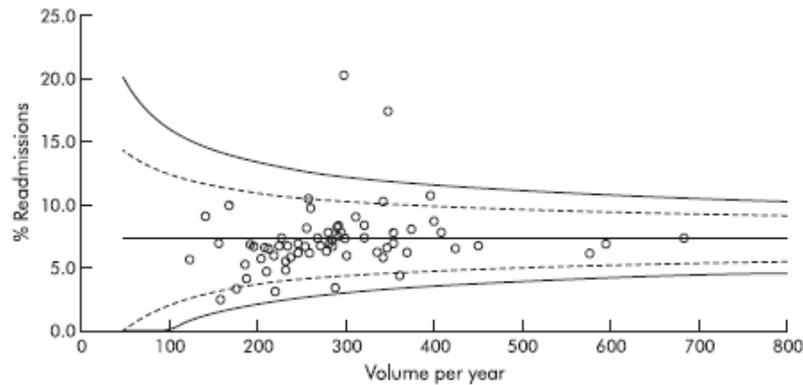


FIGURE 1. "Funnel plot" of Emergency Readmission Rates Following Treatment for a Stroke in Large Acute or Multiservice Hospitals in England and Wales in 2000–2001. Exact 95% and 99.9% Binomial Limits are Used. (From Spiegelhalter (2002). Reproduced with permission from the BMJ Publishing Group.).

Στο παραπάνω διάγραμμα, έχουν χρησιμοποιηθεί **95%** και **99,9%** όρια ελέγχου, βάσει της διωνυμικής κατανομής και απ' ό,τι μπορούμε να παρατηρήσουμε δύο νοσοκομεία παρουσιάζουν σημαντικά χαμηλή απόδοση (για αυτά τα δύο νοσοκομεία η διαδικασία τίθεται εκτός των ορίων ελέγχου).

2.2.2. Συμπεράσματα

Στον τομέα της Υγείας και της Πρόληψης, τα επιχειρήματα υπέρ της χρήσης των διαγραμμάτων Cusum, βασισμένων στην πιθανοφάνεια του Page (1954), για την ανίχνευση συνεχών αλλαγών στη Φάση II, είναι πολύ πιο πειστικά από τα αντίστοιχα της χρήσης της SPRT μεθόδου και της *Sets Method* που αναφέραμε στις προηγούμενες ενότητες. Η διαγραμματική μέθοδος Cusum, την οποία επιδοκίμασε έντονα και ο Spiegelhalter (2004), μπορεί να εφαρμοστεί για οποιαδήποτε κατανομή πιθανότητας.

Αξίζει βεβαίως να σημειωθεί ότι τα διαγράμματα Cusum επιτρέπουν την προσαρμογή του κινδύνου (**risk adjustment**), παραγόντων δηλαδή που επηρεάζουν το αποτέλεσμα μίας επέμβασης ή τη γρήγορη εξάπλωση μίας ασθένειας, όπως η ηλικία, το φύλο του ασθενούς ή ακόμη το ιατρικό ιστορικό του, οι οποίοι θα μας δώσουν μία ξεκάθαρη εικόνα για το ποσοστό του χαρακτηριστικού που μελετούμε, π.χ. το ποσοστό των επιτυχών μεταμοσχεύσεων καρδιάς σε συγκεκριμένα νοσοκομεία.

Παρόλα αυτά δεν έχουν όλη την ίδια γνώμη σχετικά με την απόδοση και την αξιοπιστία των διαγραμμάτων Cusum. Ο Rogers (2004), για παράδειγμα, στήριξε ότι το διάγραμμα

Cusum, η μέθοδος SPRT βασισμένη σε πιθανότητες σφάλματος και άλλες ακόμη διαγραμματικές μέθοδοι, είναι όλες ισοδύναμες ως προς την εγκυρότητα και την αξιοπιστία που παρέχουν και για αυτό μάλιστα ο όρος «σφωρευτικό διάγραμμα ελέγχου» (*cumulative sum chart*), αναφέρεται συνήθως σε όλες αυτές τις μεθόδους. Το γεγονός μάλιστα ότι η χρήση ενός διαγράμματος Cusum απαιτεί ένα εξωτερικό σημείο αναφοράς, δηλαδή ένα σταθερό σημείο που θα λειτουργήσει ως πρότυπο για την εφαρμογή του, θεωρείται από αρκετούς ότι η μέθοδος Cusum αποτελεί μία ειδική μορφή της μεθόδου RSPRT, που περιγράφηκε στην ενότητα 2.1.2.

Συγκεκριμένα, η μέθοδος RSPRT περιγράφεται ως μία γενίκευση του συνήθους *one-sided* διαγράμματος Cusum με μία αρχική τιμή όχι απαραίτητα ίση με το μηδέν. Αν χρησιμοποιηθούν, αυτά τα διαγράμματα πρέπει να σχεδιαστούν βάσει των κριτηρίων απόδοσης μήκους ροής, όπως το ARL, ANOS και ATS, όπως συνέστησαν οι *Reynolds* και *Stoumbos* (2000, 2001). Γενικά, συνιστάται έντομα ότι οι ιδιότητες του μήκους ροής και τα κριτήρια απόδοσης αντικαθιστούν τη χρήση των πιθανοτήτων σφάλματος στη Φάση II σχεδιασμού του διαγράμματος ελέγχου & ανάλυσης. Η εφαρμογή των πιθανοτήτων σφάλματος (error probabilities) έχει νόημα για μερικές μόνο περιπτώσεις ελέγχου και δειγματοληψίας, όπως κατά τη χρήση της *single* SPRT, ή κατά τη χρήση ενός διαγράμματος *Shewhart*, όπου με γνωστές-εντός ελέγχου-τιμές της παραμέτρου.

Για τη διασταύρωση δεδομένων, το διάγραμμα **funnel plot** προτιμάται περισσότερο σε σχέση με τον **league table**, ή τα αντίστοιχα σε αυτό διαγράμματα ελέγχου.

Οι επαγγελματίες του βιομηχανικού κλάδου και του τομέα της παραγωγής ενθαρρύνονται να εξετάσουν τη χρησιμότητα της προσαρμογής του κινδύνου και τα πιθανά οφέλη από τη χρήση χωρικών δεδομένων, με στόχο την ανίχνευση clusters σφοδρών νόσων και μεταδοτικών ασθενειών. Αν και αυτές είναι θεμελιώδεις προσεγγίσεις, στον τομέα της παρακολούθησης της υγείας, ωστόσο δεν έχουν επεκταθεί τόσο πολύ και στον τομέα της βιομηχανίας. Από την άλλη μεριά όμως, τα διαγράμματα ελέγχου σχεδιασμένα με χρήση **regression-adjusted** μεταβλητών, ίσως αποδειχθούν πολύ χρήσιμες στο χώρο της υγειονομικής περίθαλψης.

Οι επαγγελματίες του βιομηχανικού κλάδου και της εφαρμογής του Στατιστικού Ελέγχου Διεργασιών (Σ.Ε.Δ) στο χώρο της παραγωγής, ενθαρρύνονται επίσης να ερευνήσουν περαιτέρω τις εφαρμογές του Σ.Ε.Δ, προς όφελος των ασθενών και της υγειονομικής τους φροντίδας. Η ανάγκη για βελτιωμένες υπηρεσίες πρόληψης, διάγνωσης και θεραπείας είναι πλέον επιβεβλημένη και ο ρόλος της παρακολούθησης και του ελέγχου της δημόσιας υγείας

καθίσταται ολοένα και περισσότερο επιτακτικός και αναπόφευκτος. Ερευνητές και εμπλεκόμενοι στον παραγωγικό και επιχειρησιακό κλάδο της βιομηχανικής στατιστικής, έχουν την ευκαιρία να κάνουν ορισμένες επιπρόσθετες, αλλά εξίσου σημαντικές συνεισφορές στη θεωρία και την εφαρμογή αυτού του ρόλου.

ΚΕΦΑΛΑΙΟ 3

Η προσαρμογή του κινδύνου στα διαγράμματα ελέγχου

3.1. Η έννοια του κινδύνου στην εξάπλωση και τις επιπτώσεις μίας ασθένειας

Σε κλινικές και ιατρικές πρακτικές, είναι αρκετά σημαντικό να παρακολουθούμε τα αποτελέσματα που σχετίζονται με τις εκάστοτε επεμβάσεις. Στη μεταμόσχευση οργάνων, για παράδειγμα, είναι σημαντικό να παρακολουθούμε και να αξιολογούμε τα αποτελέσματα των μοσχευμάτων που προετοιμάζονται στα συγκεκριμένα κέντρα, ώστε να σημάνει έγκαιρα συναγερμός, μόλις ανιχνευθεί μία σημαντική άνοδος στους ρυθμούς αποτυχίας των μοσχευμάτων. Μία τέτοια δραστηριότητα καταγραφής και παρακολούθησης των αποτελεσμάτων, θα μπορούσε, πέραν της αποστολής προειδοποιητικών σημάτων στις επιβλέπουσες ιατρικές επιτροπές και νοσοκομειακές εγκαταστάσεις, να αποδειχθεί χρήσιμη στο συντονισμό και στη βέλτιστη απόδοση της διαδικασίας γενικότερα. Συγκεκριμένες τεχνικές που χρησιμοποιούνται κατά κόρον στην Στατιστική, όπως ο μέσος ρυθμός θνησιμότητας, η *risk – adjusted* θνησιμότητα, καθώς και η πολυπαραμετρική μοντελοποίηση, θα μπορούσαν να χρησιμοποιηθούν για την ανίχνευση αλλαγών ως προς την απόδοση και σε εθνικό επίπεδο.

Οι *Grigg & Farewell* (2003), έδωσαν εξαιρετικές αναφορές για τη δημιουργία των *risk – adjusted* διαγραμμάτων ελέγχου. Για τη δημιουργία αυτών των χαρακτηριστικών διαγραμμάτων, πρέπει να γνωρίζουμε ότι η – εντός ελέγχου – πιθανότητα ενός θανάτου, για παράδειγμα, μπορεί να ποικίλει από άτομο σε άτομο, αναλόγως του μοντέλου που έχουμε υποθέσει. Κάποιες από τις μεθόδους δημιουργίας και παρουσίασης των *risk – adjusted control charts*, περιγράφονται στο συγκεκριμένο κεφάλαιο.

Ο Lovegrove και ο Poloniecki *et al.* (1998), πρότειναν, ανεξάρτητα ο ένας από τον άλλον, σωρευτικές γραφικές παραστάσεις, για τον υπολογισμό της αναμενόμενης θνησιμότητας, οι οποίες θα μπορούσαν να εφαρμοστούν με μεγάλη επιτυχία σε ιδιωτικές κλινικές και νοσοκομειακά ιδρύματα.

Μία θετική τάση θα μπορούσε να αποδειχθεί καλύτερη, συγκριτικά με τη μέση απόδοση. Τα διαγράμματα VLAD και CRAM, στα οποία θα αναφερθούμε και σε επόμενες ενότητες, δεν εμπεριέχουν ουσιαστικό έλεγχο, ούτε και όρια απόφασης. Ο Poloniecki (1998), για παράδειγμα, παρουσίασε όρια ελέγχου, τα οποία υπολογίστηκαν με συγκεκριμένο επίπεδο σημαντικότητας, ελέγχοντας αν ο ρυθμός θανάτων σε ένα πρόσφατο σύνολο ασθενών (όπου αναμένονταν 16 θάνατοι), ήταν διαφορετικός απ' ό,τι σε όλες τις προηγούμενες περιπτώσεις. Αυτό οδήγησε σε επαναλαμβανόμενους ελέγχους υποθέσεων, χρησιμοποιώντας την κατανομή χ^2 με ένα βαθμό ελευθερίας, αλλά οι επιστημονικοί αναλυτές στήριζαν τη θέση ότι αυτό δεν ισοδυναμεί με έναν τυπικό έλεγχο σημαντικότητας, καθώς οι υπολογισμοί έγιναν έπειτα από κάθε επέμβαση, χωρίς να ληφθεί υπ' όψιν το πλήθος των ελέγχων.

Με άλλα λόγια, η πιθανότητα Σφάλματος Τύπου I για τους συγκεκριμένους ελέγχους, δεν ερμηνεύεται άμεσα, όσον αφορά το μέσο μήκος ροής της διαδικασίας (*run length performance*). Παράδειγμα του διαγράμματος τύπου CRAM που πρότεινε ο Poloniecki, απεικονίζεται στο παρακάτω διάγραμμα (*Figure 3*).



FIGURE 3. Cumulative Risk Adjusted Mortality (CRAM) Chart with 99% Control Limits for Change in Mortality in Last 16 Expected Deaths. (From Poloniecki et al. (1998). Reproduced with permission from the BMJ Publishing Group.).

Το συγκεκριμένο διάγραμμα απεικονίζει καλύτερα από ό,τι αναμενόταν, μία συνολική εικόνα για την απόδοση του - υπό παρακολούθηση - νοσοκομείου, καθώς εδώ παρατηρείται μία αυξητική τάση. Το κατώτερο όριο ελέγχου ξεπεράστηκε στην επέμβαση υπ' αριθμόν 1651 και στην επέμβαση 2189, δείχνοντας έτσι φθορές στην απόδοση, σε σχέση με πριν.

Πλήθος διαφορετικών τύπων διαγραμμάτων ελέγχου, τροποποιήθηκαν, προκειμένου να προσαρμόσουν και τον παράγοντα του κινδύνου στο εσωτερικό τους. Ο *Steiner et. al (1999, 2000)* παρουσίασε risk-adjusted διαγράμματα ελέγχου Cusum, βασισμένος στις προσεγγίσεις του *Page (1954)*. Το διάγραμμα σωρευτικής άθροισης βασίζεται σε αναλογίες πιθανοφάνειας (likelihood ratios), οι οποίες επηρεάζονται από τις ποικίλες –εντός ελέγχου– πιθανότητες θνησιμότητας για τους ασθενείς. Ένα risk-adjusted διάγραμμα τύπου RSPRT, προτάθηκε από τον *Spiegelhalter et. al (2003)* και risk-adjusted sets method προτάθηκε από τους *Grigg & Farewell (2004)*.

Χαρακτηριστικό παράδειγμα ενός risk-adjusted διαγράμματος Cusum, εικονίζεται αμέσως παρακάτω (**Figure 4**). Είναι ένα διμερές διάγραμμα Cusum (*two-sided Cusum Chart*), όπου το άνω διάγραμμα Cusum είναι σχεδιασμένο με στόχο την ανίχνευση ελαττωματικής απόδοσης της διεργασίας, ενώ το κάτω διάγραμμα σχεδιάστηκε αντίστοιχα για να ανιχνεύει βελτιώσεις στην απόδοση.

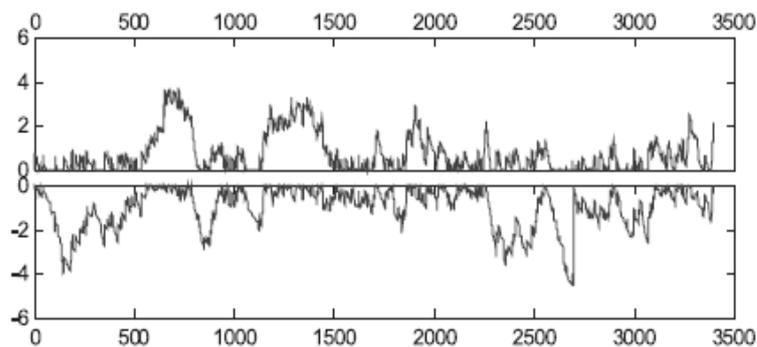


FIGURE 4. Example of a Two-Sided Risk-Adjusted CUSUM Chart (provided by Stefan H. Steiner).

Ο *Lie (1991)*, φαίνεται να είναι ο πρώτος εκ των στατιστικών αναλυτών, ο οποίος έκανε μία καθοριστική συνεισφορά στο πεδίο των Control Charts, ο οποίος προσαρμοσε τον παράγοντα του κινδύνου σε ένα διάγραμμα ελέγχου, χρησιμοποιώντας λογιστική παλινδρόμηση (logistic regression). Συγκεκριμένα, χρησιμοποιήθηκε μοντέλο λογιστικής παλινδρόμησης, για να προσαρμοστεί ο κίνδυνος για το σύνδρομο Down, θέτοντας ως βάση την ηλικία της μητέρας. Χρησιμοποιήθηκε επίσης, αλυσίδα Markov για τη μελέτη των ιδιοτήτων του ARL, ενός risk-adjusted διαγράμματος Cusum, όπως και πραγματοποιήθηκε μεταγενέστερα από τον *Steiner et. al (2000)*. Η πρόταση του *Lie* αγνοήθηκε από τον *Woodall (1997)*, ο οποίος στήριξε ότι η λογιστική παλινδρόμηση θα μπορούσε να χρησιμοποιηθεί για το σχεδιασμό διαγραμμάτων ελέγχου, αλλά δε γνώριζε καμία εφαρμογή για αυτήν.

Με risk-adjusted διαγράμματα ελέγχου, μπορεί πλέον κάποιος να χρησιμοποιήσει απλή ή πολλαπλή γραμμική παλινδρόμηση, προκειμένου να υπολογίσει την απόκλιση στην ποιότητα των αντικειμένων, καθώς αυτά βρίσκονται υπό παρακολούθηση. Η προσαρμογή του κινδύνου στα Control Charts διαδραματίζει σημαντικότερο ρόλο στην απόκλιση που παρουσιάζει η διεργασία ελέγχου στον τομέα της Υγείας, απ' ό,τι στον τομέα της βιομηχανίας και της παραγωγής. Η γνώση των ιδιοτήτων του ARL στα risk-adjusted διαγράμματα ελέγχου, είναι πολύ πιο σύνθετη και περίπλοκη, σε σχέση με τα non-risk-adjusted, διότι οι ιδιότητες των διαγραμμάτων εξαρτώνται επίσης και από τους παράγοντες κινδύνου των ασθενών. Επίσης, η επάρκεια και η ακρίβεια των *risk-adjusted* μεθόδων επηρεάζουν την απόδοση του τελικού διαγράμματος. Συνεπώς, η επάρκεια του risk-adjusted μοντέλου πρέπει να παρακολουθείται διαρκώς. Με εξαίρεση τον *Steiner et. al (2000)*, λίγη μελέτη έχει πραγματοποιηθεί για την επίδραση που ασκεί το εκτιμώμενο σφάλμα και η ανεπάρκεια του μοντέλου στην απόδοση των *risk-adjusted* διαγραμμάτων.

Η επιλογή του μοντέλου όπου θα προσαρμοστεί ο παράγοντας του κινδύνου (*risk adjustment*), είναι ένα σημαντικό στατιστικό ζήτημα. Όταν πρόκειται για καρδιοχειρουργικές επεμβάσεις, τα μοντέλα συνήθως επιλέγονται βάσει του δείκτη *Parsonnet* ή του δείκτη *Euroscore*. Αυτοί οι δείκτες βασίζονται σε χαρακτηριστικά του ασθενούς, όπως η ηλικία, το φύλο και σε παράγοντες που σχετίζονται με την υγεία * όπως διαβήτης ή νεφρική διαταραχή.

3.2. Μελέτη των risk-adjusted αποτελεσμάτων στην καρδιοχειρουργική επέμβαση

3.2.1. Επισκόπηση της ενότητας

Η προσαρμογή του κινδύνου στα αποτελέσματα καρδιοχειρουργικών επεμβάσεων, στην παιδιατρική, αποτελεί πρόκληση εξαιτίας της μεγάλης ποικιλίας των διαγνώσεων και των σχετικών διεργασιών. Είχαμε αναφερθεί στην προηγούμενη ενότητα στο γεγονός ότι τα διαγράμματα VLAD (variable life-adjusted display) παρέχουν μία αποτελεσματική γραφική απεικόνιση των risk-adjusted αποτελεσμάτων τέτοιου είδους διεργασιών και επεμβάσεων. Το ερώτημα που αναδύεται είναι εάν το μοντέλο κινδύνου παραμένει κατάλληλο, με την πάροδο του χρόνου.

Όπως θα αναλυθεί και στη συνέχεια, για να απαντηθεί επαρκώς το ερώτημα αυτό, χρησιμοποιήθηκε μία πρόσφατα ανεπτυγμένη διαγραμματική τεχνική, ώστε να αξιολογηθεί η απόδοση του υπάρχοντος *risk model* μεταξύ των ασθενών ενός συγκεκριμένου νοσοκομειακού κέντρου, για τη χρονική περίοδο από το έτος 2000 έως το 2003. Στη συνέχεια, συγκρίθηκε η κατανομή του εκτιμημένου ρίσκου μεταξύ των ασθενών αυτών, με την αντίστοιχη των ασθενών της περιόδου από το 2004 έως το 2006. Τελικά, όπως θα δούμε και στη συνέχεια, κατασκευάστηκε ένα διάγραμμα τύπου VLAD, για τα risk-adjusted αποτελέσματα της πιο πρόσφατης χρονικής περιόδου (2004-2006).

Τα αποτελέσματα της έρευνας έδειξαν ελάχιστη διαφορά στην κατανομή του προβλεπόμενου κινδύνου, μεταξύ των ασθενών της πρώτης περιόδου και των 903 ασθενών της περιόδου: Ιούνιος 2004 - Οκτώβριος του 2006. Τα αποτελέσματα της πιο πρόσφατης χρονικής περιόδου ήταν σαφώς καλύτερα από τα αναμενόμενα, βάσει του συγκεκριμένου *risk model*.

3.2.2. Εισαγωγή

Η διαστρωμάτωση του κινδύνου που διατρέχουν οι υποβληθέντες σε καρδιοχειρουργική επέμβαση, είναι σημαντικό κομμάτι της διαδικασίας ελέγχου, καθώς ελαχιστοποιεί την πιθανότητα αθέμιτης αξιολόγησης ενός χειρουργού, του οποίου το ποσοστό των θανάτων είναι υψηλό, λόγω του ότι χειρουργεί ασθενείς υψηλής ευπάθειας.

Διάφορες μέθοδοι έχουν αναπτυχθεί προς διευκόλυνση της ανάλυσης και ερμηνείας των risk-stratified (διαστρωμάτωση κινδύνου) δεδομένων, τόσο για αυτά που έχουν προκύψει από καρδιοχειρουργικές επεμβάσεις ενηλίκων, όσο και για τα αντίστοιχα παιδιών. Η προσαρμογή του κινδύνου στις καρδιακές επεμβάσεις παιδιών, όπως αναφέραμε και πριν, αποτελεί πρόκληση εξαιτίας του μεγάλου πλήθους των ασθενών, όσον αφορά τις διαγνώσεις, την ίδια την επέμβαση, την ηλικία του ασθενούς που καθορίζει την επέμβαση επιτακτική, και άλλους παράγοντες. Ένα διεθνώς αποδεκτό σχήμα ταξινόμησης διαδικασιών, γνωστό ως «*Risk Adjustment for Congenital Heart Surgery (RACHS-1)*», ομαδοποιεί 79 διαφορετικούς τύπους επεμβάσεων, σε έξι κατηγορίες, κατατάσσοντας τες σε αύξουσα σειρά ως προς το επίπεδο κινδύνου που εμπεριέχεται σε αυτές.

Τα διαγράμματα VLAD δημιουργήθηκαν με σκοπό να χρησιμοποιηθούν σε καρδιοχειρουργικές επεμβάσεις, ώστε να εκθέσουν σαφώς το ρυθμό θνησιμότητας κατά τη διάρκεια επεμβάσεων σε μία νοσοκομειακή μονάδα, παρέχοντας συνεπώς μια προσιτή επισκόπηση των σχετικών αποτελεσμάτων που παρατηρούνται.

Το διάγραμμα είναι μία γραφική παράσταση της μέσης σωρευτικής θνησιμότητας που αναμένεται στην ακολουθία των περιστατικών μείον την αντίστοιχη προβλεπόμενη σωρευτική θνησιμότητα. Ειδικότερα, το διάγραμμα VLAD δίνει μία σαφή εικόνα του πόσο καλύτερη (ή χειρότερη) είναι η σύγκριση των αποτελεσμάτων ως προς τα αντίστοιχα που αναμένονταν βάσει του *risk model*. Το «ίχνος» που διαγράφεται στο διάγραμμα VLAD, παρουσιάζει άνοδο όταν ο ασθενής επιβιώνει, ενώ κάθοδο όταν στην αντίθετη περίπτωση. Ο βαθμός ανόδου ή καθόδου καθορίζεται από τον προβλεπόμενο κίνδυνο που χαρακτηρίζει την εκάστοτε περίπτωση. Για παράδειγμα, αν ο ασθενής επιβιώσει από μόλυνση στη μεσοκοιλιακή χώρα, με προβλεπόμενο κίνδυνο θανάτου 2%, το «ίχνος» θα διαγράψει άνοδο 0,02 μονάδων. Αν ο ασθενής όμως δεν τα καταφέρει, τότε το «ίχνος» θα διαγράψει κάθοδο 0,98 μονάδων.

