

ΠΜΣ ΒΙΟΣΤΑΤΙΣΤΙΚΗ

Μάιος 2022

ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΙΑΤΡΙΚΗ ΣΧΟΛΗ

ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΡΟΜΠΟΤΗ ΓΕΩΡΓΙΑ

Reconstructing survival data from published Kaplan–Meier curves

ΑΘΗΝΑ, 2022

Η παρούσα διπλωματική εργασία εκπονήθηκε στο πλαίσιο των σπουδών για την απόκτηση του Μεταπτυχιακού Διπλώματος Ειδίκευσης στη

ΒΙΟΣΤΑΤΙΣΤΙΚΗ

που απονέμει η Ιατρική Σχολή και το Τμήμα Μαθηματικών του Εθνικού & Καποδιστριακού Πανεπιστημίου Αθηνών

Εγκρίθηκε την 09/05/2022 από την εξεταστική επιτροπή:

ΟΝΟΜΑΤΕΠΩΝΥΜΟ	ΒΑΘΜΙΔΑ	ΥΠΟΓΡΑΦΗ
1. Καρλής Δημήτριος (Επιβλέπων)	Καθηγητής
2. Μπουρνέτας Απόστολος	Καθηγητής
3. Δεμίρης Νικόλαος	Επίκουρος Καθηγητής

Ευχαριστίες

Με την ολοκλήρωση της παρούσας διπλωματικής θα ήθελα να ευχαριστήσω ολόθερμα και εγκάρδια τον επιβλέποντα καθηγητή μου κ. Καρλή Δημήτριο για την αμέριστη συμπαράσταση και τη βοήθειά του καθ' όλη τη διάρκεια πραγματοποίησης της παρούσας επιστημονικής μελέτης. Οι συμβουλές του, η υπομονή του και οι προτάσεις του με βοήθησαν να μην εγκαταλείψω ακόμα και όταν όλες οι προσπάθειες μου κατέληγαν στο κενό. Ήταν πάντα εκεί να με ενθαρρύνει και με τον τρόπο του να με ωθεί στο να γίνομαι καλύτερη.

Ακόμη αισθάνομαι έντονα την ανάγκη να ευχαριστήσω τους γονείς μου, Κωνσταντίνο και Ευσταθία καθώς και τον αδερφό μου, Λευτέρη που είναι πάντα δίπλα μου και πιστεύουν σε μένα. Η αγάπη τους και η συμπαράστασή τους έπαιξαν καθοριστικό ρόλο σε ολόκληρη την ακαδημαϊκή μου πορεία, όντας πάντα στο πλάι μου, τόσο στις όμορφες όσο και στις δύσκολες στιγμές πάντα με αστείρευτη υπομονή και υπερηφάνεια.

Επιπρόσθετα, θέλω να ευχαριστήσω τους φίλους μου που με στήριξαν όλο αυτό το διάστημα αποδίδοντας ένα ιδιαίτερο ευχαριστώ στις μοναδικές μου Σεβαστή και Θεοδώρα, στην παιδική μου φίλη, Φωτεινή και στην κουμπάρα μου, Βασιλική των οποίων η στήριξη και η ενθάρρυνση δεν έλειψε ποτέ. Τέλος, ένα μεγάλο ευχαριστώ στη συνάδελφό μου, Τερέζα, που μπορεί να τη γνωρίζω μικρό χρονικό διάστημα αλλά μου μετέδωσε απίστευτη δύναμη.

Σας ευχαριστώ πολύ όλους!

Περιεχόμενα

Κεφάλαιο 1 ^ο - Εισαγωγή.....	8
Κεφάλαιο 2 ^ο -Ανάλυση Επιβίωσης.....	12
2.1 Εισαγωγή στην Ανάλυση Επιβίωσης.....	12
2.1.1 Λογοκρίσια.....	13
2.1.2 Βασικές συναρτήσεις.....	17
2.1.3 Αναλογικότητα Κινδύνων.....	20
2.2 Κύριες παραμετρικές κατανομές επιβίωσης.....	21
2.3 Εκτίμηση της συνάρτησης επιβίωσης και της συνάρτησης αθροιστικού κινδύνου.....	23
2.3.1 Εκτιμητής Kaplan Meier.....	24
2.3.2 Μέθοδος Lifetable.....	26
2.3.3 Εκτιμητής Nelson-Aalen.....	27
2.4 Log-Rank test.....	28
2.5 Το μοντέλο του Cox.....	29
2.5.1 Αναλογικότητα κινδύνων και παλινδρόμηση Cox.....	30
2.6 Περιορισμένος Μέσος Χρόνος Επιβίωσης (RMST).....	31
2.6.1 Εκτίμηση RMST.....	33
2.6.2 Επιλογή του καθορισμένου σημείου τ.....	35
2.6.3 Πλεονεκτήματα και περιορισμοί.....	35
2.6.4 Συνοπτικά μέτρα και Περιορισμένος Μέσος Χρόνος Επιβίωσης (RMST).....	36
2.6.4.1. Λόγος Κινδύνου (Hazard Ratio).....	36
2.6.4.2. Διάμεσος (Median).....	36
Κεφάλαιο 3 ^ο -Προτεινόμενοι μέθοδοι ανακατασκευής δεδομένων.....	37
3.1 Η προσέγγιση του Parmar.....	38
3.2 Η μέθοδος του Williamson.....	42
3.2.1 Εκτίμηση με χρήση καμπυλών Kaplan Meier.....	42
3.2.2 Εκτίμηση με χρήση life-table.....	44
3.2.3 Εκτίμηση μέσω μιας συνολικής καμπύλης επιβίωσης.....	44
3.3 Η προσέγγιση των Hoyle and Henley.....	46
3.4 Ο αλγόριθμος της Guyot.....	52
3.4.1 Προσαρμογές αλγορίθμου.....	60
3.5 Μελέτη επικύρωσης των τεσσάρων μεθόδων.....	63
3. 6 Βελτιώσεις αλγορίθμου Guyot.....	64
3.7 Ανακατασκευή δεδομένων επιβίωσης ενσωματώνοντας τους σημειωμένους χρόνους	

λογοκρισίας (marked censoring times)	69
Κεφάλαιο 4^ο-Μελέτη Προσομοίωσης.....	74
4.1 Περιγραφή Προσομοίωσης	74
4.2 Αποτελέσματα	78
Κεφάλαιο 5^ο-Συζήτηση και περαιτέρω έρευνα	89
Περίληψη	93
Abstract.....	95
ΒΙΒΛΙΟΓΡΑΦΙΑ	97

ΛΙΣΤΑ ΕΙΚΟΝΩΝ

Εικόνα 1: Απεικόνιση συνάρτησης πυκνότητας πιθανότητας	18
Εικόνα 2: Απεικόνιση συνάρτησης κινδύνου	19
Εικόνα 3: Διάγραμμα ροής αλγορίθμου Guyot (Guyot, et al. 2012)	59
Εικόνα 4: Διάγραμμα ροής αλγορίθμου Rogula (Rogula, Lozano-Ortega and Johnston 2022) ...	71
Εικόνα 5: Αρχικές και ανακατασκευασμένες καμπύλες Kaplan Meier μαζί με τον πίνακα του πλήθους των ατόμων σε κίνδυνο, έχοντας χρησιμοποιήσει για την ανακατασκευή 15 σημεία αναφέροντας τον αριθμό των ατόμων σε κίνδυνο. A. Έχοντας γίνει εξαγωγή 100 σημείων από την αρχική καμπύλη KM. B. Έχοντας γίνει εξαγωγή 50 σημείων από την αρχική καμπύλη KM. C. Έχοντας γίνει εξαγωγή 25 σημείων από την αρχική καμπύλη	79
Εικόνα 6: Αρχικές και ανακατασκευασμένες καμπύλες Kaplan Meier μαζί με τον πίνακα του πλήθους των ατόμων σε κίνδυνο, έχοντας χρησιμοποιήσει για την ανακατασκευή 7 σημεία αναφέροντας τον αριθμό των ατόμων σε κίνδυνο A. Έχοντας γίνει εξαγωγή 100 σημείων από την αρχική καμπύλη KM B. Έχοντας γίνει εξαγωγή 50 σημείων από την αρχική καμπύλη KM C. Έχοντας γίνει εξαγωγή 25 σημείων από την αρχική καμπύλη KM.....	81
Εικόνα 7: Αρχικές και ανακατασκευασμένες καμπύλες Kaplan Meier έχοντας χρησιμοποιήσει μόνο το συνολικό αριθμό των ασθενών A. Έχοντας γίνει εξαγωγή 100 σημείων από την αρχική καμπύλη KM B. Έχοντας γίνει εξαγωγή 50 σημείων από την αρχική καμπύλη KM C. Έχοντας γίνει εξαγωγή 25 σημείων από την αρχική καμπύλη.....	82
Εικόνα 8: Διαγράμματα διασποράς μεταξύ των αρχικών και των ανακατασκευασμένων συνοπτικών μέτρων HR και διαφορά RMSTs κάτω από τα διαφορετικά σενάρια εξαγωγής συντεταγμένων και έχοντας χρησιμοποιήσει 15 σημεία αναφοράς των ατόμων σε κίνδυνο.	84
Εικόνα 9: Διαγράμματα διασποράς μεταξύ των αρχικών και των ανακατασκευασμένων συνοπτικών μέτρων HR και διαφορά RMSTs κάτω από τα διαφορετικά σενάρια εξαγωγής συντεταγμένων και έχοντας χρησιμοποιήσει 7 σημεία αναφοράς των ατόμων σε κίνδυνο.	85
Εικόνα 10: Διαγράμματα διασποράς μεταξύ των αρχικών και των ανακατασκευασμένων συνοπτικών μέτρων HR και διαφορά RMSTs κάτω από τα διαφορετικά σενάρια εξαγωγής συντεταγμένων και έχοντας χρησιμοποιήσει μόνο το συνολικό αριθμό των ασθενών κάθε ομάδας.....	86

ΛΙΣΤΑ ΓΡΑΦΗΜΑΤΩΝ

Γράφημα 1: Εκτιμητές Kaplan Meier από ένα σύνολο δεδομένων μη συμπεριλαμβανομένων των λογοκριμένων παρατηρήσεων και με τη συμπερίληψή τους.....	17
Γράφημα 2: RMST (η κόκκινη περιοχή κάτω από την καμπύλη επιβίωσης) όταν το $\tau = 1$	32
Γράφημα 3: Διαφορά RMST για $\tau = 1$	34
Γράφημα 4: Κατανομή Weibull με παράμετρο κλίμακας, $\lambda=1$ και παράμετρο σχήματος $\gamma=0.8$ για την ομάδα ελέγχου.....	75

ΛΙΣΤΑ ΠΙΝΑΚΩΝ

Πίνακας 1: Log-rank τεστ τη χρονική στιγμή t_j	28
Πίνακας 2: Πίνακας του μέσου όρου και της τυπικής απόκλισης των συνοπτικών μέτρων HR (λόγος κινδύνου) και RMST (περιορισμένος μέσος χρόνος επιβίωσης) για όλα τα σενάρια που πραγματοποιήθηκαν όσο αφορά τον αριθμό των εξαγόμενων μεταβλητών από το αρχικό γράφημα και το πλήθος των αριθμών σε κίνδυνο που εισήχθησαν στον αλγόριθμο ανακατασκευής.....	87

Κεφάλαιο 1^ο- Εισαγωγή

Η συνεχής αύξηση των πληροφοριών σχετικά με διάφορες επιστημονικές υποθέσεις καθώς και η ανάγκη για άμεσες και έγκυρες αποφάσεις σε θέματα που αφορούν τόσο την καθημερινή κλινική πρακτική όσο και θέματα δημόσιας υγείας, έχουν καταστήσει σαφή την αναγκαιότητα της σύνθεσης των αποτελεσμάτων που προέρχονται από την πληθώρα των μελετών που διεξάγονται. Προς αυτή την κατεύθυνση λοιπόν, όλο και περισσότερες μετα-αναλύσεις σχεδιάζονται και πραγματοποιούνται με στόχο τον υπολογισμό ενός συγκεντρωτικού αποτελέσματος για την εκτίμηση της σχέσης μεταξύ της μελετώμενης έκθεσης και του υπό μελέτη γεγονότος. Υποστηρίζεται γενικά η άποψη ότι η χρήση μετα-αναλύσεων αποτελεί την βέλτιστη επιλογή για τους ερευνητές στην προσπάθεια που καταβάλλουν για λήψη ορθολογικών αποφάσεων μεγαλύτερης ακρίβειας και εγκυρότητας, ειδικά στον κλάδο της υγείας (Pigott & Polanin, 2020). Απαιτείται όμως προσοχή στον τρόπο διεξαγωγής τους καθώς οι μετα-αναλύσεις που διενεργούνται χωρίς ιδιαίτερα αυστηρά κριτήρια μπορούν να συντελέσουν στην εξαγωγή διαφορετικών συμπερασμάτων.

Στην πράξη οι μετα-αναλύσεις αποτελούν ένα σύνολο στατιστικών τεχνικών για τη σύνθεση των αποτελεσμάτων πολλαπλών μελετών (Higgins, 2011; Borenstein, et al., 2021), οι οποίες έχουν συλλεχθεί μέσω κάποιας συστηματικής ανασκόπησης (Gough, et al., 2017). Με άλλα λόγια οι μετα-αναλύσεις στοχεύουν στο να παρέχουν μία πλήρη και περιεκτική σύνθεση των μελετών που έχουν ασχοληθεί με κάποιο παρόμοιο ερώτημα και τείνουν να γίνονται ολοένα και πιο αποδεκτές, ως το καταλληλότερο μέσο για τη συγκέντρωση των αποτελεσμάτων ενός αριθμού τυχαιοποιημένων ελεγχόμενων δοκιμών (RCTs, Randomized Clinical Trials). Ωστόσο σκεπτόμενοι την ποικιλομορφία των αναφορών στις αναλύσεις επιβίωσης στα διάφορα περιοδικά καθίσταται σαφές ότι δεν υφίσταται μία πάγια και επαρκής μέθοδος για την εξαγωγή αποτελεσμάτων σχετικά με εκβάσεις από το χρόνο έως το γεγονός (time to event) (Altman, et al., 1995). Ακόμη και οι συστάσεις τόσο του CONSORT (Moher, et al., 2010) όσο και του Cochrane (Higgins & Thomas, 2022) να χρησιμοποιείται ως μέτρο επίδρασης για το χρόνο επιβίωσης ο λόγος κινδύνου (hazard ratio) περιορίζουν τον τρόπο με τον οποίο οι RCTs (Randomize Clinical Trials) με δεδομένα χρόνου επιβίωσης μπορούν να συμπεριληφθούν σε κάθε είδους δευτερογενούς αναλύσεις ή σε αναλύσεις κόστους αποτελεσματικότητας, καθώς απαιτούνται ισχυρές παραδοχές που ενδέχεται να μην έχουν δοκιμαστεί επαρκώς. Για παράδειγμα στις κλινικές δοκιμές, μέτρα όπως ο μέσος χρόνος επιβίωσης, αποτελούν μία μεροληπτική εκτίμηση της αναμενόμενης επιβίωσης, καθώς η παρουσία λογοκριμένων (censored) παρατηρήσεων έχει ως αποτέλεσμα τα δεδομένα αυτά να μην επιδέχονται τυπικές στατιστικές διαδικασίες που χρησιμοποιούνται για την ανάλυση συνεχών τυχαίων μεταβλητών (όπως mean και standard deviation). Παρόλο λοιπόν που γίνονται μετα-αναλύσεις με χρήση συγκεντρωτικών δεδομένων (aggregate data), οι δυσκολίες που παρουσιάζονται λόγω της έλλειψης δημοσίευσης επαρκών στατιστικών

στοιχείων καθώς και ότι οι μετα-αναλύσεις με χρήση ατομικών δεδομένων (Individual Patient Data meta analysis-IPD) παρέχουν το καλύτερο επίπεδο ερευνητικών προτύπων έκανε τις τελευταίες να θεωρούνται προσεγγίσεις χρυσού προτύπου (gold standard).

Πιο αναλυτικά η IPD μετα-ανάλυση αναγνωρίζεται ως μία άριστη επιλογή για τη σύνθεση αποδεικτικών στοιχείων επειδή μπορεί να προσφερθεί σε επανάληψη χρησιμοποιώντας πιο εύρωστα και ευέλικτα μοντέλα επιβίωσης (Syn, et al., 2020). Ακόμη δίνει τη δυνατότητα διερεύνησης πρόσθετων υποθέσεων που σχετίζονται με τα μεμονωμένα χαρακτηριστικά του ασθενούς, που πιθανότατα να έλειπαν από ήδη δημοσιευμένα αποτελέσματα (Syn, et al., 2020). Επίσης προσφέρει την ευχέρεια λεπτομερούς εξέτασης των δεδομένων και αξιολόγησης της αρχικής μεθόδου ανάλυσης. Η λήψη της IPD βοηθά στη διερεύνηση της ετερογένειας των αποτελεσμάτων σε μία μετα-ανάλυση προκειμένου να γίνει προσαρμογή σε αυτή (Simmonds, et al., 2005). Η ετερογένεια εμφανίζεται συχνά σε περιστάσεις όπου η επίδραση της θεραπείας ποικίλλει με το χρόνο και οι δοκιμές στη μετα-ανάλυση έχουν διαφορετικό μήκος. Ακόμη η διαθεσιμότητα των ανακατασκευασμένων ατομικών δεδομένων από το χρόνο έως το γεγονός παρέχει τη δυνατότητα αποφυγής προκαταλήψεων που σχετίζονται με τη χρήση των συγκεντρωτικών δεδομένων. Για παράδειγμα είναι υψίστης σημασίας όταν υπάρχουν ενδείξεις παρέκκλισης από την υπόθεση των αναλογικών κινδύνων ή όταν τα συνοπτικά αποτελέσματα της δοκιμής δεν αναφέρονται. Επιτρέπει στον ερευνητή δηλαδή να εφαρμόσει εναλλακτικά μοντέλα στις δευτερεύουσες αναλύσεις, όπως τον περιορισμένο μέσο χρόνο επιβίωσης (RMST) που δεν απαιτεί την υπόθεση της αναλογικότητας, και μπορεί να έχουν μία πιο διαισθητική ερμηνεία υπό τους μη αναλογικούς κινδύνους (Parner & Andersen, 2010; Parmar, et al., 1998).

Η χρήση όμως μέτρων όπως το RMST που αναπτύχθηκαν σχετικά πρόσφατα δεν αναφέρονται σε προηγούμενες μελέτες. Ως εκ τούτου η ωφέλειά τους βασίζεται είτε στην ομαδική χρήση ατομικών δεδομένων είτε σε τεχνικές που επιτρέπουν την ανακατασκευή τέτοιων δεδομένων από τις αναφορές των δοκιμών (Messori, et al., 2021). Ωστόσο η απόκτηση των ατομικών δεδομένων των ασθενών δεν είναι μία εύκολη διαδικασία εφόσον συχνά τα δεδομένα δεν είναι διαθέσιμα, καθώς οι ερευνητές δεν είναι πρόθυμοι να παρέχουν τέτοιου είδους δεδομένα, τόσο για προσωπικούς λόγους όσο και για θέματα εμπιστευτικότητας, όπως επίσης μπορεί να μην υπάρχει ο απαιτούμενος χρόνος για διενέργεια IPD ανάλυσης, όπως σε μία ενδιάμεση ανάλυση μιας εν εξελίξει δοκιμής.

Με έναυσμα λοιπόν όλα τα προαναφερθέντα, τα τελευταία χρόνια οι ερευνητικές προσεγγίσεις ασχολούνται στενά με την ανακατασκευή δεδομένων επιβίωσης από τις δημοσιευμένες καμπύλες Kaplan Meier, μιας και είναι η πιο συχνά αναφερόμενη πληροφορία στις μελέτες ανάλυσης επιβίωσης, χρησιμοποιώντας ωστόσο και άλλες διαθέσιμες πληροφορίες όπως τον αριθμό των ατόμων σε κίνδυνο σε διάφορες χρονικές στιγμές. Σε κάποιες ανασκοπήσεις που είχαν γίνει πριν το 1995 αναφερόταν ότι το 88%

των δημοσιευμένων μελετών παρουσίαζαν καμπύλες επιβίωσης ενώ μόνο το 8% περιλάμβανε τους αριθμούς των ατόμων που βρισκόταν σε κίνδυνο (number at risk) (Williamson, et al., 2002). Μία μετέπειτα ανασκόπηση που δημοσιεύτηκε σε τέσσερα ιατρικά περιοδικά διαπίστωσε ότι σχεδόν όλες οι δοκιμές παρουσίαζαν καμπύλες επιβίωσης και στις περιπτώσεις που υπήρχε μεταβλητή διάρκεια παρακολούθησης είχαν αναφέρει και τους αριθμούς σε κίνδυνο. Γίνεται λοιπόν αισθητό ότι παρόλο που η δημοσίευση επαρκών στατιστικών στοιχείων διευκολύνει μεταγενέστερες αναλύσεις, η αναφορά των αποτελεσμάτων των μελετών σπάνια ακολουθεί αυτές τις αρχές στο έπακρο και συνεπώς επισημαίνεται η ανάγκη για ανακατασκευή των ατομικών δεδομένων επιβίωσης μέσω των καμπυλών KM. Πόσο όμως αξιόπιστα είναι αυτά τα δεδομένα άραγε; Ποιος είναι ο βαθμός συμφωνίας ανάμεσα στα ανακατασκευασμένα συνοπτικά μέτρα και στα αρχικά;

Ορμώμενοι λοιπόν από τα παραπάνω ερωτήματα, σκοπός της παρούσας διπλωματικής είναι αρχικά να αναδείξει τις διαθέσιμες μεθόδους ανακατασκευής δεδομένων από δημοσιευμένες καμπύλες Kaplan Meier και άλλα διαθέσιμα συνοπτικά μέτρα. Στη συνέχεια με τη βοήθεια της βέλτιστης μεθόδου και κάποιων προσομοιωμένων δεδομένων σκοπό αποτελεί η διαπίστωση του βαθμού απόκλισης των ανακατασκευασμένων από τα αρχικά συνοπτικά μέτρα, αν υπάρχει, με στόχο την ενίσχυση της αξιοπιστίας των ανακατασκευασμένων τιμών. Τέλος, επιχειρείται να αναδειχθεί ο ρόλος διάφορων ποσοτήτων (όπως ο αριθμός σημείων που εξάγονται από δημοσιευμένες καμπύλες και ο αριθμός των σημείων που αναφέρεται το πλήθος των ατόμων σε κίνδυνο) που εισάγονται στις μεθόδους ανακατασκευής ως προς τις αποκλίσεις (αν αυτές υπάρχουν) στα ανακατασκευασμένα δεδομένα και συνοπτικά μέτρα .

Ξεκινώντας λοιπόν από το επόμενο κεφάλαιο (Κεφάλαιο 2) που αποτελεί την εισαγωγή στην Ανάλυση Επιβίωσης, παρέχεται η θεμελιώδης ορολογία που χρήζει απαραίτητη για την καλή κατανόηση της συνέχειας. Πιο αναλυτικά διευκρινίζονται αναφερόμενοι όροι και έννοιες όπως η ανάλυση επιβίωσης, η συνάρτηση επιβίωσης, ο εκτιμητής Kaplan Meier καθώς και το μοντέλο αναλογικών κινδύνων του Cox. Ακόμη αποσαφηνίζεται ένα σχετικά νεοεμφανιζόμενο μέτρο, ο περιορισμένος μέσος χρόνος επιβίωσης (RMST) καθώς και η διαφορά του περιορισμένου χρόνου επιβίωσης (RMST difference) που χρησιμοποιείται για τη σύγκριση ομάδων.

Έχοντας λοιπόν αναφέρει όλο το απαραίτητο θεωρητικό υπόβαθρο των εννοιών που θα χρειαστούν, το Κεφάλαιο 3 παρουσιάζει αναλυτικά τις υπάρχουσες πέντε μεθόδους για ανακατασκευή δεδομένων. Πιο συγκεκριμένα οι πρώτες δύο αναφέρονται στην ανακατασκευή συγκεντρωτικών δεδομένων ενώ οι επόμενες τρεις στην ανακατασκευή ατομικών δεδομένων. Στη συνέχεια παρουσιάζεται και μία μελέτη επικύρωσης των τεσσάρων μεθόδων (απουσιάζει η τελευταία), στην οποία βέλτιστη κατά κοινή ομολογία αποτελεί η μέθοδος της Guyot και των συνεργατών της. Τέλος αναφέρονται κάποιες

βελτιώσεις του προηγούμενου αλγορίθμου που αποτρέπουν την ύπαρξη ατέρμονων βρόγχων και βελτιώνουν λίγο ακόμα την ποιότητα της ανακατασκευής των δεδομένων.

Έπειτα στο κεφάλαιο 4 παρουσιάζεται ένα πείραμα προσομοίωσης στο οποίο θα δημιουργηθούν πενήντα σετ δεδομένων και θα γίνουν τα κατάλληλα γραφήματα Kaplan Meier. Στη συνέχεια χρησιμοποιείται ο αλγόριθμος της Guyot, μαζί με τις βελτιώσεις ώστε να επιτευχθεί η ανακατασκευή των ατομικών δεδομένων. Έχοντας την εν λόγω πληροφορία στη συνέχεια γίνεται διερεύνηση στο ρόλο του πλήθους των σημείων που εξάγονται από τις αρχικές καμπύλες καθώς και στο ρόλο του πλήθους των χρονικών σημείων για τα οποία αναφέρεται ο αριθμός των ατόμων σε κίνδυνο και χρησιμοποιούνται ως εισαγωγές στον αλγόριθμο. Τέλος πραγματοποιείται έλεγχος του βαθμού συμφωνίας των αρχικών συνοπτικών μέτρων (HR, RMST) από τα δεδομένα με τα ανακατασκευασμένα συνοπτικά μέτρα από την ανακατασκευασμένη IPD, παρέχοντας ωστόσο κάποια γραφήματα για οπτική αξιολόγηση της συσχέτισής τους.

Ολοκληρώνοντας το Κεφάλαιο 5 συνοψίζει τα ευρήματα του παραπάνω κεφαλαίου και στη συνέχεια σκιαγραφεί μία προσπάθεια που έγινε (και απέτυχε) για εξαγωγή των σημείων από τις καμπύλες επιβίωσης αυτοματοποιημένα, με στόχο την εξάλειψη του σφάλματος του ερευνητή, μιας και η υπάρχουσα μέθοδος εμπλέκει τον ερευνητή στη συλλογή των σημείων, γεγονός που αυξάνει το σφάλμα (bias). Τέλος γίνονται προτάσεις για περαιτέρω έρευνα με βάση τις ανιχνευμένες αδυναμίες της μεθόδου.

Κεφάλαιο 2^ο-Ανάλυση Επιβίωσης

2.1 Εισαγωγή στην Ανάλυση Επιβίωσης

Η ανάλυση επιβίωσης (survival analysis) αποτελεί ένα κλάδο της στατιστικής, ο οποίος εισήχθη για πρώτη φορά από τον Bernoulli, γύρω στο 1700, αναδεικνύοντας τη χρησιμότητα των εμβολίων σε άτομα τα οποία ήταν προσβεβλημένα από ευλογιά και στο τέλος κατέληξαν (Kleinbaum & Klein, 1996). Αρχικά εφαρμόστηκε σε ιατρικές έρευνες και αφορούσε το χρόνο από τη θεραπεία έως το θάνατο, γεγονός που συντέλεσε στην εν λόγω ονομασία. Σήμερα η ανάλυση επιβίωσης πραγματεύεται το χρονικό διάστημα που μεσολαβεί έως την εμφάνιση ενός συγκεκριμένου συμβάντος (event). Το χρονικό διάστημα αυτό (time-to-event) καθορίζεται συναρτήσει της έναρξης της παρακολούθησης του ατόμου, μέχρι την εμφάνιση του γεγονότος και μπορεί να εκφράζει το χρόνο επιβίωσης (survival time) ή το χρόνο αποτυχίας (failure time).