Διαπιστώνουμε λοιπόν με αφορμή την προαναφερθείσα περίπτωση ότι τα διαγράμματα VLAD παρέχουν μία προσιτή γραφική αναπαράσταση της χειρουργικής δραστηριότητας, αναφορικά πάντα με τα risk-adjusted αποτελέσματα της εκάστοτε επέμβασης.

Χρησιμοποιώντας μοντέλο προβλεπόμενου κινδύνου, βάσει του RACHS-1, δείξαμε προηγουμένως ότι τα διαγράμματα VLAD παρέχουν μία αποτελεσματική γραφική αναπαράσταση της απόδοσης, όταν εφαρμόζονται σε καρδιοχειρουργικές επεμβάσεις.

Πριν όμως αυτού του είδους οι μέθοδοι εφαρμοστούν σε περιπτώσεις ρουτίνας, τόσο οι διάφοροι τύποι των κινδύνων όσο και οι μέθοδοι καταγραφής και παρακολούθησης, πρέπει

να ελεγχθούν και να δοκιμαστούν λεπτομερώς, προς αποφυγή λανθασμένων συμπερασμάτων. Ο απόλυτος στόχος είναι να καταστήσουμε έγκυρο το μοντέλο κινδύνου, χρησιμοποιώντας σύνολα δεδομένων από άλλα νοσοκομεία και ιδρύματα και επίσης να καταγράψουμε οποιαδήποτε ένδειξη αλλαγής στην καταλληλότητα του με την πάροδο του χρόνου.

Σε αυτήν την ενότητα, θα επικεντρωθούμε στην αξιολόγηση πρόσφατων αποτελεσμάτων από καρδιοχειρουργικές επεμβάσεις που πραγματοποιήθηκαν σε μία νοσοκομειακή μονάδα, χρησιμοποιώντας το προαναφερθέν μοντέλο κινδύνου βάσει του RACHS-1 (το οποίο δημιουργήθηκε χρησιμοποιώντας δεδομένα από το ίδιο νοσοκομείο). Θα ερευνήσουμε επίσης, μέχρι ποιο σημείο μπορούν οι παρατηρούμενες αλλαγές στα risk-adjusted αποτελέσματα, να αποδίδονται σε χαρακτηριστικά του μοντέλου.

3.2.2. Ασθενείς και Μέθοδοι: Μελέτη, Ανάλυση και Συμπεράσματα

Δεδομένα σχετικά με επεμβάσεις ανοικτής καρδιάς που πραγματοποιήθηκαν από τον Ιούνιο του 2004 έως τον Οκτώβριο του 2006 στο *Ormond Street Hospital*, του Λονδίνου, τέθηκαν σε αντιπαράβολή. Οι ασθενείς ταξινομήθηκαν σύμφωνα με το σχήμα ταξινόμησης RACHS-1. Σημειώθηκαν ωστόσο η ηλικία των ασθενών και ο αναμενόμενος ρυθμός επιβίωσης εντός της νοσοκομειακής μονάδας. Αναφέραμε προηγουμένως τη δημιουργία του μοντέλου του κινδύνου θανάτων κατά τη διάρκεια καρδιοχειρουργικής επέμβασης, χρησιμοποιώντας δεδομένα σχετικά με 1.083 εγχειρήσεις ανοικτής καρδιάς, οι οποίες πραγματοποιήθηκαν μεταξύ Απριλίου του 2000 και Μαρτίου του 2003, στην ίδια νοσοκομειακή μονάδα.

Από την στιγμή που διαμορφώθηκε το μοντέλο κινδύνου για τις εγχειρήσεις καρδιάς, το γραφικό εργαλείο γνωστό ως «*Mean Adjusted Deaths Compared Against Predicted (MADCAP)*», κατέστη πλέον διαθέσιμο· το εργαλείο αυτό μας προσφέρει οπτική αναγνώριση του κινδύνου ως μέρους του μοντέλου κινδύνου. Αρχικά στη συγκεκριμένη μελέτη, χρησιμοποιήθηκε η συγκεκριμένη τεχνική για να αξιολογηθεί η απόδοση του μοντέλου κινδύνου, κατά μήκος του εύρους του προβλεπόμενου κινδύνου, μεταξύ του συνόλου των 1.083 ασθενών.

Χρησιμοποιήθηκε το διάγραμμα MADCAP το οποίο τελικά συνέβαλε στο να αναγνωρίσουμε την κατανομή του προβλεπόμενου κινδύνου, σύμφωνα με το αρχικό

σύνολο δεδομένων. Το μοντέλο κινδύνου χρησιμοποιήθηκε στη συνέχεια για τον υπολογισμό του προβλεπόμενου κινδύνου εμφάνισης θανάτων εντός του νοσοκομείου, για κάθε ασθενή, από την ομάδα των ασθενών που μελετήθηκαν την πιο πρόσφατη χρονική περίοδο. Πριν διαμορφωθεί το διάγραμμα VLAD των *risk-adjusted* αποτελεσμάτων για την περίοδο Ιούνιος 2004 – Οκτώβριος 2006, υπολογίστηκε το ποσοστό των ασθενών που ήταν σε κάθε εύρος του προβλεπόμενου κινδύνου (που αναφέρθηκε μόλις πριν), σύμφωνα με την απόδοση του μοντέλου κινδύνου στο αρχικό σύνολο δεδομένων. Στη συνέχεια, κατασκευάστηκε το διάγραμμα, ώστε να μας παρέχει μία οπτική σύγκριση του κινδύνου μεταξύ των πιο πρόσφατων περιστατικών και αντίστοιχα του κινδύνου μεταξύ του αρχικού συνόλου δεδομένων.

➤ Αποτελέσματα

Από τον Ιούνιο του 2004, έως τον Οκτώβριο του 2006, πραγματοποιήθηκαν 903 το πλήθος επεμβάσεις ανοικτής καρδιάς. Αν και δεν αποτελούσε εύρημα της τρέχουσας μελέτης, ωστόσο το πρωταρχικό σύνολο δεδομένων χρησιμοποιήθηκε στη διαμόρφωση του risk model, το οποίο αποτελείται από 1.083 επεμβάσεις που έλαβαν χώρα στη χρονική περίοδο από τον Απρίλιο του 2000 έως το Μάρτιο του 2003.

Το διπλανό διάγραμμα (Figure 1) είναι ένα διάγραμμα τύπου MADCAP για το μοντέλο κινδύνου μεταξύ των 1.083 περιστατικών (το αρχικό μας σύνολο δεδομένων). Εκφράζει την προβλεπόμενη σωρευτική θνησιμότητα και την παρατηρούμενη σωρευτική θνησιμότητα μεταξύ των ασθενών οι οποίοι είχαν αυξημένο επίπεδο κινδύνου κατά τη διάρκεια της επέμβασης.

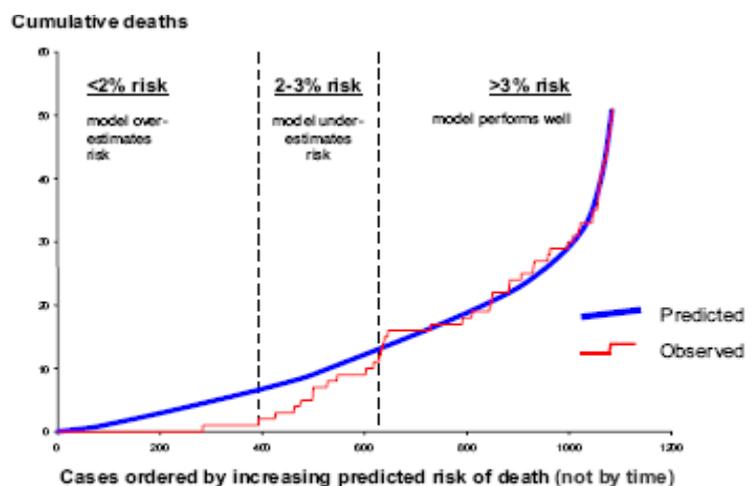


Fig 1. A chart of mean adjusted deaths compared against predicted (MADCAP) shows the performance of the risk model across the whole spectrum of risk among the 1083 procedures undertaken between April 2000 and March 2003 used in the development of the model.

Σε επίπεδο προβλεπόμενου κινδύνου μικρότερου του 2%, το «ίχνος» που εκφράζει την προβλεπόμενη σωρευτική θνησιμότητα παρουσιάζει μία πιο απότομη άνοδο, σε σχέση με την αντίστοιχη της παρατηρούμενης σωρευτικής θνησιμότητας, δείχνοντας έτσι ότι το μοντέλο υπερεκτιμάει τον κίνδυνο για αυτούς τους ασθενείς. Αντιστοίχως, για ασθενείς με επίπεδο προβλεπόμενου κινδύνου από 2% έως 3%, η αναμενόμενη σωρευτική θνησιμότητα παρουσιάζει μία λιγότερο απότομη άνοδο, σε σχέση με την παρατηρούμενη σωρευτική θνησιμότητα, δείχνοντας ότι το μοντέλο υποτιμάει τον κίνδυνο θανάτου για τους συγκεκριμένους ασθενείς. Τελικά, για επίπεδα κινδύνου που υπερβαίνουν την τάξη του 3%, το μοντέλο παρουσιάζει ικανοποιητική απόδοση, γεγονός που φαίνεται από το ότι οι δύο καμπύλες παρουσιάζουν άνοδο με τον ίδιο ρυθμό. Δεδομένου ότι οι 1.083 επεμβάσεις χρησιμοποιήθηκαν προς διαμόρφωση του μοντέλου κινδύνου, δεν είναι τυχαίο το γεγονός ότι οι δύο καμπύλες ολοκληρώνονται στο ίδιο ακριβώς κάθετο επίπεδο· τονίζουμε συνεπώς ότι αυτό είναι αναπόφευκτη συνέπεια της διαδικασίας με την οποία διαμορφώθηκε το μοντέλο μας.

➤ **Μεταβολές στο προφίλ του κινδύνου, στη Συνδυαστική Περίπτωση (Case Mix)**

Το παρακάτω διάγραμμα (**Figure 2**) απεικονίζει το ποσοστό θανάτων των ασθενών σε κάθε μία από τις 3 κατηγορίες κινδύνου, όπως ορίστηκαν στο διάγραμμα **Figure 1**: λιγότερο από 3%, μεταξύ 2% και 3% και άνω του 3%, για το αρχικό σύνολο δεδομένων και για τα πιο πρόσφατα δεδομένα.

Ένα ελαφρώς υψηλότερο ποσοστό ασθενών παρουσίασε προβλεπόμενο επίπεδο κινδύνου άνω του 3% στο πιο πρόσφατο σύνολο δεδομένων, σε σχέση με το πρωταρχικό, με ένα αντίστοιχο μικρότερο ποσοστό να παρουσιάζει κινδύνους και στο εύρος κάτω του 2% και στο εύρος μεταξύ του 2% και 3%.

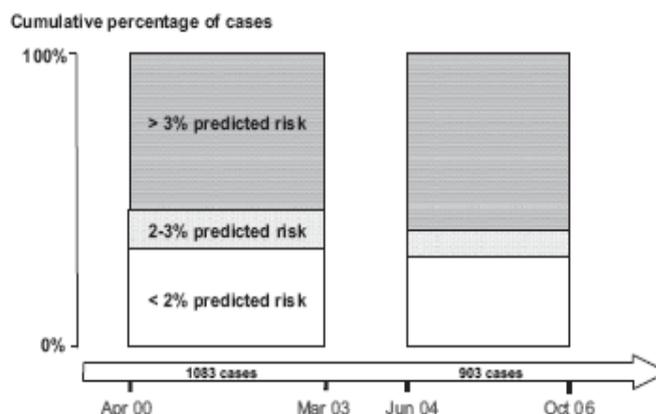
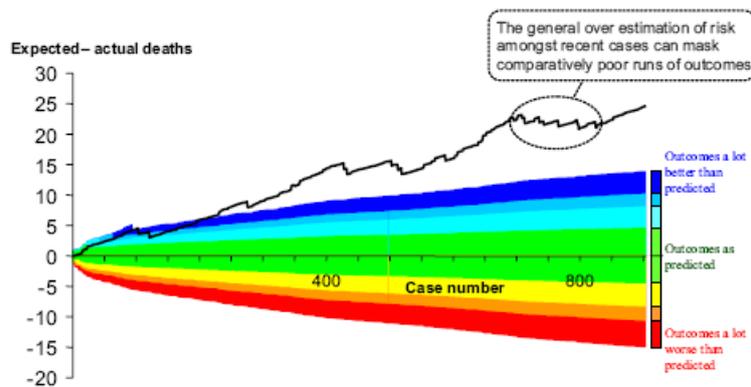


Fig 2. A chart summarizing the distribution of predicted risk in both the original data set was used to develop the risk model and in the more recent data set.

➤ Πρόσφατα Αποτελέσματα (Recent Outcomes)

Υπήρξαν 19 θάνατοι στο νοσοκομείο, μεταξύ των 903 διαδοχικών επεμβάσεων που πραγματοποιήθηκαν από τον Ιούνιο του 2004, έως τον Οκτώβριο του 2006. Ο αριθμός αυτός αντιστοιχεί σε ρυθμό θνησιμότητας της τάξεως του 2.1% (95% διάστημα εμπιστοσύνης, 1.3% έως 3.3%). Το διάγραμμα **Figure 3** είναι ένα διάγραμμα τύπου *VLAD chart*, το οποίο απεικονίζει τα *risk-adjusted* αποτελέσματα για τους συγκεκριμένους 903 ασθενείς.

Fig 3. A variable life-adjusted display chart shows the cumulative difference between expected (blue line) and observed (red line) mortality across the series of 903 operations performed between June 2004 and October 2006. The chart indicates that the outcomes achieved are appreciably better than expected on the basis of the risk model.



Όπως μπορούμε να παρατηρήσουμε από την σταθερή άνοδο του «ίχνους» στο διάγραμμα VLAD, τα πρόσφατα αποτελέσματα απεδείχθησαν σαφώς καλύτερα από τα αντίστοιχα που αναμενόταν βάσει του μοντέλου κινδύνου. Η χρωματικά τονισμένη περιοχή του διαγράμματος, αναφέρεται στη χρονική περίοδο όχι και τόσο προτιμώμενων αποτελεσμάτων.

➤ Σχόλια - Συμπεράσματα

Το μοντέλο κινδύνου παρουσίασε ικανοποιητική απόδοση, σε όλο το εύρος του αναμενόμενου κινδύνου, για τους ασθενείς που χρησιμοποιήθηκαν στη διαμόρφωση του. Το εύρημα αυτό έχει επιπτώσεις στην ερμηνεία των *risk-adjusted* αποτελεσμάτων. Παρόλα αυτά, δεδομένου ότι υπάρχει μικρή διαφορά στο προφίλ του κινδύνου της πρόσφατης συνδυαστικής περίπτωσης σε σχέση με τον κίνδυνο στο πρωταρχικό σύνολο δεδομένων, το γεγονός ότι τα πρόσφατα αποτελέσματα είναι καλύτερα από τα αναμενόμενα βάσει του μοντέλου κινδύνου δεν είναι καθόλου τυχαίο.

Το πόρισμα αυτό ενισχύει το επιχείρημα ότι η παρατηρούμενη βελτίωση στα risk-adjusted αποτελέσματα, αντανακλά μία γνήσια βελτίωση των κλινικών αποτελεσμάτων που διεξήχθησαν στη συγκεκριμένη νοσοκομειακή μονάδα.

Αν πράγματι έχει επιτευχθεί τόσο σημαντική βελτίωση, το γεγονός αυτό αναδύει στην επιφάνεια το ερώτημα αν το μοντέλο κινδύνου που διαμορφώθηκε βάσει των επεμβάσεων που έλαβαν χώρα τη χρονική περίοδο μεταξύ του 2000 και του 2003, εξακολουθεί να είναι κατάλληλο. Στόχος λοιπόν θα πρέπει να είναι η αδιάκοπη και συνεχής προσπάθεια για συνεχή βελτίωση της απόδοσης, καθώς η μελέτη και καταγραφή των *risk-adjusted* αποτελεσμάτων κατέχουν ρόλο-«κλειδί» σε αυτήν την προσπάθεια. Είναι σαφές το γεγονός ότι αν χρησιμοποιηθεί στην καταγραφή ένα μοντέλο που υπερεκτιμάει τον παράγοντα του κινδύνου, μεταξύ πρόσφατων ασθενών, σίγουρα αυτό θα φέρει εφησυχασμό στην κοινότητα τόσο της Ιατρικής, όσο και στους ίδιους τους ασθενείς.

Η συγκεκριμένη ενότητα επιβεβαιώνει ότι τα γραφικά εργαλεία παρακολούθησης, τα οποία χρησιμοποιούνται ευρέως στο πεδίο των καρδιοχειρουργικών επεμβάσεων σε ενήλικες, πιθανόν να εφαρμοστούν στην παρακολούθηση των αποτελεσμάτων των χειρουργειών καρδιάς του κλάδου της Παιδιατρικής. Είναι ενθαρρυντικό να σημειώσουμε ότι τα πρόσφατα αποτελέσματα που εξήχθησαν στο *Great Ormond Street Hospital*, είναι σαφώς καλύτερα από τα αντίστοιχα αναμενόμενα βάσει του μοντέλου παλινδρόμησης που διαμορφώθηκε το 2004 από το ίδιο νοσοκομειακό κέντρο. Προτού οι συγκεκριμένες γραφικές μέθοδοι παρακολούθησης της υγείας εφαρμοστούν στο πεδίο της παιδιατρικής καρδιοχειρουργικής, θα πρέπει να υποστούν περαιτέρω ελέγχους και δοκιμές. Καθώς τα μοντέλα προβλεπόμενου κινδύνου πρόκειται να χρησιμοποιηθούν για την παρακολούθηση της απόδοσης, οι μέθοδοι αυτές πρέπει να επανεξεταστούν προσεκτικά και να αναβαθμονομηθούν, προκειμένου να παρέχουν σωστή αξιολόγηση για την εκάστοτε διεργασία, στην πάροδο του χρόνου. Ο ρόλος των μοντέλων κινδύνου στην πρόβλεψη του κινδύνου και στη λήψη αποφάσεων σε ατομικό επίπεδο ασθενούς, είναι σημαντικά περιορισμένος.

Αρκετοί ερευνητές έχουν συνεισφέρει σημαντικά στην «αντιμετώπιση» του φλέγοντος ζητήματος της προσαρμογής του κινδύνου και της παρακολούθησης της απόδοσης στο χώρο της παιδιατρικής καρδιοχειρουργικής. Το σχήμα ταξινόμησης RACHS-1 έχει πλέον επικυρωθεί στη Βόρειο Αμερική και στην Ευρώπη, διευκολύνοντας την αξιολόγηση σημαντικών θεμάτων, όπως επιπλοκές σε επεμβάσεις, θνησιμότητα, φυλετική ανισότητα και πολλά ακόμη.

Δεδομένης μίας έκθεσης της *Bristol Infirmary Inquiry*, του Ηνωμένου Βασιλείου, ήρθε στο φως ένας εκπληκτικά υψηλός ρυθμός θνησιμότητας, όπως παρατηρήθηκε σε ένα κέντρο παιδιατρικής καρδιοχειρουργικής και πλέον έχει εκδηλωθεί τεράστιο εθνικό ενδιαφέρον για την παρακολούθηση και την αξιολόγηση χειρουργικών επεμβάσεων. Η σχετική βάση δεδομένων «*UK Central Cardiac Audit Database (CCAD)*» απαίτησε την υποχρεωτική απόδοση των λεπτομερειών που αφορούσαν όλες τις εγχειρήσεις καρδιάς που πραγματοποιήθηκαν στο Ηνωμένο Βασίλειο ήδη από το 2000 και δημοσίευσε το 2007, για πρώτη φορά στο Διαδίκτυο, συγκεκριμένα αποτελέσματα για το εν λόγω νοσοκομείο.

Αυτήν την στιγμή, δεδομένα που αφορούν την –ενός μήνα και ενός έτους- θνησιμότητα και που εξήχθησαν από 36 διαφορετικές χειρουργικές επεμβάσεις, αναφέρονται από 17 μεμονωμένα κέντρα παιδιατρικής καρδιοχειρουργικής του Λονδίνου. Η ποικιλία όλων αυτών των δεδομένων αποτελεί ένα αναπόσπαστο χαρακτηριστικό αυτής της πολύπλοκης ειδικότητας και καθιστά δύσκολη την ανασκόπηση της – σχετικής με το εκάστοτε χειρουργικό κέντρο-πληροφορίας. Επιπροσθέτως, η χρήση μίας μεθόδου καταγραφής η οποία απεικονίζει σε γράφημα τη σωρευμένη διαφορά μεταξύ του παρατηρούμενου πλήθους δυσμενών περιστατικών και του αντίστοιχου αναμενόμενου βάσει ενός *risk-adjusted* μοντέλου, ίσως αποδεικνύεται καλύτερη για την έγκαιρη ανίχνευση μίας ανησυχητικής τάσης (*trend*).

3.3. Καταγραφή και παρακολούθηση της χειρουργικής δραστηριότητας με χρήση διαγραμμάτων risk-adjusted σωρευτικών αθροισμάτων

Η ανάγκη για παρακολούθηση των αποτελεσμάτων που διεξάγονται από τις χειρουργικές επεμβάσεις, ήρθε πλέον στην πρώτη γραμμή μέσα από συγκεκριμένα περιστατικά που είδαν το φως της δημοσιότητας, στα οποία οι ανεξήγητα υψηλοί ρυθμοί εμφάνισης επιπλοκών στις επεμβάσεις δεν ανιχνεύθηκαν για ένα αρκετά μεγάλο χρονικό διάστημα. Σε αυτές τις περιπτώσεις, η ταχεία φθορά της χειρουργικής απόδοσης είναι σημαντική, καθώς θα οδηγήσει σε άμεση έρευνα της αιτίας και των αλλαγών στη διαδικασία.

Η μεθοδολογία της σωρευτικής άθροισης CUSUM, αναπτύχθηκε αρχικά από τον Page (1954) για προβλήματα βιομηχανικής φύσεως, όπου η παρακολούθηση της παραγωγικής διαδικασίας είναι μεγίστης σημασίας. Για τέτοιου είδους προβλήματα, τα διαγράμματα CUSUM είναι ιδανικά για την ανίχνευση μικρών αλλαγών στη διαδικασία (Montgomery, 1991). Όσον αφορά τον τομέα της Ιατρικής, όπως είχαμε αναφέρει και στα πρώτα κεφάλαια, τα διαγράμματα CUSUM έχουν προταθεί για παρακολούθηση διεργασιών στην Κλινική Χημεία (Nix *et. al.*, 1986) και για την παρακολούθηση σπάνιων γενετήσιων δυσμορφιών. Η εφαρμογή της μεθοδολογίας CUSUM στην παρακολούθηση της χειρουργικής διαδικασίας και απόδοσης, προτάθηκε αρχικά από τον Williams *et al.* (1992). Η πρώτη εφαρμογή διαγράμματος CUSUM στην παρακολούθηση της χειρουργικής διαδικασίας, είναι τεκμηριωμένη από τον De Leval *et al.* (1994) και τον Steiner *et al.* (1999), οι οποίοι αναλογίστηκαν το πρόβλημα παρακολούθησης των αποτελεσμάτων που εξάγονται από χειρουργικές επεμβάσεις σε παιδιά.

Στην εφαρμογή αυτή, παράγοντες όπως συγκεκριμένα χαρακτηριστικά ασθενών ή χαρακτηριστικά της διαδικασίας και η φυσική κόπωση της χειρουργικής ομάδας, δεν βρέθηκε να ασκούν σημαντική επίδραση στο ρυθμό αποτυχίας της εκάστοτε επέμβασης. Παρόλα αυτά σε αρκετά ιατρικά συμβάντα, υπάρχει αισθητή μεταβολή στα χαρακτηριστικά των –υπό μελέτη– ατόμων (ανομοιογένεια). Ασθενείς με διαφορετική κλινική κατάσταση και φυσιολογία παρουσιάζουν διαφορετικούς πρότερους κινδύνους.

Συνεπώς, ακόμη και για τους χειρουργούς (ή τις χειρουργικές ομάδες) με έναν αποδεκτό ρυθμό εμφάνισης επιπλοκών, η πιθανότητα επιτυχίας μπορεί να ποικίλει αισθητά, μεταξύ των ασθενών. Στις περισσότερες ιατρικές εφαρμογές, η ανομοιογένεια των ασθενών θα πρέπει να ληφθεί σοβαρά υπ' όψιν, για οποιαδήποτε διαδικασία παρακολούθησης. Πλήθος μεθόδων που χρησιμοποιούνται για την παρακολούθηση της χειρουργικής διαδικασίας και λαμβάνουν υπ' όψιν διαφορετικά επίπεδα κινδύνων, περιγράφηκαν πρόσφατα.

Οι Lovegrove *et al.* (1997, 1999) και Poloniecki *et al.* (1998) προτείνουν συστήματα παρακολούθησης και καταγραφής, σχεδιάζοντας το διάγραμμα της διαφοράς μεταξύ των σωρευτικών παρατηρούμενων και των αναμενόμενων θανάτων. Και στις δύο προσεγγίσεις, ο αναμενόμενος αριθμός θανάτων υπολογίζεται χρησιμοποιώντας το δείκτη Parsonnet (Parsonnet *et al.*, 1989) για κάθε ασθενή. Το σύστημα βαθμολόγησης Parsonnet χρησιμοποιείται ευρέως στην καρδιοχειρουργική και μπορεί να χρησιμοποιηθεί για την προσαρμογή κινδύνου ακόμη και σε συνδυασμό διαφορετικών περιπτώσεων. Παρόλα αυτά είναι δύσκολο να ερμηνεύσουμε τα διαγράμματα που προτάθηκαν από τον Lovegrove *et al.* (1997, 1999) καθώς δεν αποσαφηνίζουν το μέγεθος της απόκλισης που αναμένεται σε περίπτωση καλής χειρουργικής απόδοσης, και ως εκ τούτου τι μέγεθος πρέπει να έχει η εν λόγω απόκλιση ώστε να κριθεί ανησυχητική.

Ο Poloniecki *et al.* (1998) προτείνει έναν διαισθητικό τρόπο καθορισμού των ορίων ελέγχου. Απαραιτήτως το σύστημα εμπεριέχει, έπειτα από κάθε παρατήρηση, έλεγχο της υπόθεσης ότι ο ρυθμός αποτυχίας στην τελευταία σειρά των ασθενών (όπου θα περιμέναμε 16 θανάτους) είναι διαφορετικός από τον αντίστοιχο σε όλες τις προηγούμενες επεμβάσεις. Ωστόσο δεν έχει γίνει καμία επίσημη προσαρμογή για πολλαπλό έλεγχο, αν και προτείνεται ο έλεγχος σε επίπεδο σημαντικότητας 0,01. Οι συγγραφείς αναγνωρίζουν ότι αυτό δεν ισοδυναμεί με έναν τυπικό έλεγχο σημαντικότητας και συνεπώς η εξέταση του ρυθμού εσφαλμένων συναγερωμών δεν έχει απολύτως κανένα νόημα. Αφού το σύστημα ελέγχου καθιστά συνεχώς ενημερωμένη την εκτίμηση της επιθυμητής χειρουργικής απόδοσης (χρησιμοποιώντας μεγαλύτερη σειρά ιστορικών δεδομένων), τα όρια ελέγχου συνεπώς αλλάζουν στην πάροδο του χρόνου. Η ενημέρωση είναι σημαντικό πλεονέκτημα, αν έχουμε ξεκινήσει με μία «φτωχή» εκτίμηση της τρέχουσας διαδικασίας, αλλά επίσης καθιστά πιθανή τη μη ανίχνευση μικρών βαθμιαίων αλλαγών στη διαδικασία. Όμως η ενημέρωση αυτή καθιστά πολύ δύσκολο τον καθορισμό της θεωρητικής απόδοσης του διαγράμματος, σε σχέση με τα κριτήρια μήκους ροής (*run length criteria*).

Για να αντιμετωπιστεί το ζήτημα της ερμηνείας των αποτελεσμάτων, προτάθηκε η χρήση ενός νέου *risk-adjusted* διαγράμματος CUSUM, προκειμένου να χρησιμοποιηθεί στην παρακολούθηση των αποτελεσμάτων κάθε χειρουργικής επέμβασης. Η προσαρμογή του κινδύνου γίνεται με ένα δείκτη πιθανοφάνειας. Η *risk-adjusted* διεργασία CUSUM είναι εμπλουτισμένη με δεδομένα ενός καρδιοχειρουργικού κέντρου στο Ηνωμένο Βασίλειο. Το σύνολο δεδομένων αφορά 6.994 επεμβάσεις, ενός χειρουργικού κέντρου, που συνέβησαν τη χρονική περίοδο 1992 - 1998. Τα δεδομένα περιέχουν πληροφορίες για κάθε ασθενή, όπως ηλικία, στοιχεία του χειρουργού, είδος επέμβασης, καθώς και προεγχειρητικούς παράγοντες που περιλαμβάνονται στο δείκτη Parsonnet. Οι παράγοντες αυτοί σχετίζονται με την ηλικία, το φύλο, υπέρταση, επίπεδα διαβήτη, νεφρική λειτουργία και μάζα της αριστερής κοιλιακής χώρας.

Για να περιγράψουμε τη συγκεκριμένη μέθοδο, εστιάζουμε στη μετεγχειρητική θνησιμότητα που καταγράφεται, ένα μήνα μετά την επέμβαση. Στην παράγραφο **3.3.1.** θα προσδιορίσουμε τη δομή ενός τυπικού διαγράμματος CUSUM (*unadjusted*) και ενός *risk-adjusted* διαγράμματος CUSUM. Προκειμένου να επεξηγηθούν κατά το βέλτιστο, τα πλεονεκτήματα της χρήσης ενός *risk-adjusted* CUSUM, εφαρμόζουμε τη συγκεκριμένη μέθοδο στο παράδειγμα της παραγράφου **3.3.2.** Στην τελευταία ενότητα **3.3.3.**, θα προβούμε σε μία λεπτομερέστερη συζήτηση για τις ιδιότητες της νέας αυτής διαγραμματικής μεθόδου, αναφορικά πάντα με το *risk-adjusted* CUSUM.

3.3.1. Standard CUSUM / Risk – adjusted CUSUM:

Περιγραφή και σύγκριση των διαγραμματικών μεθόδων

➤ Standard CUSUM

Η διαγραμματική μέθοδος CUSUM είναι ένα εδραιωμένο σύστημα διαδοχικού ελέγχου, η οποία σχεδιάστηκε για την ανίχνευση αλλαγών σε μία συγκεκριμένη διεργασία, ως την ορίσουμε ως θ . Η αρχική διατύπωση του CUSUM, όπως έχουμε αναφέρει και σε προηγούμενα κεφάλαια, οφείλεται στον **Page**. Ένας εναλλακτικός τρόπος εφαρμογής της μεθόδου CUSUM, είναι η μέθοδος της V-μάσκας (*V-mask Cusum Charts*) που προτάθηκε από τον Barnard, το 1959. Η μέθοδος αυτή χρησιμοποιεί μία V σχήματος μάσκα πάνω σε ένα ενιαίο αμφίπλευρο Cusum, η οποία υποδεικνύει με τις πλευρές της, τα εντός και εκτός ελέγχου σημεία του διαγράμματος. Δυστυχώς, η V αποδεικνύεται άτεχνη σε πρακτικές εφαρμογές. Μία πιο εύκολη μορφή Cusum, ως προς τη χρήση της, είναι η μέθοδος Tabular Cusum, η οποία μπορεί να ανιχνεύει άνοδο ή πτώση της διεργασίας θ . Συνδυάζοντας δύο διαγράμματα τύπου Tabular Cusum, καθίσταται πλέον εφικτή η ανίχνευση οποιασδήποτε αλλαγής στη διαδικασία. Ένα σύνθητες Tabular Cusum, έχει την ακόλουθη μορφή:

$$X_t = \max(0, X_{t-1} + W_t), t = 1, 2, 3, \dots \quad (1)$$

όπου $X_0 = 0$ και W_t είναι το βάρος του δείγματος, για την υπο-ομάδα t . Υπο-ομάδες, είναι μία συλλογή μονάδων που πάρθηκαν από την παραγωγική διαδικασία, περίπου την ίδια ώρα. Κάνοντας μία συνετή επιλογή για το W_t , το Cusum μπορεί να σχεδιαστεί έτσι ώστε να ανιχνεύει ενδεχόμενη πτώση ή άνοδο της διαδικασίας θ .

Το Cusum της σχέση (1), ελέγχει διαδοχικά την υπόθεση $H_0: \theta = \theta_0$ versus $H_1: \theta = \theta_A$.

Η τιμή του θ_0 συνήθως ορίζεται βάσει της τρέχουσας απόδοσης της διαδικασίας, ενώ το θ_A εκφράζει μία εναλλακτική τιμή που αντιστοιχεί σε κατώτερες επιδόσεις. Η διαδικασία αναμένεται να βρίσκεται στην κατάσταση H_0 , για όσο ισχύει ότι $X_t < h$ και θεωρείται ότι έχει μετατοπιστεί στην κατάσταση H_A , αν $X_t \geq h$, για οποιαδήποτε χρονική στιγμή t .

Η σταθερά h είναι το όριο ελέγχου του Cusum. Σύμφωνα με την ορολογία που χρησιμοποιείται στον έλεγχο ποιότητας, το Cusum που ξεπερνάει το όριο ελέγχου, λέμε ότι παρουσιάζει *signal* (σήμα). Ο όρος *signal*, σημαίνει ότι το διάγραμμα έχει συγκεντρώσει αρκετά στοιχεία, ώστε να συμπεράνει ότι η παράμετρος της διαδικασίας έχει αλλάξει.

Ας σημειώσουμε σε αυτό το σημείο ότι αν και οι τιμές W_t μπορεί να είναι και αρνητικές, ωστόσο το Tabular Cusum βάσει των X_t , περιορίζεται μόνο σε μη αρνητικές τιμές. Τα διαγράμματα Cusum έχουν σχεδιαστεί με τέτοιο τρόπο, ώστε να παρακολουθούν τα αποτελέσματα διαδοχικά, μέχρι την ανίχνευση φθοράς στην απόδοση της διαδικασίας. Συνεπώς, το Cusum θα σημάνει «συναγερμό» (*signal*), ωστόσο υπάρχει πιθανότητα να είναι και εσφαλμένος. Καλές επιλογές επομένως για το όριο ελέγχου h , βασίζονται στο αναμενόμενο (ή μέσο) μήκος ροής (*average run length*) του Cusum, υπό τις υποθέσεις H_0 και H_A . Καθώς, η διαδικασία είναι στην κατάσταση H_0 , το μήκος ροής πρέπει να είναι μεγάλο, καθώς -σε αυτό το πλαίσιο- τα *signals* εκφράζουν εσφαλμένους συναγερμούς. Από την άλλη μεριά, αν η διαδικασία μετατοπιστεί στο θ_A , σε αυτήν την περίπτωση θα επιθυμούσαμε μικρό μήκος ροής.

Το ARL, υπό την υπόθεση H_0 , μπορεί να θεωρηθεί ανάλογο προς το ρυθμό σφάλματος Τύπου I, ενός κλασικού στατιστικού ελέγχου. Ωστόσο, δεν υπάρχει ένα γενικά αποδεκτό επίπεδο (όπως οι έλεγχοι σε επίπεδο 5%), καθώς ο ορισμός και η ερμηνεία του αποδεκτού ποικίλει από εφαρμογή σε εφαρμογή. Ομοίως, το ARL του Cusum, όταν η τιμή του θ έχει αλλάξει σημαντικά, είναι ανάλογο προς την ισχύ ενός κλασικού στατιστικού ελέγχου. Ο καθορισμός του ARL ενός Cusum, είναι υπολογιστικά εντατικός, καθώς βασίζεται σε όλα τα πιθανά αποτελέσματα ενός συνόλου χειρουργικών επεμβάσεων. Παρόλα αυτά το ARL μπορεί να προσεγγιστεί αρκετά.

Ο σχεδιασμός του Cusum προκύπτει από την επιλογή του δειγματικού βάρους W_t και του ορίου ελέγχου h . Ο Moustakides (1986) απέδειξε ότι η βέλτιστη επιλογή των W_t , ενός Tabular Cusum, βασίζεται στο λόγο της λογαριθμικής πιθανοφάνειας (*log-likelihood ratio*). Για παράδειγμα, έστω y το αποτέλεσμα της τρέχουσας διαδικασίας (το οποίο θα μπορούσε να είναι ο μέσος όρος ή κάποιο άλλο στατιστικό μέτρο, μίας υπο-ομάδας δεδομένων που συλλέχθηκαν την ίδια χρονική στιγμή). Ορίζουμε επίσης ως $f(y; \theta)$ την κατανομή πιθανότητας των πιθανών αποτελεσμάτων της υπο-ομάδας. Ο λόγος της λογαριθμικής πιθανοφάνειας εκφράζεται ως: $\ln(f(y; \theta_A) / f(y; \theta_0))$. Η επιλογή είναι η βέλτιστη, με τη λογική ότι μεταξύ όλων των συστημάτων (schemes) με το ίδιο ARL υπό την H_0 , ο λόγος της λογαριθμικής πιθανοφάνειας δίνει το μικρότερο ARL, υπό την υπόθεση H_A . Η επιλογή των ορίων ελέγχου για διεργασίες Cusum, έχει συζητηθεί για κανονικώς κατανομημένα αποτελέσματα (Woodall, 1986) και για την περίπτωση όπου τα δεδομένα ακολουθούν τη Διωνυμική Κατανομή (Gan, 1991).

Όταν εφαρμόζουμε την συνήθη μεθοδολογία Cusum στην παρακολούθηση χειρουργικών επεμβάσεων, σύμφωνα με τον Steiner *et al.* (1999), μπορούμε να είμαστε πιο συγκεκριμένοι όσον αφορά τον ορισμό των υπο-ομάδων και τη διαμόρφωση των W_t . Κατ' αρχάς λόγω της κρισιμότητας της εκάστοτε επέμβασης, το Cusum ενημερώνεται έπειτα από κάθε ασθενή, ώστε να είναι διασφαλισμένη η έγκαιρη ανίχνευση οποιασδήποτε αλλαγής στη διαδικασία. Συνεπώς, το αποτέλεσμα y αντιστοιχεί σε ένα από τα δύο πιθανά αποτελέσματα (επιτυχία ή αποτυχία) για κάθε ασθενή, και το Cusum είναι το άθροισμα των αποτελεσμάτων που έχουν ληφθεί από όλους τους ασθενείς που εγχειρήθηκαν, από την αρχή της παρακολούθησης της διαδικασίας. Υποθέτοντας ότι y_t είναι το αποτέλεσμα για κάθε ασθενή t , και ότι $y_t = 0$ αν ο ασθενής επιβιώσει, ενώ $y_t = 1$ σε αντίθετη περίπτωση, θα έχουμε ότι: $f(y_t/\theta) = p(\theta)^{y_t} [1 - p(\theta)]^{1-y_t}$, όπου: $p(\theta_0) = c_0$, ο εκτιμημένος τρέχων ρυθμός αποτυχίας και $p(\theta_A) = c_A$ μία σημαντική αλλαγή στο ρυθμό αποτυχίας.

Ας σημειωθεί ότι c_A μπορεί επίσης να εκφράζει χειροτέρευση ή βελτίωση του ρυθμού αποτυχίας των επεμβάσεων. Συνεπώς, το Cusum ελέγχει διαδοχικά την υπόθεση $H_0: p = c_0$ versus $H_1: p = c_A$ και τα αποτελέσματα του Cusum είναι:

$$W_t = \begin{cases} \log([1 - c_A] / [1 - c_0]), & \text{αν } y_t = 0 \\ \log(c_A / c_0), & \text{αν } y_t = 1 \end{cases} \quad (2)$$

Όταν σχεδιάζουμε το διάγραμμα για να ανιχνεύσουμε ενδεχόμενη αύξηση στο ρυθμό αποτυχίας χειρουργικών επεμβάσεων, τα αποτελέσματα που εκφράζουν τις αποτυχίες θα είναι θετικά, ενώ τα αντίστοιχα των επιτυχιών έχουν αρνητική τιμή. Με αυτόν τον τρόπο, η βαρύτητα κάθε ασθενούς καθορίζεται από τρεις παράγοντες: το τρέχων αποδεκτό επίπεδο της χειρουργικής απόδοσης (c_0), ένα επιλεγέν επίπεδο ρυθμού αποτυχίας που αντανακλά μία σημαντική αλλαγή στην απόδοση (c_A) και το τελικό αποτέλεσμα της επέμβασης για τον εκάστοτε ασθενή (y_t).

➤ Risk - Adjusted CUSUM

Σε διαδικασίες που αφορούν χειρουργικές επεμβάσεις, ο κίνδυνος πρόκλησης θανάτου που εκτιμάται ήδη πριν την επέμβαση, είναι πιθανόν να ποικίλει αισθητά από ασθενή σε ασθενή. Η προσαρμογή του προ-εκτιμημένου κινδύνου είναι απαραίτητη, προκειμένου να

επιβεβαιωθεί ότι οι ρυθμοί θνησιμότητας (που δείχνουν ασυνήθιστα υψηλοί) αποδίδονται δικαιολογημένα στο χειρουργό.

Είναι εφικτή η προσαρμογή του CUSUM, βάσει του πρότερου κινδύνου, με την προσαρμογή του μεγέθους των αποτελεσμάτων' χρησιμοποιώντας τον προ-εκτιμημένο χειρουργικό κίνδυνο του ασθενούς. Ο κίνδυνος αυτός είναι διαφορετικός για κάθε ασθενή και εξαρτάται από τους λεγόμενους παράγοντες κινδύνου που εμφανίζει ο κάθε ασθενής.

Ορίζουμε λοιπόν $p_t(\theta) = g(\theta, \mathbf{x}_t)$, όπου $\mathbf{x}_t = (\mathbf{x}_{t1}, \mathbf{x}_{t2}, \dots, \mathbf{x}_{tp})^T$ είναι ο δείκτης που εκφράζει τους παράγοντες κινδύνου για κάθε ασθενή t . Η συνάρτηση g μπορεί να οριστεί πριν τη επέμβαση, χρησιμοποιώντας μία μέθοδο αξιολόγησης, όπως Parsonnet παράγοντες κινδύνου (Parsonnet *et al.*, 1989), ή βασιζόμενοι σε ένα μοντέλο λογιστικής παλινδρόμησης προσαρμοσμένο στα δεδομένα του δείγματος. Από την στιγμή που καθένας ασθενής έχει και ένα διαφορετικό επίπεδο κινδύνου, ορίζουμε τις υποθέσεις H_0 και H_A , βάσει πιθανοτήτων (*odds ratio*). Έστω R_0 και R_A , τα οποία εκφράζουν τις πιθανότητες (ποσοστά), υπό τη μηδενική και την εναλλακτική υπόθεση, αντιστοίχως.

Προκειμένου να ανιχνεύσουμε ενδεχόμενες αυξήσεις στο ρυθμό θνησιμότητας των ασθενών, θέτουμε $R_0 > R_A$. Η επιλογή του R_A πραγματοποιείται, προσδιορίζοντας την ελάχιστη σημαντική επίδραση σε μία κλινική δοκιμή. Αν ο εκτιμημένος κίνδυνος p_t βασίζεται στις τρέχουσες συνθήκες, μπορούμε να θέσουμε $R_0 = 1$. Για δεδομένο κίνδυνο αποτυχίας ίσο με p_t , η πιθανότητα αποτυχίας ισούται με: $p_t / (1 - p_t)$.

Συνεπώς, για κάθε ασθενή t , υπό τη μηδενική υπόθεση H_0 , η πιθανότητα αποτυχίας είναι ίση με: $R_0 \cdot p_t / (1 - p_t)$, ενώ υπό την εναλλακτική υπόθεση H_A , η πιθανότητα αποτυχίας είναι $R_A \cdot p_t / (1 - p_t)$, η οποία αντιστοιχεί στην πιθανότητα αποτυχίας ίση με $R_A \cdot p_t / (1 - p_t + R_A \cdot p_t)$ υπό την εναλλακτική πιθανότητα H_A .

Συνεπώς, το CUSUM ελέγχει διαδοχικά τις υποθέσεις:

$$H_0: \text{odds ratios} = R_0 \quad \text{versus} \quad H_A: \text{odds ratios} = R_A$$

Συνεπώς, τα δύο αποτελέσματα λογαριθμικής πιθανοφάνειας, για κάθε ασθενή t , δίνονται από τις σχέσεις:

$$W_t = \begin{cases} \log \left[\frac{(1 - p_t + R_0 p_t) R_A}{(1 - p_t + R_A p_t) R_0} \right], & \text{αν } y_t = 1 \\ \log \left[\frac{1 - p_t + R_0 p_t}{1 - p_t + R_A p_t} \right], & \text{αν } y_t = 0 \end{cases} \quad (3)$$

Σημειώνουμε, σε αυτό το σημείο, ότι είναι πιθανό να έχουμε κι άλλα *weights*. Για παράδειγμα, έστω ότι θέλουμε να κατασκευάσουμε ένα CUSUM, χρησιμοποιώντας *weights* βάσει του παρατηρούμενου πλήθους θανάτων μείον το αναμενόμενο πλήθος των θανάτων. Η μέθοδος είναι όμοια με αυτήν που πρότειναν ο Lovegrove *et al.* (1997, 1999) και ο Poloniecki *et al.* (1998) και είναι ισοδύναμη με την παρακολούθηση των *cumulative weights* των ασθενών, όπου το βάρος κάθε ασθενή είναι ίσο με $-p_t$ αν επιζήσει και $1 - p_t$ αν ο ασθενής δεν τα καταφέρει. Παρόλα αυτά, όπως θα δούμε και σε επόμενη ενότητα αυτού του κεφαλαίου, τα *weights*, όπως υπολογίζονται βάσει της σχέσης (1), είναι τα βέλτιστα (Moustakides, 1986), για την ανίχνευση μετατοπίσεων στην πιθανότητα R_A .

Το CUSUM εκπέμπει σήμα, μόλις $\mathbf{X}_t \geq \mathbf{h}$. Την στιγμή εκείνη, έρχεται στην επιφάνεια το κρίσιμο ερώτημα σχετικά με το πώς πρέπει να ανταποκριθούμε στην εκπομπή σήματος σε ένα διάγραμμα. Καταρχήν, δεν μπορούμε να υποθέσουμε ότι στο μαθηματικό μοντέλο έχουν ληφθεί υπ' όψιν όλοι οι *risk-adjusted* παράγοντες. Οι θάνατοι που συνέβησαν πριν την εκπομπή σήματος στο CUSUM και είναι υπεύθυνοι για αυτόν το «συναγερμό», πρέπει να εξεταστούν από έναν ενημερωμένο αλλά αντικειμενικό αξιολογητή. Ίσως να υπάρχουν παράγοντες κινδύνου που είναι συνυφασμένοι με τον ασθενή και δεν ανιχνεύθηκαν στο μοντέλο πρόβλεψης κινδύνου. Το σύστημα βαθμολόγησης Parsonnet, για παράδειγμα, παραβλέπει παράγοντες, όπως τεχνικές δυσκολίες στις στεφανιαίες αρτηρίες και πνευμονικές ασθένειες. Υπάρχει καλός λόγος όμως για το ότι και οι δύο αυτοί παράγοντες παραλείφθηκαν. Και οι δύο επιδέχονται υποκειμενική «φόρτιση» από το χειρουργό που μπορεί να «φορτώσει» τόσο πολύ το επίπεδο κινδύνου, ώστε να φαίνεται τελικά ότι ο χειρουργός αυτός απέδωσε καλύτερα από ό,τι αναμενόταν. Αυτοί κι άλλοι παράγοντες μπορούν να αναγνωριστούν από άλλο χειρουργό. Περιστασιακοί παράγοντες, ίσως και να αγνοηθούν, ανάλογα με την κρίση του συγκεκριμένου χειρουργού.