Εντούτοις, η ανάλυση επιβίωσης εξαιρουμένως της ιατρικής μπορεί να εφαρμοστεί σε διάφορους άλλους τομείς όπως στη μηχανολογία, αναφερόμενη ως θεωρία αξιοπιστίας (reliability theory) αναλύοντας το χρόνο μέχρι τη δυσλειτουργία κάπου εξαρτήματος ή αντίστοιχα το χρόνο αντοχής του σε συγκεκριμένες συνθήκες. Επίσης μπορεί να εφαρμοστεί στη γεωργία προκειμένου να μελετηθεί ο χρόνος μέχρι την ανάπτυξη ενός φυτού καθώς και στην οικονομία, σε διάφορες αναλύσεις κόστους-αποτελεσματικότητας (CEA) αναφερόμενη ως «μοντελοποίηση διάρκειας», όπως για παράδειγμα για την μοντελοποίηση του χρόνου όπου κάποιος είναι άνεργος. Τέλος χρησιμοποιείται και σε ποιο θεωρητικές επιστήμες όπως στην κοινωνιολογία ως «ανάλυση ιστορίας γεγονότων» (Jenkins, 2005).

Αρα γίνεται αντιληπτό ότι το υπό μελέτη γεγονός που αντανακλάται σε θετικές τυχαίες μεταβλητές δεν περιορίζεται στο θάνατο αλλά ούτε και στους ανθρώπους. Μπορεί να αφορά το χρόνο υποτροπής (μετάστασης) μιας ασθένειας, το χρονικό διάστημα παραμονής σε μία μονάδα νοσηλείας, τη διάρκεια μια κατάληψης, το χρόνο που μεσολαβεί μέχρι τη λήψη διαζυγίου, τη διάρκεια μιας διδακτορικής διατριβής κ.ά. Γενικά λοιπόν μπορεί να χρησιμοποιηθεί για τη μοντελοποίηση οποιουδήποτε χρόνου γεγονότος ορίζεται ως αποτυχία. Σε κάθε περίπτωση όμως απαιτείται ο ορισμός του συμβάντος να είναι ακριβής και σαφής προκειμένου τα αποτελέσματα της ανάλυσης να μην είναι ασυνεπή. Γιατί ο θάνατος μπορεί να είναι ένα οριστικό γεγονός που δεν χωράει παρερμηνείες, δεν συμβαίνει όμως κάτι αντίστοιχο με τη δυσλειτουργία μιας συσκευής, καθώς το πρόβλημα μπορεί να βρίσκεται σε σημείο που να την καταστεί μη λειτουργική ή σε σημείο που να επηρεάζει μόνο την εξωτερική του οπτική. Επομένως ανάλογες περιπτώσεις χρήζουν επιπλέον διευκρινήσεις.

Είναι επιτακτική η ανάγκη να αναφερθεί ότι η μεθοδολογία που εφαρμόζεται στην ανάλυση δεδομένων επιβίωσης λόγω της ιδιαίτερης φύσης τους, δεν ταυτίζεται με τις παραδοσιακές μεθόδους που

χρησιμοποιούνται για την ανάλυση δεδομένων. Πιο αναλυτικά σε μία μελέτη ο χρόνος παρακολούθησης του κάθε ασθενή διαφέρει και μπορεί να κυμαίνεται από μερικές ώρες έως και χρόνια, ανάλογα με τη διάρκεια της μελέτης. Ακόμη το υπό μελέτη γεγονός συμβαίνει σε ακανόνιστους χρόνους στο παρατηρούμενο δείγμα καθώς επίσης παρατηρούνται και λογοκριμένοι χρόνοι επιβίωσης (για τους οποίους θα μιλήσουμε εκτενέστερα στην επόμενη ενότητα), οι οποίοι αντανακλούν τη διάρκεια ζωής των υπό μελέτη αντικειμένων, η οποία μπορεί να είναι άγνωστη ή εν μέρει καταγεγραμμένη. Τα παραπάνω γνωρίσματα καθιστούν την κατανομή των δεδομένων μη συμμετρική και πιο συγκεκριμένα συνήθως ακολουθούν μία θετική συμμετρία (πλειοψηφία παρατηρήσεων στην αριστερή κορυφή του ιστογράμματος), γεγονός που εμποδίζει την υπόθεση της κανονικής κατανομής και δυσχεραίνει την ανάλυση.

Όμως επειδή ο ορισμός του συμβάντος δεν είναι αποκλειστικά μονοσήμαντος (συχνά πολλαπλός), η ανάλυση επιβίωσης αποτελεί μία ευρέως εφαρμόσιμη μέθοδος η οποία μπορεί να περιγράψει επαρκώς τα δεδομένα επιβίωσης (μέσω των καμπυλών Kaplan Meier καθώς και άλλων μετρήσεων όπως ο λόγος κινδύνου (hazard ratio) και ο περιορισμένος μέσος χρόνος επιβίωσης (RMST)). Ακόμη επιτρέπει τη διεκπεραίωση πολλαπλών αναλύσεων χρησιμοποιώντας το ίδιο σύνολο δεδομένων ανάλογα με τον εκάστοτε ορισμό του υπό μελέτη γεγονότος, παρέχει συγκρίσεις ανάμεσα στους χρόνους επιβίωσης των παρατηρούμενων ομάδων καθώς και βοηθά στη διερεύνηση της επίδρασης μιας μεταβλητής στην επιβίωση μέσω της δημιουργίας στατιστικών (παραμετρικών, ημι-παραμετρικών ή μη-παραμετρικών) μοντέλων.

2.1.1 Λογοκρισία

Ένα σύνθηρες φαινόμενο στην ανάλυση επιβίωσης είναι η αδυναμία παρακολούθησης του πλήρους χρόνου επιβίωσης των ασθενών σε σχέση με το υπό μελέτη γεγονός. Αυτό δεν ηχεί περίεργα αν αναλογιστεί κανείς τις φορές που κάποιοι ασθενείς για τους δικούς τους προσωπικούς λόγους (πχ μετακόμιση σε άλλη πόλη, θάνατος από άλλη αιτία) εκούσια ή ακούσια εγκαταλείπουν τη μελέτη στην οποία συμμετείχαν και η μοναδική πληροφόρηση που λαμβάνεται είναι ότι μέχρι εκείνη τη δεδομένη στιγμή δεν είχαν εκδηλώσει το συμβάν. Επίσης είναι αρκετές οι περιπτώσεις όπου μετά την πάροδο της προκαθορισμένης χρονικής περιόδου διενέργειας της μελέτης, πολλά από τα άτομα της μελέτης δεν έχουν εμφανίσει το υπό μελέτη γεγονός. Οι παρατηρήσεις αυτές ονομάζονται λογοκριμένες παρατηρήσεις (censored observations) και κατά κάποιο τρόπο διακόπτουν τη διαδικασία παρακολούθησης (Hosmer & Lemeshow, 1999). Ωστόσο πρέπει να καταστεί σαφές ότι μία λογοκριμένη παρατήρηση δεν αποτελεί μία ελλιπή παρατήρηση και αυτό γιατί λαμβάνουμε την πληροφορία ότι μέχρι τη στιγμή που σταμάτησε η παρακολούθηση δεν είχε λάβει χώρα το γεγονός. Συνεπώς πρόκειται για μία γνώση που είναι άκρως απαραίτητη στην ανάλυση επιβίωσης και δεν μπορούμε να την αγνοήσουμε όπως θα κάναμε σε κάποιο άλλο στατιστικό πλαίσιο.

Πριν αναφερθούν οι διαφορετικοί τύποι λογοκρισίας των δεδομένων γίνεται μία μικρή αναφορά στα βασικά σημεία του χρόνου μέχρι το συμβάν (time to event). Ο χρόνος επιβίωσης (survival time) ή ο χρόνος μέχρι την αποτυχία (failure time) για κάθε αντικείμενο είναι ο χρόνος από ένα προκαθορισμένο σημείο εκκίνησης έως την εμφάνιση του γεγονότος ενδιαφέροντος. Γίνεται επομένως αντιληπτό ότι για να τον ορισμό του χρόνου μέχρι το συμβάν από μία τυχαία μεταβλητή απαιτείται μία σαφή χρονική προέλευση (π.χ. τυχαιοποιημένη κλινική δοκιμή, διάγνωση ασθένειας), μία χρονική κλίμακα μέτρησης (π.χ. ημέρες, μήνες, χρόνια) καθώς και ο σαφής ορισμός του γεγονότος (π.χ θάνατος, υποτροπή) όπως αναφέρθηκε και προηγουμένως. Μία σημαντική διευκρίνιση είναι ότι τα άτομα μπορεί να εισέρχονται στη μελέτη σε διαφορετικό ημερολογιακό χρόνο, πρέπει όμως η χρονική τους προέλευση να είναι κοινή (π.χ. ο χρόνος μηδέν είναι ο χρόνος έναρξης της θεραπείας).

Υπάρχουν τρεις βασικοί τύποι λογοκρισίας:

2.1.1.1 Δεξιά λογοκριμένες παρατηρήσεις (Right Censoring)

Εμφανίζονται όταν μία παρατήρηση ξεκινά από τον προκαθορισμένο χρόνο προέλευσης και σταματάει πριν πραγματοποιηθεί το γεγονός ενδιαφέροντος εξαιτίας της απώλειας στην παρακολούθηση (loss to follow-up), είτε της εγκατάλειψης, είτε του τερματισμού της μελέτης (Parner & Andersen, 2010). Πρακτικά αυτό συνεπάγεται ότι ο πραγματικός χρόνος επιβίωσης υπερβαίνει μία συγκεκριμένη τιμή, που είναι ο χρόνος παρατήρησης. Αποτελούν τις πιο συχνά εμφανιζόμενες παρατηρήσεις (με αυτές θα ασχοληθούμε).

Η συμπεριφορά ενός υποκειμένου i σε μία μελέτη μπορεί να αντικατοπτριστεί από δύο μεταβλητές (t_i, δ_i) όπου, (Pigott & Polanin, 2020)

$$t_i = \min\{T_i, C_i\} \quad \text{και} \quad \delta_i = \begin{cases} 1, & \text{αν } T_i \leq C_i \\ 0, & \text{αλλιώς} \end{cases}$$

- t_i : ο παρατηρούμενος χρόνος επιβίωσης του υποκειμένου i , με $i=1, \dots, n$
- C_i : ο χρόνος λογοκρισίας του υποκειμένου i
- T_i : ο ακριβής χρόνος πραγματοποίησης του γεγονότος

Ωστόσο αυτό το είδος λογοκρισίας διακρίνεται από τρεις κατηγορίες.

1. Δεξιά λογοκρισία τύπου I (Right Censoring: Type I)

Στον εν λόγω τύπο δεν είναι γνωστός ο αριθμός των παρατηρήσεων που θα λογοκριθούν αλλά οι χρόνοι λογοκρισίας είναι καθορισμένοι. Το γεγονός ενδιαφέροντος παρατηρείται μόνο εάν συμβεί πριν από κάποιο προκαθορισμένο χρόνο.

2. Δεξιά λογοκρισία τύπου II (Right Censoring: Type II)

Εδώ παρατηρείται η αντίστροφη κατάσταση από πριν. Δηλαδή οι χρόνοι λογοκρισίας είναι τυχαίοι ενώ ο αριθμός των λογοκριμένων περιπτώσεων είναι καθορισμένος και σταθερός. Αυτός ο τύπος λογοκρισίας παρατηρείται σπάνια στην επιδημιολογία ενώ είναι πιο σύνηθες χρησιμοποιήσιμος στον έλεγχο της διάρκειας λειτουργίας ενός μηχανήματος.

3. Τυχαία δεξιά λογοκρισία (Right Censoring: Random)

Η συγκεκριμένη κατάσταση παρατηρείται όταν ανταγωνιστικά γεγονότα επιφέρουν την απομάκρυνση των υποκειμένων από τη μελέτη και τόσο οι χρόνοι λογοκρισίας όσο και ο αριθμός των λογοκριμένων γεγονότων είναι τυχαίοι.

2.1.1.2 Αριστερά λογοκριμένες παρατηρήσεις (Left Censoring)

Εμφανίζονται όταν το γεγονός ενδιαφέροντος έχει συμβεί πριν την εισαγωγή του υποκειμένου στη μελέτη. Ο πραγματικός χρόνος επιβίωσης λοιπόν είναι μικρότερος από μια συγκεκριμένη τιμή και η ατελής φύση των παρατηρήσεων αυτών φαίνεται στην αριστερή ουρά του άξονα του χρόνου. Για παράδειγμα έστω ότι θέλουμε να μελετήσουμε το χρόνο που μεσολαβεί μέχρι ένα υποκείμενο να ξεκινήσει να εργάζεται και για το σκοπό αυτό πραγματοποιούνται κάποιες συνεντεύξεις. Αν το υποκείμενο ήδη απασχολείται κατά τη στιγμή της πρώτης συνέντευξης τότε έχουμε αριστερή λογοκρισία.

Η συμπεριφορά ενός υποκειμένου i σε μία τέτοια μελέτη μπορεί να αντικατοπτριστεί από δύο μεταβλητές (t_i, δ_i) όπου,

$$t_i = \min\{T_i, C_i\} \quad \text{και} \quad \delta_i = \begin{cases} 1, & \text{εάν } T_i \geq C_i \\ 0, & \text{αλλιώς} \end{cases}$$

- t_i : ο παρατηρούμενος χρόνος επιβίωσης του υποκειμένου i , με $i=1, \dots, n$
- C_i : ο χρόνος λογοκρισίας του υποκειμένου i
- T_i : ο ακριβής χρόνος πραγματοποίησης του γεγονότος

2.1.1.3 Λογοκριμένες παρατηρήσεις εντός διαστήματος (Interval Censoring)

Τα δεδομένα λογοκρίνονται εντός διαστήματος όταν το υπό μελέτη γεγονός συνέβη μεταξύ δύο χρονικών στιγμών και δεν δύναται να καθοριστεί η ακριβής τιμή. Εμφανίζονται σε μελέτης όπου οι χρόνοι συνάντησης (του ερευνητή και του

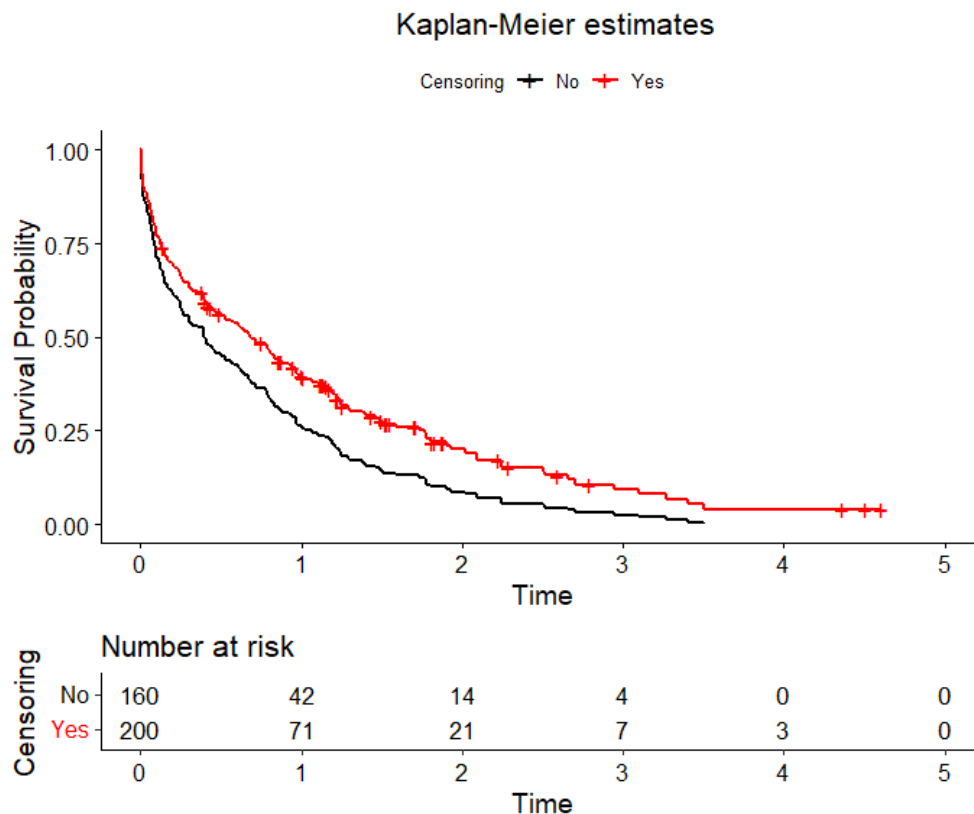
συμμετέχοντα) είναι καθορισμένοι και το συμβάν μπορεί να συμβεί μεταξύ αυτών των στιγμών.

Τέλος ένας ακόμα διαχωρισμός που πραγματοποιείται στα λογοκριμένα δεδομένα τα κατατάσσει σε πληροφοριακά (informative) και μη πληροφοριακά (non-informative) αναλόγως την εξάρτηση που εμφανίζουν με το χρόνο επιβίωσης (Kleinbaum & Klein, 1996). Πιο αναλυτικά, πληροφοριακή λογοκρισία (informative censoring) υπάρχει όταν η κατανομή του χρόνου λογοκρισίας C περιέχει οποιαδήποτε πληροφορία σχετικά με τις παραμέτρους που χαρακτηρίζουν την κατανομή του T , ενώ σε αντίθετη περίπτωση έχουμε μη-πληροφοριακή ή ανεξάρτητη λογοκρισία (independent or non-informative censoring). Με άλλα λόγια στην ανεξάρτητη λογοκρισία τα υποκείμενα οποιαδήποτε υποομάδας ενδιαφέροντος που λογοκρίνονται τη χρονική στιγμή t , είναι αναγκαίο να είναι αντιπροσωπευτικά όλων των υποκειμένων που παρέμειναν σε κίνδυνο σε αυτή την υποομάδα εκείνη τη χρονική στιγμή σε σχέση με την επιβίωσή τους. Παραδείγματα τέτοιων καταστάσεων είναι κλινικές δοκιμές με προκαθορισμένη διάρκεια όπου τα άτομα που εμφανίζουν το γεγονός μετά το πέρας της μελέτης θεωρούνται λογοκριμένες παρατηρήσεις. Ωστόσο το γεγονός αυτό οφείλεται καθαρά στην προκαθορισμένη χρονική περίοδο της μελέτης και ουδέ μία σχέση δεν έχει με την κατάσταση του υποκειμένου και συνεπώς με το χρόνο επιβίωσης του.

Προκειμένου να αποδοθεί όσο καλύτερα γίνεται η σημασία των λογοκριμένων παρατηρήσεων, προσομοιώσαμε 200 ασθενείς για τους οποίους η έκβαση ενδιαφέροντος ήταν ο θάνατος. Χρησιμοποιήθηκε η Weibull κατανομή τόσο για την μοντελοποίηση της συνάρτησης επιβίωσης [Weibull (0.7,0.9)] όσο και για την κατανομή των λογοκρισιών [Weibull (2,2.9)], χρησιμοποιώντας ωστόσο διαφορετικές παραμέτρους για να επιτευχθεί 20% λογοκρισία κατά μέσο όρο. Υπενθυμίζουμε ότι η συνάρτηση επιβίωσης αντανάκλα την πιθανότητα ένας ασθενής να μην έχει βιώσει το γεγονός μέχρι το εκάστοτε χρονικό σημείο και η κατανομή λογοκρισίας δίνει την πιθανότητα σε κάθε χρονικό σημείο ο ασθενής να μην έχει λογοκριθεί μέχρι εκείνο το σημείο. Μία παρατήρηση θεωρείται λογοκριμένη αν η κατάσταση αλλάξει πριν από την έλευση του γεγονότος. Θέλοντας ωστόσο να εκτιμήσουμε τις πιθανότητες επιβίωσης κάθε ασθενούς σχεδιάσαμε τις πιθανότητες αυτές σε ένα γράφημα Kaplan Meier μία φορά μη λαμβάνοντας υπόψη τις λογοκριμένες παρατηρήσεις και μία δεύτερη έχοντας λάβει υπόψη. Η απεικόνιση των εκτιμήσεων αυτών βρίσκεται στο Γράφημα 1.

Στη μαύρη γραμμή απεικονίζονται οι KM εκτιμήσεις μη έχοντας συμπεριλάβει τις λογοκριμένες παρατηρήσεις και έχοντας στην κόκκινη. Γίνεται επομένως άμεσα αντιληπτό ότι το σύνολο δεδομένων χωρίς τις λογοκρισίες υποεκτιμά τις πιθανότητες επιβίωσης, εφόσον η μαύρη γραμμή βρίσκεται πιο κάτω από την κόκκινη. Αυτό συμβαίνει γιατί όπως έχει ήδη αναφερθεί οι λογοκριμένες παρατηρήσεις μεταφέρουν την πληροφορία ότι μέχρι εκείνη τη στιγμή δεν έχει πραγματοποιηθεί το συμβάν. Αφαιρώντας

λοιπόν αυτές τις παρατηρήσεις αφαιρείται ένα μέρος της υπάρχουσας πληροφορίας και προφανώς επηρεάζεται και το μέγεθος του δείγματος, καθώς όπως διαφαίνεται από τον πίνακα των ατόμων σε κίνδυνο η μαύρη καμπύλη εκτιμήθηκε από 160 άτομα ενώ η κόκκινη από 200.



Γράφημα 1: Εκτιμητές Kaplan Meier από ένα σύνολο δεδομένων μη συμπεριλαμβανομένων των λογοκριμένων παρατηρήσεων και με τη συμπερίληψή τους

2.1.2 Βασικές συναρτήσεις

Για την μαθηματική περιγραφή της κατανομής της πιθανότητας μιας τυχαίας μεταβλητής επιβίωσης υπάρχουν διάφοροι ισοδύναμοι τρόποι. Μερικοί από αυτούς είναι:

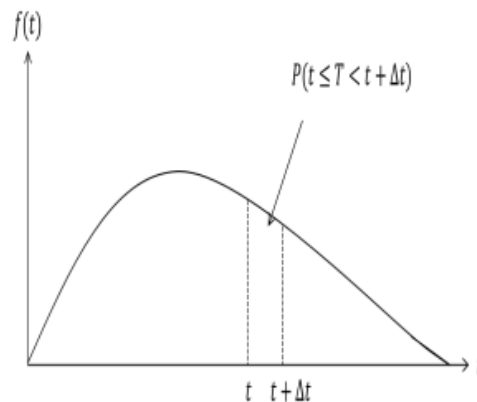
- Η συνάρτηση πυκνότητας πιθανότητας $f(t)$ [density function]
- Η συνάρτηση επιβίωσης $S(t)$ [survival function]
- Η συνάρτηση κινδύνου $\lambda(t)$ [hazard function]
- Η αθροιστική συνάρτηση κινδύνου $\Lambda(t)$ [cumulative hazard function]

Αξίζει να αναφερθεί ότι στα παραδοσιακά στατιστικά μοντέλα οι συναρτήσεις που χρησιμοποιούνται συνήθως είναι οι κατανομές πυκνότητας και οι αθροιστικές κατανομές, αλλά εξαιτίας της ιδιαίτερης φύσης των δεδομένων επιβίωσης στα οποία απαντώνται λογοκριμένες παρατηρήσεις, οι τυπικές αυτές συναρτήσεις δεν είναι προτιμητέες. Αντ' αυτού οι συναρτήσεις επιβίωσης και κινδύνου λογίζονται ως πιο κατάλληλες.

Οι εν λόγω συναρτήσεις παρουσιάζουν διαφορές ανάλογα με τον τρόπο που λαμβάνεται υπόψη ο χρόνος, ως συνεχή ή διακριτή μεταβλητή. Προκειμένου να επιτευχθεί παρακάτω η περιγραφή τους αλλά και για την υπόλοιπη εργασία ο χρόνος θα θεωρείται ως μία συνεχής μεταβλητή.

Έστω μία μεταβλητή T η οποία παίρνει μη αρνητικές τιμές και περιγράφει τον πραγματικό χρόνο επιβίωσης ενός υποκειμένου. Εκφράζει δηλαδή το διάστημα που μεσολαβεί από το προκαθορισμένο σημείο εκκίνησης έως την εμφάνιση του γεγονότος ενδιαφέροντος. Το σύνολο των τιμών που λαμβάνει η τυχαία μεταβλητή T δημιουργεί μία κατανομή πιθανότητας με την αντίστοιχη *συνάρτηση πυκνότητας πιθανότητας* $f(t)$. Όπου, για T συνεχή η συνάρτηση γίνεται:

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t}$$



Εικόνα 1: Απεικόνιση συνάρτησης πυκνότητας πιθανότητας

Έπειτα η *αθροιστική συνάρτηση κατανομής* $F(t)$ της τυχαίας μεταβλητής T , που αντιστοιχεί στο ποσοστό των ατόμων των οποίων τα γεγονότα συμβαίνουν συναρτήσει του t ορίζεται ως:

$$F(t) = P(T < t) = \int_0^t f(u) du$$

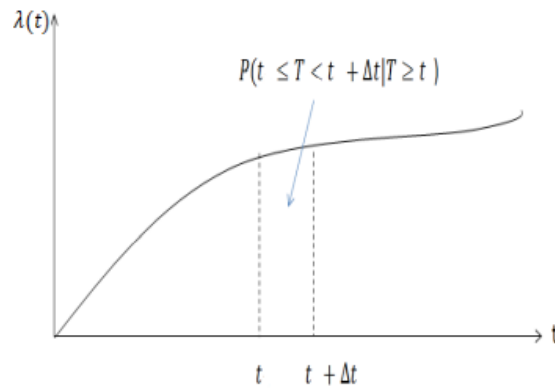
Η *συνάρτηση επιβίωσης* $S(t)$ εκφράζει την πιθανότητα ένα υποκείμενο να ζήσει περισσότερο από μία καθορισμένη χρονική στιγμή t (και προφανώς είναι συμπληρωματική της αθροιστικής συνάρτησης κατανομής).

$$S(t) = 1 - F(t) = P(T \geq t)$$

Οι κοινές ιδιότητες όλων των συναρτήσεων επιβίωσης είναι ότι αρχικά με την πάροδο του χρόνου δεν αυξάνεται η τιμή τους, είτε παραμένει σταθερή είτε μειώνεται. Δεύτερον το χρονικό σημείο 0 είναι η στιγμή που ξεκινάει η παρακολούθηση και επομένως ισχύει ότι $S(0) = 1$ και τέλος η πιθανότητα ενός ατόμου να επιζήσει περισσότερο από t , καθώς το t τείνει στο άπειρο είναι μηδέν [$\lim_{t \rightarrow \infty} S(t) = 0$].