Έστω κι αν η απόδοση του χειρουργού είναι απαλλαγμένη σε τεχνικό επίπεδο, ωστόσο κάποιος μπορεί να αναρωτηθεί εάν ο κλινικός κίνδυνος έχει πλήρως εκτιμηθεί. Η αποτυχία να ληφθούν σωστές κλινικές αποφάσεις είναι ένα σημαντικό πρόβλημα τεχνικής ειδικευσης. Αν η χειρουργική απόδοση δείχνει διαφορετική από το προβλεπόμενο, η οποία εξηγείται από την παρουσία των νέων παραγόντων κινδύνου, οι παράγοντες θα πρέπει να ενσωματωθούν στο μοντέλο που χρησιμοποιήθηκε για την εκτίμηση του πρότερου κινδύνου. Εναλλακτικά, ο αξιολογητής μπορεί να ανακαλύψει ότι το λάθος δεν αποδίδεται στον ίδιο το χειρουργό, αλλά σε κάποιο άλλο μέλος της ομάδας.

Πολύ πιθανόν να έχει εκδηλωθεί μηδενική ή τουλάχιστον ανεπαρκής μηδενική φροντίδα του ασθενούς, πριν την εγχείρηση. Το CUSUM εκπέμπει σήμα, όταν εντοπίσει πλήθος θανάτων πάνω από το αναμενόμενο, αλλά δεν μπορεί ούτε να αναζητήσει ούτε την αιτία του προβλήματος, ούτε και την επίλυση του. Μόλις εντοπιστεί η λύση του προβλήματος, πρέπει στη συνέχεια να ξεκινήσει η διαδικασία επίλυσης του προβλήματος, η οποία περιλαμβάνει επανεκπαίδευση, καθοδήγηση και περαιτέρω παρακολούθηση.

3.3.2. Εφαρμογή του risk-adjusted CUSUM σε καρδιοχειρουργική επέμβαση

Προκειμένου να παρουσιάσουμε τα χαρακτηριστικά του *risk-adjusted* CUSUM, θα χρησιμοποιηθεί ως παράδειγμα αυτό της καρδιοχειρουργικής επέμβασης. Αφού οι διεργασίες CUSUM σχεδιάστηκαν με σκοπό την άμεση ανίχνευση αλλαγών στη χειρουργική απόδοση, θα πρέπει πρώτα να εκτιμήσουμε το τρέχον επίπεδο της απόδοσης. Στην εγχείρηση καρδιάς για παράδειγμα, τα δεδομένα αποτελούνται από χαρακτηριστικά των ασθενών και χειρουργικά αποτελέσματα για τον καθένα ξεχωριστά, όπως παρατηρήθηκαν σε ένα συγκεκριμένο χειρουργικό κέντρο, την περίοδο 1992-1998. Την περίοδο αυτή, δεν έγινε καμία επίσημη παρακολούθηση της όλης διαδικασίας. Ωστόσο προς εξήγηση όλου του μηχανισμού, υποθέτουμε ότι η παρακολούθηση ξεκίνησε το 1994 και θα χρησιμοποιήσουμε τα δεδομένα που αφορούν τα πρώτα δύο χρόνια (1992 και 1993) προκειμένου να αναγνωρίσουμε τους παράγοντες κινδύνου και να εκτιμήσουμε τις επιδράσεις που ασκούν. Όλα αυτά θα γίνουν μέσω ενός μοντέλου λογιστικής παλινδρόμησης. Στα δύο πρώτα χρόνια, στις 2.218 εγχειρήσεις που πραγματοποιήθηκαν, καταγράφηκαν 413 θάνατοι για ρυθμό θνησιμότητας της τάξης του 6,5%.

Χρησιμοποιείται λοιπόν στη συνέχεια το μοντέλο λογιστικής παλινδρόμησης, το οποίο περιγράφεται από την ακόλουθη σχέση:

$$\text{logit}(p_t) = - 3,68 + 0,77 X_t \quad (4)$$

όπου X_t υποδηλώνει την τιμή του δείκτη *Parsonnet* για κάθε ασθενή t , όπου $t = 1, 2, 3, \dots$. Αυτό δε χρειάζεται να μας προκαλεί ιδιαίτερη έκπληξη, καθώς ο δείκτης *Parsonnet* είναι βασισμένος σε ένα συνδυασμό πολλών επεξηγηματικών μεταβλητών σημαντικών για καρδιοχειρουργικές επεμβάσεις. Βάσει του μοντέλου αυτού, για τους ασθενείς χαμηλότερου κινδύνου (*Parsonnet* score = 0), εκτιμήθηκε ότι ο κίνδυνος θανάτου στην επόμενη εγχείρηση είναι 2,5%, ενώ ο αντίστοιχος κίνδυνος για τους ασθενείς υψηλότερου

κινδύνου (Parsonnet score = 71) είναι 86%. Αυτό σημαίνει ότι η προσαρμογή για το σύνολο των ασθενών (*patient mix*), είναι κρίσιμη.

Στην εφαρμογή αυτή χρησιμοποιούμε δύο διεργασίες CUSUM. Η πρώτη σχεδιάστηκε με σκοπό την ανίχνευση διπλασιασμού του ποσοστού θανάτων (έστω ότι επιλέγουμε *odds ratios* $R_0 = 1$ και $R_A = 2$) και η δεύτερη σχεδιάστηκε αντίστοιχα για την ανίχνευση υποδιπλασιασμού του ποσοστού (με $R_0 = 1$ και $R_A = 0,5$).

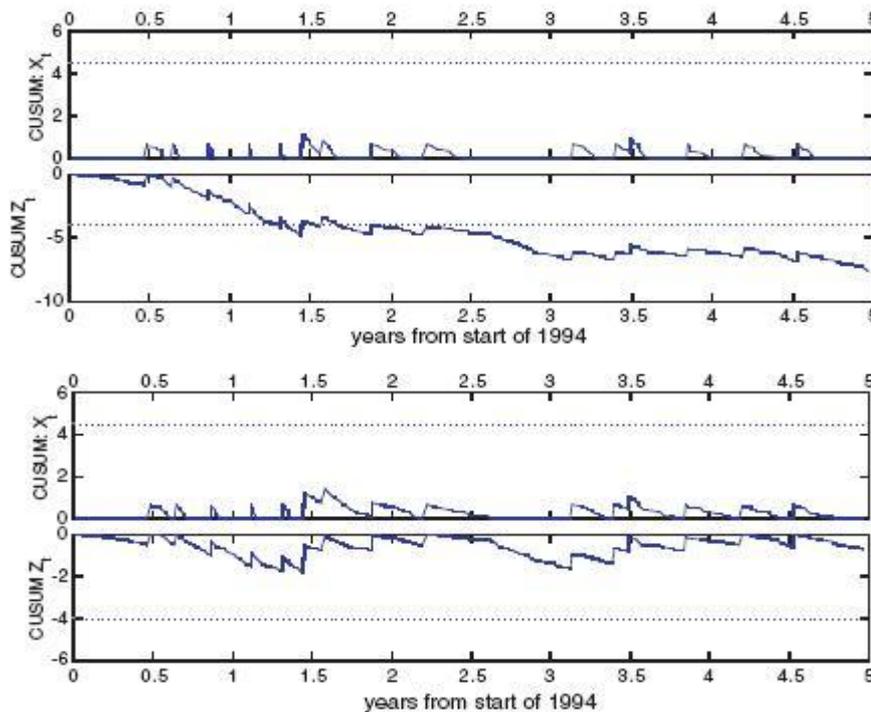


Fig. 1. Trainee surgeons CUSUM; unadjusted CUSUMs on top, risk-adjusted CUSUMs on the bottom.

Με αυτόν τον τρόπο, θα πρέπει να είμαστε σε άμεση ετοιμότητα ώστε να εντοπίσουμε γρήγορες αλλαγές στο ρυθμό αποτυχίας της επέμβασης. Φυσικά, αν το επιθυμούσε κάποιος, θα μπορούσε να παρακολουθήσει μόνο την περίπτωση ενδεχόμενης αύξησης στο ρυθμό αποτυχίας. Παρόλα αυτά το CUSUM που σχεδιάστηκε για ανίχνευση μείωσης στο ρυθμό αποτυχίας χειρουργικών επεμβάσεων είναι χρήσιμο, διότι σε περίπτωση υπερεκτιμημένου ρυθμού αποτυχίας ή βελτιωμένης χειρουργικής απόδοσης, το CUSUM που ανιχνεύει αυξήσεις θα είναι λιγότερο ευαίσθητο. Αν το CUSUM που σχεδιάστηκε για ανίχνευση βελτίωσης της χειρουργικής απόδοσης, εκπέμψει σήμα, τότε πρέπει να επανεκτιμήσουμε το ρυθμό αποτυχίας ώστε να διασφαλιστεί η προστασία των ασθενών απέναντι σε μελλοντικές αλλαγές του ρυθμού αυτού.

Τα αποτελέσματα του CUSUM υπολογίζονται από τη σχέση (4) σε συνδυασμό με τη σχέση (3), όπου εκχωρείται θετικό αποτέλεσμα σε περίπτωση θανάτου, ενώ αρνητικό σε περίπτωση επιβίωσης. Για παράδειγμα, το CUSUM που ανιχνεύει αύξηση στο ρυθμό αποτυχίας, δίνει τα ακόλουθα αποτελέσματα:

0,67 και **-0,024** για ασθενείς με *Parsonnet score* = 0, και **0,26** και **-0,43** για ασθενείς υψηλότερου κινδύνου με *Parsonnet score* = 50.

Σημειώνουμε ότι τα προαναφερθέντα αποτελέσματα εκφράζουν τον προ-εγχειρητικό κίνδυνο ανά ασθενή, καθώς η «τιμωρία» για το θάνατο ενός χαμηλού κινδύνου-ασθενούς, είναι πολύ πιο αυστηρή απ' ό,τι στο θάνατο ενός υψηλού κινδύνου-ασθενούς.

Χάριν εύκολης εκτέλεσης και παρουσίασης της διαδικασίας, το CUSUM που ανιχνεύει μειώσεις στο ρυθμό αποτυχίας των χειρουργικών επεμβάσεων, θα «συσσωρεύσει» αρνητικές τιμές, δηλαδή ο νέος τύπος δίνεται από τη σχέση:

$$Z_t = \min(0, Z_{t-1} - W_t) \quad (5)$$

όπου $Z_0 = 0$ και W_t που προσδιορίζεται από τη σχέση (3) και όχι από την (1). Για το ίδιο CUSUM, θέτουμε επίσης το όριο ελέγχου να είναι μία αρνητική τιμή. Με αυτόν τον τρόπο, μπορούμε να σχεδιάσουμε το CUSUM το οποίο ανιχνεύει μειώσεις στο ρυθμό χειρουργικής αποτυχίας, κάτω από το αντίστοιχο CUSUM που ανιχνεύει αυξήσεις (βλ. *Figure 1*)

Για την εφαρμογή στην καρδιοχειρουργική επέμβαση, ο προσδιορισμός των ορίων ελέγχου για τα δύο διαγράμματα CUSUM (τα οποία δίνονται από τις σχέσεις (1) και (5)), με τιμές **4,5** και **-4** αντίστοιχα, δίνει μέσο μήκος ροής περίπου ίσο με 9.600 ασθενείς για κάθε διάγραμμα, όταν η χειρουργική απόδοση είναι αποδεκτή.

Δοθείσης της συχνότητας των χειρουργικών επεμβάσεων σε αυτό το παράδειγμα, αυτό υπονοεί ότι η διαδικασία θα εκπέμπει σήμα κάθε εννέα χρόνια, κατά μέσο όρο, ακόμη κι αν δεν αλλάξει το ποσοστό θνησιμότητας. Φυσικά, τα όρια ελέγχου μπορούν να αντικατασταθούν από άλλες τιμές, ανάλογα με τους στόχους που έχουμε θέσει προς υλοποίηση στο διάγραμμα μας.

Σε αυτό το παράδειγμα, το «μείγμα» των ασθενών αναφέρεται στην κατανομή των *Parsonnet scores* του εξεταζόμενου χειρουργού. Το «μείγμα» των ασθενών διέφερε για κάθε χειρουργό, με πιο έμπειρους χειρουργούς να λαμβάνουν υψηλότερου κινδύνου κατάστασεις. Το διάγραμμα *Figure 1* απεικονίζει το CUSUM, στο οποίο περιλαμβάνονται μόνο οι ασθενείς οι οποίοι χειρουργήθηκαν από ασκούμενους χειρουργούς, καθ' όλη τη διάρκεια

της περιόδου. Οι ασκούμενοι είχαν ποσοστό θνησιμότητας περίπου 2,5%, σημαντικά χαμηλότερο σε σχέση με το συνολικό ποσοστό, οι οποίοι ασχολήθηκαν μόνο με τις σχετικά απλές περιπτώσεις. Επίσης, αν κατά τη διάρκεια της επέμβασης προκύψουν σοβαρές επιπλοκές, τότε τον έλεγχο θα αναλάβει κάποιος έμπειρος χειρουργός και συνεπώς το περιστατικό θα αποδοθεί σε αυτόν.

Συνεπώς, κάποιος μπορεί να περιμένει λιγότερους – σε σχέση με το αναμενόμενο - θανάτους σε επεμβάσεις που πραγματοποιούνται από ασκούμενους, βάσει ενός κριτηρίου κινδύνου σαν το *Parsonnet score*. Με την ίδια λογική, κάποιος μπορεί να περιμένει περισσότερους – από το αναμενόμενο - θανάτους από έναν έμπειρο χειρουργό ο οποίος επιβλέπει τρεις ασκούμενους. Η χειρουργική εκπαίδευση συζητείται εκτενέστερα από τον Anderson *et al.* (1996). Χρησιμοποιώντας μία συνήθη (*unadjusted*) διεργασία CUSUM, θα λάβουμε τα δύο διαγράμματα CUSUM, που απεικονίζονται στο διάγραμμα **Figure 1**.

Σύμφωνα με το διάγραμμα, η απόδοση των ασκούμενων χειρουργών δείχνει σημαντικά καλύτερη, αφότου το CUSUM (που ανιχνεύει βελτιώσεις στην απόδοση) εκπέμπει σήμα κοντά στο Μάρτιο του 1995. Λόγω του μεγέθους του ορίου ελέγχου, είμαστε αρκετά βέβαιοι πως η εκπομπή του σήματος δε συνέβη από καθαρή τύχη. Παρατηρώντας ένα τέτοιο σήμα στο CUSUM, θα επιχειρούσαμε να προσδιορίσουμε το λόγο για τον οποίον οι ασκούμενοι απέδωσαν τόσο καλά, με την ελπίδα ότι την ίδια βελτίωση θα δείξουν κι άλλοι χειρουργοί. Για παράδειγμα, άλλοι χειρουργοί ίσως προσπαθήσουν να αντιγράψουν την ικανοποιητική απόδοση των «καλών» χειρουργών. Βάσει αλλαγών που παρατηρούνται στη χειρουργική διαδικασία, θα επαναφέρουμε όλες τις τιμές του CUSUM στο μηδέν για κάθε χειρουργό, ώστε να εξετάσουμε αν οι αλλαγές ήταν ενεργές.

Όταν παρατηρούμε τα *risk-adjusted* CUSUM (το ζεύγος των CUSUM στο κάτω μέρος του διαγράμματος Figure 1), διαπιστώνουμε ότι οι ασκούμενοι χειρουργοί αποδίδουν τόσο καλά όσο αναμενόταν, χωρίς αισθητά σημαντική βελτίωση. Το σήμα που παρατηρούμε στο σύνηθες CUSUM, οφείλεται στο «μείγμα» των ασθενών χαμηλού κινδύνου που δόθηκε στους ασκούμενους. Συνεπώς, το συμπέρασμα που προέκυψε από το *unadjusted* CUSUM ήταν αναληθές και θα ήταν μάταια ο χρόνος και η προσπάθεια που αφιερώνονται στην εύρεση του παράγοντα που βελτίωσε την απόδοση.

Στη συνέχεια εξετάζεται η απόδοση ενός έμπειρου χειρουργού. Στο διάγραμμα **Figure 2** της επόμενης σελίδας, το άνω ζεύγος διαγραμμάτων CUSUM προκύπτει από ένα σύνηθες CUSUM με *unadjusted* αποτελέσματα, ενώ το κάτω ζεύγος διαγραμμάτων CUSUM προκύπτει από τη χρήση *risk-adjusted* σταθμών (*weights*). Σε αυτήν την περίπτωση, βάσει

των *unadjusted* αποτελεσμάτων, ο έμπειρος χειρουργός φαίνεται να έχει αισθητά χαμηλή απόδοση. Το *unadjusted* CUSUM εκπέμπει σήμα, καθώς εντοπίζει αύξηση στο ρυθμό θνησιμότητας ακριβώς πριν το τέλος του έτους 1994. Παρατηρώντας το σήμα αυτό, προσπαθούμε να εντοπίσουμε την αιτία του «προβλήματος».

Παρόλα αυτά όταν κάνουμε την προσαρμογή για το σύνολο των ασθενών που έχει αναλάβει ο έμπειρος χειρουργός, το σήμα «χάνεται» (βλ. στο κάτω μέρος του Figure 2). Και πάλι το συμπέρασμα που προέκυψε από το *unadjusted* CUSUM ήταν λανθασμένο και όλη η προσπάθεια που αφιερώνεται στην εύρεση της αιτίας του προβλήματος θα ήταν μάταιη.

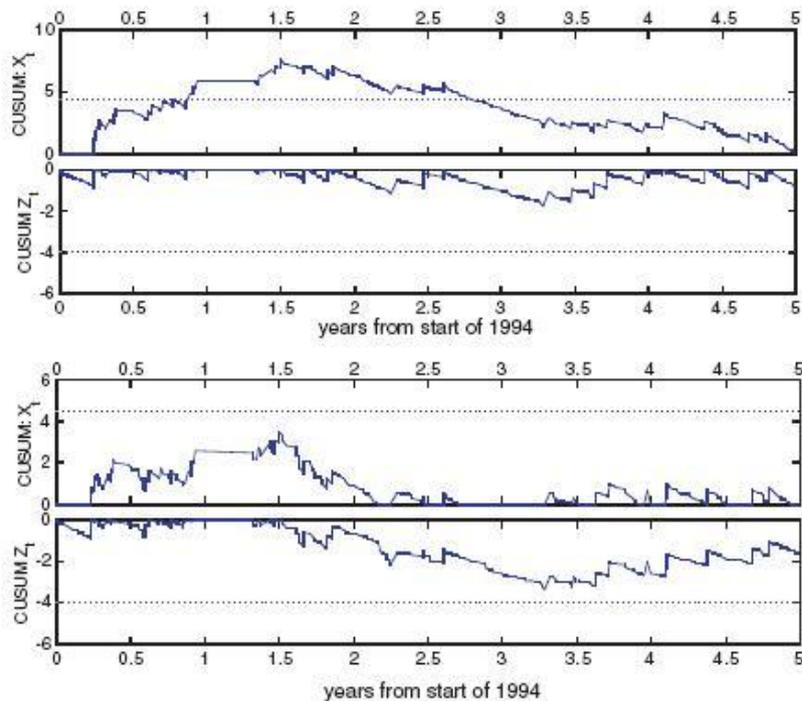


Fig. 2. Experienced surgeon CUSUM; unadjusted CUSUMs on the top, risk-adjusted CUSUMs on the bottom.

3.3.3. Χαρακτηριστικά της *risk-adjusted* διεργασίας CUSUM

Βήμα-κλειδί στο σχεδιασμό διαγραμμάτων CUSUM, είναι ο προσδιορισμός των ορίων ελέγχου. Χρησιμοποιώντας τον κανόνα αλυσίδας Markov μπορούμε να προσεγγίσουμε αρκετά το ARL ενός διαγράμματος ελέγχου για οποιοδήποτε όριο ελέγχου, ακριβές επίπεδο απόδοσης και το «μείγμα» των ασθενών. Χρησιμοποιώντας αυτήν την προσέγγιση, επιλέγουμε το όριο ελέγχου με τέτοιον τρόπο, ώστε το ARL, δοθέντων του τρέχοντος (εκτιμημένου) επιπέδου απόδοσης και του «μείγματος των ασθενών», να είναι σχετικά μεγάλο. Όρια ελέγχου με τιμή μεγαλύτερη του μηδενός, θα οδηγήσουν σε μεγαλύτερες τιμές του ARL. Εδώ όμως χρειάζεται προσοχή, καθώς ένα όριο ελέγχου μεγαλύτερο του μηδενός οδηγεί σε μεγαλύτερα ARL, όταν έχει συμβεί αλλαγή στη χειρουργική απόδοση.

Για να το μελετήσουμε αυτό, μπορούμε να ποσοτικοποιήσουμε την ικανότητα της διεργασίας CUSUM να ανιχνεύει αυξήσεις (ή μειώσεις) στα ποσοστά θνησιμότητας. Το διάγραμμα *Figure 3* δείχνει τη γραφική παράσταση του ARL ως προς τη χειρουργική απόδοση (εκφρασμένης ως *odds ratio*) για διαφορετικά «μείγματα» ασθενών. Οι τιμές του ARL δίνονται για το CUSUM που σχεδιάστηκε για να ανιχνεύει αυξήσεις στο ποσοστό αποτυχίας. Σημειώνουμε ότι τα ARLs δίνονται σε μεγάλη κλίμακα.

Στο διάγραμμα **Figure 3** της επόμενης σελίδας, οι συμπαγείς γραμμές δίνουν τα αποτελέσματα για το τρέχων μείγμα των ασθενών για όλους τους χειρουργούς που παρατηρήθηκαν κατά τα πρώτα δύο χρόνια των δεδομένων. Για λόγους σύγκρισης, απεικονίζεται και η καμπύλη ARL που προκύπτει από τη διαφορά των σωρευτικών προβλεπόμενων θανάτων μείον τους σωρευτικούς εκτιμημένους θανάτους. Όπως συζητήθηκε, το προτεινόμενο *weighting scheme* το οποίο βασίζεται σε *log-likelihood ratios* είναι ανώτερο, υπό την έννοια ότι για μεγάλες αλλαγές στα ποσοστά το ARL - του CUSUM που χρησιμοποιούμε - είναι πολύ μικρότερο. Σχεδιάζουμε επίσης δύο διακεκομμένες γραμμές στο *Figure 3*, οι οποίες δίνουν τα αποτελέσματα για το «μείγμα» των ασθενών, δύο συγκεκριμένων χειρουργών του παραδείγματος μας. Οφείλουμε να υπενθυμίσουμε σε αυτό το σημείο, ότι αναφερόμενοι στο «μείγμα» ασθενών, εννοούμε τους ασθενείς υψηλότερου και τους ασθενείς χαμηλότερου κινδύνου. Προκειμένου να ποσοτικοποιήσουμε την ευαισθησία της συγκεκριμένης διεργασίας CUSUM, με την αρχική εκτίμηση του «μείγματος» των ασθενών και τις παραμέτρους παλινδρόμησης, χρησιμοποιούμε την εξής διαδικασία η οποία προτάθηκε από τους Efron & Tibshirani *et. al* (2001).

Παράγονται τα δείγματα (*bootstrap samples*) των 2.218 παρατηρήσεων, από τα δεδομένα που συλλέχθηκαν τα δύο πρώτα χρόνια, προβαίνοντας ουσιαστικά σε δειγματοληψία με αντικατάσταση. Στη συνέχεια εκτιμώνται οι παράμετροι του μοντέλου που περιγράφεται από τη σχέση (4). Έτσι καθορίζονται τα αποτελέσματα λογαριθμικής πιθανοφάνειας, όπως εκφράζονται στη σχέση (3). Το διάγραμμα **Figure 4**, απεικονίζει το 5ο και το 95ο εκατοστημόριο των CUSUM «μονοπατιών», βάσει 1.000 δειγμάτων (*bootstrap samples*), όπου το άνω διάγραμμα αντιστοιχεί στους ασκούμενους, ενώ το κάτω στους έμπειρους χειρουργούς.

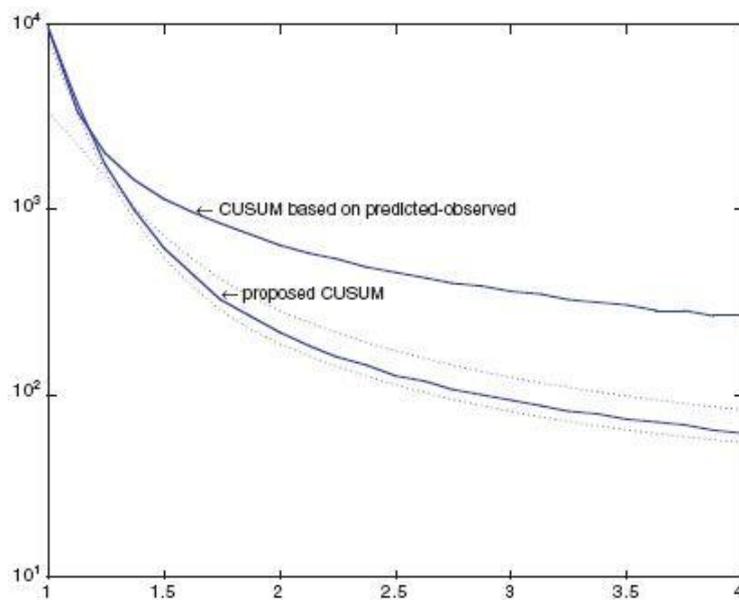


Fig. 3. Average run length (on a log scale) for different actual odds ratios. Solid line shows performance with current patient mix. Dotted lines—surgeon with the lowest and highest risk patients.