Η *συνάρτηση κινδύνου* $\lambda(t)$ δίνει το στιγμιαίο κίνδυνο εμφάνισης του υπό μελέτη γεγονότος ανά μονάδα χρόνου, δεδομένου ότι το υποκείμενο έχει επιβιώσει έως τη χρονική στιγμή t , και μαθηματικά ορίζεται ως

$$\lambda(t) = \lim_{\Delta t \rightarrow \infty} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}$$



Εικόνα 2: Απεικόνιση συνάρτησης κινδύνου

Ωστόσο επειδή δεν είναι άμεσα προφανείς η σύνδεση της συνάρτησης κινδύνου με τη συνάρτηση πυκνότητας, την αθροιστική κατανομή και τη συνάρτηση επιβίωσης, έπειτα από λίγες μαθηματικές πράξεις προκύπτει η ισότητα (2.1).

$$\begin{aligned} \lambda(t) &= \lim_{\Delta t \rightarrow \infty} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \\ &= \lim_{\Delta t \rightarrow \infty} \frac{P(t \leq T < t + \Delta t)}{\Delta t \cdot P(T \geq t)} \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{P(T \geq t)} \lim_{\Delta t \rightarrow \infty} \frac{P(t \leq T < t + \Delta t)}{\Delta t} \\
&\Rightarrow \lambda(t) = \frac{f(t)}{S(t)}. \tag{2.1}
\end{aligned}$$

Στη συνέχεια η αθροιστική συνάρτηση κινδύνου $\Lambda(t)$ ορίζεται ως το ολοκλήρωμα από το 0 έως το t της συνάρτησης κινδύνου και ισούται με την περιοχή κάτω από την καμπύλη κινδύνου μέχρι τη στιγμή t .

$$\Lambda(t) = \int_0^t \lambda(u) du.$$

Η οποία μέσω της (2.1) $\left[\lambda(t) = -\frac{d}{dt} \log S(t) \right]$ παίρνει τη μορφή $\Lambda(t) = -\log S(t)$ και τελικά

$$S(t) = e^{-\Lambda(t)}.$$

Άρα γίνεται αντιληπτό ότι αν είναι γνωστή μία εκ των συναρτήσεων $\lambda(t)$, $S(t)$ και $\Lambda(t)$ τότε ο υπολογισμός και τον υπολοίπων είναι μία άμεση διαδικασία.

Η θεμελίωση αυτών των συναρτήσεων είχαν ως αποτέλεσμα τη δημιουργία διάφορων μοντέλων προσαρμογής στα δεδομένα. Το κάθε ένα κάνει τις δικές του υποθέσεις και παραδοχές για τις παραπάνω συναρτήσεις, ενώ έχουν αναπτυχθεί και μη παραμετρικές μέθοδοι, όπως ο εκτιμητής Kaplan Meier, όπου για τον καθορισμό της συνάρτησης επιβίωσης χρησιμοποιείται ένας πιο εξειδικευμένος τύπος ο οποίος θα αναφερθεί διεξοδικά στη συνέχεια.

2.1.3 Αναλογικότητα Κινδύνων

Από τους βασικούς σκοπούς της ανάλυσης επιβίωσης είναι η σκιαγράφηση της επιβίωσης των υπό μελέτη πληθυσμών καθώς και η σύγκρισή τους με άλλους. Μια κοινή υπόθεση που χρησιμοποιείται συχνά για τη σύγκριση διαφορετικών πληθυσμών είναι η αναλογικότητα των κινδύνων. Με άλλα λόγια θεωρείται ότι η συνάρτηση κινδύνου ενός πληθυσμού είναι ανάλογη της συνάρτησης κινδύνου του πληθυσμού σύγκρισης. Συνεπώς ανάμεσα σε δύο ομάδες ο στιγμιαίος ρυθμός αποτυχίας θεωρείται σταθερός. Αναμφίβολα πρόκειται για μία πολύ ισχυρή υπόθεση η οποία όμως εφαρμόζεται ευρέως λόγω της απλότητάς της.

2.2 Κύριες παραμετρικές κατανομές επιβίωσης

Προκειμένου να μοντελοποιηθούν διάφορα δεδομένα επιβίωσης έχουν προταθεί κάποιες ευρέως χρησιμοποιούμενες κατανομές τόσο για το χρόνο επιβίωσης όσο και για το χρόνο λογοκρισίας. Μερικές από αυτές που είναι η εκθετική κατανομή, η κατανομή Weibull, η (piecewise) εκθετική κατανομή, η Γάμμα κατανομή, η λογαριθμολογιστική κατανομή (log-logistic distribution), η κατανομή Pareto καθώς και η λογαριθμοκανονική κατανομή (log-normal distribution).

Επειδή ο χρόνος είναι πάντα θετικός οι παραπάνω κατανομές αντιστοιχούν σε μη αρνητικές τυχαίες μεταβλητές που στην πλειονότητά τους είναι συνεχείς.

Εκθετική Κατανομή

Εάν ο χρόνος επιβίωσης (ή αποτυχίας) ακολουθεί την εκθετική κατανομή, $T \sim \text{Exp}(\lambda)$, τότε για $t \geq 0$ οι βασικές συναρτήσεις γίνονται:

$$f(t) = \lambda e^{-\lambda t},$$

$$F(t) = 1 - e^{-\lambda t},$$

$$S(t) = 1 - F(t) = \int_t^{\infty} f(u) du = e^{-\lambda t},$$

$$\lambda(t) = \frac{f(t)}{S(t)} = \lambda \text{ (σταθερός κίνδυνος),}$$

$$\Lambda(t) = \int_0^t \lambda(u) du = \int_0^t \lambda du = \lambda t.$$

Η τιμή λ αντιπροσωπεύει το ποσοστό κινδύνου (hazard rate), δηλαδή την πιθανότητα αποτυχίας τη χρονική στιγμή t , δεδομένου ότι μέχρι εκείνη τη στιγμή δεν έχει συμβεί το γεγονός. Διαπιστώνεται ότι η τιμή αυτή είναι μία σταθερά και συνεπώς ο κίνδυνος δεν αλλάζει με την πάροδο του χρόνου, γεγονός που αποδεικνύεται εύκολα.

Έστω $T_1 \sim \text{Exp}(\lambda_1)$ για έναν πληθυσμό A και $T_2 \sim \text{Exp}(\lambda_2)$ για ένα πληθυσμό B τότε,

$$\frac{\lambda_2(t)}{\lambda_1(t)} = \frac{\lambda_2}{\lambda_1}$$

που είναι ανεξάρτητος από το χρόνο.

Χαρακτηριστικά εκθετικής κατανομής:

$$\text{Διάμεση τιμή } (\tau): 0.5 = S(\tau) = e^{-\lambda\tau} \implies \tau = -\frac{\log(0.5)}{\lambda}$$

$$\text{Μέση τιμή: } E(T) = \int_0^{\infty} u\lambda e^{-\lambda u} du = \frac{1}{\lambda}$$

Κατανομή Weibull

Η κατανομή Weibull αποτελεί μία γενίκευση της εκθετικής κατανομής.

Εάν ο χρόνος επιβίωσης ακολουθεί την κατανομή Weibull, $T \sim \text{Weibull}(\lambda, \gamma)$, τότε για $t \geq 0$ οι βασικές συναρτήσεις παρουσιάζονται ως εξής:

$$f(t) = \lambda\gamma t^{\gamma-1} e^{-\lambda t^\gamma},$$

$$F(t) = 1 - e^{-\lambda t^\gamma},$$

$$S(t) = 1 - F(t) = e^{-\lambda t^\gamma},$$

$$\lambda(t) = \frac{f(t)}{S(t)} = \frac{\lambda\gamma t^{\gamma-1} e^{-\lambda t^\gamma}}{e^{-\lambda t^\gamma}} = \lambda\gamma t^{\gamma-1},$$

$$\Lambda(t) = \int_0^t \lambda(u) du = \int_0^t \lambda\gamma u^{\gamma-1} du = \lambda t^\gamma.$$

Οι παράμετροι λ και γ ονομάζονται παράμετρος κλίμακας (scale parameter) και παράμετρος σχήματος (shape parameter) αντιστοίχως. Αν $\gamma=1$ παίρνουμε την εκθετική κατανομή. Η Weibull κατανομή είναι πιο ευέλικτη και βολική από την εκθετική και περιλαμβάνει διάφορα μοτίβα κινδύνου, δηλαδή

- Αν $\gamma = 1$ τότε ο κίνδυνος (hazard) παραμένει σταθερός
- Αν $0 < \gamma < 1$ τότε ο κίνδυνος μειώνεται και
- Αν $\gamma > 1$ τότε ο κίνδυνος αυξάνεται

Ακόμη η σχέση του $\log(-\log(S(t)))$ με το $\log(t)$ είναι γραμμική [$S(t) = e^{-\lambda t^\gamma} \rightarrow \log(-\log(S(t))) = \log(\lambda) + \gamma \log(t)$] γεγονός που επιτρέπει την αξιολόγηση της καταλληλότητας του μοντέλου Weibull μέσω της γραφικής παράστασης της $\log(-\log \hat{S}(t))$ ως προς $\log(t)$, όπου $\hat{S}(t)$ είναι η εκτίμηση της συνάρτησης επιβίωσης μέσω της μεθόδου Kaplan Meier. Έτσι στην περίπτωση όπου προκύψει μία κατά προσέγγιση ευθεία γραμμή τότε υπάρχουν ενδείξεις ότι η κατανομή Weibull μπορεί να περιγράψει επαρκώς τα δεδομένα.

Μοντέλο Σταθερού κατά Τμήματα Κινδύνου (Piecewise Exponential Distribution)

Το μοντέλο της εκθετικής κατανομής είναι δύσκολο να αντανakλά ικανοποιητικά τα παρατηρούμενα δεδομένα καθώς περιέχει μόνο μία παράμετρο, αλλά ακόμα και η Weibull κατανομή αδυνατεί πολλές φορές να τα περιγράψει καθώς επιτρέπει είτε μία σταθερή είτε μία μονότονη συνάρτηση κινδύνου. Η piecewise εκθετική κατανομή είναι λοιπόν μία επέκταση της εκθετικής κατανομής όπου ο κίνδυνος παραμένει σταθερός σε προκαθορισμένα χρονικά διαστήματα. Υιοθετείται δηλαδή για κάθε χρονικό διάστημα ένα ξεχωριστό εκθετικό μοντέλο.

Η πιθανότητα επιβίωσης έπειτα του χρόνου t , αν το t ανήκει στο i -στο διάστημα ($s_{i-1} \leq t < s_i$) είναι:

$$S(t) = \exp \left\{ -\lambda_i(t - s_{i-1}) - \sum_{j=1}^{i-1} \lambda_j(s_j - s_{j-1}) \right\}$$

και η συνάρτηση κινδύνου ορίζεται ως

$$\lambda(t) = \sum_{j=1}^J \lambda_j I_j(t), \quad \text{όπου } I_i(t) = 1 (s_{j-1} \leq t < s_j)$$

2.3 Εκτίμηση της συνάρτησης επιβίωσης και της συνάρτησης αθροιστικού κινδύνου

Η εκτίμηση της συνάρτησης επιβίωσης (ή κινδύνου) μπορεί να επιτευχθεί είτε καθορίζοντας ένα παραμετρικό μοντέλο για την συνάρτηση κινδύνου $\lambda(t)$ με βάση μία συγκεκριμένη συνάρτηση πυκνότητας πιθανότητας $f(t)$ και στη συνέχεια υπολογίζοντας τις παραμέτρους μέσω του εκτιμητή μεγίστης πιθανοφάνειας, είτε αναπτύσσοντας μία εμπειρική εκτίμηση της συνάρτησης επιβίωσης (μη-παραμετρική εκτίμηση). Όταν δεν υπάρχουν λογοκριμένες παρατηρήσεις στα δεδομένα τότε η εμπειρική εκτίμηση της συνάρτησης επιβίωσης $\hat{S}(t)$ είναι η αναλογία των ατομικών συναρτήσεων επιβίωσης με χρόνους συμβάντων μεγαλύτερους από t . Ωστόσο όταν υπάρχουν λογοκριμένες παρατηρήσεις η εκτίμηση $\hat{S}(t)$ δεν αποτελεί μία καλή εκτίμηση της πραγματικής συνάρτησης επιβίωσης. Επομένως άλλες μη παραμετρικές μέθοδοι χρησιμοποιούνται ώστε να ληφθεί υπόψη η λογοκρισία (μέθοδος πίνακας επιβίωσης και εκτιμητής Kaplan Meier).

Σε αντίθεση με τις παραμετρικές μεθόδους όπου γίνεται η υπόθεση ότι τα δεδομένα ακολουθούν μία συγκεκριμένη ειδική κατανομή, στις μη-παραμετρικές δεν υπάρχει καμία τέτοια προϋπόθεση. Γενικά οι μη-παραμετρικές μέθοδοι χρησιμοποιούνται σε μεγαλύτερη συχνότητα στην ανάλυση των δεδομένων

επιβίωσης καθώς είναι λιγότερο περιορισμένες και μπορούν να εφαρμοστούν στα δεδομένα οποιασδήποτε κατανομής.

2.3.1 Εκτιμητής Kaplan Meier

Αναμφίβολα η πιο διαδεδομένη και ευρέως χρησιμοποιούμενη μέθοδο μη παραμετρικής εκτίμησης της συνάρτησης επιβίωσης είναι ο εκτιμητής Kaplan Meier. Συχνά αναφέρεται και ως εκτιμητής ορίου γινομένου (product limit estimator). Αρχικά υπενθυμίζοντας ότι η συνάρτηση επιβίωσης είναι απλώς η πιθανότητα ένα υποκείμενο να επιβιώσει πέραν της χρονικής στιγμής t , η εμπειρική συνάρτηση επιβίωσης όταν δεν τίθενται θέματα λογοκρισίας είναι

$$S(t) = \frac{\text{αριθμός των ατόμων όπου } T \geq t}{\text{συνολικό μέγεθος δείγματος}}$$

Ωστόσο η χαμένη πληροφορία σε δεδομένα επιβίωσης αποτελεί μία συνηθισμένη κατάσταση, αν αναλογιστεί κανείς ότι τα δεδομένα αυτά στην πλειοψηφία των περιπτώσεων προέρχονται από μεγάλες περιόδους παρακολούθησεων όπου για ποικίλους λόγους χάνεται πληροφορία σχετικά με το γεγονός ενδιαφέροντος. Το 1958 οι Kaplan και Meier πρότειναν ένα μη-παραμετρικό τρόπο υπολογισμού της πιθανότητας επιβίωσης (Kaplan & Meier, 1958). Η ιδέα τους αυτή στηριζόταν στη θεμελιώδη έννοια της δεσμευμένης πιθανότητας καθώς και στον πολλαπλασιαστικό νόμο των πιθανοτήτων. Σύμφωνα με αυτό για n γεγονότα A_1, A_2, \dots, A_n ισχύει

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1)P(A_2|A_1) \dots P(A_n|A_1, A_2, \dots, A_{n-1})$$

Τα δεδομένα στην ανάλυση επιβίωσης περιλαμβάνουν πληροφορίες που αφορούν το χρόνο που πραγματοποιείται το συμβάν ή η λογοκρισία καθώς και ένα δείκτη (indicator) που καθορίζει τι από τα δύο έχει συμβεί. Πώς όμως υπολογίζεται ο εν λόγω εκτιμητής;

Έστω ότι m είναι το μέγεθος δείγματος και n είναι τα παρατηρούμενα γεγονότα σε μία δοκιμή. Τότε

t_1, \dots, t_n είναι οι ακριβείς χρόνοι που συνέβησαν τα γεγονότα,

r_j είναι ο αριθμός των ατόμων σε κίνδυνο τη χρονική στιγμή t_j , $j = 1, \dots, n$ (αυτό σημαίνει ότι τα άτομα αυτά είχαν επιβιώσει μέχρι τότε και είναι πιθανό να αποτύχουν στο χρόνο t_j ή αργότερα) καθώς και

d_j είναι ο αριθμός των αποτυχιών τη χρονικά στιγμή t_j

Η πιθανότητα επιβίωσης για το πρώτο χρονικό διάστημα είναι εύκολο να υπολογιστεί καθώς ισούται με την αναλογία των ατόμων που επιβίωσαν τουλάχιστον έως τη χρονική στιγμή t_1 . Επομένως

$$\hat{S}(t) = P(T \geq t_1) = 1, \quad \text{για } t \in [0, t_1]$$

Γενικά για $t \in (t_k, t_{k+1}]$ ισχύει

$$\begin{aligned}\hat{S}(t) &= P(T \geq t_{k+1}) \\ &= P(T \geq t_1, \dots, T \geq t_k, T \geq t_{k+1}) \\ &= P(T \geq t_1) \prod_{j=1}^k P(T \geq t_{j+1} | T \geq t_j) \\ &= \prod_{j=1}^k \{1 - P(T = t_j | T \geq t_j)\}\end{aligned}$$

όπου $P(T \geq t_1) = 1$ και

$$P(T = t_j | T \geq t_j) = \frac{d_j}{r_j} \quad \text{για κάθε } j \in \{1, \dots, n\}$$

Άρα

$$\hat{S}(t) = \prod_{j=1}^k \left(1 - \frac{d_j}{r_j}\right), \quad \text{για } t \in (t_k, t_{k+1}], \quad k \in \{1, \dots, n\}$$

Η μέθοδος Kaplan Meier χρησιμοποιείται όταν είναι γνωστοί οι ακριβείς χρόνοι πραγματοποίησης των γεγονότων ή της λογοκρισίας, σε αντίθετη περίπτωση όπου είναι γνωστά μόνο τα διαστήματα των συμβάντων χρησιμοποιείται μία άλλη μέθοδος που θα περιγραφεί στη συνέχεια (μέθοδος lifetable). Για το σκοπό αυτά τα διαστήματα που επιλέγονται είναι πολύ μικρά ώστε να διακρίνουν κάθε μεμονωμένο συμβάν. Ακόμη πρέπει να αναφερθεί ότι εάν οι χρόνοι αποτυχίας μετρηθούν ακριβώς, οι αποτυχίες θα συμβούν όλες σε διαφορετικούς χρόνους, αλλά αν μετρηθούν π.χ με βάση τον πλησιέστερο μήνα τότε μπορεί να υπάρχουν πολλές αστοχίες ταυτόχρονα. Στην περίπτωση αυτή η πιθανότητα επιβίωσης υπό όρους υπολογίζεται διαιρώντας τον αριθμό των αποτυχιών σε αυτό το χρονικό διάστημα με το συνολικό αριθμό των ατόμων που βρίσκονται σε κίνδυνο λίγο πριν το χρόνο αποτυχίας. Τέλος πρέπει να γίνει αντιληπτό ότι οι λογοκρισίες δεν επηρεάζουν την εκτίμηση του $S(t)$ αλλά συμβάλλουν μειώνοντας τον αριθμό των ατόμων που κινδυνεύουν στον επόμενο χρόνο αποτυχίας.

Συνοπτικά τα βήματα του αλγορίθμου για την εκτίμηση της πιθανότητας επιβίωσης περιορίζονται στα εξής:

Βήμα 1

Ταξινόμηση των n χρόνων που παρατηρήθηκαν σε αύξουσα σειρά. Εάν υπάρχουν επικαλύψεις, οι λογοκρισίες τοποθετούνται μετά τους θανάτους.

Βήμα 2

Ανάθεση βαρών $\left(\frac{1}{n}\right)$ σε κάθε χρόνο

Βήμα 3

Προχωρώντας από αριστερά προς τα δεξιά, κάθε φορά που εμφανίζεται λογοκριμένη παρατήρηση προσδίδεται η συνεισφορά της στις επόμενες παρατηρήσεις στα δεξιά.

Βήμα 4

Εκτιμάται η συνάρτηση επιβίωσης $\hat{S}(t)$, αφαιρώντας το τελικό βάρος για το χρόνο j από το \hat{S}_{j-1} .

Για την ανάλυση αυτή γίνονται οι παρακάτω τρεις παραδοχές

- 1) Γίνεται η υπόθεση ότι σε κάθε χρονική στιγμή οι ασθενείς που λογοκρίνονται έχουν τις ίδιες προοπτικές επιβίωσης με εκείνους που συνεχίζουν να παρακολουθούνται
- 2) Θεωρείται ότι οι πιθανότητες επιβίωσης είναι οι ίδιες για άτομα που εισήχθησαν είτε νωρίς είτε αργά στη μελέτη
- 3) Γίνεται η υπόθεση ότι το γεγονός συμβαίνει σε προκαθορισμένη χρονική στιγμή. Έτσι δημιουργούνται κάποια προβλήματα σε ορισμένες συνθήκες που είναι γνωστό μόνο το διάστημα στο οποίο πραγματοποιήθηκε. Αυτό συμβαίνει συχνά σε μελέτες όπου τα γεγονότα εμφανίζονται σε τακτικές εξετάσεις (και μία πιθανή λύση για τη μεγαλύτερη ακρίβεια των αποτελεσμάτων θα ήταν η συχνότερη παρακολούθηση των υποκειμένων).

Τέλος η μόνη προϋπόθεση της μεθόδου KM που πρέπει να αναφερθεί και να λαμβάνεται υπόψη είναι ότι οι λογοκριμένες παρατηρήσεις πρέπει να είναι μη πληροφοριακές καθώς και να υπάρχουν αρκετά σε πλήθος δεδομένα σε κάθε μία καμπύλη.

Ο διάμεσος χρόνος επιβίωσης (**median survival time**), είναι ο χρόνος t τέτοιος ώστε $S(t) = \frac{1}{2}$.

Ωστόσο η συνάρτηση επιβίωσης μπορεί να μην είναι συνεχής στο $\frac{1}{2}$. Τότε η διάμεσος είναι η μικρότερη t τέτοια ώστε $S(t) \leq \frac{1}{2}$. Όταν ο υπολογισμός του μέσου χρόνου επιβίωσης γίνεται με τη μη-παραμετρική μέθοδο Kaplan Meier, η τιμή του αντιστοιχεί στο χρόνο όπου η καμπύλη KM διασχίζει τη γραμμή του 50%. Στις περιπτώσεις όπου η καμπύλη δεν έχει φτάσει σε αυτό το σημείο, τότε ο εκτιμητής δεν μπορεί να υπολογιστεί.

2.3.2 Μέθοδος Lifetable

Η μέθοδος life table (πίνακες επιβίωσης) αποτελεί μία από τις αρχικές μεθόδους περιγραφής χρόνων δεδομένων επιβίωσης που εφαρμόστηκε ευρέως στον αναλογισμό και στη δημογραφία και

χαρακτηρίζεται ως μία από τις παλαιότερες τεχνικές στη μέτρηση της θνησιμότητας. Στηρίζεται στο διαχωρισμό της περιόδου παρακολούθησης σε ίσα χρονικά διαστήματα και στην εκτίμηση των πιθανοτήτων επιβίωσης (εξ ου η αθροιστική επιβίωση) σε κάθε ένα από τα διαστήματα. Η αθροιστική συνάρτηση επιβίωσης $S(i)$ (η οποία ορίζεται ως η πιθανότητα επιβίωσης στο τέλος του i -στου διαστήματος) στο τέλος ενός καθορισμένου διαστήματος δίνεται από τα ποσοστά επιβίωσης των καθορισμένων διαστημάτων για όλα τα διαστήματα συμπεριλαμβανομένου και αυτού. Ελλείψει λογοκρισίας, η αναλογία των επιζώντων για κάθε διάστημα υπολογίζεται ως εξής

$$p = \frac{I - d}{I}$$

όπου d : είναι ο αριθμός των συμβάντων που παρατηρήθηκαν κατά τη διάρκεια του διαστήματος και

I : είναι ο αριθμός των ασθενών που ζουν στην αρχή του διαστήματος

Από την άλλη όταν υπάρχει λογοκρισία θεωρείται ότι αυτή εκδηλώνεται ομοιόμορφα σε όλο το διάστημα, έτσι ώστε κάθε άτομο με λογοκριμένο χρόνο επιβίωσης να βρίσκεται σε κίνδυνο για το μισό διάστημα κατά μέσο όρο. Ο πραγματικός αριθμός ασθενών που βρίσκονται σε κίνδυνο κατά τη διάρκεια του διαστήματος δίνεται από τον τύπο $I^* = I - \frac{W}{2}$ όπου W είναι ο αριθμός των λογοκριμένων παρατηρήσεων στη διάρκεια του εν λόγω διαστήματος. Έπειτα η εκτιμώμενη αναλογία επιβίωσης για κάθε διάστημα δίνεται από τον τύπο $p = \frac{I^* - d}{I^*}$.

Συμπερασματικά οι πίνακες επιβίωσης είναι μία «επέκταση» των κλασικών πινάκων συχνότητας (frequency tables) όταν στα δεδομένα υπάρχουν λογοκριμένες παρατηρήσεις. Στην ουσία η Kaplan Meier μέθοδος είναι η life table μέθοδος, όπου το μέγεθος του διαστήματος μειώνεται στο μηδέν, έτσι ώστε ο αριθμός των διαστημάτων να τείνει στο άπειρο. Κάθε διάστημα έχει αρκετό μήκος για ένα συμβάν. Και οι δύο λαμβάνουν υπόψη το χρόνο και τη λογοκρισία ενώ η μόνη τους διαφορά είναι ότι στην Kaplan Meier χρησιμοποιούνται οι ακριβείς χρόνοι αποτυχίας ενώ στην μέθοδο των πινάκων επιβίωσης, οι χρόνοι επιβίωσης ομαδοποιούνται σε διαστήματα.

2.3.3 Εκτιμητής Nelson-Aalen

Το 1972 προτάθηκε από τους Nelson-Aalen ένας μη παραμετρικός εκτιμητής, ο οποίος εκτιμά τη αθροιστική συνάρτηση κινδύνου για δεξιά λογοκριμένες παρατηρήσεις (Hosmer & Lemeshow, 1999). Ορίζεται ως το άθροισμα των συμβάντων ως τη χρονική στιγμή t_j προς τον αριθμό των ατόμων που βρισκόταν σε κίνδυνο πριν τη χρονική στιγμή j και μαθηματικά εκφράζεται ως εξής

$$\hat{\Lambda} = \sum_{j:\tau_j < t} \frac{d_j}{r_j}$$

Ακόμη συχνά χρησιμοποιείται για την αξιολόγηση της εφαρμογής διάφορων παραμετρικών μοντέλων Πιο αναλυτικά για να υπάρχουν ενδείξεις καλής προσαρμογής του μοντέλου στα δεδομένα θα πρέπει η γραφική απεικόνιση της εκτίμησης της αθροιστικής συνάρτησης κινδύνου με το χρόνο να είναι γραμμική.

2.4 Log-Rank test

Εφόσον έχουν περιγραφεί διάφοροι παραμετρικοί και μη-παραμετρικοί τρόποι εκτίμησης της πιθανότητας επιβίωσης ενός πληθυσμού, είναι αναγκαίο να αναφερθούν και τρόποι που προβαίνουν στην σύγκριση διαφορετικών πληθυσμών. Ωστόσο η αναφορά θα γίνει σε μη-παραμετρική σύγκριση καθώς είναι αρκετά στιβαρή (fairly robust), πιο αποτελεσματική σε σχέση με τις παραμετρικές δοκιμές και συχνά είναι απλή και διαισθητική. Πιο συγκεκριμένα η υπόθεση προς έλεγχο είναι η

$$H_0: S_1(t) = S_2(t)^\theta \quad [\leftrightarrow \text{ισοδύναμα } H_0: h_1(t) = \theta h_2(t)]$$

vs

$$H_1: S_1(t) \neq S_2(t)^\theta \quad [\leftrightarrow \text{ισοδύναμα } H_1: h_1(t) \neq \theta h_2(t)]$$

Αν λοιπόν ο ισχυρισμός της εναλλακτικής υπόθεσης είναι αληθής, τότε ο πληθυσμός της δεύτερης ομάδας θα έχει ομοιόμορφα μεγαλύτερους χρόνους επιβίωσης από αυτούς της πρώτης ομάδας. Ο έλεγχος αυτός μπορεί να πραγματοποιηθεί με το Log-rank τεστ (Peto, et al., 1977). Ωστόσο πριν τη χρησιμοποίηση της μεθόδου πρέπει απαραίτητα να ελεγχθεί η υπόθεση της αναλογικότητας των κινδύνων. Πως λειτουργεί όμως το συγκεκριμένο τεστ;

Το log-rank τεστ λαμβάνεται κατασκευάζοντας έναν πίνακα 2×2 (βλ. Πίνακα 2.1) σε κάθε διακριτό χρόνο αποτυχίας και συγκρίνοντας τα ποσοστά θανάτου μεταξύ των δύο ομάδων, υπό τον όρο του αριθμού των ατόμων που βρίσκονται σε κίνδυνο στις ομάδες. Η σύγκριση αρχικά ανάγεται στον έλεγχο της ανεξαρτησίας του αριθμού των αποτυχιών ανάμεσα στους δύο πληθυσμούς.

Πληθυσμός	Αριθμός γεγονότων	Αριθμός επιζώντων πριν τη χρονική στιγμή t_j	Σύνολο
1 ^{ος}	d_{1j}	$r_{1j} - d_{1j}$	r_{1j}
2 ^{ος}	d_{2j}	$r_{2j} - d_{2j}$	r_{2j}
Σύνολο	d_j	$r_j - d_j$	r_j

Πίνακας 1: Log-rank τεστ τη χρονική στιγμή t_j

Όπου

d_{1j} : ο αριθμός των αποτυχιών στον πληθυσμό 1 τη χρονική στιγμή t_j

d_{2j} : ο αριθμός των αποτυχιών στον πληθυσμό 2 τη χρονική στιγμή t_j

d_j : συνολικός αριθμός αποτυχιών

r_{1j} : ο αριθμός των ατόμων σε κίνδυνο στον πληθυσμό 1 τη χρονική στιγμή t_j

r_{2j} : ο αριθμός των ατόμων σε κίνδυνο στον πληθυσμό 2 τη χρονική στιγμή t_j

r_j : συνολικός αριθμός ατόμων σε κίνδυνο

Αν είναι όντως ανεξάρτητοι και δοθέντων των r_j, r_{1j}, r_{2j}, d_j τότε το d_{1j} ακολουθεί την υπεργεωμετρική κατανομή, καθώς όταν άτομο στον πρώτο πληθυσμό αποτυγχάνει είναι αδύνατο να ξανασυμβεί αυτό αργότερα. Συνεπώς οι υπολογισμοί της μέσης τιμής και της διακύμανσης του d_{1j} είναι εύκολοι.

$$e_{1j} = d_j \frac{r_{1j}}{r_j} \quad v_{1j} = \frac{d_j r_{1j} r_{2j} (r_j - d_j)}{r_j^2 (r_j - 1)}$$

Στη συνέχεια αθροίζεται η διαφορά μεταξύ των αναμενόμενων και των παρατηρούμενων τιμών

$$U_0 = \sum (d_{1j} - e_{1j}), \quad \text{με διακύμανση } Var(U_0) = \sum v_{1j} = V_0$$

Υπό τη μηδενική υπόθεση λοιπόν το στατιστικό τεστ αυτό ασυμπτωτικά ακολουθεί τη χ^2 κατανομή με ένα βαθμό ελευθερίας (d.f.)

$$\frac{U_0^2}{V_0} \sim \chi_1^2$$

Συμπερασματικά το log-rank τεστ χρησιμοποιείται για να ελέγξει εάν η διαφορά μεταξύ των χρόνων επιβίωσης μεταξύ δύο ομάδων είναι στατιστικά διαφορετική ή όχι, αλλά δεν επιτρέπει τη δοκιμή της επίδρασης άλλων επεξηγηματικών μεταβλητών.

2.5 Το μοντέλο του Cox

Το μοντέλο αναλογικού κινδύνου του Cox (Cox's proportional hazard model) είναι ανάλογο με ένα μοντέλο παλινδρόμησης και επιτρέπει τον έλεγχο της διαφοράς της επίδρασης άλλων ανεξάρτητων μεταβλητών μεταξύ των χρόνων επιβίωσης συγκεκριμένων ομάδων (Cox, 1979). Σε αυτό το μοντέλο η μεταβλητή απόκρισης (response variable) είναι ο κίνδυνος (hazard). Η συνάρτηση κινδύνου στο μοντέλο είναι παραγόμενο μιας αυθαίρετης συνάρτησης κινδύνου (baseline hazard function) με ένα σταθερό όρο (εκθετικός όρος), η οποία είναι ανεξάρτητη από το χρόνο t . Συνεπώς ο λόγος κινδύνου (hazard ratio) δεν εξαρτάται από το χρόνο. Ο υπολογισμός των παραμέτρων παλινδρόμησης δεν προϋποθέτουν τη γνώση ή

την εκτίμηση της αυθαίρετης συνάρτησης κινδύνου καθώς υπολογίζονται μέσω της μεθόδου της μεγίστης πιθανοφάνειας (MLE). Η μέθοδος αυτή λοιπόν δεν κάνει κάποια συγκεκριμένη υπόθεση για την κατανομή των χρόνων επιβίωσης γεγονός που την καθιστά πιο ευσταθής. Η μαθηματική έκφραση του μοντέλου γράφεται ως εξής:

$$h(t; Z) = h(t)e^{(bZ)} = h_0(t)e^{(b_1Z_1+b_2Z_2+\dots+b_pZ_p)} , \text{ όπου}$$

$h(t)$: είναι ο κίνδυνος τη χρονική στιγμή t

$Z_1, Z_2 \dots Z_p$: είναι οι επεξηγηματικές μεταβλητές και

$h_0(t)$: είναι η βασική συνάρτηση κινδύνου (baseline hazard function) η οποία δεν χρειάζεται να είναι πλήρως καθορισμένη και αντανακλά τον κίνδυνο που έχει κάθε ένα από τα δεδομένα, να υποστεί σε χρόνο t το προκαθορισμένο γεγονός, χωρίς καμία εξάρτηση από τις τιμές της μεταβλητής Z .

Ένας εναλλακτικός τρόπος γραφής του παραπάνω μοντέλου που χρησιμοποιείται ευρέως όταν δεν προμηνύεται η εκτίμηση του $h_0(t)$ είναι :

$$h_i(t) = h_0(t)e^{\beta_1x_{i1}+\dots+\beta_px_{ip}}$$

Ωστόσο όπως αναφέρθηκε προηγουμένως η ιδέα αυτή του Cox στηρίχθηκε στην μερική πιθανοφάνεια (partial likelihood), η οποία προϋποθέτει διακριτούς χρόνους γεγονότων ή λογοκρισίας, όπου αυτό δεν αληθεύει πάντα στην πραγματικότητα καθώς συχνά υπάρχουν επικαλυπτόμενοι (ties) χρόνοι επιβίωσης (Cox, 1975). Οι πιο συχνοί τρόποι προσαρμογής της πιθανοφάνειας για επικαλυπτόμενα δεδομένα είναι αυτοί του Efron (Efron, 1977) και του Breslow (Breslow, 1974). Όταν υπάρχουν λίγες επικαλύψεις η μέθοδος του Breslow εφαρμόζει αρκετά καλά. Τέλος πρέπει να αναφερθεί ότι ενώ το μοντέλο αναλογικού κινδύνου χειρίζεται συνεχείς μεταβλητές, δεν μπορεί να ανταπεξέλθει σε μεταβλητές που αλλάζουν στο χρόνο (time varying covariates), και χρειάζεται να επεκταθεί για να καλύψει αυτές τις περιπτώσεις..

2.5.1 Αναλογικότητα κινδύνων και παλινδρόμηση Cox

Το μοντέλο παλινδρόμησης του Cox και η υπόθεση της αναλογικότητας των κινδύνων είναι άρρηκτα συνδεδεμένα γι αυτό άλλωστε πολύ συχνά στη βιβλιογραφία αναφέρεται ως μοντέλο αναλογικών κινδύνων. Πράγματι η εγκυρότητα και η αξιοπιστία του μοντέλου ελέγχεται από τη συγκεκριμένη υπόθεση η οποία αναφέρει ότι ο λόγος των κινδύνων είναι σταθερός και δεν επηρεάζεται από το χρόνο. Το hazard

ratio (ο λόγος κινδύνου) ορίζεται ως κίνδυνος ενός υποκειμένου A σε σχέση με το υποκείμενο B και δίνεται από τη σχέση

$$HR = \frac{h_1(t)}{h_2(t)} = \frac{h_0(t)e^{\beta x_1}}{h_0(t)e^{\beta x_2}} = e^{\beta(x_1 - x_2)}$$

Πρέπει πάντα να ελέγχεται η υπόθεση της αναλογικότητας των κινδύνων πριν τη χρήση του μοντέλου καθώς εγκυμονούν κίνδυνοι οι εκτιμήσεις να είναι ασταθείς και μεροληπτικές εάν αυτή δεν επιβεβαιωθεί, έχοντας ως επακόλουθο τα λανθασμένα και παραπλανητικά συμπεράσματα τα οποία πρέπει να αποφεύγονται ειδικά σε ζητήματα που αφορούν την υγεία. Ο έλεγχος της αναλογικότητας μπορεί να ελεγχθεί με γραφικές μεθόδους καθώς και με τεστ καλής προσαρμογής.

Ανακεφαλαιώνοντας τόσο το log-rank τεστ όσο και το μοντέλο αναλογικού κινδύνου του Cox υποθέτουν ότι ο ρυθμός κινδύνου (hazard ratio) παραμένει σταθερός με την πάροδο του χρόνου.

2.6 Περιορισμένος Μέσος Χρόνος Επιβίωσης (RMST)

Όπως έχει αναφερθεί και προηγουμένως στην ανάλυση επιβίωσης συχνά (αν όχι πάντα) δεν παρατηρούνται οι πλήρεις χρόνοι επιβίωσης όλων των συμμετεχόντων. Χαρακτηριστικό παράδειγμα αποτελεί μία κλινική δοκιμή η οποία εξετάζει το χρόνο έως την επέλευση του θανάτου από μία ασθένεια, όπου είναι σχεδόν αδύνατο να παρατηρηθεί ο θάνατος όλων των συμμετεχόντων, λόγω του ότι οι συμμετέχοντες μπορούν να αποχωρήσουν οποιαδήποτε στιγμή και λόγω των χρονικών περιορισμών που εφαρμόζονται. Γίνεται αντιληπτό λοιπόν ότι σε πολλές περιπτώσεις καθίσταται αδύνατη η εκτίμηση του μέσου χρόνου επιβίωσης και συχνά χρησιμοποιείται η διάμεσος ή κάποιο εκατοστημόριο.

Ακόμη στις περιπτώσεις όπου δεν πληρείται η αναλογικότητα των κινδύνων η χρήση εκτιμήσεων σταθμισμένων παραμέτρων για την εξαγωγή ενός συνοπτικού μέτρου έχει προκαλέσει διάφορες αντιδράσεις ανάμεσα στους ερευνητές, με κύριους τους Royston και Parmar (Royston & Parmar, 2011).

Σε συνέχεια των παραπάνω έχει αρχίσει να συζητείται ξανά μία ποσότητα που είχε προταθεί για πρώτη φορά το 1949 από τον Irwin και έκτοτε είχε αποσιωπηθεί (Irwin, 1949). Πρόκειται για τον περιορισμένο μέσο χρόνο επιβίωσης (RMST) όπου είναι ο αναμενόμενος χρόνος που δαπανάται χωρίς συμβάντα για ένα υποκείμενο που παρακολουθείται για ένα καθορισμένο χρονικό διάστημα. Με άλλα λόγια το RMST σε ένα χρονικό σημείο ερμηνεύεται ως ο μέσος χρόνος επιβίωσης ενός ασθενούς μέχρι το συγκεκριμένο σημείο. Μέσω της μαθηματικής μοντελοποίησης εάν $R = \min(T, \tau)$ μία μη αρνητική τυχαία

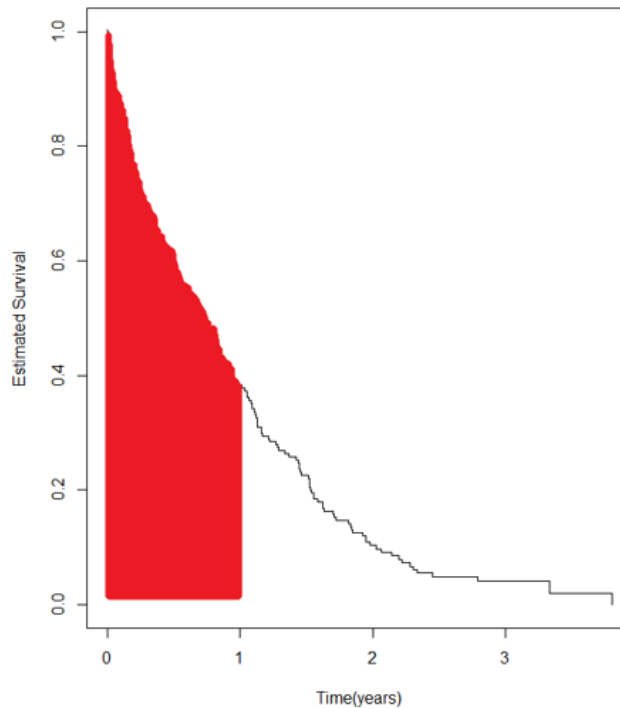
μεταβλητή, όπου T είναι η μεταβλητή που καθορίζει το χρόνο επιβίωσης ενός υποκειμένου και τ είναι το καθορισμένο χρονικό σημείο, τότε το RMST δίνεται από τη σχέση:

$$RMST(\tau) = E[R] = E[\min(T, \tau)]$$

και εφόσον το $R \in [0, \tau]$ ισχύει

$$\begin{aligned} E[R] &= \int_0^{\tau} [1 - F_R(u)] du = \int_0^{\tau} [1 - P(R \leq u)] du \\ &= \int_0^{\tau} [1 - P(\min(T, \tau) \leq u)] du = \int_0^{\tau} [1 - P(T \leq u)] du \\ &= \int_0^{\tau} [1 - F_T(u)] du = \int_0^{\tau} S_T(u) du \end{aligned}$$

Συνεπώς η συνάρτηση επιβίωσης και ο περιορισμένος μέσος χρόνος επιβίωσης είναι άμεσα συνδεδεμένοι αφού ισούται με το εμβαδό κάτω από την καμπύλη έως το καθορισμένο σημείο τ (Εικόνα 4).



Γράφημα 2: RMST (η κόκκινη περιοχή κάτω από την καμπύλη επιβίωσης) όταν το $\tau = 1$

Επιπρόσθετα όταν το $\tau \rightarrow \infty$ τότε το $RMST_{\tau} \rightarrow \mu$, όπου μ είναι ο μέσος χρόνος επιβίωσης που δίνεται από τη σχέση $\mu = \int_0^{\infty} t f(t) dt$ και πολλές φορές είναι δύσκολο να εκτιμηθεί λόγω ύπαρξης της λογοκρισίας. Εφόσον λοιπόν το RMST είναι μία αυξανόμενη συνάρτηση του καθορισμένου σημείου τ , έχει ως

αποτέλεσμα ο περιορισμένος μέσος χρόνος επιβίωσης να είναι πάντα μικρότερος από το μέσο χρόνο επιβίωσης.

2.6.1 Εκτίμηση RMST

Ένας φυσικός εκτιμητής του RMST δίνεται αν εκτιμήσουμε τη συνάρτηση επιβίωσης $S(t)$ με τη μέθοδο Kaplan Meier. Ο εκτιμητής δίνεται από τη σχέση

$$RMST = \int_0^{\tau} \hat{S}(t) dt, \text{ όπου } \hat{S}(t) \text{ είναι ο εκτιμητής Kaplan Meier.}$$

Μία άμεση μέθοδο υπολογισμού έχει προταθεί από τους Wei και συνεργάτες (Wei, et al., 2015) με τον ακόλουθο τύπο, όπου ενσωματώνουν τις εκτιμήσεις KM μέχρι τη χρονική στιγμή τ

$$RMST_{\tau} = \sum_{j=1}^k \hat{S}(t_{j-1})[t_j - t_{j-1}] + \hat{S}(t_m)[\tau - t_m]$$

όπου k είναι ο αριθμός των γεγονότων,

t_j : είναι ο χρόνος μέχρι το j -στο γεγονός, με $t_0 = 0$ και

$\hat{S}(t_{j-1})$: είναι ο εκτιμητής Kaplan-Meier τη χρονική στιγμή $j - 1$

Ο παραπάνω τρόπος υπολογισμού αντανακλά μία μη-παραμετρική μέθοδο, ωστόσο υπάρχουν και άλλες μέθοδοι όπως η Jackknife μέθοδος (Andersen & Perme, 2010) με χρήση ψευδο-τιμών καθώς και η ευέλικτη παραμετρική μέθοδος των Royston και Parmar που είναι εύκολο να επεκταθεί και σε περιπτώσεις μη ισχύς της αναλογικότητας (Royston & Parmar, 2011). Ωστόσο με βάση το γεγονός ότι αν εκτιμηθεί η συνάρτηση επιβίωσης τότε ο υπολογισμός του RMST είναι μία εύκολη διαδικασία έχουν προταθεί πολλές ακόμη προσεγγίσεις.

ΔΙΑΦΟΡΑ RMST (RMST difference)

Όταν θέλουμε να πραγματοποιήσουμε κάποια σύγκριση, μεταξύ δύο θεραπειών για παράδειγμα, μπορεί να χρησιμοποιηθεί η διαφορά των RMSTs τα οποία έχουν υπολογιστεί στο ίδιο χρονικό σημείο τ για τα δύο σκέλη.

$$RMST_{difference} = RMST_{new\ trt} - RMST_{standard}$$

όπου $RMST_{new\ trt}$: ο περιορισμένος μέσος χρόνος επιβίωσης για τη νέα θεραπεία μέχρι το προκαθορισμένο χρονικό σημείο τ

$RMST_{standard}$: ο περιορισμένος μέσος χρόνος επιβίωσης για την τυπική θεραπεία μέχρι το προκαθορισμένο χρονικό σημείο τ

Η γενική μορφή της δοκιμής παίρνει τη μορφή

$$H_0: RMST_1(\tau) = \dots = RMST_k(\tau) \quad \text{vs} \quad H_1: RMST_{i_1}(\tau) \neq RMST_{i_2}(\tau) \text{ για } i_1, i_2 \in \{1, \dots, k\}$$

Το στατιστικό τεστ που χρησιμοποιείται για τον έλεγχο, όπου $RMST(\tau) = \xi$ είναι το

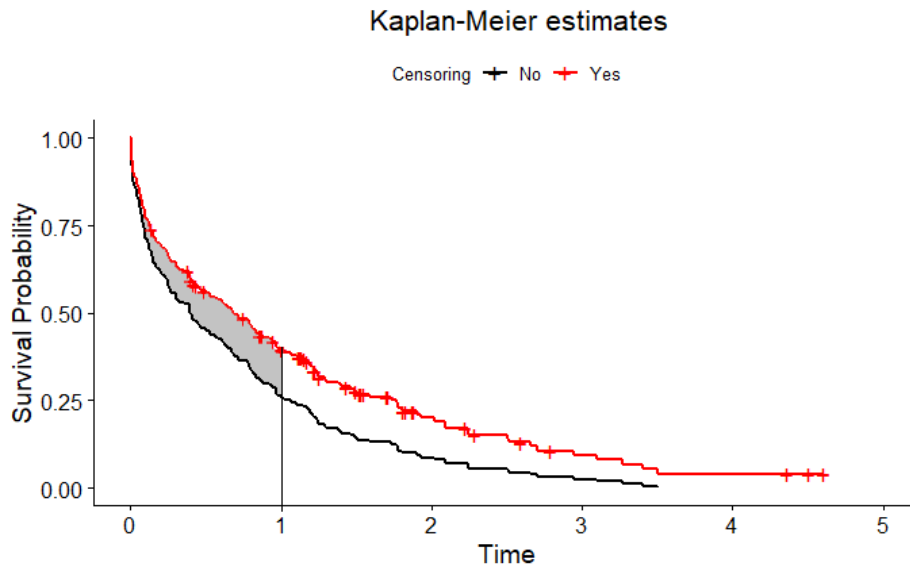
$$\xi'[D'(D\Sigma D')^{-1}D]\xi$$

και ασυμπτωτικά ακολουθεί τη χ^2 κατανομή με τους βαθμούς ελευθερίας να ισούνται με τη βαθμίδα του $D\Sigma D'$ [$\text{rank}(D\Sigma D')$]. Όπου Σ είναι ο πίνακας συνδιακύμανσης του διανύσματος $\xi(\tau) = (\xi_1(\tau), \dots, \xi_k(\tau))'$ και D επίσης ένα πίνακας $(K - 1) \times K$ του οποίου η j -στη γραμμή αντιστοιχεί στο $e_j - e_{j+1}$, με e_j ένα διάνυσμα με μονάδα στη θέση j και μηδενικά στις υπόλοιπες θέσεις.

Ακόμη αν οι εκτιμήσεις του RMST γίνουν με τη χρήση του εκτιμητή Kaplan Meier (Πανάγου, 2014) τότε ισχύει ότι

$$\sqrt{n}(\hat{\xi} - \xi) \rightarrow N(0, \sigma^2)$$

Με άλλα λόγια η διαφορά RMST συγκρίνει τις περιοχές κάτω από τις δύο καμπύλες επιβίωσης για τις ομάδες παρέμβασης και ελέγχου για ένα καθορισμένο διάστημα. Για παράδειγμα στο παρακάτω γράφημα απεικονίζεται με γκρι σκιαγράφηση η διαφορά RMST ανάμεσα στις δύο ομάδες όταν ο προκαθορισμένος χρόνος ισούται με τη μονάδα.



Γράφημα 3: Διαφορά RMST για $\tau = 1$

Όσο πιο μεγάλη είναι αυτή η διαφορά, τόσο πιο πιθανό είναι να απορριφθεί η μηδενική υπόθεση. Αυτή η αντίθεση αντιστοιχεί στη μέση χρονική αναβολή του αποτελέσματος στη μία ομάδα σε σύγκριση με την άλλη, όπου κάθε RMST μιας ομάδας ποσοτικοποιεί τη μέση καθυστέρηση στο συμβάν σε συγκεκριμένο χρονικό ορίζοντα (Kloecker, et al., 2020).

2.6.2 Επιλογή του καθορισμένου σημείου τ

Το συνοπτικό μέτρο RMST καθώς και το τεστ διαφοράς RMST έχουν αρκετά πλεονεκτήματα κάποια από τα οποία αναφέρονται στη συνέχεια, ωστόσο και τα δύο έρχονται αντιμέτωπα με τη δυσκολία καθορισμού του χρονικού σημείου τ . Γενικά δεν υπάρχει ένας γενικός κανόνας για την επιλογή του και αυτό θέλει προσοχή γιατί μπορεί να οδηγήσει σε μεροληπτικά αποτελέσματα και αποφάσεις. Ωστόσο αξίζει να αναφερθεί ότι τα πιο συχνά επιλεγόμενα τ βρίσκονται στο τέλος της παρακολούθησης και πιο συγκεκριμένα ορίζονται ως το ελάχιστο του μέγιστου γεγονότος που παρατηρήθηκε στις δύο ομάδες. Συνεπώς για τη μία ομάδα θα χρησιμοποιηθούν στο έπακρο οι διαθέσιμες πληροφορίες ενώ στην άλλη θα υπάρχει κάποια απώλεια. Το γεγονός αυτό είναι αναπόφευκτο αν δεν ληφθεί υπόψη η εκτίμηση της άλλης ομάδας. Εν τούτοις η προσέγγιση αυτή χρησιμοποιεί όσο το δυνατό περισσότερη πληροφορία από αυτή που είναι διαθέσιμη, χωρίς τον κίνδυνο της παρέκτασης (extrapolation). Ακόμη ένας τρόπος επιλογής του σημείου είναι μέσω της σχέσης του με το υπό μελέτη γεγονός (Uno, et al., 2015). Τέλος μία ακόμα προσέγγιση που αναφέρεται στη βιβλιογραφία είναι η επιλογή του τ με τέτοιο τρόπο («data maturity analysis») ώστε να ελαχιστοποιείται το μέγεθος του δείγματος που απαιτείται για συγκεκριμένο επίπεδο ισχύος και στατιστικής σημαντικότητας (Royston & Parmar, 2013).

2.6.3 Πλεονεκτήματα και περιορισμοί

Η χρήση του RMST αναμφίβολα έχει αρκετά πλεονεκτήματα. Μερικά από αυτά συνοψίζονται στο ότι παρέχει μία κλινικά σημαντική περίληψη μεταξύ των ομάδων σύγκρισης καθώς δίνει τη δυνατότητα σύγκρισης δύο καμπυλών επιβίωσης οποιαδήποτε στιγμή. Ακόμη δεν απαιτεί καμία υπόθεση για την συμπεριφορά των καμπυλών προκειμένου να υπολογιστεί, γεγονός που του προσδίδει ένα μεγάλο πλεονέκτημα αν αναλογιστεί κανείς ότι ο λόγος κινδύνου (HR), το πιο δημοφιλές στατιστικό μέτρο χρειάζεται την υπόθεση των αναλογικών κινδύνων για την εξαγωγή έγκυρων και αξιόπιστων αποτελεσμάτων. Επιπλέον προσφέρει μια πιο σταθερή εκτίμηση από τη διάμεση τιμή στις μελέτες όπου το ενδιαφέρον στρέφεται στο χρόνο επιβίωσης. Τέλος ο περιορισμένος μέσος χρόνος επιβίωσης είναι χρήσιμος στην αξιολόγηση της ασφάλειας μιας παρέμβασης ή του αποτελέσματος της θεραπείας σε δοκιμές μη-κατωτερότητας (Uno, et al., 2015).

Από την άλλη μεριά βέβαια έχει έναν ισχυρό περιορισμό όπως αναφέρθηκε και προηγουμένως. Είναι απαραίτητος ο προκαθορισμός του χρονικού σημείου ενδιαφέροντος. Ιδανικά αυτό θα προκαθορίζεται μέσω κλινικών εκτιμήσεων κατά το σχεδιασμό της μελέτης, ενώ πολύ συχνά επιλέγεται και το τελευταίο γεγονός που παρατηρήθηκε ή η λογοκρισία (McCaw, Yin, & Wei, 2019). Επιπλέον η διαφορά RMST εξαρτάται από τη διάρκεια παρακολούθησης και τα ποσοστά εκβάσεων στις δύο ομάδες τα οποία καθορίζονται από τα χαρακτηριστικά των συμμετεχόντων και την επίδραση της παρέμβασης (Kloecker, et al., 2020). Συνεπώς δεν πρέπει να χρησιμοποιείται σε ετερογενής μελέτες.