Τα αποτελέσματα του *Figure 4*, δείχνουν ότι το αρχικό συμπέρασμα που διεξήχθη, ότι δηλαδή οι δύο χειρουργοί έχουν αναμενόμενη απόδοση (την στιγμή που θα γίνει η προσαρμογή για το «μείγμα» των ασθενών), είναι σχετικά ισχυρό. Στην περίπτωση των ασκούμενων χειρουργών, ακόμη και στη χειρότερη περίπτωση, το μονοπάτι CUSUM δε διασχίζει τα όρια ελέγχου. Στην περίπτωση όμως του έμπειρου χειρουργού, η αβεβαιότητα είναι μεγαλύτερη. Αν και το 95ο εκατοστημόριο του CUSUM υπερβαίνει το άνω όριο ελέγχου πολύ κοντά στα μέσα του 1995, αμέσως γρήγορα το CUSUM φθίνει και πάλι, υποδηλώνοντας έτσι μια ανωμαλία. Συγκρίνοντας αυτό το διάγραμμα με το CUSUM του *Figure 2*, όπου το γράφημα παρέμεινε πάνω από το όριο ελέγχου για μεγάλη χρονική

περίοδο, θα συμπεράνουμε ότι η απόδοση του έμπειρου χειρουργού ήταν αισθητά ικανοποιητική. Στα διαγράμματα CUSUM του κάτω μέρους στο *Figure 4*, οι «ακραίες» περιπτώσεις στο μονοπάτι CUSUM, διασχίζουν επίσης το κατώτερο όριο ελέγχου. Εδώ, το CUSUM παραμένει κάτω από το όριο ελέγχου για περισσότερο χρόνο, υποδηλώνοντας ότι ο έμπειρος χειρουργός αποδίδει τελικά καλύτερα από το αναμενόμενο.

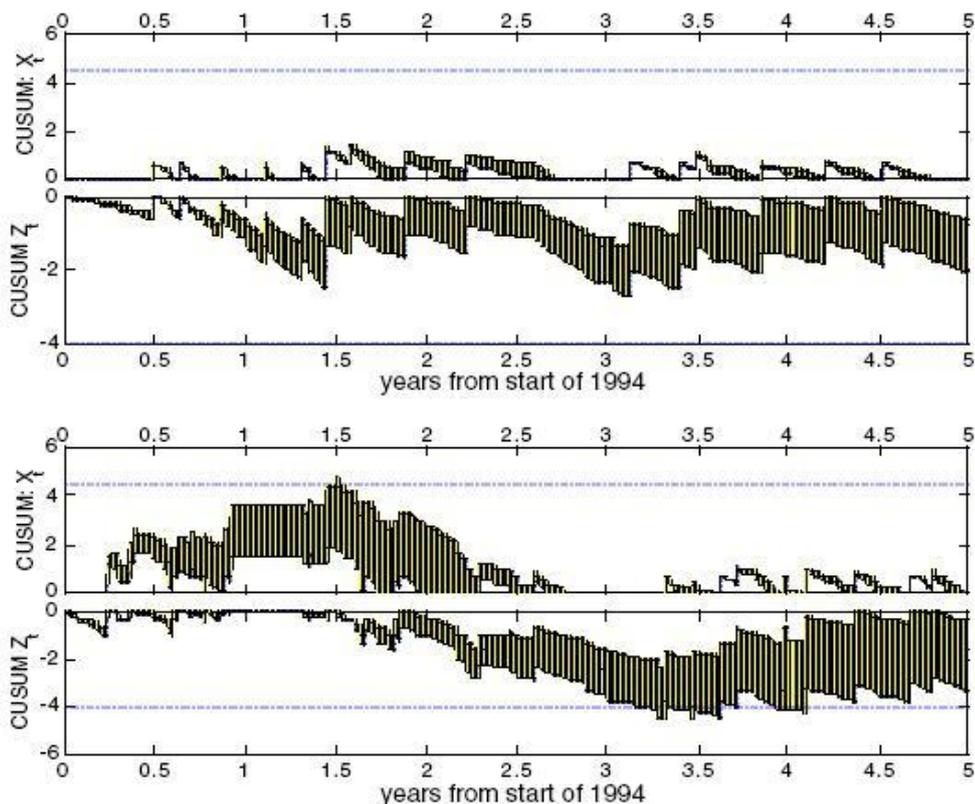


Fig. 4. Plot of the 5th and 95th percentile of the CUSUM paths based on 1000 bootstrap samples; trainee surgeons on the top, experienced surgeon on the bottom.

ΚΕΦΑΛΑΙΟ 4

ΔΙΑΓΡΑΜΜΑ CUSUM ΣΕ ΣΥΝΕΧΗ ΧΡΟΝΟ, ΒΑΣΕΙ ΤΟΥ ΜΟΝΤΕΛΟΥ COX: ΕΦΑΡΜΟΓΗ ΣΤΗ ΜΕΤΑΜΟΣΧΕΥΣΗ ΟΡΓΑΝΩΝ

4.1. Το *risk-adjusted* CUSUM βάσει του μοντέλου Cox: Εισαγωγή, Υποθέσεις

Σε κλινικές και ιατρικές εφαρμογές, είναι αρκετά σημαντικό να παρακολουθούμε τα αποτελέσματα που σχετίζονται με τις εκάστοτε επεμβάσεις. Στη μεταμόσχευση οργάνων, για παράδειγμα, είναι σημαντικό να παρακολουθούμε και να αξιολογούμε τα αποτελέσματα που σχετίζονται με τα συμμετέχοντα ιατρικά κέντρα. Στη μεταμόσχευση οργάνων, για παράδειγμα, είναι σημαντική η παρακολούθηση της έκβασης των μεταμοσχεύσεων στα συγκεκριμένα κέντρα, καθώς και η αποστολή σήματος σε περίπτωση ανοδικής τάσης στο ποσοστό αποτυχίας.

Όσον αφορά τον τομέα της βιομηχανίας και τη διαδικασία ελέγχου, η τεχνική της σωρευτικής άθροισης (CUSUM), χρησιμοποιείται ως σύστημα διαδοχικής παρακολούθησης για ορισμένο χρονικό διάστημα. Αρκετά πρόσφατα, το CUSUM έχει επίσης προταθεί για χρήση του σε εφαρμογές της Ιατρικής. Σε αυτό το κεφάλαιο, αξιολογείται το *risk-adjusted* CUSUM βάσει του μοντέλου Cox, για την περίπτωση ανεπιτυχούς έκβασης στη μεταμόσχευση. Συνεπώς, λαμβάνονται θεωρητικές προσεγγίσεις στο μέσο μήκος ροής για τη νέα αυτή πρόταση και για ορισμένες διεργασίες διακριτού χρόνου που προτάθηκαν στη βιβλιογραφία.

Σε ιατρικά προγράμματα όπου συμμετέχουν αρκετά ιατρικά κέντρα, είναι σημαντική τόσο η παρακολούθηση όσο και η παροχή έγκαιρης πληροφόρησης ως προς την έκβαση της εκάστοτε ιατρικής επέμβασης. Μία τέτοια δραστηριότητα παρακολούθησης, μπορεί να αποδειχθεί αρκετά χρήσιμη ως προς το να παρέχει προειδοποιητικά σήματα σε ελεγκτικές επιτροπές και νοσοκομειακά ιδρύματα και ιατρικά κέντρα. Συνήθεις στατιστικές τεχνικές, όπως η μέση θνησιμότητα, η *risk-adjusted* θνησιμότητα και η πολυπαραμετρική μοντελοποίηση, μπορεί να χρησιμοποιηθεί στον εντοπισμό αλλαγών στην απόδοση της

διαδικασίας, σε εθνικό επίπεδο. Παρόλα αυτά ίσως να αποδειχθεί «αδύναμη» στον εντοπισμό μικρών αλλαγών, σε διαδικασίες που λαμβάνουν χώρο σε νοσοκομειακό επίπεδο.

Η διαδικασία της σωρευτικής άθροισης (CUSUM) η οποία συζητήθηκε αργότερα από τον Van Dobben de Bruyn *et. al* (1999) και πολλούς ακόμη, είναι μία γραφική διαδικασία για συνεχή και διαδοχική παρακολούθηση. Όπως αναφέρθηκε και σε προηγούμενα κεφάλαια, αρχικά είχε προταθεί ως τεχνική στατιστικού ελέγχου στη βιομηχανική και παραγωγική διαδικασία και έπειτα συζητήθηκε η χρήση του στον τομέα της Υγείας.

Ως προς την εφαρμογή του CUSUM, ή άλλων μεθόδων παρακολούθησης και αξιολόγησης, είναι σημαντική η χρήση *risk-adjusted* μεθόδων, καθώς μπορεί να συμβάλλουν στην εξήγηση θεσμικών διαφορών που προκύπτουν εξαιτίας διαφορών στο «μείγμα» των ασθενών (*mix of patients*). Προσφάτως, οι Axelrod, Steiner (2000), καθώς και οι Grigg & Farewell (2000), πρότειναν τη χρήση *risk-adjusted* διαγραμμάτων CUSUM, για την παρακολούθηση της νοσοκομειακής απόδοσης. Στο συγκεκριμένο κεφάλαιο λοιπόν, θα περιοριστούμε σε προτεινόμενες μεθόδους για την αξιολόγηση των αποτελεσμάτων των μεταμοσχεύσεων, σε νοσοκομειακό επίπεδο. Αν και οι μέθοδοι εφαρμόζονται σε εφαρμόζονται σε άλλους τομείς, εδώ θα χρησιμοποιηθεί αυτός της μεταμόσχευσης, όπως θα δούμε σε ακόλουθες ενότητες.

Όσον αφορά τις παραπάνω εφαρμογές του CUSUM στην Υγεία, κάθε μεταμόσχευση ακολουθείται για ένα προκαθορισμένο χρονικό διάστημα, ώστε να αποφασιστεί αν συμβαίνει ή όχι ένα αποτέλεσμα Bernoulli, αν δηλαδή θα επιβιώσει ή όχι ο ασθενής. Μία βαθμολογική προσέγγιση βασισμένη σε λόγο πιθανοφάνειας, χρησιμοποιείται στη συνέχεια για να συσσωρεύσει αυτά τα αποτελέσματα σε ένα CUSUM. Μία συνέπεια αυτής της προσέγγισης είναι ότι ένας θάνατος (αντίστοιχα επιβίωση) επηρεάζει το CUSUM, μόνο ένα χρόνο από τότε που πραγματοποιήθηκε η μεταμόσχευση. Αν και οι προσεγγίσεις αυτές αποδίδουν ικανοποιητικώς, αυτή η -ενός χρόνου- καθυστέρηση στην αξιολόγηση, είναι ανεπιθύμητη: θα μπορούσε να παραχθεί ένα πιο αποτελεσματικό CUSUM αν υπήρχε μία συνεχής ενημέρωση των πραγμάτων. Αξίζει βεβαίως να αναφέρουμε ότι αν τα δεδομένα αναφέρονται με μία σταθερή χρονική καθυστέρηση ενός έτους από την ημερομηνία μεταμόσχευσης, τότε το CUSUM διακριτού χρόνου θα είναι απολύτως κατάλληλο.

Στο συγκεκριμένο κεφάλαιο θα παρουσιάσουμε στη συνέχεια ένα *risk-adjusted* CUSUM το οποίο είναι βασισμένο στο μοντέλο Cox και είναι ορισμένο σε συνεχή χρόνο. Θα υποθέσουμε ότι οι αναφορές είναι έγκαιρες: συνεπώς δεν υπάρχει καθόλου καθυστέρηση ανάμεσα στην εμφάνιση του αποτελέσματος και στη γνωστοποίησή του.

Ως αναμενόμενο, οι προσομοιώσεις και οι θεωρητικές προσεγγίσεις δείχνουν ότι αυτή η μέθοδος παρέχει μία πιο άμεση ανίχνευση επιδείνωσης της απόδοσης, σε σχέση με τις μεθόδους διακριτού χρόνου που περιγράφονται παραπάνω.

Το παρόν κεφάλαιο δομείται ως εξής: στην παρούσα ενότητα, την 4.1. θα καθοριστούν οι υποθέσεις και το πλαίσιο ορισμού του CUSUM σε συνεχή χρόνο. Η Ενότητα 4.2. θέτει στο επίκεντρο της, το διάγραμμα CUSUM σε συνεχή και σε διακριτό χρόνο, ενώ η 4.3. συζητά αρκετές μελέτες προσομοιώσεων. Τέλος, στην ενότητα 4.4. εφαρμόζονται οι διεργασίες CUSUM σε δεδομένα μεταμόσχευσης νεφρού όπως περιέχονται στη βάση δεδομένων στατιστικής μεταμοσχεύσεων: «*Scientific Registry of Transplant Recipients (SRTR)*»

► Σημειογραφία και Υποθέσεις

Ας θεωρήσουμε ένα συγκεκριμένο νοσοκομειακό ίδρυμα και ας υποθέσουμε ότι οι μεταμοσχεύσεις ξεκινούν από μία συγκεκριμένη χρονική αφετηρία, έστω $t = 0$. Ορίζουμε ως S_1, S_2, \dots , τις επιτυχείς φορές μεταμοσχεύσεων για τους ασθενείς $1, 2, \dots$, αντίστοιχα. Ως μονάδα μέτρησης του χρόνου θεωρούμε τα έτη.

Έστω X_i ο χρόνος που μεσολαβεί από την στιγμή της μεταμόσχευσης μέχρι την ανεπιτυχή έκβαση της επέμβασης ή το θάνατο και ας υποθέσουμε ότι X_i και S_i είναι ανεξάρτητες μεταξύ τους. Οι μεταβλητές Z_i υπολογίζονται κατά τη χρονική στιγμή της μεταμόσχευσης. Η χρονολογική στιγμή της αποτυχίας στη μεταμόσχευση, δίνεται από τη σχέση: $T_i = S_i + X_i$. Υποθέτουμε ότι δεδομένων των μεταβλητών Z_i , υπάρχει μία γνωστή, *null* κατανομή των X_i , η οποία προσδιορίζεται από τη συνάρτηση επικινδυνότητας $\alpha_i(x)$. Στις εφαρμογές που μας απασχολούν, αυτή η κατανομή εκτιμάται από τη συνδυασμένη εμπειρία όλων των κέντρων μεταμόσχευσης.

Χρησιμοποιήσαμε ένα μοντέλο παλινδρόμησης Cox βάσει των μεταβλητών Z_i , που εκφράζεται με τη σχέση: $\alpha_i(x) = \lambda_0(x) \cdot e^{Z_i^T \beta}$, για $x > 0$. Συνεπώς, η συνάρτηση $\alpha_i(x)$ εκφράζει τον εθνικό μέσο ρυθμό αποτυχίας για κάθε ασθενή. Φυσικά, θα μπορούσαν να χρησιμοποιηθούν άλλα μοντέλα παλινδρόμησης. Σε ό,τι ακολουθεί, υποθέτουμε ότι οι εκτιμήσεις που λαμβάνονται για το α_i , δεν εμπεριέχουν σφάλμα. Στην πραγματικότητα, τα σφάλματα εκτίμησης θα είναι πολύ μικρά, καθώς οι εκτιμήσεις προκύπτουν από ένα μεγάλο σύνολο δεδομένων. Σε άλλες εφαρμογές, κάποιος θα μπορούσε να λάβει υπ' όψιν τον παράγοντα της αβεβαιότητας κατά την εκτίμηση του α_i .

Είναι σχετικά σύνηθες να χρησιμοποιούνται ενός έτους – προμεταμοσχευτικά- αποτελέσματα, ως ένα τουλάχιστον κριτήριο επιτυχίας. Ορίζουμε επομένως ως αποτυχία μεταμόσχευσης την περίπτωση θανάτου του ασθενούς, ή την αποτυχία δράσης του εμβολίου εντός ενός έτους από την ημερομηνία μεταμόσχευσης. Έστω η χρονική στιγμή t και μια διαδικασία υπολογισμού, $N_i^D(t)$, η οποία αντιστοιχεί στην i -οστή μεταμόσχευση. Συγκεκριμένα, έστω $N_i^D(t) = 1$, για κάθε $t \geq T_i$, αν η χρονική στιγμή αποτυχίας $T_i \in (S_i, S_i + 1]$ και $N_i^D(t) = 0$, σε κάθε άλλη περίπτωση.

Έστω: $dN_i^D(t) = N_i^D(t + dt) - N_i^D(t)$, όπου dt είναι μία στοιχειώδης ποσότητα. Στη συνέχεια, έχουμε:

$$P(dN_i^D(t) = 1/T_i \geq t, S_i, Z_i) = \begin{cases} \{e^{-\alpha_i(t - S_i)} dt, \text{αν } 0 \leq t - S_i \leq 1, T_i \geq t\} \\ 0, \text{αλλιώς} \end{cases} \quad (4.1)$$

Ορίζουμε το δεξί μέλος της (4.1.) ως $e^{-\alpha_i} \cdot d\Lambda_i(t)$ και να σημειώσουμε ότι εκφράζει τον στιγμιαίο κίνδυνο της αποτυχίας της μεταμόσχευσης. Εδώ η ποσότητα $e^{-\alpha_i}$ είναι ο συντελεστής με τον οποίον ο κίνδυνος στο υπό μελέτη νοσοκομειακό ίδρυμα, διαφέρει από τον αντίστοιχο σε εθνικό επίπεδο ($d\Lambda_i(t)$). Σημειώνουμε ότι εδώ χρησιμοποιούμε δύο διαφορετικές χρονικές βαθμίδες: η διαδικασία $N_i^D(t)$ ορίζεται σε χρονική σειρά, ενώ η στιγμή αποτυχίας X_i υπολογίζεται από την στιγμή της μεταμόσχευσης, S_i .

Για το υπό μελέτη νοσοκομειακό ίδρυμα, έστω ότι η ποσότητα $N^D(t) = \sum_{i \geq \mathcal{D}} N_i^D(t)$ είναι ο συνολικός αριθμός των αποτυχιών στη μεταμόσχευση που συνέβησαν μέχρι τη χρονική στιγμή t .

Η μελέτη έχει στηριχθεί πάνω στην υπόθεση αναλογικών κινδύνων για την εναλλακτική στην (4.1.), ωστόσο θα μπορούσαν να χρησιμοποιηθούν κι άλλα μοντέλα. Σημειώνουμε ότι η αναλογικότητα στη σχέση (4.1.) ισχύει μόνο για το διάστημα $[0, 1]$. Η απόδοση του CUSUM θα ποικίλει σε κάποιο βαθμό, αν η υπόθεση της αναλογικότητας δεν ακολουθείται με ακρίβεια.

4.2. Το διάγραμμα CUSUM σε συνεχή και διακριτό χρόνο

➤ Το Cusum σε Συνεχή Χρόνο

Ως συνήθως, το CUSUM που ορίζουμε, είναι βασισμένο στη μέθοδο διαδοχικών ελέγχων, SPRT, όπως την εισήγαγε ο Wald (1997). Θεωρούμε τη διαδικασία $N_D(t)$, $t > 0$, έτσι ώστε η πιθανοφάνεια σε οποιοδήποτε μικρό διάστημα $(t, t+dt]$ να λαμβάνει υπ' όψιν τις πληροφορίες για όλες τις μεταμοσχεύσεις, με κίνδυνο αποτυχίας τη χρονική στιγμή t .

Όπως αναφέρθηκε παραπάνω, έστω ότι έχουμε *risk-adjusted* τιμές για τη συνάρτηση κινδύνου, $a_i(t)$, για κάθε μεταμόσχευση i , με κίνδυνο αποτυχίας τη χρονική στιγμή $t > 0$. Για μια επιλεγμένη σταθερά $\theta > 0$, θεωρούμε ένα στατιστικό λόγου πιθανοφάνειας για τον έλεγχο της υπόθεσης $\mu = 0$ versus $\mu = \theta$.

Θεωρούμε το στατιστικό πιθανοφάνειας βάσει της ποσότητας $dN_i^D(t)$, με ένταση $e^\theta d\Lambda_i(t)$ versus $d\Lambda_i(t)$, το οποίο είναι:

$$\prod_{i \geq 1} \left[\frac{e^{\theta dN_i^D(t)} [1 - e^\theta d\Lambda_i(t)]^{1-dN_i^D(t)}}{[1 - d\Lambda_i(t)]^{1-dN_i^D(t)}} \right]$$

όπου χρησιμοποιήθηκε η σχέση (4.1.) και ένα μοντέλο Bernoulli για τη διαφορική προσαύξηση $dN_i^D(t)$. Ο λογάριθμος του λόγου πιθανοφάνειας μέχρι τη χρονική στιγμή t , δίνεται από τη σχέση:

$$U_t = \theta N^D(t) - (e^\theta - 1)A(t), \quad t > 0 \quad (4.2.)$$

όπου: $A(t) = \sum_{i \geq 1} \Lambda_i(t)$. Είναι εύκολο να παρατηρήσουμε ότι το $\min_{0 \leq s \leq t} U_s$ δεν αυξάνει

και αλλάζει ως προς την τιμή μόνο όταν $G_t = 0$. Στη σχέση (4.2.) το $(e^\theta - 1)$ μπορεί να ερμηνευθεί ως ο επιπλέον κίνδυνος υπό την εναλλακτική υπόθεση, ενώ το $A(t)$ ως το αναμενόμενο σωρευτικό ποσοστό κινδύνου υπό τη μηδενική.

Προκειμένου να ενημερώνουμε συνεχώς το CUSUM για κάθε χρονική στιγμή t , εργαζόμαστε με τις προσαυξήσεις dU_t του U_t , οι οποίες ορίζονται ως εξής:

$$dU_t = \theta dN^D(t) - (e^\theta - 1)dA(t) \quad (4.3.)$$

Είναι σημαντικό σε αυτό το σημείο να σημειώσουμε το διαχωρισμό που γίνεται από το διακριτό διωνυμικό CUSUM, όπου η μονάδα του χρόνου είναι μία μεμονωμένη μεταμόσχευση. Στην παρούσα μαθηματική διατύπωση, οι χρονικές προσαυξήσεις εκφράζουν τις δοκιμές Bernoulli και όχι μεμονωμένες μεταμοσχεύσεις. Αυτό μας δίνει το πλεονέκτημα να ενημερώνουμε το CUSUM, αμέσως μόλις συμβαίνει η αποτυχία. Αν θεωρήσουμε αντίστοιχα ένα διμερές CUSUM, θα υπολογίζουμε το CUCUM, με τη σχέση

$$U_t = \int_0^t dU_s \quad \text{και οι έντονες τάσεις (ανοδικές ή καθοδικές) θα μας παρέχουν}$$

προειδοποιητικό σήμα. Στις εφαρμογές όπου εστιάζουμε σε αυτό το κεφάλαιο, αποκλειστικά το μονομερές CUSUM το οποίο παρουσιάζει ανοδική τάση, βρίσκεται στο επίκεντρο της προσοχής μας. Αυτή είναι μία συνεχής χρονική επέκταση του συνήθους CUSUM σε διακριτό χρόνο, όπως ορίστηκε προηγουμένως. Το μονομερές CUSUM συμβολίζεται ως $\{G_t\}$, όπου $G_0 = 0$ και $G_{t+dt} = \max(0, G_t + dU_t)$, για $t > 0$.

Ισοδύναμα, μπορούμε να έχουμε και την ακόλουθη διατύπωση:

$$G_t = U_t - \min_{0 \leq s \leq t} U_s, \quad t > 0$$

Η διαδικασία U_t μπορεί επίσης να εκληφθεί ως μία συνεχής επέκταση του τυχαίου περιπάτου (random walk) στο διακριτό χρόνο. Στο διακριτό χρόνο, η τάση ορίζεται ως το μέσο μέγεθος βήματος, ενώ στο συνεχή χρόνο ως ο μέσος ρυθμός αλλαγών στη διαδικασία U_t .