2.6.4 Συνοπτικά μέτρα και Περιορισμένος Μέσος Χρόνος Επιβίωσης (RMST)

2.6.4.1. Λόγος Κινδύνου (Hazard Ratio)

Όπως έχει ήδη αναφερθεί το RMST σε σχέση με το λόγο κινδύνου (hazard ratio) δεν απαιτεί την υπόθεση της ύπαρξης αναλογικών κινδύνων. Ακόμη ενώ για το HR υπάρχουν στοιχεία για εσφαλμένη ερμηνεία που καταλήγουν σε υπερβολική ερμηνεία του κινδύνου ή της ωφέλειας, το RMST είναι λιγότερο επιρρεπές στην αβεβαιότητα (Kloecker, et al., 2020). Με άλλα λόγια όταν ο αριθμός των γεγονότων σε μία μελέτη είναι μικρός τότε το RMST παρουσιάζει στενότερα όρια στο διάστημα εμπιστοσύνης από το HR, καθώς επίσης έχει και καλύτερη αιτιολογική ερμηνεία. Τέλος ο λόγος κινδύνου δεν αναφέρεται συχνά σε σχέση με τη διάρκεια της παρακολούθησης, με αποτέλεσμα να οδηγεί σε εσφαλμένα συμπεράσματα ότι το όφελος εκτείνεται και πέρα από τα χρονικά όρια διεκπεραίωσης της μελέτης ενώ από την άλλη το RMST δεν δύναται να υπολογιστεί εάν δεν καθοριστεί ένα συγκεκριμένο χρονικό σημείο.

2.6.4.2. Διάμεσος (Median)

Σε γενικές γραμμές το RMST δεν έχει μεγάλη απόκλιση από τη γνωστή διάμεσο, ωστόσο διακρίνονται τρία βασικά πλεονεκτήματα (Messori, et al., 2021). Αρχικά το RMST είναι αριθμητικά πολύ πιο σταθερό από τη διάμεσο αφού εξετάζει ολόκληρη την καμπύλη, ενώ η διάμεσος εξετάζει τα χρονικά σημεία εκείνα όπου η υπολειπόμενη επιβίωση μειώνεται από μεγαλύτερη σε μικρότερη του 50%. Ακόμη το RMST σε αντίθεση με τη διάμεσο λαμβάνει υπόψη το τμήμα της καμπύλης επιβίωσης που έπεται της επίτευξης της ενδιάμεσης τιμής και συνεπώς εξηγεί την ύπαρξη της σχεδόν επίπεδης εικόνας στη δεξιά ουρά της καμπύλης. Τέλος, στις καμπύλες επιβίωσης όπου στο τελευταίο χρονικό σημείο της παρακολούθησης παραμένουν πάνω από 50% άτομα στην υπολειπόμενη επιβίωση το RMST μπορεί να τις χειριστεί ενώ η διάμεσος όχι.

Κεφάλαιο 3^ο-Προτεινόμενοι μέθοδοι ανακατασκευής δεδομένων

Έχοντας ως γνώμονα την εποχή των μεγάλων δεδομένων (big data) όπου διανύουμε αλλά και την πληθώρα πληροφοριών που υπάρχει, η ανάγκη για διεξαγωγή μετα-αναλύσεων ολοένα και μεγαλώνει. Το γεγονός αυτό εξηγείται ως απόρροια της ασταίρευτης ποσότητας πληροφοριών που από τη μία είναι εύκολα προσβάσιμη στις πλείστες των περιπτώσεων και από τη άλλη είναι συγκεχυμένη, με αποτέλεσμα να δημιουργεί μεγαλύτερη σύγχυση στους ερευνητές σχετικά με την εξαγωγή συμπερασμάτων. Προς αυτή την κατεύθυνση λοιπόν, οι ορθά διεξαγόμενες μετα-αναλύσεις στοχεύουν στην ανάλυση των δεδομένων όλων των ασθενών που συμμετέχουν στις επιλεγόμενες προς ανάλυση μελέτες, αυξάνοντας την ισχύ, με στόχο την επίτευξη της μέγιστης ακρίβειας και σταθερότητας των αποτελεσμάτων. Ωστόσο οι περισσότερες μετα-αναλύσεις βασίζονται στην εξαγωγή συνοπτικών δεδομένων από τη δημοσιευμένη βιβλιογραφία μιας και πραγματοποιούνται σε πολύ πιο σύντομο χρονικό διάστημα, δεν έχουν μεγάλο κόστος και αποτελούν πρόδρομο για το σχεδιασμό νέων δοκιμών. Τα κύρια συνοπτικά δεδομένα άμεσου ενδιαφέροντος είναι ο λογαριθμισμένος λόγος κινδύνου (log hazard ratio), που χρησιμοποιείται για τη σύγκριση δύο καμπυλών επιβίωσης (εφόσον λαμβάνει υπόψη τόσο τη λογοκρισία όσο και το χρόνο μέχρι την έλευση του γεγονότος), καθώς και η διακύμανσή του ($var(\log HR)$).

Στο σημείο αυτό είναι αναγκαίο να αναφερθεί ότι παρόλο που υπάρχει η θεώρηση ότι για τη χρήση του $\log HR$ είναι απαραίτητη η τήρηση των αναλογικών κινδύνων, δεν είναι απόλυτα σωστή. Αυτό γιατί ο λόγος κινδύνου αποτυπώνει τη διαφορά μεταξύ δύο καμπυλών Kaplan Meier και αντανakλά τη συνολική μείωση του κινδύνου για έλευση θανάτου ανάμεσα στην ομάδα που λαμβάνει μια θεραπεία σε σύγκριση με την ομάδα ελέγχου κατά την περίοδο της παρακολούθησης (Parmar, et al., 1998). Γίνεται λοιπόν αντιληπτό ότι σαφώς είναι πιο εύκολα ερμηνεύσιμο όταν ισχύει η αναλογικότητα των κινδύνων ωστόσο αποτελεί ένα χρήσιμο στατιστικό μέτρο και όταν δεν ισχύει αυτό.

Στόχος λοιπόν των μετα-αναλύσεων είναι η συγκέντρωση των εκτιμήσεων όλων αυτών των στατιστικών από κάθε δοκιμή και μετέπειτα ο συνδυασμός τους μέσω μιας στρωματοποιημένης ανάλυσης. Περιγράφοντας το μαθηματικά και για λόγους απλότητας γίνεται η παραδοχή ότι σε κάθε δοκιμή συγκρίνεται μία ομάδα θεραπείας (research arm) με μία ομάδα ελέγχου (control arm). Έτσι, έστω ότι σε μία μετα-ανάλυση έχουν επιλεγθεί k δοκιμές, τότε για κάθε δοκιμή i , όπου $i = 1, \dots, k$ λαμβάνονται ο λογαριθμισμένος λόγος κινδύνου ($\log HR_i$) καθώς και η διακύμανσή του ($var(\log HR_i)$). Στη συνέχεια δύναται να υπολογιστούν ο συνολικός λογαριθμισμένος λόγος κινδύνου ($\log HR$), που είναι στην ουσία ένας σταθμισμένος μέσος των μεμονωμένων $\log HR_i$, με βάρη αντιστρόφως ανάλογα της διακύμανσής τους και η συνολική διακύμανση ($var(\log HR)$) ως εξής:

$$\log(HR) = \frac{\sum_{i=1}^k \frac{\log HR_i}{\text{var}(\log HR_i)}}{\sum_{i=1}^k \frac{1}{\text{var}(\log HR_i)}} \quad \text{και} \quad \text{var}[\log(HR)] = \left[\sum_{i=1}^k \frac{1}{\text{var}[\log(HR_i)]} \right]^{-1} \quad (3.1)$$

Ωστόσο δεν είναι λίγες οι δοκιμές όπου δεν αναφέρουν αυτά τα μέτρα και η αναζήτηση μεθόδων για την εκτίμησή τους θεωρείται αναγκαία.

3.1 Η προσέγγιση του Parmar

Ο Parmar και οι συνεργάτες του το 1998 παρουσίασαν μία σειρά μεθόδων για την εκτίμηση αυτών των στατιστικών μέτρων και μάλιστα εισηγήθηκαν μία σαφής ιεραρχία για τη χρησιμοποίηση των εν λόγω μεθόδων (Parmar, Torri, & Stewart, 1998).

- Αρχικά λοιπόν αναφέρουν ότι αν υπάρχουν διαθέσιμα δεδομένα, όπως ο αριθμός των παρατηρούμενων και των αναμενόμενων γεγονότων τόσο στην ομάδα θεραπείας όσο και στην ομάδα ελέγχου καθώς και η διακύμανση του λόγου κινδύνου, τότε η εκτίμηση των συνοπτικών μέτρων μπορεί να γίνει με χρήση άμεσων μεθόδων, όπου είναι και η προτιμητέα μέθοδος. Ωστόσο είναι ελάχιστες οι δοκιμές οι οποίες προσφέρουν όλη αυτή την πληροφορία.
- Στη συνέχεια και εφόσον δεν υπάρχουν διαθέσιμες οι παραπάνω τιμές κάνουν λόγο για χρήση έμμεσων μεθόδων εκτίμησης των ποσοτήτων που απουσιάζουν, χρησιμοποιώντας όσο το δυνατόν περισσότερες πληροφορίες προσφέρονται. Οι πληροφορίες αυτές μπορεί να αφορούν το διάστημα εμπιστοσύνης του λογαριθμισμένου λόγου κινδύνου για τον υπολογισμό της διακύμανσης, το συνολικό αριθμό των γεγονότων που παρατηρήθηκαν και στις δύο ομάδες καθώς και την τιμή σημαντικότητας (p-value) από ένα στατιστικό έλεγχο logrank ή Mantel-Haenszel, όπου έρχεται σε εξισορρόπηση με το εκτιμώμενο στατιστικό χ^2 , με σκοπό την εκτίμηση τόσο του logHR όσο και της διακύμανσής του.
- Τέλος γίνεται λόγος για χρήση της τιμής σημαντικότητας (p-value) με ανάλογο τρόπο από ένα μοντέλο αναλογικών κινδύνων όπου έχει γίνει στάθμιση για διάφορους προγνωστικούς παράγοντες.

Οι άμεσες και έμμεσες μέθοδοι όμως δύσκολα μπορούν να εφαρμοστούν στην πράξη καθώς σπάνια υπάρχουν επαρκή στατιστικά στοιχεία που να επιτρέπουν την εκτίμηση αυτών των συνοπτικών μέτρων και γι αυτό δεν θα αποτελέσουν θέματα περαιτέρω ανάλυσης στην εν λόγω εργασία σε αντίθεση με την τρίτη προτεινόμενη μέθοδο. Αυτή αναφέρεται στην εκτίμηση των συνοπτικών μέτρων από τις

καμπύλες επιβίωσης. Συνοπτικά γίνεται η εκτίμηση του log hazard ratio σε κάθε ένα από τα μη επικαλυπτόμενα διαστήματα στα οποία έχει χωριστεί ο άξονας του χρόνου και στη συνέχεια μέσω μιας στρωματοποιημένης ανάλυσης λαμβάνεται το συνολικό log hazard ratio.

Για να γίνει πιο κατανοητή αυτή η διαδικασία, μπορεί να χωριστεί στα εξής τέσσερα βήματα:

- i. Αρχικά για κάθε δοκιμή i ($i=1, \dots, k$), μέσω των καμπυλών Kaplan Meier καθορίζονται οι πιθανότητες επιβίωσης σε κάθε ένα από τα προκαθορισμένα μη επικαλυπτόμενα χρονικά σημεία t ($t = 1, \dots, T$). Ακόμη μέσω των διαθέσιμων πληροφοριών που υπάρχουν στο κείμενο γίνεται μία εκτίμηση για το ελάχιστο και το μέγιστο χρόνο παρακολούθησης των ασθενών. Στη συνέχεια είναι αναγκαία η εκτίμηση του αριθμού των ατόμων που βρίσκονται σε κίνδυνο τόσο στην αρχή όσο και στο τέλος σε κάθε ένα από τα εν λόγω διαστήματα. Ως προς θέματα συμβολισμού το t_s : αναφέρεται στην αρχή κάθε διαστήματος, το t_e στο τέλος του, ενώ το t αντιπροσωπεύει ολόκληρο το διάστημα $(t - 1, t)$.
- ii. Θεωρώντας ένα μοντέλο λογοκρισίας, για κάθε διάστημα υπολογίζεται ο αριθμός των ατόμων που βρίσκονται σε κίνδυνο στην ομάδα που λαμβάνει κάποια θεραπεία (R_r) (με ανάλογο τρόπο βρίσκεται και στην ομάδα ελέγχου - R_c) με τη βοήθεια της σχέσης (3.2)

$$R_r(t) = R_r(t - 1) - D_r(t - 1) - C_r(t) \quad (3.2)$$

όπου $R_r(t - 1)$ είναι ο αριθμός των ατόμων που βρισκόταν σε κίνδυνο στην ομάδα θεραπείας στο διάστημα $(t - 2, t - 1)$

$D_r(t - 1)$ είναι ο αριθμός των ατόμων που βίωσαν το γεγονός στην ομάδα θεραπείας στο διάστημα $(t - 2, t - 1)$ και

$C_r(t)$ είναι ο αριθμός των ατόμων που λογοκρίθηκαν στην ομάδα θεραπείας στο διάστημα $(t - 2, t - 1)$

Ακόμη $R_r(0)$, $R_c(0)$ αντιστοιχούν στον συνολικό αριθμό των ασθενών που περιλαμβάνονται στην ανάλυση στις ομάδες θεραπείας και στην ομάδα ελέγχου αντίστοιχα. Επίσης γίνονται οι παρακάτω συμβάσεις

$$D_r(0) = D_c(0) = C_r(0) = C_c(0) = 0$$

Ωστόσο για να υπολογιστεί ο αριθμός των λογοκριμένων παρατηρήσεων στο διάστημα $(t - 1, t)$ πρέπει πρώτα να υπολογιστεί ο αριθμός των ατόμων που βρισκόταν σε κίνδυνο στην αρχή του διαστήματος $(t - 1, t)$, ο οποίος ισούται με τον αριθμό των ατόμων που βρισκόταν σε κίνδυνο στο προηγούμενο διάστημα αφαιρώντας αυτούς που βίωσαν το γεγονός.

$$R_r(t_s) = R_r(t - 1) - D_r(t - 1)$$

Έτσι η σχέση (3.1) μετασχηματίζεται ως εξής

$$R_r(t) = R_r(t_s) - C_r(t)$$

Υποθέτοντας ότι οι ασθενείς λογοκρίνονται με σταθερό ρυθμό κατά τη διάρκεια ενός διαστήματος θεωρούμε ότι συνεισφέρουν τον κίνδυνο μισού ατόμου και έτσι ο υπολογισμός των ατόμων που λογοκρίθηκαν στο διάστημα πραγματοποιείται άμεσα μέσω της επόμενης σχέσης

$$C_r(t) = R_r(t_s) \left\{ \frac{1}{2} \frac{(t_e - t_s)}{(F_{max} - t_s)} \right\}$$

όπου F_{max} είναι η μέγιστη διάρκεια παρακολούθησης που έχει εκτιμηθεί στην αντίστοιχη δοκιμή i. Έχοντας λοιπόν τις παραπάνω πληροφορίες δύναται να υπολογιστεί ο αριθμός των γεγονότων στο συγκεκριμένο διάστημα ως εξής

$$D_r(t) = \left[R_r(t) * \left(\frac{S_r(t_s) - S_r(t_e)}{S_r(t_s)} \right) \right]$$

όπου $S_r(t_s)$ είναι η εκτίμηση της πιθανότητας επιβίωσης μέσω της καμπύλης Kaplan Meier στην αρχή του διαστήματος $(t - 1, t)$ και $S_r(t_e)$ είναι η αντίστοιχη εκτίμηση στο τέλος του διαστήματος $(t - 1, t)$.

- iii. Έπειτα εκμεταλλευόμενοι το γεγονός ότι όταν η λογοκρισία και ο χρόνος έως το γεγονός δεν εισάγονται επίσημα στον υπολογισμό, ο λόγος κινδύνου μπορεί να εκτιμηθεί από το σχετικό κίνδυνο, προκύπτει για κάθε δοκιμή ότι

$$\log HR(t) = \log \left(\frac{D_r(t)/R_r(t)}{D_c(t)/R_c(t)} \right)$$

και

$$\text{var}(\log HR(t)) = \frac{1}{D_r(t)} - \frac{1}{R_r(t)} + \frac{1}{D_c(t)} - \frac{1}{R_c(t)}$$

Για να αποφευχθούν προβλήματα υπολογισμού προτείνεται όταν $D_r(t)$ ή $D_c(t)$ είναι μηδέν, η αντικατάστασή του από μία μικρή ποσότητα της τάξης του 10^{-6} .

- iv. Τέλος έχοντας κατά νου ότι οι λόγοι κινδύνου για κάθε δοκιμή είναι ανεξάρτητοι ο συνολικός λόγος κινδύνου (overall log hazard ratio) και η διακύμανση μπορούν εύκολα να υπολογιστούν με βάση τον τύπο (3.1).

Συνολικά λοιπόν, για τη συγκεκριμένη μέθοδο έχουν γίνει οι υποθέσεις της ομοιόμορφης λογοκρισίας με την πάροδο του χρόνου καθώς και ότι οι πληροφορίες για τους ασθενείς έχουν γίνει σε τακτά χρονικά διαστήματα και συνεπώς ελάχιστες πληροφορίες απουσιάζουν λόγω κακής διεξαγωγής.

Ακόμη, σε μία μελέτη σύγκρισης 48 δοκιμών στις οποίες έχει εφαρμοστεί η συγκεκριμένη μέθοδος καθώς και άλλες πιο άμεσες, τα αποτελέσματα δείχνουν ότι η εκτίμηση της διακύμανσης από μία καμπύλη επιβίωσης υπό την αποδοχή της σταθερής λογοκρισίας σε ολόκληρη την περίοδο παρακολούθησης έχει την τάση να υποτιμά την αλήθεια και η απόκλιση αυτή διευρύνεται όσο μεγαλώνει η πραγματική τιμή της διακύμανσης, με αποτέλεσμα την πιθανή απόδοση μεγαλύτερου βάρους στην εν λόγω μελέτη σε μία μετα-ανάλυση.

Μερικά προβλήματα που εντοπίστηκαν είναι αρχικά ότι το δημοσιευμένο μέγεθος και η ανάλυση των καμπυλών δεν βοηθούν στην εξαγωγή αξιόπιστων τιμών με αποτέλεσμα να εισάγεται κάποιο σφάλμα στο συγκεκριμένο στάδιο. Ακόμη ένας περιορισμός που κατέχει σημαντικό ρόλο στη συγκεκριμένη προσέγγιση είναι ότι η επιλογή του μήκους των διαστημάτων πρέπει να γίνεται με τέτοιο τρόπο ώστε ο ρυθμός με τον οποίο διαδραματίζονται τα γεγονότα να είναι μικρός. Εμπειρικά αναφέρεται ότι δεν πρέπει να ξεπερνά το ποσοστό της τάξης του 20% από τον αρχικό ρυθμό. Ένα ακόμα σημείο που χρήζει ιδιαίτερης προσοχής είναι η επιλογή του μήκους του τελευταίου διαστήματος ώστε το συνολικό μήκος να μην ξεπερνά το μέγιστο χρόνο παρακολούθησης των υποκειμένων της δοκιμής. Τέλος ενώ δεν είναι απαραίτητο τα χρονικά διαστήματα σε κάθε δοκιμή να είναι ισομήκη, όταν αυτό δεν συμβαίνει δημιουργούνται προβλήματα στην κατασκευή των συνολικών καμπυλών επιβίωσης για κάθε ομάδα.

Αξίζει να αναφερθεί ότι μέχρι τότε και άλλοι ερευνητές είχαν ασχοληθεί με πτυχές του συγκεκριμένου θέματος χωρίς όμως να λαμβάνουν πολλές παραμέτρους υπόψιν τους. Για παράδειγμα ο A'hern και οι συνεργάτες του είχαν προτείνει τεχνικές για εκτίμηση συνοπτικών μέτρων από καμπύλες επιβίωσης αλλά δεν παρείχαν μοντέλο υπολογισμού της λογοκρισίας και της διακύμανσης (A'hern, et al., 1993). Επίσης οι Hunik & Wong παρουσίασαν τρόπους συνδυασμού των δεδομένων αλλά έπρεπε απαραίτητως να πληρείται η υπόθεση των αναλογικών κινδύνων σε όλες τις δοκιμές καθώς και η εκθετική κατανομή σε κάθε χρονικό διάστημα (Hunink & Wong, 1994). Ένας ακόμη ερευνητής, ο Dear παρουσίασε μία επαναληπτική γενικευμένη μέθοδο ελαχίστων τετραγώνων προκειμένου να επιτύχει το συνδυασμό δεδομένων επιβίωσης χωρίς όμως να αναφέρεται στις δυσκολίες για την εξαγωγή των δεδομένων από τις καμπύλες ή το πρόβλημα ενσωμάτωσης της λογοκρισίας (Dear, 1994). Ίσως βέβαια να μην υπάρχει και μία συγκεκριμένη μέθοδος που να καλύπτει τα πάντα καθώς όπως αναφέρει και ο Altman και οι συνεργάτες του καμία μεμονωμένη μέθοδος για την εκτίμηση του logHR δεν είναι συνήθως επαρκής για όλες τις δοκιμές σε μία μετα-ανάλυση (Altman, et al., 1995).

3.2 Η μέθοδος του Williamson

Στη συνέχεια οι Williamson, et al. 2002 αφού έκαναν την γνωστή διαπίστωση ότι οι μετα-αναλύσεις με συγκεντρωτικά δεδομένα σε πολλές περιπτώσεις είναι η μόνη επιλογή, παρόλο που οι μετα-αναλύσεις μεμονωμένων δεδομένων ασθενών θεωρούνται ως το χρυσό πρότυπο, πρότειναν έναν πιο βελτιωμένο τρόπο υπολογισμού των συνοπτικών αυτών μέτρων.

Η αναφερόμενη βελτίωση αναφέρεται στον τρόπο εκτίμησης των γεγονότων και των λογοκριμένων παρατηρήσεων σε κάθε διάστημα για τα συγκεντρωτικά δεδομένα επιβίωσης. Στόχο αποτελεί η εκτίμηση του λογαριθμισμένου λόγου κινδύνου και της διακύμανσης, λαμβάνοντας πιο έντονα υπόψη το μοτίβο λογοκρισίας που παρατηρείται σε κάθε δοκιμή (Williamson, et al., 2002). Είναι σημαντικό να αναφερθεί ότι το μοτίβο λογοκρισίας μπορεί να διαφέρει αρκετά αναλόγως του τελικού σημείου ενδιαφέροντος και συνεπώς κρίθηκε σκόπιμη η χρησιμοποίησή του. Πληροφορίες για το εν λόγω μοτίβο παρέχονται από τον αριθμό των ατόμων που βρίσκονται σε κίνδυνο σε κάθε χρονική στιγμή και μπορούν να εξαχθούν από μία καμπύλη επιβίωσης ή από ένα πίνακα επιβίωσης (lifetable).

3.2.1 Εκτίμηση με χρήση καμπυλών Kaplan Meier

Σε αυτή τη μέθοδο, για κάθε συμπεριλαμβανόμενη δοκιμή i ($i=1, \dots, k$), θεωρείται γνωστός ο αριθμός των ατόμων που βρίσκονται σε κίνδυνο (R) και για τις δύο ομάδες θεραπείας συμβολίζοντας R_r για την ομάδα ελέγχου (research arm) και R_c για την ομάδα ελέγχου (control arm). Οι πληροφορίες αυτές βρίσκονται είτε στην καμπύλη επιβίωσης είτε αναφέρονται μέσα στο κείμενο. Στη συνέχεια αντλούνται οι πιθανότητες επιβίωσης από την καμπύλη KM για κάθε ένα χρονικό σημείο ($t = 1, \dots, T$) και συμβολίζονται με $S^*(t)$ για $t = 1, \dots, T$. Γενικά ισχύει $S^*(0) = 1$ και $R_r(0), R_c(0)$ ισούται με τον αριθμό των ατόμων που τυχαιοποιήθηκαν στην ομάδα θεραπείας και στην ομάδα ελέγχου, αντίστοιχα.

Η γενική εικόνα της μεθόδου ταιριάζει με αυτή που περιεγράφηκε παραπάνω. Εν συντομία γίνεται εκτίμηση του $\log HR$ και της διακύμανσης σε κάθε χρονικό διάστημα $[t - 1, t]$, όπου $t = 1, \dots, T$ και στη συνέχεια συνδυασμός αυτών των εκτιμήσεων με την προσέγγιση της αντίστροφης σταθμισμένης διακύμανσης. Ξεκινώντας λοιπόν απαραίτητη είναι η εκτίμηση του αριθμού των γεγονότων και των λογοκριμένων παρατηρήσεων σε κάθε διάστημα. Οι υπολογισμοί που ακολουθούν αφορούν την ομάδα που λαμβάνει τη θεραπεία, ισχύουν όμως τα ακριβώς αντίστοιχα και για την ομάδα ελέγχου. Επιπρόσθετα χρησιμοποιείται η αναλογιστική προσέγγιση (actuarial approach) σύμφωνα με την οποία η λογοκρισία θεωρείται σταθερή εντός των διαστημάτων αλλά όχι απαραίτητως και μεταξύ αυτών. Ωστόσο η μέθοδος θεωρεί μία ομοιόμορφη κατανομή της λογοκρισίας στα διαστήματα που ορίζονται από τους αριθμούς των ατόμων που βρίσκονται σε κίνδυνο. Μαθηματικά αυτό μοντελοποιείται ως εξής:

$$S_r^*(t) = S_r^*(t-1) \left[1 - \frac{D_r^*(t)}{R_r(t-1) - \frac{C_r^*(t)}{2}} \right]$$

$$R_r(t) = R_r(t-1) - D_r^*(t) - C_r^*(t)$$

Όπου $D_r^*(t)$ είναι ο αριθμός των γεγονότων στο διάστημα $(t-1, t)$ και

$C_r^*(t)$ είναι ο αριθμός των λογοκριμένων παρατηρήσεων στο διάστημα $(t-1, t)$

Κάνοντας αναδιάταξη οι παραπάνω τύποι δίνουν

$$D_r^*(t) = \frac{[R_r(t-1) + R_r(t)][S_r^*(t-1) - S_r^*(t)]}{(S_r^*(t-1) + S_r^*(t))}$$

$$C_{r_i}^*(t) = \frac{2(R_r(t-1)S_r^*(t) - R_{r_i}(t)S_r^*(t-1))}{(S_r^*(t-1) + S_r^*(t))}$$

$$R_r^*(t) = \frac{(R_r(t-1) + R_r(t))S_r^*(t-1)}{(S_r^*(t-1) + S_r^*(t))}$$

Όπου $R_r^*(t)$: είναι ο αριθμός των ατόμων που βρίσκονται σε κίνδυνο κατά το διάστημα $(t-1, t)$

Στη συνέχεια λαμβάνοντας υπόψη όπως και στη μέθοδο του Parmar ότι ο λόγος κινδύνου μπορεί να εκτιμηθεί από το σχετικό κίνδυνο όταν η λογοκρισία και ο χρόνος έως το γεγονός δεν περιλαμβάνονται στους υπολογισμούς, η εκτίμηση του λογαριθμισμένου λόγου κινδύνου για κάθε δοκιμή επιτυγχάνεται ως εξής:

$$\log \widehat{HR}(t) = \log \left(\frac{D_r^*(t)/R_r^*(t)}{D_c^*(t)/R_c^*(t)} \right)$$

$$\text{var}(\log \widehat{HR}(t)) = \frac{1}{D_r^*(t)} - \frac{1}{R_r^*(t)} + \frac{1}{D_c^*(t)} - \frac{1}{R_c^*(t)}$$

Τελικό στάδιο αποτελεί η εκτίμηση του συνολικού λογαριθμισμένου λόγου κινδύνου όλων των δοκιμών i ($i=1, \dots, k$), η οποία δίνεται μέσω των παρακάτω σχέσεων

$$\log(\widehat{HR}) = \frac{\sum_{i=1}^k w_i \log \widehat{HR}_i}{\sum_{i=1}^k w_i} \quad \text{και} \quad \text{var}[\log(\widehat{HR})] = \frac{1}{\sum_{i=1}^k w_i}$$

όπου $w_i^{-1} = \text{var}[\log(\widehat{HR}_i)]$

3.2.2 Εκτίμηση με χρήση life-table

Σε περιπτώσεις όπου οι χρόνοι επιβίωσης αναφέρονται σε διαστήματα (π.χ. έξι μηνών, ενός έτους κτλ) και δεν παρέχεται πιο λεπτομερειακή πληροφόρηση, είναι προτιμότερη η χρήση των πινάκων επιβίωσης (life table) παρά της προσέγγισης του εκτιμητή Kaplan Meier (product limit approach). Οι πληροφορίες που είναι διαθέσιμες στην περίπτωση αυτή είναι ο αριθμός των ατόμων σε κίνδυνο καθώς και οι πιθανότητες επιβίωσης στην αρχή κάθε διαστήματος. Γίνεται επομένως άμεσα αντιληπτό ότι και σε αυτή την περίπτωση μπορεί να εφαρμοστεί η μέθοδος που περιγράφηκε παραπάνω.

3.2.3 Εκτίμηση μέσω μιας συνολικής καμπύλης επιβίωσης

Πολλές είναι οι περιπτώσεις στις οποίες υπάρχουν διαθέσιμες πληροφορίες όπως το $\log HR$ και η διακύμανση για ένα πλήθος ξεχωριστών διαστημάτων σε κάθε δοκιμή και τότε είναι εφικτός ο υπολογισμός της εκτίμησης του συνολικού λογαριθμισμένου λόγου κινδύνου (overall $\log HR$) των δοκιμών για κάθε διάστημα. Γι αυτό προτείνεται η συγκέντρωση και ο σχηματισμός μιας συνολικής καμπύλης επιβίωσης όπου τα διαστήματα θα ορίζονται από τα ίδια χρονικά σημεία $t'_l (l = 1, \dots, r)$ και θα πρέπει να χρησιμοποιούνται από όλες τις δοκιμές. Εν τούτοις χρειάζονται περεταίρω τροποποιήσεις στην προαναφερθείσα μέθοδο καθώς είναι απίθανο τα άτομα σε κίνδυνο να αναφέρονται στις ίδιες χρονικές στιγμές για κάθε δοκιμή.

Ξεκινώντας όπως περιγράφηκε νωρίτερα, όταν δίνονται οι αριθμοί των ατόμων σε κίνδυνο στις χρονικές στιγμές $t_j, j = 1, \dots, T$ τότε υποθέτουμε ότι οι πιθανότητες για επιβίωση στα συγκεκριμένα σημεία είναι $S^*(t_j)$ για $j = 1, \dots, T$, οι οποίες έχουν παρθεί από τις καμπύλες Kaplan Meier. Έπειτα με την παραπάνω διαδικασία εκτιμώνται για κάθε διάστημα $[t_{j-1}, t_j]$ με $j=1, \dots, T$ ο αριθμός των γεγονότων που συνέβησαν $D^*(t_j)$, ο αριθμός των λογοκριμένων παρατηρήσεων $C^*(t_j)$ καθώς και ο αριθμός των ατόμων που βρίσκονται σε κίνδυνο $R^*(t_j)$. Ακόμη θεωρούνται οι πιθανότητες επιβίωσης που και πάλι εξάγονται από τις καμπύλες επιβίωσης $S^*(t'_l)$, στα απαιτούμενα χρονικά σημεία t'_l (με $l = 1, \dots, r$) που είναι απαραίτητα για κάθε δοκιμή. Έτσι για κάθε διάστημα $[t'_{l-1}, t'_l]$ με $l = 1, \dots, r$ εκτιμώνται οι

αναγκαίες ποσότητες $D^{*'}(t'_l), R^{*'}(t'_l), C^{*'}(t'_l)$ για τον υπολογισμό των συνοπτικών μέτρων. Ακόμη για λόγους απλότητας θα χρησιμοποιούνται οι εξής συντομεύσεις: $S^*(t_j)=S_j^*(t), D^*(t_j)=D_j^*(t), R^*(t_j)=R_j^*(t), C^*(t_j)=C_j^*(t)$ και αντιστοίχως $S^{*'}(t'_l) = S_l^{*'}(t), D^{*'}(t'_l) = D_l^{*'}(t), R^{*'}(t'_l) = R_l^{*'}(t), C^{*'}(t'_l) = C_l^{*'}(t)$.

Για το υποδιάστημα $[t_{j-1}, t'_l]$, όπου $t'_{l-1} < t_{j-1} < t'_l < t_j$ με $j = 1, \dots, T$ και $l = 1, \dots, r$

$$D^*(t) = 0.5 \left[\frac{S_l^{*'}(t)((C_j^*(t) + 2D_j^*(t))(S_l^{*'}(t) - S_{j-1}^*(t)) - (S_j^*(t)C_j^*(t))) + (S_j^*(t)S_{j-1}^*(t)C_j^*(t))}{S_l^{*'}(t)(S_j^*(t) - S_{j-1}^*(t))} \right]$$

$$C^*(t) = \left[\frac{S_l^{*'}(t) \left((2S_j^*(t)R_{j-1}(t)) - S_{j-1}^*(t) (C_j^*(t) + 2R_j(t)) \right) - (S_j^*(t)S_{j-1}^*(t)C_j^*(t))}{S_l^{*'}(t)(S_j^*(t) - S_{j-1}^*(t))} \right]$$

Για το υποδιάστημα $[t'_l, t_j]$,

$$D^*(t) = 0.5 \left[\frac{S_l^{*'}(t)((C_j^*(t) + 2D_j^*(t))(S_j^*(t) - S_l^{*'}(t)) + (S_{j-1}^*(t)C_j^*(t))) - (S_j^*(t)S_{j-1}^*(t)C_j^*(t))}{S_l^{*'}(t)(S_j^*(t) - S_{j-1}^*(t))} \right]$$

$$C^*(t) = - \left[\frac{S_l^{*'}(t) \left(S_j^*(t) (2R_{j-1}(t) - C_j^*(t)) - 2S_{j-1}^*(t)R_j(t) \right) - (S_j^*(t)S_{j-1}^*(t)C_j^*(t))}{S_l^{*'}(t)(S_j^*(t) - S_{j-1}^*(t))} \right]$$

Ο εκτιμώμενος αριθμός των ατόμων σε κίνδυνο τη χρονική στιγμή t'_l υπολογίζεται ως εξής:

$$R_l'(t) = R_{j-1}(t) - D_{([j-1,l])}^*(t) - C_{([j-1,l])}^*(t)$$

Τελικά, για ολόκληρο το διάστημα $[t'_{l-1}, t'_l]$ ο λογαριθμισμένος λόγος κινδύνου και η διακύμανση μπορεί να υπολογιστεί προσεγγιστικά χρησιμοποιώντας τις ακόλουθες εκτιμήσεις

$$D_l^{*'}(t) = D_{([l-1,j-1])}^*(t) + D_{([j-1,l])}^*(t)$$

$$C_l^{*'}(t) = C_{([l-1,j-1])}^*(t) + C_{([j-1,l])}^*(t)$$

$$R_i^{*'}(t) = R_{i-1}^{*'}(t) - \frac{C_i^{*'}(t)}{2}$$

Στη συνέχεια και κατ' αντιστοιχία με την (3.1) υπολογίζεται ο συνολικός λογαριθμισμένος λόγος κινδύνου και η διακύμανση.

Με τον τρόπο αυτό λοιπόν υπολογίστηκε ο αριθμός των συμβάντων και των λογοκριμένων παρατηρήσεων σε διαστήματα διαφορετικά από αυτά που αντιστοιχούσαν στους αριθμούς των ατόμων που βρισκόταν σε κίνδυνο και αναφέρονταν στη δοκιμή. Το κίνητρο για την εν λόγω διαδικασία ήταν ο καθορισμός κοινών χρονικών διαστημάτων μεταξύ των δοκιμών προκειμένου να επιτευχθεί η εκτίμηση του συνολικού λόγου κινδύνου (overall) pooled hazard ratio.

Ωστόσο αναφέρεται ότι η χρήση συγκεντρωτικών δεδομένων από καμπύλες επιβίωσης για την εκτίμηση του *logHR* μπορεί να δώσει μόνο προσεγγιστικά αποτελέσματα καθώς εναπόκεινται σε δύο βασικούς περιορισμούς.

- Αρχικά ο τρόπος με τον οποίο παίρνει κανείς δεδομένα από την καμπύλη δεν γίνεται αυτοματοποιημένα αλλά εξαρτάται από την ευχέρεια του ερευνητή, συνεπώς υπάρχει μία διακύμανση μεταξύ των τιμών που λαμβάνονται, που ονομάζεται inter-reader διακύμανση και δεν μπορεί να προσδιοριστεί.
- Επιπρόσθετα είναι απαραίτητες οι υποθέσεις για τον ρυθμό λογοκρισίας των παρατηρήσεων.

Ιδανικά αναφέρουν ότι θα έπρεπε να γίνεται ανάλυση ευαισθησίας με σκοπό τη διερεύνηση οποιασδήποτε υπόθεσης γιατί αν για παράδειγμα ο ρυθμός των συμβάντων σε μία μελέτη είναι γρήγορος αυτό θα έχει ως αποτέλεσμα η λογοκρισία να είναι μικρή και έτσι δεν θα δημιουργηθεί μεγάλη μεροληψία, αν όμως στην αρχή κάθε διαστήματος υπάρχει έντονη λογοκρισία θα μειώσει τον αριθμό των ατόμων που βρίσκονται σε κίνδυνο σε κάθε διάστημα και πιθανώς τα αποτελέσματα της μετα-ανάλυσης θα είναι πλήρως αναξιόπιστα. Στις περιπτώσεις αυτές κρίνεται επιτακτική η συλλογή των ατομικών δεδομένων των ασθενών (IPD).

3.3 Η προσέγγιση των Hoyle and Henley

Με εφιαλτήριο την άμεση και έντονη επηρεασμένη σχέση κόστους αποτελεσματικότητας των τεχνολογιών υγείας ή των παρεμβάσεων υγείας από την επιλογή της καμπύλης επιβίωσης, το 2011 οι Hoyle και Henley ανέπτυξαν μια πιο ακριβή μέθοδο προσαρμογής καμπυλών επιβίωσης σε συνοπτικά δεδομένα επιβίωσης (Hoyle & Henley, 2011). Η προαναφερθείσα εξάρτηση του εκτιμώμενου αναμενόμενου κόστους και οφέλους για κάθε θεραπεία προκύπτει καθώς αυτές είναι συναρτήσεις του αναμενόμενου χρόνου

παραμονής των ασθενών σε μια δεδομένη κατάσταση υγείας. Συνεπώς ακόμα και μικρές αλλαγές στον τρόπο εφαρμογής των καμπυλών μπορούν να επιφέρουν σοβαρές επιπτώσεις, οι οποίες γίνονται πιο έντονες σε περιπτώσεις όπου απαιτείται κάποια παρέκταση (extrapolation).

Η προτεινόμενη μέθοδος εκτελείται σε δύο στάδια. Στο πρώτο γίνεται η εκτίμηση των υποκείμενων μεμονωμένων δεδομένων των ασθενών και στο δεύτερο πραγματοποιείται η προσαρμογή των καμπυλών επιβίωσης στο εκτιμώμενο IPD με τη μέθοδο της μεγίστης πιθανοφάνειας. Ωστόσο στη συγκεκριμένη εργασία δεν θα αναφερθεί λεπτομερειακά ο τρόπος προσαρμογής των καμπυλών καθώς δεν αποτελεί άμεσο αντικείμενο μελέτης. Παρ' όλ' αυτά αξίζει να αναφερθεί ότι σε σύγκριση με τις παραδοσιακές μεθόδους προσαρμογής καμπυλών (μέθοδος ελαχίστων τετραγώνων και παλινδρόμησης) η προτεινόμενη έδωσε πιο ακριβείς προσαρμογές δίνοντας και τις αντίστοιχες εκτιμήσεις αβεβαιότητας, γεγονός που ήταν αδύνατο με τις παραδοσιακές. Επιπλέον ένα ακόμα αδύνατο σημείο των παραδοσιακών μεθόδων, το οποίο προσπάθησε να λάβει υπόψη η νέα προσέγγιση, είναι ότι ενώ οι προσαρμογές επηρεάζονται εξίσου από όλα τα σημεία των καμπυλών KM, οι ουρές τους είναι αρκετά αβέβαιες εξαιτίας του μικρού αριθμού ασθενών σε κίνδυνο. Το γεγονός αυτό έχει άμεσο αντίκτυπο στην εκτίμηση της μέσης επιβίωσης όπου και αυτή με τη σειρά της καθοδηγεί τη σχέση κόστους αποτελεσματικότητας.

Επιστρέφοντας στο πρώτο στάδιο, που ενδιαφέρει έντονα την παρούσα διατριβή, όσο αφορά την εκτίμηση των ατομικών δεδομένων των ασθενών πρέπει να αναφερθεί ότι η διαδικασία αυτή στηρίχθηκε στη μοντελοποίηση του εύρους των γεγονότων που ως επί το πλείστον λαμβάνονται υπόψη στις αναλύσεις κόστους αποτελεσματικότητας.

Πιο αναλυτικά, λαμβάνοντας υπόψη τους περιορισμούς που είχαν ανιχνευτεί από τη μέθοδο που περιγράφηκε από τους Parmar και Tierney (Duchateau, et al., 2000 ; Tierney, et al., 2007) στην οποία ο υπολογισμός των ατόμων που είχαν το γεγονός καθώς και των ατόμων που λογοκρίθηκαν μπορούσαν να υπολογιστούν μόνο σε διαστήματα στα οποία ήταν γνωστός ο αριθμός των ατόμων σε κίνδυνο, γεγονός που συχνά οδηγούσε σε λίγα χρονικά σημεία από τα οποία μπορούσε να εκτιμηθεί η καμπύλη επιβίωσης (συνήθως 5-12 σημεία), η νέα μέθοδος χρησιμοποιεί πιθανότητες επιβίωσης και σε ενδιάμεσους χρόνους ώστε να βελτιώσει τις εκτιμήσεις εντός των διαστημάτων. Είναι φυσιολογικό στο σημείο αυτό να αναρωτηθεί κανείς ότι αφού ο Williamson και οι συνεργάτες του έκαναν κάτι αντίστοιχο, ποια είναι η υπεροχή της συγκεκριμένης μεθόδου; Η απάντηση έρχεται από το γεγονός ότι ενώ στη μέθοδο που αναφέρει ο Williamson και οι συνεργάτες του χρησιμοποιούν πιθανότητες επιβίωσης σε ενδιάμεσους χρόνους με κίνητρο τον καθορισμό κοινών χρονικά διαστημάτων, η μέθοδος των Holey και Henley χρησιμοποίησε πρόσθετες πιθανότητες (additional probabilities) προκειμένου να βελτιώσουν τις εκτιμήσεις του αριθμού των γεγονότων σε κάθε διάστημα.

Πιο συγκεκριμένα θεωρώντας ξανά ότι ο ρυθμός λογοκρισίας παραμένει σταθερός εντός κάθε διαστήματος, χρησιμοποιήθηκαν οι πιθανότητες επιβίωσης σε ενδιάμεσους χρόνους $S(t + \frac{1}{2})$, ώστε να υπολογιστεί ο αριθμός των γεγονότων και των λογοκριμένων παρατηρήσεων σε κάθε διάστημα μήκους $\frac{1}{2}$. Οι υπολογισμοί όπως και στις προηγούμενες μεθόδους πραγματοποιούνται ξεχωριστά για την ομάδα ελέγχου και την ομάδα θεραπείας και για λόγους απλότητας παραλείφθηκαν οι αντίστοιχοι δείκτες. Γενικά ισχύουν:

$$S\left(t + \frac{1}{2}\right) = S(t) \left(\frac{R(t) - \frac{C(t, t+1)}{4} - D(t, t + \frac{1}{2})}{R(t) - \frac{C(t, t+1)}{4}} \right)$$

$$S(t+1) = S\left(t + \frac{1}{2}\right) \left(\frac{R(t) - \frac{3}{4}C(t, t+1) - D\left(t, t + \frac{1}{2}\right) - D\left(t + \frac{1}{2}, t+1\right)}{R(t) - \frac{3}{4}C(t, t+1) - D\left(t, t + \frac{1}{2}\right)} \right)$$

$$R(t+1) = R(t) - D\left(t, t + \frac{1}{2}\right) - D\left(t + \frac{1}{2}, t+1\right) - C(t, t+1)$$

όπου

$R(t+1)$: αριθμός ατόμων που βρίσκονται σε κίνδυνο τη χρονική στιγμή $t+1$

$D\left(t, t + \frac{1}{2}\right)$: είναι ο αριθμός των ατόμων που έπαθαν το γεγονός στο διάστημα $[t, t + \frac{1}{2})$

$D\left(t + \frac{1}{2}, t+1\right)$: είναι ο αριθμός των ατόμων που έπαθαν το γεγονός στο διάστημα $[t + \frac{1}{2}, t+1)$

Ωστόσο το άθροισμα αυτών των δύο δεν είναι αναγκαίο να ισούται με το $D(t, t+1)$ που υπολογιζόταν στη μέθοδο του Williamson καθώς εδώ έχει χρησιμοποιηθεί προσθετική πληροφορία από την πιθανότητα επιβίωσης στο μεσοδιάστημα.

$$D\left(t + \frac{1}{2}, t+1\right) = \left[S\left(t + \frac{1}{2}\right) - S(t+1) \right] \left(\frac{S\left(t + \frac{1}{2}\right)R(t) + S\left(t + \frac{1}{2}\right)R(t+1) + 2S(t)R(t+1)}{S\left(t + \frac{1}{2}\right)S(t) + S\left(t + \frac{1}{2}\right)S(t+1) + 2S(t)S(t+1)} \right)$$

$$D\left(t, t + \frac{1}{2}\right) = R(t) + 3R(t + 1) - \left[3S(t + 1) + S\left(t + \frac{1}{2}\right)\right] \left(\frac{S\left(t + \frac{1}{2}\right)R(t) + S\left(t + \frac{1}{2}\right)R(t + 1) + 2S(t)R(t + 1)}{S\left(t + \frac{1}{2}\right)S(t) + S\left(t + \frac{1}{2}\right)S(t + 1) + 2S(t)S(t + 1)} \right)$$

Κάτι αντίστοιχο συμβαίνει και με την εκτίμηση του αριθμού των λογοκριμένων παρατηρήσεων. Για λόγους απλότητας όμως γίνεται η υπόθεση ότι ο ρυθμός με τον οποίο λογοκρίνονται οι παρατηρήσεις εντός του διαστήματος $[t, t + 1)$ παραμένει σταθερός (σημείωση: αν δεν είχαν υποθέσει σταθερή λογοκρισία στο διάστημα $[t, t + 1)$ τότε με 3 εξισώσεις και 4 αγνώστους δεν θα υπήρχε μοναδική λύση). Έτσι το πλήθος των λογοκριμένων παρατηρήσεων στο διάστημα $[t, t + 1)$ δίνεται από τον τύπο:

$$C(t, t + 1) = [R(t) - R(t + 1) - D\left(t, t + \frac{1}{2}\right) - D\left(t + \frac{1}{2}, t + 1\right)]$$

Ακόμη ο αριθμός των ατόμων σε κίνδυνο στο μεσοδιάστημα υπολογίζεται ως εξής:

$$R\left(t + \frac{1}{2}\right) = R(t) - D\left(t, t + \frac{1}{2}\right) - \frac{C(t, t + 1)}{2}$$

Στη συνέχεια και προκειμένου να βελτιωθούν περισσότερο οι εκτιμήσεις για τον αριθμό των γεγονότων και τη λογοκρισία, χρησιμοποιήθηκαν οι ενδιάμεσες πιθανότητες επιβίωσης των προηγούμενων διαστημάτων. Με άλλα λόγια χρησιμοποιήθηκαν οι $S\left(t + \frac{1}{4}\right)$ και $S\left(t + \frac{3}{4}\right)$, όπου επέτρεψαν την εκτίμηση του αριθμού των γεγονότων και των λογοκριμένων παρατηρήσεων σε κάθε διάστημα μήκους $\frac{1}{4}$. Ομοίως με πριν

$$D\left(t + \frac{1}{4}, t + \frac{1}{2}\right) = \left[S\left(t + \frac{1}{4}\right) - S\left(t + \frac{1}{2}\right) \right] \times \left(\frac{S\left(t + \frac{1}{4}\right)R(t) + S\left(t + \frac{1}{4}\right)R\left(t + \frac{1}{2}\right) + 2S(t)R\left(t + \frac{1}{2}\right)}{S\left(t + \frac{1}{4}\right)S(t) + S\left(t + \frac{1}{4}\right)S\left(t + \frac{1}{2}\right) + 2S(t)S\left(t + \frac{1}{2}\right)} \right)$$

$$D\left(t + \frac{3}{4}, t + 1\right) = \left[S\left(t + \frac{3}{4}\right) - S(t + 1) \right] \times \left(\frac{S\left(t + \frac{3}{4}\right)R\left(t + \frac{1}{2}\right) + S\left(t + \frac{3}{4}\right)R(t + 1) + 2S\left(t + \frac{1}{2}\right)R(t + 1)}{S\left(t + \frac{3}{4}\right)S\left(t + \frac{1}{2}\right) + S\left(t + \frac{3}{4}\right)S(t + 1) + 2S\left(t + \frac{1}{2}\right)S(t + 1)} \right)$$

Μερικές επιπλέον διευκρινήσεις για τις ισότητες σε αντιστοιχία με αυτές του διαστήματος $[t, t + \frac{1}{2})$:

$$D(t, t + \frac{1}{4}) = D(t, t + \frac{1}{2}) - D(t + \frac{1}{4}, t + \frac{1}{2})$$

$$D(t + \frac{1}{2}, t + \frac{3}{4}) = D(t + \frac{1}{2}, t + 1) - D(t + \frac{3}{4}, t + 1)$$

$$C(t, t + \frac{1}{4}) = C(t + \frac{1}{4}, t + \frac{1}{2}) = C(t + \frac{1}{2}, t + \frac{3}{4}) = C(t + \frac{3}{4}, t + 1) = \frac{C(t, t + 1)}{4}$$

Επιπλέον είναι σημαντικό να αναφερθεί ότι έχει γίνει η υπόθεση της κατά διαστήματα λογοκρισίας (interval censored). Πιο απλά, εφόσον δεν είναι γνωστός ο ακριβής χρόνος που διαδραματίστηκαν τα γεγονότα, χρησιμοποιείται η μόνη γνώση ότι έγιναν στα συγκεκριμένα διαστήματα, χωρίς ωστόσο να γίνει κάποια επιπλέον θεώρηση όπως για παράδειγμα ότι συνέβησαν όλα στο μέσο του διαστήματος.

Στις περιπτώσεις όπου ο αριθμός των ατόμων σε κίνδυνο δεν αναφέρεται, οι εκτιμήσεις των παραπάνω ποσοτήτων μπορούν να εκτιμηθούν εύκολα από ένα υπολογιστικό φύλλο που αναπτύχθηκε από τους Tierney και συνεργάτες, κάνοντας την υπόθεση ότι η λογοκρισία δεν συμβαίνει στον ημερολογιακό χρόνο αποκοπής αλλά λαμβάνει χώρα με σταθερό ρυθμό (Tierney, et al., 2007).

ΑΞΙΟΛΟΓΗΣΗ ΜΕΘΟΔΟΥ

Στη συνέχεια οι ερευνητές προκειμένου να αξιολογήσουν την ακρίβεια της μεθόδου διενέργησαν μία προσομοίωση στην οποία η επιβίωση των ασθενών ακολουθούσε την Weibull κατανομή και επιλέχθηκαν διάφοροι συνδυασμοί των παραμέτρων της (λ , γ). Η λογοκρισία θεωρήθηκε μη ενημερωτική (non-informative censoring), ο μέγιστος χρόνος παρακολούθησης καθορίστηκε στις 10 μονάδες και η αναφορά του αριθμού των ασθενών που ήταν σε κίνδυνο επιλέχθηκε να γίνει συντηρητικά σε 6 χρονικά σημεία. Έπειτα έγινε η καταγραφή των πιθανοτήτων επιβίωσης ανά $\frac{1}{2}$ χρονικό σημείο. Για κάθε συνδυασμό των παραμέτρων προσομοιώθηκαν 1000 ανεξάρτητες δοκιμές. Στη συνέχεια υπολογίστηκαν οι αριθμοί των ατόμων που λογοκρίθηκαν και των γεγονότων σε κάθε διάστημα μήκους $\frac{1}{2}$ και συγκρίθηκαν με τις πραγματικές τιμές. Ακόμη για τον έλεγχο της απόδοσης της προτεινόμενης μεθόδου δημιούργησαν παραλλαγές της και τη συνέκριναν με την αρχική με απώτερο σκοπό να μετρηθεί η σημασία της λογοκρισίας του διαστήματος καθώς και η χρήση των αναφερόμενων πιθανοτήτων επιβίωσης σε ενδιάμεσα τελικά σημεία. Επιπρόσθετα έγινε αξιολόγηση των διαφορετικών μεθόδων συγκρίνοντας την μεροληψία (bias) και το απόλυτο σφάλμα (absolute error) των μέσων χρόνων επιβίωσης καθώς η αβεβαιότητα στην εκτίμηση του μέσου χρόνου αποτελεί καθοριστικό παράγοντα της σχέσης κόστους-αποτελεσματικότητας.

Τα αποτελέσματα έδειξαν ότι η προτεινόμενη μέθοδος προβλέπει με ακρίβεια τον αριθμό των γεγονότων και των λογοκριμένων παρατηρήσεων και είναι ιδιαίτερα ακριβής όταν δεν υπάρχει επιπλέον

λογοκρισία και ο κίνδυνος με το χρόνο συνδέονται με μία μονότονα φθίνουσα συνάρτηση. Πιστεύεται ακόμα ότι η ακρίβεια του εκτιμώμενου αριθμού γεγονότων και λογοκρισιών αυξάνεται με την αύξηση του συνολικού αριθμού γεγονότων. Όσο αφορά τις εκτιμήσεις του μέσου χρόνου δεν υπήρξε ένδειξη μεροληψίας σε καμία δοκιμή 100 ή 500 ατόμων όταν δεν υπήρχε πρόσθετη λογοκρισία, ενώ όπως ήταν αναμενόμενο το σφάλμα ήταν μεγαλύτερο στις δοκιμές των λιγότερων ατόμων ειδικά όταν εισαγόταν πρόσθετη λογοκρισία. Με άλλα λόγια για αυτό το συνδυασμό των παραμέτρων, η αβεβαιότητα στις εκτιμήσεις της μέσης τιμής που βρέθηκε από την προτεινόμενη μέθοδο ήταν περίπου τόσο ακριβής όσο και αυτή που εκτιμήθηκε χρησιμοποιώντας την πραγματική IPD.