Για να εφαρμόσουμε τη διαδικασία CUSUM, καταγράφουμε G_t σε συνεχή χρόνο, μέχρι την στιγμή που θα διασχίσει ένα καθορισμένο άνω φράγμα στο $h > 0$. Την στιγμή που θα συμβεί αυτό, θα δοθεί ένα προειδοποιητικό σήμα. Η τιμή του h ορίζεται με τέτοιο τρόπο, ώστε να προσδώσει ορισμένες κατάλληλες ιδιότητες στο σύστημα παρακολούθησης. Έστω ότι η χρονική στιγμή μέχρι να φθάσει στο όριο στο h , είναι:

$$\tau_h = \inf \{t > 0: G_t \geq h\}$$

Στην ορολογία του CUSUM, η ποσότητα $E(\tau_h)$ αναφέρεται ως το μέσο μήκος ροής (average run length – ARL) και διαδραματίζει σημαντικό ρόλο στο σχεδιασμό τιμών για το CUSUM. Είναι πιθανό να εξασφαλιστεί μία θεωρητική προσέγγιση για το $E(\tau_h)$, υπό ορισμένες προϋποθέσεις για τη διαδικασία $N^D(t)$. Ενώ είναι πιθανό να εκτιμήσουμε την $E(\tau_h)$ με χρήση μεθόδων προσομοίωσης, θα ήταν χρήσιμος ένας εύκολα υπολογίσιμος τύπος για το $E(\tau_h)$ ως προς το σχεδιασμό των παραμέτρων του CUSUM και του ορίου

ελέγχου. Ανεξαρτήτως του θεωρητικού του ενδιαφέροντος, μία τόσο γρήγορη εκτίμηση μπορεί επίσης να χρησιμοποιηθεί για την επιλογή κατάλληλων τιμών για το h .

➤ Το Cusum σε Διακριτό Χρόνο

Τα διαγράμματα CUSUM σε διακριτό χρόνο, βασίζονται επίσης στη μέθοδο SPRT, αλλά προσδιορίζονται αναφορικά με τις επιτυχείς μεταμοσχεύσεις S_1, S_2, S_3, \dots ανά χρονική στιγμή. Ο δείκτης $i = 1, 2, 3, \dots$ αντιστοιχεί στην i -οστή μεταμόσχευση. Κάθε μεταμόσχευση δημιουργεί κι ένα πείραμα Bernoulli, όπου η επιτυχία ή αποτυχία εξαρτάται από το εάν η μεταμόσχευση είναι ή όχι βιώσιμη για ένα έτος. Για να ορίσουμε το CUSUM σε συνεχή χρόνο, χρησιμοποιήσαμε ένα στατιστικό λόγου πιθανοφάνειας βάσει ενός υποτιθέμενου κινδύνου ο οποίος συνέδεσε το ποσοστό αποτυχίας στο νοσοκομείο με αυτόν στον πληθυσμό. Μία παρόμοια προσέγγιση θα ακολουθήσουμε και στο διακριτό χρόνο, όπου θα χρησιμοποιηθούν ποσοστά πιθανοτήτων, ή διακριτός σχετικός κίνδυνος.

Η σημειογραφία που χρησιμοποιούμε είναι παρόμοια με αυτή που χρησιμοποιήσαμε και πριν. Ορίζουμε λοιπόν:

$$\xi_i = I(X_i \leq I), \quad i = 1, 2, \dots$$

Έστω $\pi_{i0} = E(\xi_i = 1 / Z_i)$ είναι η πιθανότητα να συμβεί, μέσα σε ένα έτος, αποτυχία στον πληθυσμό. Υποθέτουμε ότι οι δεδομένες μεταβλητές Z_i είναι γνωστές ή έχουν εκτιμηθεί με ακρίβεια από ένα μεγάλο δείγμα. Όπως πριν, υποθέτουμε ότι τα Z_1, Z_2, \dots στο υπό μελέτη νοσοκομειακό ίδρυμα, είναι ανεξάρτητες και τυχαία κατανομημένες τυχαίες μεταβλητές.

Ελέγχοντας την αναλογία πιθανοτήτων

Καταρχήν θεωρούμε ένα μοντέλο αναλογίας πιθανοτήτων:

$$\text{logit}\pi_{i1} = \text{logit}\pi_{i0} + \text{logOR} \quad (4.4)$$

όπου η αναλογία πιθανοτήτων (*odds ratio*), **OR**, υπολογίζει τη διαφορά του ποσοστού θνησιμότητας ανάμεσα στο νοσοκομειακό ίδρυμα που μελετούμε και στη χώρα γενικότερα. Θεωρούμε έλεγχο αναλογίας πιθανοφάνειας (*likelihood ratio test*) για τη μηδενική υπόθεση $H_0: OR = 1$ versus $H_0: OR = OR_A > 1$.

Έχουμε λοιπόν τη λογαριθμική πιθανοφάνεια, με τη συμβολή του ξ_i , η οποία εκφράζεται με την ακόλουθη σχέση:

$$g_i = \log \left[\left(\frac{\pi_{i1}}{\pi_{i0}} \right)^{\xi_i} \left(\frac{1 - \pi_{i1}}{1 - \pi_{i0}} \right)^{1 - \xi_i} \right] = \xi_i (\alpha_i + b_i) - b_i$$

όπου: $\alpha_i = \log(OR_A) - \log(1 - \pi_{i0} + \pi_{i0} \cdot OR_A)$ και $b_i = \log(1 - \pi_{i0} + \pi_{i0} \cdot OR_A)$.

Το αντίστοιχο CUSUM ορίζεται ως εξής: $G_0 = 0$ και $G_{i+1} = \max(0, G_i + g_i)$, $i = 1, 2, \dots$

Για επιλεχθείσα τιμή του h , η διαδικασία παράγει ένα σήμα σε χρόνο κ_h όπου:

$$\kappa_h = \min\{k \geq 1: G_k \geq h\}$$

Εδώ, το ARL ορίζεται ως $E(\kappa_h)$ που είναι ο μέσος αριθμός των μεταμοσχεύσεων που πραγματοποιήθηκαν και αξιολογήθηκαν πριν παραχθεί ένα σήμα. Εδώ εξετάζονται οι ιδιότητες του CUSUM για το υπό μελέτη ίδρυμα, υπό την υπόθεση ότι η πραγματική αναλογία πιθανοτήτων είναι $OR = e^\mu$. Για την περίπτωση αυτή, έχουμε ανεξάρτητες «υλοποιήσεις» των $\{\xi_i\}_{i \geq 1}$ με μέση τιμή $\rho_\mu = E(\xi_i / OR = e^\mu)$.

Λόγω του ότι $\alpha_i + b_i = \log(OR_A)$, το g_i έχει την ίδια δομή με τη διαφορική προσαύξηση dU_t της (4.3) για το CUSUM (συνεχούς χρόνου). Η μεταβλητή Bernoulli ξ_i είναι διακριτή ποσότητα αντίστοιχη του $dN^D(t)$ και το $\log(OR_A)$ διαδραματίζει τον ίδιο ρόλο με το σχετικό κίνδυνο θ της σχέσης (4.3).

Συμπεραίνουμε ότι η τάση ζ ισούται με:

$$\zeta = E(g_i) = E[\xi_i \cdot \log(OR_A) - b_i] = \rho_\mu \cdot \log(OR_A) - E(b_i)$$

Ελέγχοντας το σχετικό κίνδυνο

Ας θεωρήσουμε τώρα ένα μοντέλο διακριτού σχετικού κινδύνου (*relative risk – RR*):

$$\text{logit}\pi_{i1} = \text{logit}\pi_{i0} + \text{logRR} \quad (4.5)$$

και τον έλεγχο της υπόθεσης $H_0: RR = 1$ versus $H_0: RR = RR_A > 1$.

Έχουμε λοιπόν τη λογαριθμική πιθανοφάνεια, με τη συμβολή του ξ_i , η οποία εκφράζεται με την ακόλουθη σχέση:

$$g_i = \log \left[\left(\frac{\pi_{i1}}{\pi_{i0}} \right)^{\xi_i} \left(\frac{1 - \pi_{i1}}{1 - \pi_{i0}} \right)^{1 - \xi_i} \right] = \xi_i (\alpha_i + b_i) - b_i$$

όπου: $\alpha_i = \alpha = \log(RR_A)$ και $b_i = \log(1 - \pi_{i0}) - \log(1 - RR_A \cdot \pi_{i0})$.

Εδώ ομοίως με πριν, θεωρούνται ανεξάρτητες «υλοποιήσεις» των $\{\xi_i\}_{i \geq 1}$ υπό την υπόθεση ότι ο πραγματικός σχετικός κίνδυνος για το ίδρυμα που μελετούμε, είναι $RR = e^\mu$.

Συμπεραίνουμε ότι η τάση ζ ισούται με: $\zeta = E(g_i) = E[(\alpha_i + b_i) \cdot F_\mu(1/Z_i) - b_i]$, όπου οι τυχαίες μεταβλητές Z_i ακολουθούν την Κανονική Κατανομή.

4.3. Έλεγχος και αξιολόγηση της μεθόδου, με χρήση προσομοιώσεων

Στις προσομοιώσεις, θεωρήσαμε ένα νοσοκομειακό ίδρυμα, όπου οι μεταμοσχεύσεις συνέβησαν σύμφωνα με μία ομογενή διαδικασία Poisson με ρυθμό $\psi = 100$ μεταμοσχεύσεις/έτος. Έστω ότι ο προ-μεταμοσχευτικός χρόνος αποτυχίας (σε εθνικό ή πληθυσμιακό επίπεδο) ακολουθεί την εκθετική κατανομή με ρυθμό λ_0 και επιλέχθηκε η πιθανότητα να έχουμε αποτυχία της διαδικασίας εντός ενός έτους, να ισούται με:

$$1 - e^{-\lambda_0} = 10\% \quad (4.6.)$$

Κατά την αξιολόγηση των διαδικασιών, υποτέθηκε ότι στο συγκεκριμένο ίδρυμα ο προμεταμοσχευτικός χρόνος αποτυχίας X_i , έχει κίνδυνο $e^\mu \lambda_0$. Έγινε λοιπόν σύγκριση ως προς την απόδοση των διαδικασιών CUSUM που συζητήθηκαν προηγουμένως, σε σχέση με το ARL, και επίσης αξιολογήθηκε η ακρίβεια των προσεγγίσεων που λήφθηκαν.

Στο συνεχές CUSUM, επιλέχθηκε $\theta = \log 2.0$. Ομοίως, στα CUSUMs διακριτού χρόνου των σχέσεων (4.4.) και (4.5.), επιλέχθηκαν $OR_A = 2.0$ και $RR_A = 2.0$, αντίστοιχα, εξασφαλίζοντας έτσι ότι και οι τρεις διαδικασίες CUSUM ελέγχουν παρόμοιες υποθέσεις.

Από το σχεδιασμό του διακριτού CUSUM, το αποτέλεσμα (αποτυχία ή επιτυχία) μίας μεταμόσχευσης δεν λαμβάνεται υπ' όψιν πριν την πάροδο ενός έτους από τότε που πραγματοποιήθηκε. Επομένως, στην πραγματικότητα το CUSUM έχει αξιολογηθεί και είναι ικανό να στείλει ένα σήμα σε χρονικές στιγμές $\{S_1 + 1, S_2 + 1, \dots\}$. Για αυτό λοιπόν, προστέθηκε ένα έτος στα ARLs βάσει των χρονικών στιγμών $\{S_1, S_2, \dots\}$. Αντιστοίχως, οι θεωρητικές προσεγγίσεις που έγιναν στην ενότητα «*Το Cusum σε Συνεχή Χρόνο*», αυξήθηκαν επίσης κατά 1.

Παραδείγματα του CUSUM συνεχούς χρόνου, υπό των εντάσεων α_i και $e^\theta \alpha_i$, απεικονίζονται στο διάγραμμα *Figure 1*.

Για τις προσομοιώσεις, τα δεδομένα παράχθηκαν από τη χρονική στιγμή $t = 0$ και δεδομένου ότι οι αποτυχίες πρέπει να είναι εντός ενός έτους, οι διαδικασίες CUSUM είναι σε ισορροπία από τη χρονική στιγμή $t = 1$. Ως εκ τούτου, ξεκίνησε η παρακολούθηση όλων των CUSUMs από τη χρονική στιγμή $t = 1$ μέχρι την παραγωγή ενός σήματος (σε περίπτωση παραβίασης ενός ορίου ελέγχου, $h > 0$). Η επιλογή του ορίου ελέγχου για καθένα από τα τρία CUSUMs συντελέστηκε με τέτοιον τρόπο ώστε να εξασφαλιστεί ότι και τα τρία θα έχουν το ίδιο ARL όταν $\mu = 0$. Αν και θα μπορούσε κάποιος να χρησιμοποιήσει διαφορετικά κριτήρια για την επιλογή των ορίων ελέγχου, το συγκεκριμένο κριτήριο είναι απαραίτητο για μία σαφή σύγκριση του ARL μεταξύ διαφορετικών διαγραμμάτων CUSUM.

Έτσι, όταν ένα νοσοκομειακό ίδρυμα επιτελεί έργο σε εθνικό μέσο όρο, επελέγη το h με σκοπό να επιτευχθεί ARL περίπου ίσο με 30 χρόνια, το οποίο αντιστοιχούσε σε $h = 4.35$, 4.19 και 4.47 για τα συνεχή, τα διακριτά λογιστικά και τα διακριτά-σχετικού κινδύνου-CUSUMs, αντίστοιχα. Στη συνέχεια, θα δούμε ότι η συγκεκριμένη επιλογή για το ARL οδηγεί επίσης σε ορισμένες ενδιαφέρουσες ιδιότητες για το συνεχές CUSUM.

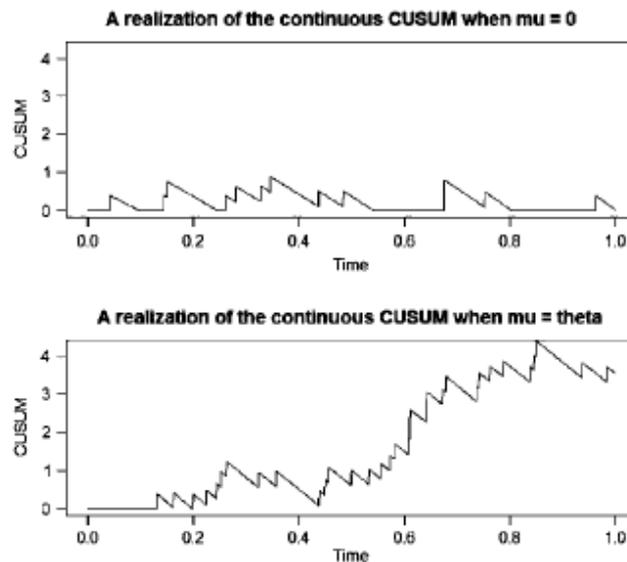


Figure 1. Realizations of the continuous CUSUM under intensities $\alpha_i(\cdot)$ and $e^\theta \alpha_i(\cdot)$.

Ο πίνακας *Table I* δείχνει τα αποτελέσματα 1.000 επαναλήψεων για διάφορες τιμές του μ , για καθένα από τα τρία CUSUMs. Η δεύτερη, τρίτη και πέμπτη στήλη δίνουν το δειγματικό μέσο, τη διάμεσο και την τυπική απόκλιση, αντίστοιχα. Η τέταρτη στήλη δίνει την προσεγγιστική τιμή για το ARL, βάσει των σχέσεων που διατυπώσαμε στις ενότητες «Σημειογραφία και Υποθέσεις» και «*To Cusum σε Συνεχή Χρόνο*». Υπό τη γενικότερη έννοια, η θεωρητική προσέγγιση φαίνεται να λειτουργεί ικανοποιητικά όταν η διαδικασία τυχαίου περιπάτου U_t έχει μία θετική τάση ($e^\mu > 1.22$).

Table I. ARLs for the three CUSUM procedures at different choices of facility relative risk e^μ .

e^μ	Average τ_h	Median τ_h	$E(\tau_h)$	$\widehat{SD}(\tau_h)$
<i>Continuous-time CUSUM</i> ($h=4.35, e^\theta=2.0$)				
0.69	518.51	428.52	410.32	394.26
0.82	142.00	91.62	105.12	143.94
1.00	29.92	20.92	23.62	29.13
1.22	8.07	5.76	6.59	7.66
1.50	2.72	2.08	2.43	2.34
1.73	1.50	1.22	1.45	1.09
2.00	0.98	0.83	0.97	0.62
<i>Discrete-time CUSUM, OR</i> ($h=4.19, OR_A=2.0$)				
0.69	469.09	347.08	499.45	383.53
0.82	146.76	102.00	119.40	142.70
1.00	29.75	21.38	26.24	27.79
1.22	8.63	6.20	7.62	7.23
1.50	3.75	3.15	3.43	2.19
1.73	2.48	2.19	2.46	1.06
2.00	1.97	1.86	1.98	0.58
<i>Discrete-time CUSUM, RR</i> ($h=4.47, RR_A=2.0$)				
0.69	466.53	317.22	372.19	469.81
0.82	119.19	85.24	107.66	114.64
1.00	29.64	22.18	24.83	27.44
1.22	8.86	6.71	7.83	7.41
1.50	3.85	3.06	3.51	2.49
1.73	2.52	2.19	2.47	1.17
2.00	2.02	1.84	1.97	0.70

Συγκριτικά με τα CUSUMs διακριτού χρόνου, το CUSUM συνεχούς χρόνου δίνει μικρότερα μήκη ροής όταν $e^\mu > 1$ (το νοσοκομείο αποδίδει χειρότερα σε σχέση με τον εθνικό μέσο όρο) και μεγαλύτερα μήκη ροής όταν $e^\mu < 1$. Τα δύο CUSUMs διακριτού χρόνου παρουσιάζουν παρόμοια απόδοση για υψηλές τιμές του σχετικού κινδύνου ή της αναλογίας πιθανοτήτων. Η μονάδα του χρόνου σε όλες αυτές τις προσομοιώσεις είναι τα έτη: συνεπώς η τιμή $ARL = 30$, σημαίνει ότι θα περιμένουμε κατά μέσο όρο 30 χρόνια για να παραχθεί ένα σήμα. Δεδομένου ότι πραγματοποιούνται κατά μέσο όρο 100 μεταμοσχεύσεις/έτος, αυτό αντιστοιχεί σε μέσο όρο 3000 μεταμοσχεύσεων, μέχρις ότου παραχθεί

το πρώτο σήμα. Μπορεί εδώ να σημειωθεί ότι ο αριθμός των μεταμοσχεύσεων μέχρι την παραγωγή του σήματος, παραμένει αναλλοίωτος. Αυτό σημαίνει ότι αν το ποσοστό των μεταμοσχεύσεων, ψ , υποδιπλασιάζεται ετησίως, τότε το ARL (όπως βρίσκεται από τις προσομοιώσεις και τις σχέσεις της ενότητας 4.1.) θα διπλασιαστεί, διατηρώντας ωστόσο σταθερό τον αριθμό των μεταμοσχεύσεων μέχρι την παραγωγή σήματος.

Δεδομένου ότι ξεκινά η παρακολούθηση όλων των CUSUMs ταυτόχρονα και για τα διακριτά CUSUMs μελετάται η έκβαση της μεταμόσχευσης ένα χρόνο μετά, συνεπώς είναι δίκαιη η σύγκριση μεταξύ συνεχούς και διακριτής μεθοδολογίας υπό την έννοια ότι χρησιμοποιείται η ίδια ποσότητα δεδομένων. Όλα τα CUSUMs υπολογίζουν τις αποτυχίες εντός του πρώτου έτους, ωστόσο ενώ το συνεχές CUSUM λογαριάζει την αποτυχία εγκαίρως, ενώ το διακριτό CUSUM ανταποκρίνεται στην αποτυχία αφού περάσει ένα έτος από τη διεξαγωγή της μεταμόσχευσης.

Στον πίνακα *Table II*, παρουσιάζεται μία μελέτη προσομοίωσης που δείχνει πώς εμπειρικά ARLs συγκρίνονται με τις θεωρητικές προσεγγίσεις των κεφαλαίων 4.1. και 4.2., για διαφορετικές επιλογές του ορίου ελέγχου h και της παραμέτρου μ . Μπορεί λοιπόν κάποιος να διαπιστώσει πώς η ακρίβεια των θεωρητικών προσεγγίσεων αυξάνει καθώς αυξάνουν το μ και το h . Στην περίπτωση του συνεχούς CUSUM, το ποσοστό της διαφοράς ανάμεσα στις εμπειρικές και θεωρητικές εκτιμήσεις, κυμαίνεται από 11% για $h = 3.0$ έως 3% για $h = 6.0$ (για $\mu = 0.0$). Σε αυτό ίσως να επέδρασε το σφάλμα που δημιουργήθηκε από υπέρβαση των ορίων των προσεγγίσεων του Wald, το οποίο μειώνεται με την αύξηση του h .

Table II. Empirical and theoretical ARLs for discrete (OR) and continuous CUSUMs for different choices of control limit h and facility relative risk e^μ .

e^μ	$h=3.0$		$h=4.0$		$h=5.0$		$h=6.0$	
	ARL	$E(\tau_h)$	ARL	$E(\tau_h)$	ARL	$E(\tau_h)$	ARL	$E(\tau_h)$
<i>Continuous-time CUSUM, $e^\theta = 2.0$</i>								
1.00	19.07	17.16	59.21	52.69	161.60	151.12	407.81	420.48
1.22	5.21	4.81	9.08	8.63	13.93	13.65	20.63	19.93
1.50	2.11	2.02	3.03	2.94	4.04	3.88	4.79	4.82
1.73	1.31	1.31	1.81	1.82	2.33	2.34	2.93	2.85
2.00	0.93	0.94	1.25	1.28	1.61	1.62	1.97	1.96
<i>Discrete (OR) CUSUM, $OR_A = 2.0$</i>								
1.00	24.74	21.18	67.40	62.86	174.36	178.25	477.63	493.99
1.22	6.81	6.23	10.74	10.14	16.04	15.11	21.91	21.12
1.50	3.15	3.15	4.08	4.23	4.93	5.03	6.00	5.99
1.73	2.41	2.39	2.91	2.92	3.50	3.45	3.96	3.98
2.00	1.98	2.00	2.34	2.35	2.68	2.71	3.12	3.06

Είναι συνεπώς ενδιαφέρον να σημειωθεί ότι σε περίπτωση που κάποιος ήθελε να σχεδιάσει τα CUSUMs, προς επίτευξη ορισμένων ARLs για $e^u > 1$, οι σχέσεις στις ενότητες 4.1. και 4.2. είναι ακριβείς και θα μπορούσαν να χρησιμοποιηθούν για την εύρεση ενός κατάλληλου h .

Οι τιμές του ARL για τις οποίες γίνεται αναφορά, εξαρτώνται επίσης σε μεγάλο βαθμό από το λ_0 της σχέσης (4.6.), σε αντιστοιχία με το ποσοστό αποτυχίας της τάξης του 10% σε εθνικό επίπεδο.

Σε αυτό το σημείο, στη θεωρία προσομοίωσης γίνεται η υπόθεση ότι $\beta = 0$ και ότι η κατανομή των τυχαίων μεταβλητών στο νοσοκομειακό ίδρυμα είναι ίδια με αυτήν του πληθυσμού. Εκτελέστηκε λοιπόν μελέτη προσομοίωσης με $\beta = 0.2$. και με την υπόθεση ότι $Z_i \sim N(0,1)$ ανεξάρτητες τυχαίες μεταβλητές στον πληθυσμό.

Η τιμή του λ_0 επιλέχθηκε έτσι ώστε να ικανοποιείται η ακόλουθη σχέση:

$$E \left[1 - e^{-\lambda_0 e^{\beta Z_i}} \right] = 10\% \quad (4.7.)$$

Στο νοσοκομειακό ίδρυμα που μελετούμε, υποθέσαμε ότι $Z_i \sim N(\nu, \rho^2)$ και υπολογίσαμε εμπειρικά τα ARLs για διάφορες τιμές του μ , με το συνεχές CUSUM στο $h = 2.30$.