Ακόμη καθώς το χρονικό διάστημα χωριζόταν σε όλο και περισσότερα διαστήματα τόσο η μεροληψία των εκτιμήσεων μίκραινε και συνεπώς η μέθοδος βελτιωνόταν. Το γεγονός αυτό ήταν πιο αισθητό όταν ο κίνδυνος μειωνόταν με το χρόνο και υπήρχε και πρόσθετη λογοκρισία. Επιβεβαιώθηκε ωστόσο η υπεροχή της επιλογής της κατά διαστήματα λογοκρισίας σε σχέση με την θεώρηση ότι όλα συμβαίνουν στη μέση κάθε διαστήματος. Τέλος, βρέθηκε ότι οι μέθοδοι της παλινδρόμησης και των ελαχίστων τετραγώνων υπερεκτιμούν αρκετά το μέσο των εκτιμήσεων ειδικά υπό την παρουσία λογοκρισίας.

Εξάλλου η προσομοίωση εκτελέστηκε χρησιμοποιώντας τις ακριβείς πιθανότητες επιβίωσης στις εκάστοτε χρονικές στιγμές, χωρίς να δοκιμαστεί η διαδικασία εξαγωγής των πιθανοτήτων από τις δημοσιευμένες καμπύλες Kaplan Meier. Εν τούτοις από τη μία πιστεύεται ότι μπορούν να διαβαστούν με μεγάλη ακρίβεια και από την άλλη ότι η ίδια μορφή ανακρίβειας εισάγετε σε όλες αυτές τις μεθόδους.

Συγκεντρωτικά στην προτεινόμενη αυτή μέθοδο λογίζονται δύο βασικές συνεισφορές:

- Αρχικά κατασκευάστηκε υποκατάστατο των ατομικών δεδομένων των ασθενών μέσω των εκτιμήσεων του αριθμού των γεγονότων και των λογοκρισιών σε κάθε χρονικό διάστημα, από το οποίο μπόρεσε να εκτιμηθεί με μεγάλη ακρίβεια η υποκείμενη κατανομή επιβίωσης.
- Και επιπλέον με τη χρήση των πιθανοτήτων επιβίωσης σε επιπρόσθετα χρονικά σημεία βελτιώθηκαν ακόμα περισσότερο οι εκτιμήσεις της υποκείμενης IPD. Ένα ακόμα προτέρημα είναι η μη απαραίτητη υπόθεση της αναλογικότητας των κινδύνων, αφού οι εκτιμήσεις για κάθε ομάδα γίνονται ανεξάρτητα.

Τέλος σημαντική είναι και η ικανότητα εκτίμησης της αβεβαιότητας των εκτιμήσεων, πράγμα που δεν μπορεί να συμβεί με τις παραδοσιακές μεθόδους. Βέβαια στο σημείο αυτό θα υπάρχει μία υποεκτίμηση επειδή θεωρείται ότι η εκτίμηση της IPD γίνεται με απόλυτη βεβαιότητα, ωστόσο η αβεβαιότητά της θα είναι αρκετά μικρή δεδομένου της καλής εκτίμησης της IPD.

Θα ήταν παράληψη να μην αναφερθεί ότι στην εν λόγω μέθοδο δεν έγινε αξιολόγηση της ακρίβειας των αποτελεσμάτων σε καταστάσεις όπου ο αριθμός των ατόμων σε κίνδυνο δεν είναι διαθέσιμος όπως επίσης και η επιλογή περισσότερων χρονικών σημείων. Επιπλέον δεδομένου του ότι μόνο η Weibull κατανομή χρησιμοποιήθηκε για να περιγράψει την κατανομή επιβίωσης είναι αβέβαιο το πως θα συμπεριφερόταν η μέθοδος αυτή σε κάποια άλλη κατανομή όπως και στην περίπτωση της ενημερωτικής λογοκρισίας.

3.4 Ο αλγόριθμος της Guyot

Όπως έχει ήδη αναφερθεί και σε συμφωνία με διεθνής οδηγίες (CONSORT guidelines, Cochrane handbook) το προτεινόμενο μέτρο της επίδρασης για δεδομένα χρόνου επιβίωσης είναι ως επί το πλείστον ο λόγος κινδύνου (hazard ratio) και η διαφορά των μέσων χρόνων επιβίωσης (difference in median survival time) (Parmar, et al., 1998; Deeks, et al., 2008). Το γεγονός αυτό πέρα από ότι περιορίζει τον τρόπο με τον οποίο αυτές οι μελέτες μπορούν να συμπεριληφθούν σε κάποιου είδους δευτερογενούς ανάλυσης, στην περίπτωση των μελετών κόστους αποτελεσματικότητας, όπου απαιτείται η εκτίμηση της διαφοράς της αναμενόμενης επιβίωσης μεταξύ των δύο ομάδων, καθιστά αδύνατη την ανακατασκευή της μόνο από τον αναφερόμενο λόγο κινδύνου ή τη διάμεσο, μη πραγματοποιώντας περαιτέρω υποθέσεις, καθώς απαιτείται κάποιου είδους παρέκτασης (extrapolation) των καμπυλών.

Γίνεται λοιπόν σαφές ότι τα δημοσιευμένα αποτελέσματα των αναλύσεων επιβίωσης είναι ασυνεπή και περιορισμένα. Με βάση αυτό το πρόβλημα αρκετοί ήταν οι ερευνητές (Dear, 1994; Altman, et al., 1995; Duchateau, et al., 2000 ; Williamson, et al., 2002), όπως οι Holey και Henley, που προσπάθησαν να εξάγουν δεδομένα επιβίωσης από καμπύλες KM, οι οποίες πέρα από λίγες εξαιρέσεις κατασκευάζονται και αναφέρονται σε κάθε ανάλυση επιβίωσης.

Εντούτοις οι μέχρι τώρα προσπάθειες εξαγωγής πιθανοτήτων επιβίωσης από γραφήματα έχουν γίνει χωρίς τη χρήση των διαθέσιμων πληροφοριών του κειμένου ή των διαγραμμάτων για εκτίμηση του μοτίβου της λογοκρισίας και σε σχετικά σύντομα διαστήματα παρακολούθησης με σκοπό την προσέγγιση συγκεντρωτικών δεδομένων ή πινάκων επιβίωσης.

Με έναυσμα λοιπόν τη φτωχή καταγραφή των αποτελεσμάτων της ανάλυσης επιβίωσης η Guyot και οι συνεργάτες της το 2012 ανέπτυξαν έναν αλγόριθμο για την ανακατασκευή των δεδομένων Kaplan Meier πάνω στα οποία στηρίζονται οι καμπύλες επιβίωσης, ώστε να είναι διαθέσιμη η εξαγωγή οποιονδήποτε στατιστικών κριθεί απαραίτητο (Guyot, et al., 2012). Κύριο χαρακτηριστικό του αλγορίθμου, που βοηθά στην εξαγωγή συνεπών αποτελεσμάτων είναι η χρήση επαναληπτικών

αριθμητικών μεθόδων για την επίλυση των ανεστραμμένων εξισώσεων KM. Προς αυτή την κατεύθυνση χρησιμοποιείται επίσης, όταν αναφέρονται, ο αριθμός των ατόμων που βρίσκονται σε κίνδυνο και ο συνολικός αριθμός των γεγονότων.

Πριν την αναλυτική περιγραφή του αλγορίθμου κρίνεται σκόπιμη η αναφορά των στοιχείων που απαιτείται να εισαχθούν προκειμένου να εκτελεσθεί ο αλγόριθμος.

Πρώτον είναι απαραίτητη η εισδοχή ενός αρχείου που περιέχει τις εξαγόμενες, μέσω κάποιου ψηφιακού λογισμικού χ -συντεταγμένων (t_k) που αντιπροσωπεύουν το χρόνο καθώς και των y -συντεταγμένων (S_k) που αποτυπώνουν τις πιθανότητες επιβίωσης για $k = 1, \dots, N$ σημεία της καμπύλης KM. Το πιο συχνά αναφερόμενο και χρησιμοποιούμενο λογισμικό που κάνει καλή εφαρμογή είναι το DigitizeIt. Βέβαια υπάρχουν πολλά ακόμα λογισμικά που εξάγουν τις συντεταγμένες των καμπυλών κάνοντας κλικ πάνω στα σημεία (mouse-clicks). Ωστόσο αυτή η διαδικασία απαιτεί τεράστια προσοχή καθώς τα δεδομένα πρέπει να είναι επαρκή, να καταγράφονται δηλαδή όλα τα βήματα που αποτυπώνονται στην καμπύλη. Ακόμα ο αριθμός των κλικ και το σημείο που θα γίνει έχει σημασία, αφού τα δεδομένα πρέπει να είναι και συνεπή. Αυτό σημαίνει ότι εφόσον η πιθανότητα πραγματοποίησης ενός συμβάντος μειώνεται με την πάροδο του χρόνου θα πρέπει αυτό να επαληθεύεται και για τα εξαγόμενα σημεία δεδομένων. Δεν είναι σπάνιο το γεγονός να προκύπτουν διάφορες αποκλίσεις σε αυτό το σημείο είτε λόγω της ποιότητας της εικόνας στη δημοσίευση είτε λόγω κάποιου ανθρώπινου λάθους στην συλλογή των σημείων κάνοντας απλά κλικ πάνω στην καμπύλη. Αυτές όμως θα πρέπει να διορθώνονται πριν την υλοποίηση του αλγορίθμου. Επιπλέον είναι πολύ σημαντικό σε αυτά τα αρχικά δεδομένα να περιλαμβάνονται οι χρόνοι για τους οποίους είναι διαθέσιμος ο αριθμός των ατόμων σε κίνδυνο. Τέλος, κατά κανόνα το πρώτο σημείο θεωρείται το $t_1 = 0$ και η πιθανότητα να συμβεί ένα γεγονός εκείνη τη χρονική στιγμή ισούται με τη μονάδα $S_1 = 1$.

Το επόμενο απαιτούμενο αρχείο περιέχει πληροφορίες για τους αναφερόμενους αριθμούς σε κίνδυνο (number at risk). Καθώς λοιπόν η μέθοδος Kaplan Meier λειτουργεί συνοψίζοντας το IPD με τη μορφή τ χρονικών διαστημάτων, τα οποία έχουν σχεδιαστεί με τέτοιο τρόπο ώστε να διαδραματίζεται τουλάχιστον ένα συμβάν στην αρχή τους, για κάθε ένα από αυτά ($i=1, \dots, n_\tau$) θα αναφέρονται τα:

n_{risk_i} : αριθμός ατόμων σε κίνδυνο στην αρχή του διαστήματος

t_{risk_i} : ο χρόνος κατά τον οποίο παρέχεται ο αριθμός σε κίνδυνο

l_i : ο αριθμός της πρώτης σειράς των εξαγόμενων συντεταγμένων γι' αυτό το διάστημα

u_i : ο αριθμός της τελευταίας σειράς των εξαγόμενων συντεταγμένων γι' αυτό το διάστημα

Το n_{risk_i} και t_{risk_i} προέρχονται από το κείμενο τη δημοσίευσης, ενώ τα l_i και u_i σχετίζονται με τον αριθμό των κλικ σε κάθε διάστημα. Για κάθε διάστημα i , το l_i ισούται με το k όταν $t_k = t_{risk_i}$ και αντίστοιχα το u_i ισούται με το k όταν $t_{k+1} = t_{risk_{i+1}}$.

Το τελευταίο απαιτούμενο δεδομένο για τον αλγόριθμο είναι ο συνολικός αριθμός των γεγονότων (totevents).

Σύμφωνα λοιπόν με τις απαιτούμενες (ή καλύτερα θεμιτές) εισαγωγές που περιγράφηκαν παραπάνω ο αλγόριθμος που ανέπτυξαν η Guyot και οι συνεργάτες της θα σκιαγραφηθεί για τέσσερα διαφορετικά ενδεχόμενα. Τα ενδεχόμενα αυτά καλύπτουν τις κατηγορίες «όλες οι πληροφορίες-all information», «όχι αριθμοί σε κίνδυνο-no number at risk», «όχι συνολικά γεγονότα-no total events» και «καμία πληροφορία-neither». Ξεκινώντας λοιπόν με την πρώτη κατηγορία όλων των πληροφοριών, θεωρείται γνωστός ο αριθμός των ατόμων σε κίνδυνο στην αρχή της μελέτης και σε ένα ακόμα σημείο τουλάχιστον καθώς και ο συνολικός αριθμός συμβάντων. Επίσης δεν θεωρείται γνωστός ο αριθμός των λογοκριμένων παρατηρήσεων και συνεπώς θα χρησιμοποιηθούν οι αναφερόμενοι αριθμοί σε κίνδυνο n_{risk_i} για να προσεγγιστεί ο αριθμός των λογοκρισιών σε κάθε διάστημα. Γενικά επειδή δεν δύναται ο προσδιορισμός του ακριβούς μοτίβου λογοκρισίας γίνεται η υπόθεση ότι η λογοκρισία λαμβάνει χώρα με σταθερό ρυθμό. Η εικασία αυτή είναι λογική στην περίπτωση της μη ενημερωτικής λογοκρισίας (non-informative censoring), όπου δηλαδή ο χρόνος λογοκρισίας ενός ασθενή είναι στατιστικά ανεξάρτητος από το χρόνο εμφάνισης του υπό μελέτη γεγονότος.

Ο αλγόριθμος περιλαμβάνει οχτώ βήματα.

All information

ΒΗΜΑ 1

Αρχικά γίνεται μία εικασία για τον αριθμό των λογοκριμένων παρατηρήσεων στο διάστημα i .

Αν δεν υπάρχουν λογοκριμένοι ασθενείς στο διάστημα αυτό, τότε ο αριθμός των ατόμων σε κίνδυνο στην αρχή του επόμενου διαστήματος θα είναι $n_{risk_{i+1}}^{nocensor}$ και θα ισούται με τον αριθμό των ατόμων σε κίνδυνο στην αρχή του διαστήματος i πολλαπλασιασμένο με την πιθανότητα να συμβεί το γεγονός στο διάστημα i δεδομένου ότι ήταν ζωντανός στην αρχή του διαστήματος αυτού. Μαθηματικά αυτό μεταφράζεται ως εξής:

$$n_{risk_{i+1}}^{nocensor} = n_{risk_i} S_{i+1} | S_i \quad (3.4.1)$$

Στη συνέχεια η συγκεκριμένη ποσότητα στρογγυλοποιείται στον πλησιέστερο ακέραιο.

Η αρχική εικασία για τη λογοκρισία του διαστήματος i υπολογίζεται από τη διαφορά του αναφερόμενου αριθμού ατόμων σε κίνδυνο στην αρχή του διαστήματος $i+1$ και του αριθμού σε κίνδυνο κάτω από την υπόθεση της μη λογοκρισίας έτσι:

$$\hat{n}_{censor_i} = n_{risk_{i+1}}^{nocensor} - n_{risk_{i+1}}$$

Και μέσω της (3.4.1) γίνεται:

$$\hat{n}_{censor_i} = n_{risk_i} S_{i+1} | S_{i+1} - n_{risk_{i+1}}$$

ΒΗΜΑ 2

Ο εκτιμώμενος αριθμός λογοκρισίας $c = 1, \dots, \hat{n}_{censor_i}$, κατανέμεται ομοιόμορφα στο διάστημα i με εκτιμώμενους χρόνους λογοκρισίας \hat{t}_{t_c} , οι οποίοι υπολογίζονται ως εξής:

$$\hat{t}_{t_c} = t_{l_i} + c(t_{l_{i+1}} - t_{l_i}) / (\hat{n}_{censor_i} + 1), \quad \text{όπου } c = 1, \dots, \hat{n}_{censor_i}$$

Ο εκτιμώμενος αριθμός των λογοκριμένων παρατηρήσεων μεταξύ δύο εξαγόμενων συνεταγμένων Kaplan Meier k και $k+1$ υπολογίζεται αθροίζοντας τον αριθμό των εκτιμώμενων χρόνων λογοκρισίας που βρίσκονται ανάμεσα στους χρόνους t_k και t_{k+1} . Μοντελοποιώντας

$$\hat{c}_k = \sum_{c=1}^{\hat{n}_{censor_i}} \left(\hat{t}_{t_c} I_{\{\hat{t}_{t_c} \in [t_k, t_{k+1}]\}} \right)$$

$$\text{όπου } I_{\{\hat{t}_{t_c} \in [t_k, t_{k+1}]\}} = \begin{cases} 1, & \text{αν } \hat{t}_{t_c} \in [t_k, t_{k+1}] \\ 0, & \text{αλλιώς} \end{cases}$$

ΒΗΜΑ 3

Εφόσον εκτιμήθηκε ο αριθμός των λογοκριμένων παρατηρήσεων, υπολογίζεται στη συνέχεια ο αριθμός των γεγονότων σε κάθε εξαγόμενη συνεταγμένη k , καθώς και ο αριθμός των ατόμων σε κίνδυνο για την επόμενη συνεταγμένη (\hat{n}_{k+1}).

Γνωρίζοντας ότι η εκτίμηση Kaplan Meier για τη συνάρτηση επιβίωσης στο χρόνο m περιγράφεται από την ακόλουθη σχέση

$$S^{KM}(t_m) = \prod_{j=1}^m \frac{n_j - d_j}{n_j} = S^{KM}(t_{m-1}) \left(1 - \frac{d_m}{n_m} \right), \quad \text{όπου } m = 1, \dots, r$$

λαμβάνεται ότι ο αριθμός των γεγονότων \hat{d}_k ισούται με τον αριθμό των ατόμων που βρίσκονται σε κίνδυνο στην εξαγόμενη KM συντεταγμένη k , πολλαπλασιασμένο με ένα μείον την πιθανότητα να συμβεί το γεγονός στην εξαγόμενη συντεταγμένη k , διαιρεμένο με την εκτίμηση της πιθανότητας επιβίωσης $\hat{S}_{last(k)}^{KM}$ στην προηγούμενη συντεταγμένη που εκτιμήθηκε ότι έλαβε χώρα το γεγονός $last(k)$.

$$\hat{d}_k = \hat{n}_k \left(1 - \frac{S_k}{\hat{S}_{last(k)}^{KM}} \right), \text{ όπου } k = l_i, \dots, u_i$$

στρογγυλοποιημένο στον πλησιέστερο ακέραιο.

Εδώ πρέπει να αναφερθεί ότι παρόλο που τα διαστήματα των εκτιμήσεων KM έχουν σχεδιαστεί έτσι ώστε να συμβαίνει τουλάχιστον ένα γεγονός στην αρχή του διαστήματος, εδώ λόγω του μη καλά καθορισμένου τρόπου εξαγωγής των συντεταγμένων, δε ισχύει απαραίτητα και επομένως πρέπει να λαμβάνουμε υπόψη τη χρονική στιγμή που πραγματοποιήθηκε το τελευταίο γεγονός και έτσι ισχύει:

$$last(k) = \begin{cases} 1, & \text{αν } k = 1 \\ k', & \text{αλλιώς} \end{cases}, \text{ όπου } k' \text{ είναι τέτοιο ώστε } \hat{d}_{k'} > 0, \text{ αλλά } \hat{d}_j = 0 \text{ για } j = k' + 1, \dots, k - 1$$

Και η συνάρτηση της πιθανότητας επιβίωσης γίνεται

$$\hat{S}_k^{KM} = \begin{cases} 1, & \text{εάν } k = 1 \\ \hat{S}_{last(k)}^{KM} \left(1 - \frac{\hat{d}_k}{\hat{n}_k} \right), & \text{αλλιώς} \end{cases}$$

Ο αριθμός των ατόμων σε κίνδυνο για κάθε εξαγόμενη συντεταγμένη k υπολογίζεται ως εξής:

$$\hat{n}_{k+1} = \hat{n}_k - \hat{d}_k - \hat{c}_k, \text{ όπου } k = l_i, \dots, u_i$$

και ακόμα στην αρχή του διαστήματος ισχύει $\hat{n}_{l_i} = n_{risk_{l_i}}$.

Μέσω αυτής της προσέγγισης γίνεται και μία εκτίμηση του αριθμού των ατόμων σε κίνδυνο στην αρχή του επόμενου διαστήματος, όπου $\hat{n}_{risk_{i+1}} = \hat{n}_{u_{i+1}}$

ΒΗΜΑ 4

Εάν $\hat{n}_{risk_{i+1}} \neq n_{risk_{i+1}}$ τότε προσαρμόζεται εκ νέου ο αριθμός των λογοκριμένων παρατηρήσεων στο διάστημα i

$$\hat{n}_{censor_i} = \hat{n}_{censor_i} + (\hat{n}_{u_{i+1}} - n_{risk_{i+1}})$$

Στη συνέχεια επαναλαμβάνονται τα βήματα 2-3 έως ότου ο εκτιμώμενος και ο δημοσιευμένος αριθμός των ατόμων σε κίνδυνο να ταυτίζονται ($\hat{n}_{risk_{i+1}} = n_{risk_{i+1}}$).

ΒΗΜΑ 5

Εάν το $i+1$ δεν είναι το τελευταίο διάστημα, τότε τα βήματα 1-4 επαναλαμβάνονται και για τα ακόλουθα διαστήματα.

ΒΗΜΑ 6

Γενικά στις δημοσιευμένες μελέτες δεν αναφέρεται συχνά ο αριθμός των ατόμων που βρίσκονται σε κίνδυνο στο τέλος του τελευταίου διαστήματος n_τ .

Έτσι γίνεται αρχικά η υπόθεση ότι ο αριθμός των λογοκριμένων παρατηρήσεων στο τελευταίο διάστημα ισούται με το συνολικό αριθμό των εκτιμώμενων λογοκρισιών πριν από το τελευταίο διάστημα $\sum_{i=1}^{n_\tau-1} \hat{n}_{censor_i}$, σταθμισμένο με τον υπολειπόμενο χρόνο σε σχέση με τον χρόνο που έχει ήδη παρέλθει. Έπειτα γίνεται στρογγυλοποίηση στον πλησιέστερο ακέραιο. Ωστόσο αν αυτός ο αριθμός είναι μεγαλύτερος από τον αριθμό των ατόμων που είναι ακόμα σε κίνδυνο στην αρχή του τελευταίου διαστήματος, τότε θα επιλεγεί αυτός ο αριθμός των ασθενών που είναι ακόμα σε κίνδυνο. Μαθηματικά μιλώντας

$$\hat{n}_{censor_{n_{int}}} = \min \left(\frac{T_{u_{n_\tau}} - T_{l_{n_\tau}}}{T_{u_{n_\tau-1}} - T_{l_1}} \sum_{i=1}^{n_\tau-1} \hat{n}_{censor_i}; n_{risk_{n_\tau}} \right)$$

Έπειτα επανεκτελούνται τα βήματα 2-3.

ΒΗΜΑ 7

Στο βήμα αυτό χρησιμοποιείται ο δημοσιευμένος συνολικός αριθμός γεγονότων (totevents).

Αρχικά υπολογίζεται ο εκτιμώμενος συνολικός αριθμός γεγονότων που λαμβάνεται στην αρχή του τελευταίου διαστήματος.

$$\sum_{k=1}^{u_{n_\tau-1}} \hat{d}_k$$

Εάν ο εκτιμώμενος συνολικός αριθμός γεγονότων που λαμβάνεται στην αρχή του τελευταίου διαστήματος είναι μεγαλύτερος ή ίσος από το συνολικό αριθμό γεγονότων γίνεται η υπόθεση ότι δεν συμβαίνουν άλλα γεγονότα ή λογοκρισίες και συνεπώς

$$\hat{d}_k = 0, \hat{c}_k = 0, \hat{n}_k = n_{u_{n_\tau-1}}, \text{ όπου } k = l_{n_\tau}, \dots, u_{n_\tau}.$$

ΒΗΜΑ 8

Εάν ο εκτιμώμενος συνολικός αριθμός γεγονότων που λαμβάνεται στην αρχή του τελευταίου διαστήματος είναι μικρότερος από το συνολικό αριθμό γεγονότων, προσαρμόζεται εκ νέου με τη βοήθεια της διαφοράς ο εκτιμώμενος αριθμός των λογοκριμένων παρατηρήσεων του τελευταίου διαστήματος (n_τ), $\hat{n}_{censor_{n_\tau}}$.

Έτσι:

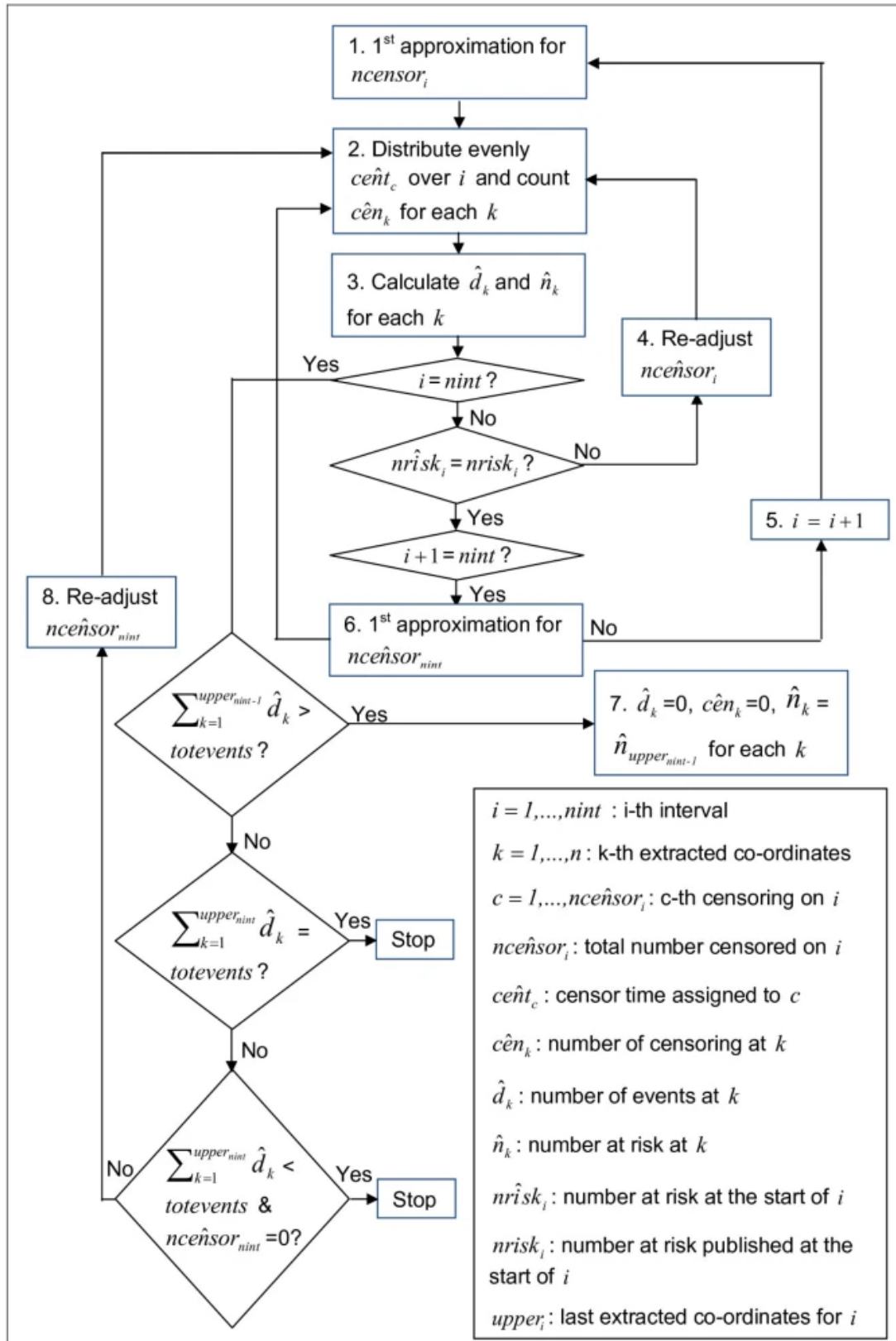
$$\hat{n}_{censor_{n_\tau}} = \hat{n}_{censor_{n_\tau}} + \left(\sum_{k=1}^{u_{n_\tau}} \hat{d}_k - totevents \right)$$

Στη συνέχεια επανεκτελούνται τα βήματα 2,3 και 8 για το τελευταίο διάστημα, n_τ , μέχρι

$$\sum_{k=1}^{u_{n_\tau}} \hat{d}_k = totevents$$

ή

$$\sum_{k=1}^{u_{n_\tau}} \hat{d}_k < totevents \quad \text{και} \quad \hat{n}_{censor_{n_\tau}} = 0$$



Εικόνα 3: Διάγραμμα ροής αλγορίθμου Guyot (Guyot, et al., 2012)

3.4.1 Προσαρμογές αλγορίθμου

Ο αλγόριθμος που περιγράφηκε παραπάνω εκμεταλλεύεται όλα τα δεδομένα που μπορεί να παρέχονται με στόχο την καλύτερη απόδοση των εκτιμήσεων. Ωστόσο στις περισσότερες των περιπτώσεων δεν είναι γνωστές όλες αυτές οι πληροφορίες. Για το σκοπό αυτό στη συνέχεια αναφέρονται οι απαιτούμενες προσαρμογές του αλγορίθμου σε περιστάσεις όπου κάποια ή όλες οι πληροφορίες απουσιάζουν.