Ο πίνακας *Table III* απεικονίζει τα αποτελέσματα. Η δεύτερη στήλη δείχνει τις τιμές της ποσότητας $E_{\nu, \rho} [F_{\mu}(1/Z_i)]$. Μπορεί να δειχθεί ότι ισχύει το εξής:

$$E_{\nu, \rho} [F_{\mu}(1/Z_i)] \approx F_{\nu, \rho} \times E_{0,1} [F_{\mu}(1/Z_i)], \text{ όπου } F_{\nu, \rho} = \exp\left\{\nu + \frac{1}{2}\beta^2(\rho^2 - 1)\right\}.$$

Μέσω της σχέσης (4.7.), θέτοντας $\mu = 0$, προκύπτει ότι: $E_{\nu, \rho} [F_{\mu}(1/Z_i)] \approx F_{\nu, \rho} \times 0.10$.

Η τρίτη στήλη δείχνει τη συγκεκριμένη προσέγγιση για $\mu = 0$ και παρατηρούμε ότι από-δίδει ικανοποιητικώς. Η πρώτη γραμμή στον πίνακα *Table III* απεικονίζει την περίπτωση όπου οι μεταβλητές στο νοσοκομειακό ίδρυμα που μελετούμε, έχουν την ίδια κατανομή με αυτήν του πληθυσμού. Οι επόμενες γραμμές δείχνουν μετατροπές στα εμπειρικά ARLs, όσον αφορά το μέρος και τις παραμέτρους κλίμακας, για την κατανομή των τυχαίων μεταβλητών του συγκεκριμένου νοσοκομείου. Είναι προφανές ότι η επίδραση του μέρους ν μπορεί να αποδειχθεί σημαντικότερη από την κλίμακα ρ , ως προς τη διαμόρφωση των ARLs. Μπορεί να σημειωθεί εδώ ότι η $F_{\mu}(1/Z_i)$ είναι η πιθανότητα αποτυχίας της μεταμόσχευσης εντός έτους από τη διεξαγωγή της.

Κατά συνέπεια, μετατοπίσεις στην κατανομή των τυχαίων μεταβλητών πρέπει να αξιολογούνται μόνο μέχρι την στιγμή της μετατόπισης στην προαναφερθείσα πιθανότητα.

Η σταθερά $F_{v,\rho}$ μπορεί να θεωρηθεί ως μία «γρήγορη» εκτίμηση της ποσότητας $E_{v,\rho}[F_{\mu}(1/Z_i)]$, όπως επίσης και ο λόγος $1/F_{v,\rho}$ μπορεί να ερμηνευθεί ως ο προσεγγιστικός παράγοντας με τον οποίον τα θεωρητικά ARLs στο νοσοκομειακό ίδρυμα αυξάνονται ή μειώνονται σε σχέση με τον πληθυσμό. Το σημαντικό εδώ που πρέπει να τονιστεί είναι ότι ο παράγοντας αυτός είναι ανεξάρτητος της παραμέτρου μ , όπου μ εκφράζει την επίδραση του αποτελέσματος της μετα-μόσχευσης.

Table III. Variation in ARLs for the continuous CUSUM across facility covariate distribution parameters at $h=2.30$, $e^{\theta}=2.0$.

(v, ρ^2)	$E_{v,\rho}[F_{\mu=0}(1 Z_i)]$	$F_{v,\rho} \times 0.10$	e^{θ}				
			1.00	1.22	1.50	1.73	2.00
(0, 1)	0.100	0.100	8.21	3.04	1.44	0.98	0.69
(0.5, 1)	0.110	0.110	7.80	2.73	1.27	0.90	0.65
(1, 1)	0.120	0.122	7.22	2.50	1.22	0.80	0.58
(0, 0.25)	0.098	0.098	8.53	3.01	1.47	0.96	0.72
(0.5, 0.25)	0.108	0.109	7.58	2.81	1.35	0.89	0.64
(1, 0.25)	0.119	0.120	7.15	2.71	1.21	0.78	0.59

4.3.1. Ελέγχοντας το ποσοστό των εσφαλμένων «συναγερμών»

Μέχρι τώρα, έχουμε συζητήσει το ARL μόνο ως κριτήριο απόδοσης μιας διαδικασίας. Σε αυτό το κεφάλαιο θα συζητήσουμε για τα ποσοστά αληθών και ψευδών συναγερμών για τις CUSUM διαδικασίες και θα αξιολογήσουμε την απόδοσή τους. Για τον σκοπό αυτό, θεωρούμε $\mu = 0$ που αντιστοιχεί σε εντός ελέγχου κατάσταση και $\mu = \theta$ για την αντίστοιχη εκτός ελέγχου. Ομοίως, για το διακριτό CUSUM θέτουμε $OR[RR]=1$ και $OR[RR]=OR_A[RR_A]$, για τις καταστάσεις εντός και εκτός ελέγχου, αντίστοιχα.

Επιλέγοντας ένα μικρό όριο ελέγχου θα έχει ως συνέπεια το CUSUM να παράγει σήμα αρκετά συχνά, αυξάνοντας έτσι την πιθανότητα λανθασμένων συναγερμών. Αντιθέτως, θέτοντας αρκετά ψηλά το όριο ελέγχου, θα οδηγήσει σε μεγάλη αναμονή μέχρι την εκπομπή ενός σήματος, οδηγώντας έτσι σε καθυστερημένη ανίχνευση μιας -εκτός ελέγχου- κατάστασης. Ως εκ τούτου, η επιλογή ενός ορίου ελέγχου είναι μία «συναλλαγή» ανάμεσα στο ποσοστό των εσφαλμένων συναγερμών και στο αντίστοιχο των αληθών.

Είναι προφανές ότι πρέπει να κρατάμε χαμηλό το ποσοστό των εσφαλμένων συναγερμών, σε αντίθεση με αυτό των αληθών.

Για διαδικασίες στατιστικού ελέγχου διεργασιών, όπως είναι το CUSUM, το ποσοστό σφαλμάτων Τύπου-I εξαρτάται από το πόσο χρόνο η διεργασία βρίσκεται υπό έλεγχο και ισοδύναμα από το μέγεθος του δείγματος. Καθώς αυξάνεται ο χρόνος, η πιθανότητα σφάλματος Τύπου-I θα προσεγγίσει τη μονάδα, μόλις το CUSUM διασχίσει ένα όριο ελέγχου. Για αυτόν το λόγο, είναι χρήσιμο να θεωρούμε ποσότητες, όπως τα εντός ελέγχου και εκτός ελέγχου ARLs. Στα επόμενα, θα συζητήσουμε τα αποτελέσματα προσομοίωσης δείχνοντας όχι μόνο την απόκλιση στα εντός και εκτός ελέγχου ARLs, αλλά και εκτιμήσεις του ποσοστού των ψευδών και αληθών συναγερμών. Λόγω του ότι το διακριτό (RR) CUSUM παρουσιάζει πανομοιότυπη απόδοση με το διακριτό (OR) CUSUM, θα μελετήσουμε μόνο το συνεχές και το διακριτό (OR) CUSUM.

Υπενθυμίζουμε ότι ελέγχουμε την υπόθεση $H_0: \mu = 0$ έναντι της $H_1: \mu = \theta$. Θέτουμε το όριο ελέγχου, h , και ορίζουμε ως ARL_0^h το εντός ελέγχου ARL (υπό την H_0) και ως ARL_1^h το αντίστοιχο εκτός ελέγχου (υπό την H_1). Έστω ένα καθορισμένο χρονικό διάστημα απόφασης $[0, T]$ και ορίζουμε το σφάλμα Τύπου-I μέχρι τη χρονική στιγμή T , ως το ποσοστό των εσφαλμένων «συναγερμών» (υπό την H_0) μέχρι τη χρονική στιγμή T . Εκφράζει με άλλα λόγια, το ποσοστό των στιγμών όπου $ARL_0^h < T$. Ορίζουμε επίσης την ισχύ (*power*) μέχρι τη χρονική στιγμή T , ως το ποσοστό των επιτυχών συναγερμών/ανιχνεύσεων (υπό την H_1) μέχρι τη χρονική στιγμή T . Είναι δηλαδή, αντίστοιχα με πριν, το ποσοστό των στιγμών όπου $ARL_1^h < T$. Συνεπώς, για να ορίσουμε ένα όριο ελέγχου h βάσει του ότι το σφάλμα Τύπου-I δε θα ξεπερνάει μία καθορισμένη τιμή α ($0 < \alpha < 1$), θα πρέπει να επιλέξουμε το h έτσι ώστε η κατώτερη τιμή α της κατανομής του ARL_0^h να είναι το πολύ T .

Στον πίνακα *Table IV*, της επόμενης σελίδας, παρουσιάζονται τα αποτελέσματα από μία μελέτη προσομοίωσης η οποία ορίζει τα όρια ελέγχου για δύο CUSUMs, χρησιμοποιώντας την προαναφερθείσα διαδικασία όπου $T = 5$ έτη και το σφάλμα Τύπου I να έχει μία τιμή περίπου $\alpha = 15\%$. Η τιμή $\alpha = 15\%$ προκύπτει από το ότι οι διαδικασίες παρακολούθησης χρησιμοποιούνται πλέον στα νοσοκομεία, όπου ο έλεγχος σημαντικότητας για την H_0 (σε επίπεδο 5%) διεξάγεται κάθε 6 μήνες, βάσει δεδομένων χρονικού εύρους 2,5 ετών. Με αυτή τη διαδικασία παρακολούθησης, η πιθανότητα να λάβουμε ένα σήμα, υπό τη μηδενική υπόθεση, είναι περίπου 10-15% , σε μία περίοδο 5 ετών. Το απαιτούμενο επίπεδο h εξαρτάται από το ποσοστό των μεταμοσχεύσεων, ψ , στο υπό μελέτη κέντρο.

Θεωρήσαμε ότι $\psi = 10, 25, 40, 60, 80$ και 100 μεταμοσχεύσεις/ έτος, όπως επίσης και τις ίδιες υποθέσεις για κάθε μεταμόσχευση και την κατανομή του χρόνου αποτυχίας, που θεωρήσαμε και στην ενότητα 4.3. Ομοίως και εδώ, θεωρείται ότι $e^\theta = 2.0$. όπως σημειώθηκε και προηγουμένως, έγινε η παρακολούθηση όλων των CUSUMs υπό κατάσταση ισορροπίας.

Από τον πίνακα *Table IV*, παρατηρούμε ότι τα όρια ελέγχου είναι κάπως υψηλότερα για το συνεχές CUSUM και αποδίδει ελάχιστα υψηλότερη ισχύ για το σφάλμα Τύπου I, σε σχέση με το διακριτό (OR) CUSUM. Το όριο ελέγχου που απαιτείται, ώστε να επιτύχουμε σφάλμα Τύπου-I ίσο με 15%, αυξάνει καθώς αυξάνει το σύνολο των μεταμοσχεύσεων (ανά έτος). Τα αποτελέσματα αυτά εξήχθησαν βάσει καθορισμένης περιόδου $T = 5$ έτη. Αν το T μειωθεί (αυξηθεί) τότε μία μικρότερη (μεγαλύτερη) τιμή του h θα επιφέρει ένα παρόμοιο σφάλμα Τύπου-I, αποδίδοντας έτσι χαμηλότερη (υψηλότερη) ισχύ.

Table IV. Variation in h , Type-I error and power across facility size for the continuous and the discrete (OR) CUSUMs.

Facility size (transplants per year)	Continuous CUSUM ($e^\theta=2.0$)			Discrete (OR) CUSUM ($OR_A=2.0$)		
	h	Type-I error	Power	h	Type-I error	Power
10	2.25	0.153	0.761	1.80	0.151	0.664
25	3.00	0.155	0.922	2.50	0.148	0.898
40	3.45	0.151	0.963	2.96	0.151	0.954
60	3.81	0.149	0.990	3.36	0.153	0.984
80	4.20	0.154	0.998	3.65	0.152	0.997
100	4.35	0.154	1.000	3.87	0.155	0.998

Όπως παρατηρείται στη συγκεκριμένη εφαρμογή, τα διαγράμματα CUSUM χρησιμοποιούνται σε ορισμένες περιπτώσεις για την παρακολούθηση της απόδοσης πολλαπλών κέντρων, καθιστώντας έτσι αναγκαίο τον επαναλαμβανόμενο έλεγχο της ίδιας υπόθεσης. Ας θεωρήσουμε M , το πλήθος, διαφορετικά κέντρα και ας υποθέσουμε ότι $0 < m_0 < M$ κέντρα είναι εντός ελέγχου και τα υπόλοιπα $M - m_0 = m_1$ κέντρα είναι εκτός ελέγχου. Το CUSUM εφαρμόζεται σε κάθε κέντρο, με απώτερο στόχο της διαδικασίας να μελετηθεί το συνολικό ποσοστό των εσφαλμένων συναγερωμών. Στην πραγματικότητα, αυτό σχετίζεται με τα ποσοστά εσφαλμένων ανιχνεύσεων (*false discovery rates – FDRs*). Με τον όρο FDR, εννοείται η αναλογία των εσφαλμένων ανιχνεύσεων (ή συναγερωμών) ως προς το σύνολο των ανιχνεύσεων γενικά. Η ιδέα να χρησιμοποιηθεί η FDR στο πεδίο των πολλαπλών διαγραμμάτων CUSUM, εισήχθη αρχικά από τους Marshall *et al.*, Aylin *et al.*, Grigg & Spiegelhalter και από πολλούς ακόμη.

Το ποσοστό των εσφαλμένων ανιχνεύσεων, FDR , μέχρι τη χρονική στιγμή T , ορίζεται ακολούθως ως εξής:

$$FDR_T^h = \frac{q_0[\text{Type-I error}]}{q_0[\text{Type-I error}] + (1 - q_0)[\text{Power}]}$$

όπου $q_0 = m_0/M$ είναι το ποσοστό (ή αναλογία) των κέντρων που βρίσκονται εντός ελέγχου. Διεξήχθησαν μελέτες προσομοίωσης, όπου θεωρήθηκε $M = 100$ κέντρα με $m_0 \sim B(M, p_0)$, $T = 5$ έτη. Επίσης, θεωρήθηκαν πέντε διαφορετικές τιμές ψ , όπου $\psi = 10, 40, 60, 100$ μεταμοσχεύσεις/ έτος και διαφορετικές τιμές για το p_0 , δηλαδή $p_0 = 0.7, 0.8$ και 0.9 . Τέλος, βάσει των ορίων ελέγχου που παρουσιάζονται στον πίνακα *Table IV*, υπολογίστηκαν οι εκτιμήσεις Monte-Carlo της FDR_T^h . Για μία ακόμη φορά, εστιάστηκε η προσοχή στα διαγράμματα CUSUM, των υποενοτήτων της 4.2. «Το CUSUM σε συνεχή χρόνο» και «Ελέγχοντας την αναλογία πιθανοτήτων». Ο πίνακας *Table V* περιλαμβάνει τα αποτελέσματα, όπου παρατηρούμε ότι η FDR αυξάνει καθώς αυξάνει το πλήθος των κέντρων που είναι εντός ελέγχου. Για το διακριτό (OR) CUSUM, η FDR είναι σχετικά σταθερή σε σχέση με το σύνολο των μεταμοσχεύσεων/ έτος, συγκριτικά με το συνεχές CUSUM.

Table V. Variation in FDR_T^h across h for the continuous and discrete CUSUMs.

Facility size (transplants per year)	Continuous CUSUM ($e^\theta = 2.0$)				Discrete (OR) CUSUM ($OR_A = 2.0$)			
	h	FDR_T^h			h	FDR_T^h		
		$p_0=0.7$	$p_0=0.8$	$p_0=0.9$		$p_0=0.7$	$p_0=0.8$	$p_0=0.9$
10	2.25	0.05	0.44	0.64	1.80	0.35	0.48	0.67
40	3.45	0.12	0.19	0.33	2.96	0.28	0.40	0.60
60	3.81	0.26	0.38	0.58	3.36	0.27	0.38	0.58
100	4.35	0.25	0.37	0.57	3.87	0.27	0.39	0.59

Όσον αφορά τη διαδικασία παρακολούθησης της παρακολούθησης πολλαπλών κέντρων, υπάρχει η πεποίθηση ότι το σφάλμα Τύπου-I και η αντίστοιχη ισχύς ($power$), όπως αναφέρονται στον πίνακα *Table IV*, σχετίζονται περισσότερο με θέματα απόφασης και αξιολόγησης, σε σχέση με τα FDR αποτελέσματα το πίνακα *Table V*. Το σύνολο των εσφαλμένων ανιχνεύσεων FDR, σχετίζεται κυρίως με την αξιολόγηση της ακρίβειας της διαδικασίας, από την στιγμή που σημαίνεται ένα κέντρο. Ωστόσο, δεν μπορεί να συμβάλλει η μέθοδος αυτή, ως προς το να αποφασιστεί σε ποια κέντρα σημάνει συναγερμός. Για την επιλογή ενός κατάλληλου ορίου ελέγχου, προτείνεται η χρήση υπολογισμών, όμοιων με αυτούς που οδήγησαν στο σχηματισμό του πίνακα *Table IV*.

4.4. Εφαρμογή σε επέμβαση μεταμόσχευσης νεφρού

Το συνεχές και το διακριτό (OR) CUSUM, εφαρμόστηκαν σε δεδομένα που λήφθηκαν από την SRTR. Η SRTR βρίσκεται υπό τη διαχείριση της «*Arbor Research Collaborative for Health*» σε συνεργασία με το κέντρο «*Kidney Epidemiology Cost Center*», του Πανεπιστημίου του Michigan. Τα δεδομένα εμπεριέχουν πληροφορίες για τις συνολικά 59.650 μεταμοσχεύσεις νεφρού που πραγματοποιήθηκαν σε 258 κέντρα στις Ηνωμένες Πολιτείες Αμερικής, από τον Ιανουάριο του 1997 έως το Δεκέμβριο του 2001. Τα αποτελέσματα της μελέτης ήταν απλή αποτυχία, συμπεριλαμβανομένου και θανάτου. Συνολικά, 5.502 (ποσοστό 9.22%) μεταμοσχεύσεις απέτυχαν μέσα σε μία χρονική περίοδο ενός έτους.

4.4.1. Προσαρμογή κινδύνου

Για το διακριτό CUSUM, χρησιμοποιήθηκε ένα μοντέλο λογιστικής παλινδρόμησης, για την προσαρμογή ενός αριθμού παραγόντων κινδύνου (βλ. πίνακα *Table VI*), ώστε να προβλεφθεί η *risk-adjusted* πιθανότητα αποτυχίας, π_{i0} , $i \geq 1$, για μία προ-μεταμοσχευτική περίοδο ενός έτους.

Table VI. Risk factors and associated relative risks in the fitted logistic model.

Risk factor	Odds ratio	P
Donor cause of death		
Anoxia	0.97	0.599
Cerebrovascular/stroke	1.12	0.011
Head trauma	1.00	Ref
CNS tumor	0.63	0.027
Other	1.10	0.416
Missing	0.73	0.028
Donor age (years)		
<18	1.08	0.225
18–34	0.89	0.005
35–49	1.00	Ref
50–64	1.13	0.012
≥65	1.43	<0.0001
Donor race		
White	1.00	Ref
African-American	1.20	0.0001
Asian	0.95	0.631
Other	1.03	0.842
Donor to recipient weight ratio		
Q1 (0–0.75)	1.18	0.004
Q2 (0.75–0.90)	1.20	0.001
Q3 (0.90–1.50)	0.99	0.961
Q4 (>1.5)	1.00	Ref
Missing	1.20	0.001

Recipient age (years)		
18–34	1.07	0.144
35–49	1.00	Ref
50–64	1.38	<0.0001
≥ 65	1.43	<0.0001
Recipient ethnicity		
Hispanic	0.73	<0.0001
Non-Hispanic	1.00	Ref
Missing	1.09	0.446
Recipient race		
White	1.00	Ref
African-American	0.96	0.282
Asian	0.70	<0.0001
Other	0.86	0.178
Deceased donor hypertension	1.12	0.016
Expanded criteria donor	1.19	0.007
Cause of ESRD		
Tubular/interstitial disease	1.18	0.020
Polycystic kidneys	0.80	0.001
Congenital/familial disorders	1.04	0.767
Diabetes	1.11	0.034
Renovascular diseases	1.12	0.127
Neoplasms	1.14	0.633
Hypertensive nephrosclerosis	1.19	0.001
Retransplant/graft failure	1.20	0.027
Glomerular disease	1.00	Ref
Other	1.10	0.131
Missing	1.17	0.459
Number of B mismatches		
0	1.00	Ref
1	1.30	<0.0001
2	1.40	<0.0001
Number of DR mismatches		
0	1.00	Ref
1	1.15	0.001
2	1.25	<0.0001
Peak PRA		
0–9 per cent	1.00	Ref
10–79 per cent	1.11	0.007
>80 per cent	1.34	<0.0001
PRA missing	0.93	0.741
Previous transplant	1.15	0.053
Recipient BMI		
<20	1.16	0.035
20–24.9	1.00	Ref
25–29.9	1.06	0.203
>30	1.12	0.030
BMI missing	1.15	0.009
Symptomatic PVD	1.22	0.009
Symptomatic PVD missing	0.95	0.483
Dialysis status		
No dialysis	0.94	0.333
Peritoneal dialysis	1.09	0.043
Hemodialysis	1.00	Ref
Unknown type	0.64	0.001
Angina/coronary artery disease	1.10	0.045
Any previous transfusions	1.12	0.0008
Unknown or missing transfusions	1.05	0.248
No previous transfusions	1.00	Ref
Drug treated systemic hypertension	0.90	0.011
Missing	1.01	0.890

Στη συνέχεια, το μοντέλο αυτό χρησιμοποιήθηκε για την κατασκευαστεί το διάγραμμα CUSUM για κάθε κέντρο ξεχωριστά. Θέσαμε λοιπόν $OR_A = 2.0$. Για το συνεχές CUSUM, χρησιμοποιήθηκε το μοντέλο Cox προσαρμοσμένο στους ίδιους παράγοντες κινδύνου (βλ. πίνακα *Table VII*), ώστε να προσεγγιστεί το *adjusted* ποσοστό αποτυχίας μετά τον πρώτο χρόνο της μεταμόσχευσης. Έπειτα αυτό χρησιμοποιήθηκε για τον υπολογισμό της συνάρτησης $a_i(x)$ για κάθε ενήλικα $i = 1, 2, 3, \dots$ και τελικά ενσωματώθηκε στο διάγραμμα CUSUM για κάθε κέντρο μεταμοσχεύσεων, όπως συζητήθηκε στην ενότητα 4.2. Και εδώ, επελέγη η τιμή $e^\theta = 2.0$ ως η εναλλακτική υπόθεση, κατά την κατασκευή του CUSUM.

Table VII. Risk factors and hazard ratios in the fitted Cox model.

Risk factor	Hazard ratio	P
Donor cause of death		
Anoxia	1.02	0.560
Cerebrovascular/stroke	1.09	0.002
Head trauma	1.00	Ref
CNS tumor	0.88	0.284
Other	0.94	0.436
Missing	0.63	<0.0001
Donor age (years)		
<18	0.95	0.187
18–34	0.90	<0.0001
35–49	1.00	Ref
50–64	1.16	<0.0001
≥65	1.39	<0.0001
Donor race		
White	1.00	Ref
African-American	1.15	<0.0001
Asian	0.96	0.516
Other	1.06	0.546
Donor to recipient weight ratio		
Q1 (0–0.75)	1.19	<0.0001
Q2 (0.75–0.90)	1.14	0.0004
Q3 (0.90–1.50)	1.02	0.497
Q4 (>1.5)	1.00	Ref
Missing	1.58	<0.0001
Recipient age (years)		
18–34	1.17	<0.0001
35–49	1.00	Ref
50–64	1.15	<0.0001
≥65	1.47	<0.0001
Recipient ethnicity		
Hispanic	0.79	<0.0001
Non-Hispanic	1.00	Ref
Missing	1.19	0.010
Recipient race		
White	1.00	Ref

African-American	1.11	<0.0001
Asian	0.72	<0.0001
Other	0.81	0.005
Deceased donor hypertension	1.08	0.008
Expanded criteria donor	1.15	0.0005
Cause of ESRD		
Tubular/interstitial disease	1.09	0.049
Polycystic kidneys	0.77	<0.0001
Congenital/familial disorders	1.17	0.055
Diabetes	1.14	<0.0001
Renovascular diseases	1.14	0.004
Neoplasms	1.14	0.434
Hypertensive nephrosclerosis	1.18	<0.0001
Retransplant/graft failure	0.94	0.270
Glomerular disease	1.00	Ref
Other	1.01	0.889
Missing	0.98	0.888
Number of B mismatches		
0	1.00	Ref
1	1.13	0.001
2	1.16	<0.0001
Number of DR mismatches		
0	1.00	Ref
1	1.13	<0.0001
2	1.18	<0.0001
Peak PRA		
0–9 per cent	1.00	Ref
10–79 per cent	1.10	<0.0001
>80 per cent	1.29	<0.0001
PRA missing	1.07	0.554
Previous transplant	1.33	<0.0001
Recipient BMI		
<20	1.22	<0.0001
20–24.9	1.00	Ref
25–29.9	0.99	0.837
>30	1.02	0.650
BMI missing	1.04	0.247
Symptomatic PVD	1.23	<0.0001
Symptomatic PVD missing	0.92	0.106
Dialysis status		
No dialysis	0.85	0.0001
Peritoneal dialysis	1.02	0.353
Hemodialysis	1.00	Ref
Unknown type	0.55	<0.0001
Angina/coronary artery disease	1.16	<0.0001
Any previous transfusions	1.18	<0.0001
Unknown or missing transfusions	1.08	0.002
No previous transfusions	1.00	Ref
Drug treated systemic hypertension	0.97	0.349
Missing	1.08	0.169

Το διάγραμμα *Figure 2* απεικονίζει την καμπύλη CUSUM για ένα συγκεκριμένο κέντρο μεταμοσχεύσεων, όπου στη χρονική βαθμίδα έχει προστεθεί ένα έτος, διότι τα αποτελέσματα στο CUSUM καταγράφονται αφού έχει παρέλθει ένα έτος από την ημερομηνία της μεταμόσχευσης. Το διάγραμμα *Figure 3* δείχνει το CUSUM σε συνεχή χρόνο, για το ίδιο κέντρο μεταμοσχεύσεων, ενώ στο *Figure 4* εικονίζεται το συνεχές CUSUM όπου παρατηρείται μία ξαφνική αλλαγή στο ποσοστό αποτυχίας. Τέλος, στο *Figure 5* έχουμε ένα κέντρο το οποίο παρουσιάζει απόδοση πολύ κοντά ή κάτω του εθνικού μέσου όρου. Και για τα δύο παραπάνω *risk-adjusted* μοντέλα, χρησιμοποιήθηκαν δεδομένα από όλα τα νοσοκομειακά ιδρύματα, ώστε να εκτιμηθεί το εθνικό ποσοστό αποτυχίας ή οι πιθανότητες αποτυχίας.

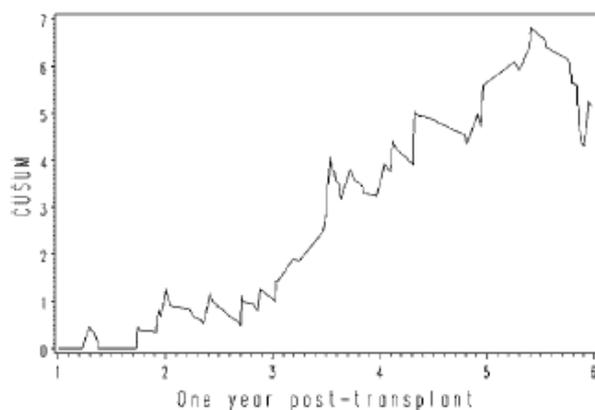


Figure 2. The discrete (OR) CUSUM chart for a facility with a failure rate higher than the national average.

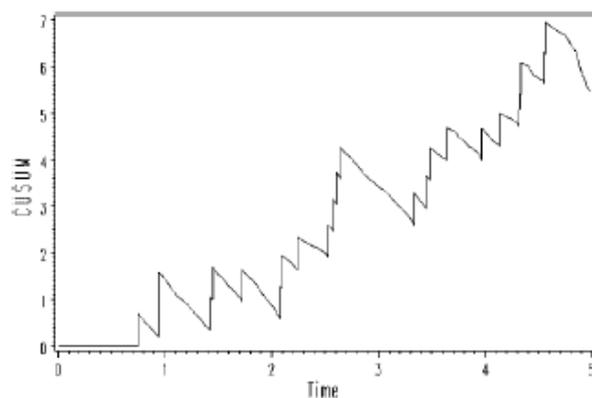


Figure 3. The continuous CUSUM chart for a facility with a failure rate higher than the national average.

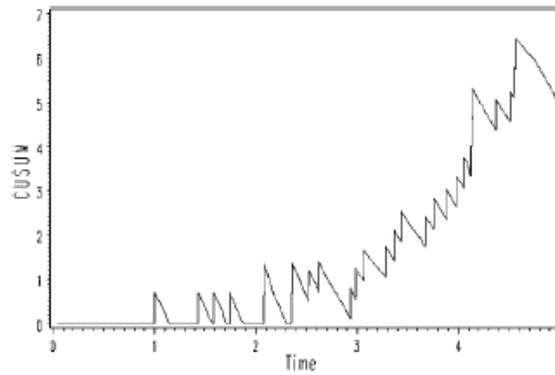


Figure 4. The continuous CUSUM chart showing a sudden possible change in the failure rate.

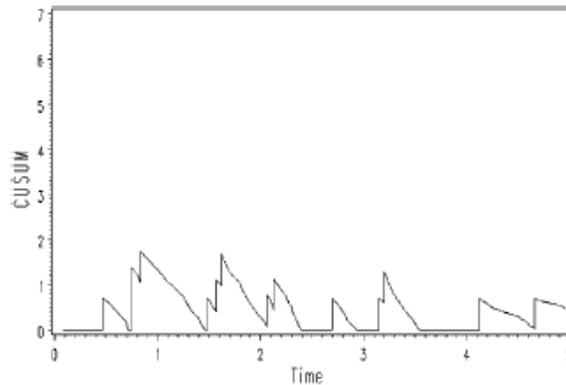


Figure 5. The continuous CUSUM chart showing a facility with failure rates near the national average.

Ας σημειωθεί εδώ ότι και στα δύο μοντέλα, παράγοντες όπως η ηλικία του λήπτη μοσχεύματος, ο λόγος των βαρών δωρητή και λήπτη, η ηλικία του δωρητή και η παρουσία ασθενειών, αυξάνουν περισσότερο το βαθμό επικινδυνότητας και το ποσοστό αποτυχίας της μεταμόσχευσης, σε σχέση με άλλους παράγοντες.

4.4.2. Επιλογή Ορίων Ελέγχου

Στην παρούσα υποενότητα, σημειώνονται μερικά ζητήματα προκειμένου να αποφασιστούν τα όρια ελέγχου για τα διαγράμματα CUSUM που σχεδιάζονται με τα παραπάνω δεδομένα. Ας σημειωθεί εδώ ότι δεν αντιμετωπίζουμε πρόβλημα καθορισμού των ορίων, όπως συζητήθηκε στην υποενότητα 4.3.1, με $T = 5$. Τα 258 διαφορετικά κέντρα μεταμοσχεύσεων παρουσιάζουν διαφορετικά μεγέθη, με το πλήθος των μεταμοσχεύσεων να ποικίλει από 1 έως 1.533, μέσα σε μία πενταετία.

Αφότου το σφάλμα Τύπου-I και το πλήθος των μεταμοσχεύσεων που παρατηρείται μέχρι την ανίχνευση σήματος εξαρτώνται από το σύνολο των μεταμοσχεύσεων σε ένα

κέντρο, ομαδοποιήθηκαν λοιπόν τα 258 κέντρα σε 5 κατηγορίες ως προς το σύνολο των μεταμοσχεύσεων που διεξάγονται σε 5 χρόνια. Εντοπίστηκαν 50 κέντρα με λιγότερες από 8 μεταμοσχεύσεις και 28 κέντρα με περισσότερες από 100 μεταμοσχεύσεις ετησίως, κατά μέσο όρο. Για πολύ μικρά κέντρα, τα CUSUMs ίσως να μην επιδεικνύουν μεγάλη ισχύ: συνεπώς 50 μικρά κέντρα δε συμπεριλήφθησαν στην ανάλυση. Τα 28 κέντρα με ετήσιο μέσο όρο μεταμοσχεύσεων πάνω από 100, ποικίλουν ευρέως ως προς τον ετήσιο αριθμό μεταμοσχεύσεων. Για την επιλογή ορίου ελέγχου, ο πίνακας *Table IV* (σελ. 95) είναι χρήσιμος, όταν $10 \leq \psi \leq 100$.

Ως εκ τούτου, τα υπόλοιπα 180 κέντρα διαιρέθηκαν σε τρεις κατηγορίες και για την κάθε μία ορίστηκαν διαφορετικά όρια ελέγχου χρησιμοποιώντας τα αριθμητικά δεδομένα του *Table IV*. Για εφαρμογή στα 20 μεγάλα κέντρα, προτείνεται η προσομοίωση των δεδομένων του *Table IV* για ένα μεγάλο εύρος τιμών του ψ . Ο πίνακας *Table VIII* δείχνει τα όρια ελέγχου για τις τρεις κατηγορίες σε σχέση με το πλήθος των κέντρων που ανήκουν σε κάθε κατηγορία, βάσει του συνεχούς και του διακριτού (OR) CUSUM. Τελικά, 37 κέντρα μεταμόσχευσης σημάνθηκαν από το συνεχές CUSUM και 41 κέντρα από το διακριτό (OR) CUSUM. Η ανάλυση πραγματοποιήθηκε χωρίς να είναι γνωστή η ταυτότητα του κάθε κέντρου.

Συνοψίζοντας, η επιλογή ενός ορίου ελέγχου για ένα νοσοκομειακό ίδρυμα, εξαρτάται από το μέσο σύνολο των μεταμοσχεύσεων που διεξάγονται σε αυτό και το επίπεδο του σφάλματος Τύπου-I.

Table VIII. Number of facilities flagged by the continuous and discrete (OR) CUSUMs.

Facility size (transplants per year)	Total number of facilities	Continuous CUSUM ($e^{\theta}=2.0$)		Discrete (OR) CUSUM ($OR_A=2.0$)	
		Control limit h	Number flagged	Control limit h	Number flagged
8–20	33	2.25	7	1.80	8
20–50	89	3.45	16	2.96	19
50–100	58	4.20	14	3.65	14

Βάσει αυτής της πληροφορίας, ένα όριο ελέγχου μπορεί να επιλεγεί, προσομοιώνοντας το CUSUM υπό τη μηδενική και χρησιμοποιώντας μία τεχνική βαθμονόμησης όπως στον πίνακα *Table IV*. Ωστόσο, για μικρά κέντρα που διεξάγουν λιγότερες από 8 ή 10 μεταμοσχεύσεις ετησίως, είναι δύσκολο να επιλεγεί ένα όριο ελέγχου που να διατηρεί μικρό το σφάλμα Τύπου-I.

ΒΙΒΛΙΟΓΡΑΦΙΑ

➤ ΕΛΛΗΝΟΓΛΩΣΣΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] Κουκουβίνος Χ., (2008). Στατιστικός Έλεγχος Ποιότητας.
- [2] Αντζουλάκος, Δ. (2006). Στατιστικός Έλεγχος Ποιότητας, Πανεπιστημιακές Σημειώσεις, ΠΜΣ «Εφαρμοσμένη Στατιστική».
- [3] Δαμιανού, Χ.Χ. (1996). Στατιστικός Έλεγχος Ποιότητας και Αξιοπιστία, Πανεπιστημιακές Σημειώσεις, Τμήμα Μαθηματικών, Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών.
- [4] Κούτρας, Μ. Β. (2001). Εισαγωγή στη Συνδυαστική, Εκδόσεις Α.Σταμούλης .
- [5] □Κούτρας, Μ. Β. (2002). Εισαγωγή στις Πιθανότητες, Θεωρία και Εφαρμογές: Μέρος Ι, Εκδόσεις Α. Σταμούλης .
- [6] □Κούτρας, Μ. Β. (2004). Εισαγωγή στις Πιθανότητες, Θεωρία και Εφαρμογές: Μέρος ΙΙ, Εκδόσεις Α. Σταμούλης .
- [7] Κούτρας, Μάρκος Β. (2007). Στατιστική Θεωρία αξιοπιστίας και έλεγχοι χρόνων ζωής, Πανεπιστημιακές Σημειώσεις, ΠΜΣ «Εφαρμοσμένη Στατιστική».

➤ ΞΕΝΟΓΛΩΣΣΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

- [8] Adab, P. (2002). “Performance League Tables: The NHS Deserves Better”. *British Medical Journal*, **324**, pp. 95-98.
- [9] Adams, B. and Woodall W. (1992). “The Use (and Misuse) of False Alarm Probabilities in Control Chart Design”. *Frontiers in Statistical Quality Control*, **4**, pp. 155-168.
- [10] Albert, A.A. and Arnrich, B. (2004). “On-line Variable Live-Adjusted (sic) Displays with Internal and External Risk-adjusted Mortalities. A Valuable Method for Benchmarking and Early Detection of Unfavorable Trends in Cardiac Surgery”. *European Journal of Cardio-thoracic Surgery*, **25**, pp.312-319.
- [11] Albert, A.A. and Walter, J.A. (2004). “On-line Variable Live-Adjusted (sic) Displays with Internal and External Risk-adjusted Mortalities. A Valuable Method for Benchmarking and Early Detection of Unfavorable Trends in Cardiac Surgery”. *European Journal of Cardio-thoracic Surgery*, **25**, pp.312-319.

- [12] Alemi, F.; and Neuhauser, D. (2004). "Time-Between Control Charts for Monitoring Asthma Attacks". *Joint Commission Journal on Quality and Patient Safety*, **2**, pp. 95–102.
- [13] Alemi, F. and Eisenstein, E. (1996). "Risk-adjusted Control Charts for Health Care Assessment". *Annals of Operations Research*, **67**, pp.45-60.
- [14] Arnkeldottir, H. (1995). "Surveillance of Rare Events". Research Report 1995:1 (ISSN 0349-8034), Goteborg University, Department of Statistics.
- [15] Aylin P. and Marshall, C. (2003). "Following Shipman: A Pilot System for Monitoring Mortality Rates in Primary Care". *The Lancet*, **362**, pp. 485-491.
- [16] Barbujani, G. (1987). "A Review of Statistical Methods for Continuous Monitoring of Malformation Frequencies". *European Journal of Epidemiology*, **3**, pp. 67-77.
- [17] Beiles, C. B., and Morton, A. P. (2004). "Cumulative Sum Control Charts for Assessing Performance in Arterial Surgery". *Australian-New Zealand Journal of Surgery*, **74**, pp. 146-151.
- [18] Benneyan, J.C. (1998a). "Statistical Quality Control Methods in Infection Control and Hospital Epidemiology, Part I: Introduction and Basic Theory". *Infection Control and Hospital Epidemiology*, **19**, pp.194-214.
- [19] Benneyan, J.C. (1998b). "Statistical Quality Control Methods in Infection Control and Hospital Epidemiology, Part II: Chart Use, Statistical Properties and Research Issues". *Infection Control and Hospital Epidemiology*, **19**, pp.265-283.
- [20] Benneyan, J.C. (2001). "Performance of Number-Between g-Type Statistical Control Charts for Monitoring Adverse Events". *Health Care Management Science*, **4**, pp. 319-336.
- [21] Benneyan, J.C. and Borgman, A. D. (2003). "Risk-adjusted Sequential Probability Ratio Tests and Longitudinal Surveillance Methods". *International Journal for Quality in Health Care*, **15**, pp. 5-6.
- [22] Benneyan, J.C. and Plsek, P.E. (2003). "Statistical Process Control as a Tool for Research and Healthcare Improvement". *Quality & Safety in Health Care*, **12**, pp. 458-464.
- [23] Blackstone, E. H. (2004). "Monitoring Surgical Performance". *Journal of Thoracic and Cardiovascular Surgery*, **128**, pp. 807-810.
- [24] Bolsin, S. and Colson, M. (2000). "The Use of the Cusum Technique in the Assessment of Trainee Competence in New Procedures". *International Journal for Quality in Health Care*, **12**, pp. 433-438.
- [25] Borrer, C. M. and Champ, C. W. (1998). "Poisson EWMA Control Charts". *Journal of Quality Technology*, **30**, pp. 352-361.

- [26] Bourke, P. D. (1991). "Detecting a Shift in Fraction Nonconforming Using Run-Length Control Charts with 100% Inspection". *Journal of Quality Technology*, **23**, pp. 225-238.
- [27] Carey, R.G. (2003). *Improving Healthcare with Control Charts: Basic and Advanced SPC Methods and Case Studies*. ASQ Quality Press, Milwaukee, WI.
- [28] Champ, C. W. and Woodall, W. H. (1987). "Exact Results for Shewhart Control Charts with Supplementary Runs Rules". *Technometrics*, **29**, pp. 393-399.
- [29] Chen, R. (1978). "A Surveillance System for Congenital Malformations". *Journal of the American Statistical Association*, **73**, pp. 323-327.
- [30] Chen, R. (1985). Letter to the Editor on "Comparison of Two Statistical Techniques for the Surveillance of Birth Defects through a Monte Carlo Simulation" by G. Barbujani and E. Calzonari (with reply). *Statistics in Medicine*, **4**, pp. 389-391.
- [31] Chen, R. (1986). "Revised Values for the Parameters of the Sets Technique for Monitoring the Incidence Rate of a Rare Disease". *Methods of Information in Medicine*, **25**, pp. 47-49.
- [32] Chen, R. (1987). "The Relative Efficiency of the Sets and the Cusum Techniques in Monitoring the Occurrence of a Rare Event". *Statistics in Medicine*, **6**, pp. 517-525.
- [33] Christensen, A. and Thyregod, P. (2003). "Environmental Monitoring Based on a Hierarchical Poisson-Gamma Model". *Journal of Quality Technology*, **35**, pp. 275-285.
- [34] Cook, D.A. and Morton, A.P. (2003). "Monitoring the Evolutionary Process of Quality: Risk-Adjusted Charting to Track Outcomes in Intensive Care". *Critical Care in Medicine*, **31**, pp. 1676-1682.
- [35] De Leval, M. R. and Spiegelhalter, D. (1994). « Analysis of a Cluster of Surgical Failures: Application to a Series of Neonatal Arterial Switch Operations". *Journal of Thoracic and Cardiovascular Surgery*, **107**, pp. 914-924.
- [36] Ewan, W.D. and Kemp, K.W. (1960). "Sampling Inspection of Continuous Processes with No Autocorrelation Between Successive Results". *Biometrika*, **47**, pp. 363-380.
- [37] Fang, Y. (2003). "c-Charts, X-Charts, and the Katz Family of Distributions". *Journal of Quality Technology*, **35**, pp. 104-114.
- [38] Farrington, C.P. and Beale, A.D. (1998). "The Detection of Outbreaks of Infectious Disease". In *GEOMED '97: Proceedings of the International Workshop on Geomedical Systems* (eds. L. Gierl, A. Cliff, A.-J. Valleron, C. P. Farrington, and Mr. Bull), Leipzig: Teuber, pp. 97-117.

- [39] Fasso, A. (1998). “One-Sided Multivariate Testing and Environmental Monitoring”. *Austrian Journal of Statistics*, **27**, pp. 17-38.
- [40] Fasso, A. (1999). “One-Sided MEWMA Control Charts”. *Communications in Statistics-Simulation and Computation*, **28**, pp. 381-401.
- [41] Friedman, D. J. (1993). “Some Considerations in the Use of Quality Control Techniques in Integrated Circuit Fabrication”. *International Statistical Review*, **61**, pp. 97-107.
- [42] Frisen, M. (1992). “Evaluations of Methods for Statistical Surveillance”. *Statistics in Medicine*, **11**, pp. 1489-1502.
- [43] Gallivan, S. and Sherlaw-Johnson, C. (1998). “Detection of Changes in Mortality After Heart Surgery (Letter)”. *British Medical Journal*, **317**, p.1453.
- [44] Gallus, G. and Radaelli, G. (1986). « On Surveillance Methods for Congenital Malformations ». *Statistics in Medicine*, **5**, pp. 565-571.
- [45] Gan, F.F. (1994). “Design of Optimal Exponential CUSUM Charts”. *Journal of Quality Technology*, **26**, pp. 109-124.
- [46] Gan, F.F. (1994). “Designs of One- and Two-Sided Exponential EWMA Charts”. *Journal of Quality Technology*, **30**, pp. 55-69.
- [47] Grigg, O. and Farewell, V.T. (2003). “Use of Risk-adjusted CUSUM and RSPRT Charts for Monitoring in Medical Contexts”. *Statistical Methods in Medical Research*, **12**, pp. 147-170.
- [48] Grunkemeier, G.L. and Furnary, A.P. (2003). „Cumulative Sum Techniques for Assessing Surgical Results“. *Annals of Thoracic Surgery*, **76**, pp. 663-667.
- [49] Gustafson, T.L. (2000). “Practical Risk-adjusted Quality Control Charts for Infection Control”. *American Journal of Infection Control*, **28**, pp. 406-414.
- [50] Hart, M. K. and Hart, R. F. (2002). *Statistical Process Control for Health Care*. Duxbury, Pacific Grove, CA.
- [51] Hawkins, D.M. (1991). “Multivariate Quality Control Using Regression Adjusted Variables”. *Tecnometrics*, **33**, pp. 61-75.
- [52] Hunter, J.S. (1990). Discussion of “Exponentially Weighted Moving Average Control Schemes – Properties and Enhancements” by J. M. Lucas and M.S. Saccucci, *Tecnometrics*, **32**, pp. 21-22.
- [53] Hutwagner, L. and Treadwell, T. (2003). “The Bioterrorism Preparedness and Response Early Aberration Reporting System (EARS)”. *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, **80**, Supplement 1, pp. 89-96.

- [54] Iezzoni, L.I. (1997). “The Risks of Risk Adjustment”. *Journal of the American Medical Association*, **278**, pp. 1600-1607.
- [55] Ismail, N. A. and Webster, R. A. (2003). “On-Line Monitoring and Retrospective Analysis of Hospital Outcomes Based on a Scan Statistic”. *Statistics in Medicine*, **22**, pp. 2861-2876.
- [56] Institute of Medicine. (2000). *To Error is Human*. Washington, D.C.
- [57] Jarpe, E. (1999). “Surveillance of the Interaction Parameter of the Ising MODEL”, *Communications in Statistics – Theory and Methods*, **28**, pp. 3009-3027.
- [58] Joner, M. and Reynolds, M. R. (2005). “The Use of Multivariate Control Charts to Detect Changes in the Spatial Patterns of Disease”. Presented at the 2005 Joint Statistical Meetings, Minneapolis, Minnesota.
- [59] Kemp, K. W. (1961). “The Average Run Length of the Cumulative Sum Chart When a V-Mask is Used”. *Journal of the Royal Statistical Society, Series B*, **23**, pp. 149-153.
- [60] Kenett, R. and Pollak, M. (1983). “On Sequential Detection of a Shift in the Probability of a Rare Event”. *Journal of the American Statistical Association*, **78**, pp. 389-395.
- [61] Khan, R. A. (1984). “On Cumulative Sum Procedures and the SPRT with Applications”. *Journal of the Royal Statistical Society, Series B*, **46**, pp. 79-85.
- [62] Knox, G. (1964). “The Detection of Space-Time Interactions”. *Applied Statistics*, **13**, pp. 25-29.
- [63] Kulldorff, M. (2001). “Prospective Time Periodic Geographical Disease Surveillance Using a Scan Statistic”. *Journal of the Royal Statistical Society, Series A*, **164**, pp. 61-72.
- [65] Lawson, A. B. (2001). *Statistical Methods in Spatial Epidemiology*. John Wiley & Sons, New York.
- [66] Lee, K. Y. and McGreevey, C. (2002a). “Using Control Charts to Assess Performance Measurement Data”. *Journal on Quality Improvement*, **28**, pp. 90-101.
- [67] Leung, C. S. and McGilchrist, C. A. (1999). “A Distribution-Free Regional Cumulative Sum for Identifying Hyperendemic Periods of Disease Incidence”. *The Statistician*, **48**, pp. 215-225.
- [68] Lie, R. T. and Irgens, L. M. (1993). “A New Sequential Procedure for Surveillance of Down’s Syndrome”. *Statistics in Medicine*, **12**, pp. 13-25.

[69] Lilford, R. and Thomson, R. (2004). “Use and Misuse of Process and Outcome Data in Managing Performance of Acute Medical Care: Avoiding Institutional Stigma”. *The Lancet*, **363**, pp. 1147-1154.

[70] Lim, T. O. and Morad, Z. (2002). “Assessing Doctors’ Competence: Application of CUSUM Technique in Monitoring Doctors’ Performance”. *International Journal for Quality in Health Care*, **14**, pp. 251-258.