I) Δεν είναι γνωστός ο αριθμός των ατόμων σε κίνδυνο – “**no number at risk**”

Στην περίπτωση αυτή ο χρόνος δεν διαιρείται και υπάρχει ένα και μόνο διάστημα, $n_{\tau}=1$. Το μοτίβο λογοκρισίας παραμένει σταθερό στο χρόνο και γίνεται η αρχική υπόθεση ότι ο συνολικός αριθμός που λογοκρίθηκε είναι ίσος με το μηδέν. Στη συνέχεια εκτελούνται όλα τα βήματα του αλγορίθμου (1-8).

II) Δεν είναι γνωστός ο συνολικός αριθμός των γεγονότων – “**no total events**”

Στην δεδομένη περίπτωση το μοτίβο λογοκρισίας εξαρτάται από τον αριθμό των ατόμων που βρίσκονται σε κίνδυνο, ενώ παραμένει σταθερό σε κάθε διάστημα. Ακόμη δεν δύναται να γίνει προσαρμογή του αλγορίθμου για τον συνολικό αριθμό συμβάντων και συνεπώς ο αλγόριθμος τερματίζει στο έκτο βήμα.

III) Δεν είναι διαθέσιμη καμία πληροφορία – “**neither**”

Στην περίπτωση όπου ούτε ο συνολικός αριθμός συμβάντων αναφέρεται αλλά ούτε και οι αριθμοί σε κίνδυνο μετά την έναρξη της μελέτης τότε γίνεται η υπόθεση ότι δεν υπάρχουν λογοκριμένες παρατηρήσεις. Πρόκειται για μία πολύ ισχυρή υπόθεση, όμως εξίσου ισχυρή θα ήταν οποιαδήποτε υπόθεση και να γινόταν χωρίς καμία πληροφόρηση. Γενικά αναμένεται χαμηλότερη ποιότητα αποτελεσμάτων ελλείψει πληροφοριών.

Εφαρμόζοντας λοιπόν τον συγκεκριμένο αλγόριθμο χρησιμοποιώντας το αντίστοιχα διαθέσιμο πλήθος πληροφοριών, εξάγονται τα ανακατασκευασμένα δεδομένα Kaplan Meier από τα οποία γίνεται εύκολα η λήψη των ατομικών δεδομένων ασθενούς (IPD). Πιο συγκεκριμένα για κάθε εξαγόμενη συντεταγμένη KM, $k=1, \dots, N$ δίνονται οι ανακατασκευασμένοι παράμετροι \hat{d}_k , \hat{c}_k , \hat{n}_k και μπορεί να γίνει η λήψη των δεδομένων (IPD) από τα οποία θα μπορούσαν να είχαν προέλθει, με τον εξής τρόπο. Κάθε αναφερόμενο γεγονός ή λογοκρισία αντιστοιχίζεται με το χρόνο διενέργειας καθώς και με ένα δείκτη συμβάντων, όπου συνήθως η μονάδα δηλώνει την ύπαρξη συμβάντος και το μηδέν τη λογοκρισία.

Αξίζει να αναφερθεί ότι αποτελεί μία από τις λίγες μεθόδους για τις οποίες έχει πραγματοποιηθεί αξιολόγηση των ανακατασκευασμένων δεδομένων τους. Τέτοιες διαδικασίες έχουν αναφερθεί ξανά στο παρελθόν από τον Earle όπου είχε αξιολογήσει την αναπαραγωγικότητα της χρήσης ψηφιακών λογισμικών για την εξαγωγή δεδομένων υπολογίζοντας ένα συντελεστή συσχέτισης (Earle & Wells, 2000) και από τον Parmar όπου μέσω του υπολογισμού του μέσου σφάλματος (ME) 48 μελετών μεταξύ του ανακατασκευασμένου λόγου κινδύνου από τις δημοσιευμένες καμπύλες και από άλλους πιο άμεσους τρόπους εξαγωγής, δεν βρήκε κάποια συστηματική μεροληψία στον υπολογισμό του HR παρά μία μικρή υποεκτίμηση της διακύμανσής του. Προκειμένου λοιπόν να αξιολογηθεί η αναπαραγωγικότητα και η ακρίβεια της εν λόγω μεθόδου και για τα τέσσερα προαναφερθέντα διαφορετικά επίπεδα πληροφοριών, επιλέχθηκαν από τους συγγραφείς έξι καμπύλες Kaplan Meier, από τέσσερις δημοσιευμένες μελέτες οι οποίες είχαν χρησιμοποιηθεί σε μία ανασκόπηση των μεθόδων της ανάλυσης επιβίωσης. Μέσω αυτών συλλέχθηκαν μερικά συνοπτικά στατιστικά (22 πιθανότητες επιβίωσης (survival probabilities), 7 διάμεσοι χρόνοι επιβίωσης (median survival times), 6 λόγοι κινδύνου (hazard ratios) και 4 τυπικά σφάλματα αυτών (standard errors of the log hazard ratios)), κάθε ένα από τα οποία κατασκευάστηκε εκ νέου κάτω από δύο διαφορετικές περιπτώσεις διαθέσιμων πληροφοριών. Μόνο για τις πιθανότητες επιβίωσης χρησιμοποιήθηκε η φυσική κλίμακα, για τα υπόλοιπα μέτρα χρησιμοποιήθηκε η λογαριθμική κλίμακα. Στη συνέχεια εφαρμόστηκε ένα μοντέλο two-way ANOVA για τη μέτρηση των διαφορών μεταξύ των ανακατασκευασμένων και των αρχικών δημοσιευμένων μέτρων.

Πιο αναλυτικά, η αναπαραγωγικότητα αντιπροσωπεύει το σφάλμα που μπορεί να υποπέσει ένας ερευνητής στην ανακατασκευή ενός συγκεκριμένου στατιστικού μέτρου. Η εκτίμησή της πραγματοποιήθηκε από το άθροισμα του εντός και μεταξύ σφάλματος του ερευνητή (within and between observer error). Ακόμη χρησιμοποιήθηκε προσομοίωση Monte Carlo (από το μοντέλο της ANOVA) για τη λήψη διαστήματος εμπιστοσύνης των τυπικών αποκλίσεων. Θεωρήθηκε ότι οι βαθμοί ελευθερίας της εντός και μεταξύ μεταβλητότητας των ερευνητών ακολουθούν την χ^2 κατανομή και προκειμένου να διασφαλιστεί η αξιοπιστία των συμπερασμάτων λήφθηκαν 150000 δείγματα βαθμών ελευθερίας για κάθε πηγή μεταβλητότητας. Στη συνέχεια οι εκτιμήσεις των μέσων τετραγώνων για κάθε ένα από τα δείγματα και για κάθε μία παρατηρούμενη μεταβλητότητα υπολογίστηκαν χρησιμοποιώντας το άθροισμα των τετραγώνων που προέκυψε από την ANOVA και το καθορισμένο δείγμα. Τέλος εκτιμήθηκαν οι τυπικές αποκλίσεις της εντός και μεταξύ μεταβλητότητας καθώς και της έκβασης και εξήχθησαν τα 2.5 και 97.5 εκατοστημόρια που καθορίζουν τα διαστήματα εμπιστοσύνης.

Όσο αφορά την αξιολόγηση της ακρίβειας της μεθόδου εξετάστηκε η μέση διαφορά (mean difference) μεταξύ των ανακατασκευασμένων και αρχικών στατιστικών. Από το εξαγόμενο μέσο σφάλμα (ME-mean error) προέκυψε συστηματική υπερ-ή υπό εκτίμηση. Τα διαστήματα εμπιστοσύνης πάρθηκαν από το

μοντέλο της ANOVA. Ακόμη υπολογίστηκε το μέσο απόλυτο σφάλμα (Mean Absolute Error) το οποίο είναι ένα μέτρο απόλυτης ακρίβειας των αποτελεσμάτων καθώς δίνει πληροφορίες για την ακρίβεια της ανακατασκευής που μπορεί να αναμένεται σε μία συγκεκριμένη περίπτωση. Γενικά στην άσκηση επικύρωσης που εκτελέστηκε βρέθηκε ότι όσο αφορά τις πιθανότητες επιβίωσης και τους διάμεσους χρόνους επιβίωσης τα αποτελέσματα ήταν ικανοποιητικά, με μικρά μέσα σφάλματα (ME) και δεν ήταν τόσο ευαίσθητα στην επιδείνωση του επιπέδου των πληροφοριών. Από την άλλη οι λόγοι κινδύνου κρίθηκαν επαρκείς για τις τρεις πρώτες περιπτώσεις πληροφοριών ενώ αντιμετωπίζουν μεγάλα προβλήματα στην περίπτωση όπου δεν είναι διαθέσιμη καμία πληροφορία-“neither”. Αυτό προκύπτει καθώς το μέσο σφάλμα των τυπικών τους αποκλίσεων έγινε στατιστικά αρνητικό υποδηλώνοντας την υποεκτίμηση της αβεβαιότητας και αυτό φυσικά προέρχεται από την υπόθεση της μη λογοκρισίας. Ανακατασκευάστηκαν δηλαδή οι λόγοι κινδύνου όντας πιο ευάλωτοι στο επίπεδο πληροφοριών και με λιγότερη ακρίβεια. Ουσιαστικά οι υποθέσεις για το βαθμό λογοκρισίας αντανακλούν τη σχετική στάθμιση των διάφορων τμημάτων της καμπύλης και όσο το επίπεδο των πληροφοριών μειώνεται τόσο οι υποθέσεις απομακρύνονται από την πραγματικότητα. Η διαφορά αυτή στα εν λόγω συνοπτικά μέτρα πιθανώς οφείλεται στο γεγονός ότι ο λόγος κινδύνου είναι στην πραγματικότητα ένας σταθμισμένος μέσος όρος των λόγων σε ολόκληρη την περίοδο κινδύνου, ενώ οι πιθανότητες επιβίωσης και οι διάμεσοι χρόνοι επιβίωσης αποτελούν σημειακές εκτιμήσεις.

Συμπερασματικά λοιπόν, η μέθοδος αυτή αποτελεί μία βελτιωμένη προσέγγιση από τις μέχρι τώρα υπάρχουσες, αφού είναι η μοναδική που χρησιμοποιεί επαναληπτικό αριθμητικό αλγόριθμο για την επίλυση των ανεστραμμένων εξισώσεων Kaplan Meier, καθώς επίσης λαμβάνει υπόψη τη χρήση των ατόμων σε κίνδυνο και τον συνολικό αριθμό συμβάντων, προκειμένου να διασφαλιστεί η συνέπεια μεταξύ των διαδοχικών εκτιμήσεων. Επιπρόσθετα έδωσε πρόσβαση στα μεμονωμένα δεδομένα ασθενών (IPD) παρέχοντας μεγάλη απελευθέρωση στις αναλύσεις επιβίωσης, αφού επιτρέπει μία πιο κατατοπιστική ανάλυση των δεδομένων που εξαρτώνται από το χρόνο, και ειδικότερα στις μετα-αναλύσεις αυτών.

Από την άλλη πλευρά ωστόσο διακατέχεται και από κάποιους περιορισμούς. Ο πρώτος αφορά ένα κοινό πρόβλημα των μετα-αναλύσεων, όπου όσο οι δημοσιευμένες καμπύλες συγκεντρώνουν δεδομένα σε διαφορετικές συμμεταβλητές που ενδέχεται να επηρεάσουν την επιβίωση, τόσο το χάσμα μεταξύ του πραγματικού και του ανακατασκευασμένου IPD μεγαλώνει. Ο δεύτερος περιορισμός αφορά τα στοιχεία που επηρεάζουν την αξιοπιστία και εγκυρότητα των ανακατασκευασμένων δεδομένων. Αυτά κατά κύριο λόγο συνοψίζονται στην ποιότητα της αρχικής εισαγωγής των δεδομένων όπου πρέπει να είναι επαρκή και συνεπή καθώς και στο επίπεδο των πληροφοριών που παρέχεται από την εκάστοτε δημοσιευμένη μελέτη.

3.5 Μελέτη επικύρωσης των τεσσάρων μεθόδων

Το 2019 η Saluja και οι συνεργάτες της πραγματοποίησαν μία μελέτη επικύρωσης των τεσσάρων μεθόδων που είχαν προταθεί από τους Parmar, Williamson, Holey and Henley και Guyot αντίστοιχα (Saluja, et al., 2019). Στόχος της μελέτης αποτελούσε ο προσδιορισμός της βέλτιστης μεθοδολογικά μεθόδου στην εκτίμηση των λόγων κινδύνου (HRs) εκμεταλλευόμενοι τις δημοσιευμένες καμπύλες Kaplan-Meier.

Εκτός από την πρωτογενή ανάλυση διενεργήθηκε και μία δευτερεύουσα με σκοπό την αύξηση της αξιοπιστίας της μελέτης. Οι καμπύλες KM για την πρώτη ανάλυση συλλέχθηκαν από τυχαιοποιημένες ελεγχόμενες κλινικές δοκιμές που δημοσιεύθηκαν από την παν-Καναδική ογκολογική ανασκόπηση φαρμάκων (pCODR) για αντικαρκινικά φάρμακα και για τη δευτερογενή ξανά από τυχαιοποιημένες κλινικές δοκιμές που χρησιμοποιούνται για εγκρίσεις φαρμάκων της υπηρεσίας τροφίμων και φαρμάκων (FDA). Κριτήρια επιλεξιμότητας αποτελούσαν η δημοσίευση καμπυλών KM συνολικής επιβίωσης (OS, overall survival) και/ή επιβίωσης χωρίς εξέλιξη (PFS, progression free survival), η αναφορά του HR καθώς και η παράθεση των αριθμών των ατόμων που ήταν σε κίνδυνο σε διάφορα χρονικά διαστήματα.

Αφού λοιπόν συλλέχθηκαν οι καμπύλες, ψηφιοποιήθηκαν από δύο ανεξάρτητους αξιολογητές οι οποίοι δεν γνώριζαν τα αναφερόμενα συνοπτικά μέτρα HRs. Έπειτα ανακατασκεύασαν τους λόγους κινδύνου σύμφωνα με τις προαναφερθείσες τέσσερις μεθόδους. Η αξιοπιστία μεταξύ των αναθεωρητών (Inter-rater reliability) αξιολογήθηκε μέσω μιας δημοφιλούς μεθόδου, υπολογίζοντας το συντελεστή συσχέτισης Intra-class (Intra-class Correlation Coefficient-ICC). Όταν $ICC < 0.4$ υποδηλώνει κακή αξιοπιστία, όταν $0.4 \leq ICC < 0.75$ αρκετά καλή αξιοπιστία και όταν $ICC \geq 0.75$ τέλεια αξιοπιστία.

Στη συνέχεια όσο αφορά τον έλεγχο της αξιοπιστίας και της ακρίβειας των μεθόδων πραγματοποιήθηκε μέσω διαγραμμάτων Bland-Altman καθώς και με τη χρήση συνοπτικών στατιστικών (Bland & Altman, 1986; Bland & Altman, 1999; Liu, et al., 2014). Οι Bland και Altman θέλοντας να περιγράψουν τη συμφωνία μεταξύ δύο ποσοτικών μετρήσεων εισήγαγαν μία μέθοδο για την ποσοτικοποίηση της συμφωνίας τους κατασκευάζοντας κάποια όρια. Αυτά τα στατιστικά όρια υπολογίζονται χρησιμοποιώντας το μέσο όρο και την τυπική απόκλιση των διαφορών μεταξύ των μετρήσεων. Στη συνέχεια μέσω μιας γραφικής προσέγγισης γίνεται διερεύνηση της κανονικότητας. Το γράφημα που λαμβάνεται αφορά ένα διάγραμμα διασποράς, στο οποίο ο y άξονας αντιπροσωπεύει τη διαφορά μεταξύ των δύο μετρήσεων ($X_1 - X_2$) ενώ ο x άξονας το μέσο όρο αυτών των μέτρων $((X_1+X_2)/2)$. Επιπρόσθετα για να διευκολυνθεί η σύγκριση των ανακατασκευασμένων λόγων κινδύνου με τους αρχικά αναφερόμενους υπολογίστηκαν το μέσο σφάλμα (ME), το μέσο απόλυτο σφάλμα (MAE), το μέσο

τετραγωνικό σφάλμα (MSE) και η ρίζα του μέσου τετραγωνικού σφάλματος (RMSE) του logHR για κάθε γραφική μέθοδο χρησιμοποιώντας τα αναφερόμενα HRs.

Τα αποτελέσματα τόσο από την πρώτη όσο από τη δευτερεύουσα ανάλυση έδωσαν τις χαμηλότερες τιμές στα στατιστικά μέτρα με τη χρήση της μεθόδου Guyot, υποδεικνύοντας την απουσία οποιασδήποτε συστηματικής μεροληψίας. Ωστόσο ούτε με τη μέθοδο των Hoyle και Henley εμφανίστηκε κάποια μεροληψία σε αντίθεση με τις μεθόδους του Parmar και του Williamson όπου υπήρξε μεροληψία προς τα πάνω, δηλαδή υπερεκτίμηση των ανακατασκευασμένων λόγων κινδύνου. Επιπλέον στην δευτερεύουσα ανάλυση η μέθοδος του Williamson έδωσε υψηλότερο ME αλλά χαμηλότερες τιμές για MAE, MSE και RMSE σε σύγκριση με τη μέθοδο Parmar, καθώς επίσης έδωσε χαμηλότερες τιμές MSE και RMSE σε σύγκριση με τη μέθοδο Hoyle και Henley. Στη συνέχεια μέσω της αξιολόγησης της συμφωνίας των μετρήσεων από τα γραφήματα Bland-Altman και πάλι η μέθοδος της Guyot επέδειξε την καλύτερη συμφωνία ενώ σε όλες τις περιπτώσεις η αξιοπιστία διαβαθμιστή ήταν εξαιρετική (ICC >0.75).

Τέλος πραγματοποιήθηκαν και αναλύσεις ευαισθησίας με σκοπό την αξιολόγηση τόσο των επίπεδων έναντι των απότομων καμπυλών όσο και της αναλογικότητας των κινδύνων. Όσο αφορά την αξιολόγηση των καμπυλών η μέθοδος της Guyot επέδειξε για ακόμη μία φορά τις χαμηλότερες τιμές ME, MAE, MSE & RMSE του logHR και την καλύτερη συμφωνία με το χρυσό πρότυπο. Οι αμέσως καλύτερες μέθοδοι για επίπεδες καμπύλες μοιάζουν να είναι των Hoyle και Henley έπειτα του Williamson και τέλος του Parmar, ενώ για τις απότομες καμπύλες η μέθοδος του Williamson ξεπέρασε αυτή των Hoyle και Henley έχοντας μικρότερο ME και μικρότερη μεροληψία. Αντίστοιχη εικόνα φάνηκε και με την αξιολόγηση των αναλογικών κινδύνων με τη μέθοδο της Guyot να επιδεικνύει το μικρότερο μέσο σφάλμα στο λογαριθμισμένο λόγο κινδύνου καθώς και την μικρότερη μεροληψία ενώ ακολούθως είναι η μέθοδος των Hoyle και Henley και έπειτα του Williamson και Parmar.

Εν κατακλείδι η μέθοδος Guyot κρίθηκε πιο ακριβής στην παραγωγή των εκτιμήσεων του λόγου κινδύνου με τα στενότερα διαστήματα εμπιστοσύνης ενώ αν υπάρχουν προβλήματα πόρων και απαιτείται εναλλακτική λύση, η μέθοδος των Hoyle και Henley είναι η αμέσως επόμενη καλύτερη συνολικά, η οποία δεν απαιτεί τόσο μεγάλη προετοιμασία δεδομένων όσο της Guyot.

3. 6 Βελτιώσεις αλγορίθμου Guyot

1)Βελτιώσεις από τους Wei και Royston

Όπως έχει αναφερθεί και προηγουμένως παρουσιάζονται αρκετές ελλείψεις στα μέτρα που εξάγονται από τις αναλύσεις επιβίωσης. Το γεγονός αυτό δυσκολεύει αφενός τη διεξαγωγή συμπερασμάτων καθώς και τη διενέργεια οποιασδήποτε δευτερογενούς ανάλυσης. Ωστόσο όπως

περιγράφηκε παραπάνω έχουν γίνει προσπάθειες ανάκτησης των ατομικών δεδομένων επιβίωσης από τις δημοσιευμένες καμπύλες Kaplan Meier με επικρατέστερη μέθοδο αυτή της Guyot.

Σε συνέχεια της πρότασης της Guyot και των συνεργατών της, οι Wei και Royston (2017) υλοποίησαν εκ νέου τον υπάρχοντα αλγόριθμο κάνοντας κάποιες βελτιώσεις και στη συνέχεια εφάρμοσαν κάποια παραδείγματα με τη βοήθεια ενός ψηφιακού λογισμικού (Parner & Andersen, 2010). Πριν από την παράθεση των παρεμβάσεών τους, αξίζει να αναφερθεί η διευκρίνηση που γίνεται ότι στην εισαγωγή των δεδομένων δεν είναι αναγκαία η χρήση κατά αποκλειστικότητα των πιθανοτήτων επιβίωσης αλλά μπορεί να αναφέρονται σε πιθανότητες αποτυχίας ή ακόμα σε ποσοστά επιβίωσης και αποτυχίας ανάλογα με το εννοιολογικό περιεχόμενο της μελέτης. Ακόμη αναφέρεται η σημαντικότητα της εξαγωγής του συνολικού αριθμού των ασθενών που τυχαιοποιούνται σε κάθε ομάδα μιας δοκιμής καθώς και στη βελτίωση που παρατηρείται στα ανακατασκευασμένα δεδομένα ενσωματώνοντας πληροφορίες που αφορούν τον αριθμό των ασθενών σε κίνδυνο σε τακτά χρονικά διαστήματα κατά τη διάρκεια της παρακολούθησης.

Τα επιπρόσθετα στοιχεία που εισάγονται σε αυτή την προσέγγιση αφορούν αρχικά την ικανότητα προσδιορισμού της αντιστοίχισης του κατώτερου και ανώτερου τελικού σημείου κάθε χρονικού διαστήματος του πίνακα κινδύνου (risk table). Πιο αναλυτικά, υπολογίζονται το l_i και u_i χρησιμοποιώντας τα δεδομένα εισόδου t_k και $trisk_i$ ως εξής:

$$l_i = \min\{k: t_k \geq trisk_i\}$$

$$u_i = \max\{k: t_k \leq trisk_{i+1}\}$$

όπου $k = 1, \dots, K$ με K : ο συνολικός αριθμός των σημείων που εξάγονται από τις καμπύλες και $i = 1, \dots, T$ με T : ο αριθμός των διαστημάτων όπου το πλήθος των ατόμων σε κίνδυνο αναφέρονται. Πλέον δηλαδή το l_i και u_i δεν απαιτούνται ως δεδομένα εισαγωγής στον αλγόριθμο αλλά υπολογίζονται από άλλα προτευντα δεδομένα που έχουν είδη εισαχθεί.

Έπειτα γνωρίζοντας ότι η συνάρτηση επιβίωσης μειώνεται εξ ορισμού με το χρόνο εισάγεται στον αλγόριθμο ένας τρόπος διόρθωσης των παραβατικών δεδομένων αυτής της αρχής. Η μη ικανοποίηση αυτού του περιορισμού μπορεί να οφείλεται τόσο στην ποιότητα των δημοσιευμένων καμπυλών Kaplan Meier όσο και σε σφάλματα που υπόκεινται οι ερευνητές κατά την εξαγωγή των συντεταγμένων κάνοντας απλά κλικ σε διάφορα σημεία της καμπύλης KM (μέθοδος των κλικ). Επειδή λοιπόν αυτές οι παραβάσεις είναι συχνές και η διόρθωσή τους επιτακτική οι Wei και Royston πρότειναν δύο μεθόδους διόρθωσης.

- i. Ο πρώτος τρόπος αφορά την ισοτονική παλινδρόμηση (Barlow et al, 1972). Σε αυτή γίνεται χρήση του αλγορίθμου pool-adjacent-violators ο οποίος βρίσκει και αντικαθιστά τις

παραβατικές εισαγωγές με τη μέση τιμή τους, έτσι ώστε τα δεδομένα να ικανοποιούν τον περιορισμό της μονοτονίας (van Putten και Royston, 2017).

- ii. Ο δεύτερος από την άλλη προτείνει την αντικατάσταση της τιμής του παραβάτη με του διπλανού του έτσι ώστε και πάλι να διατηρείται η μονοτονία

Και τις δύο αυτές περιπτώσεις διορθώνεται η απόκλιση από την μονοτονία και τα αποτελέσματα είναι παρόμοια καθώς η απόλυτη διαφορά μεταξύ των τιμών των παραβάσεων είναι συχνά μικρή για να μπορέσει να προξενήσει ιδιαίτερη επίδραση στα δεδομένα.

Τέλος, επέτρεψαν στον αλγόριθμο να λαμβάνει υπόψη του τον αριθμό των ατόμων σε κίνδυνο στο τελευταίο χρονικό διάστημα, όταν αυτό αναφέρεται. Με αυτό τον τρόπο έδωσαν τη δυνατότητα υπολογισμού του αριθμού των λογοκριμένων δεδομένων στο τελευταίο διάστημα με χρήση των μεθόδων που εφαρμόζονται και για τα υπόλοιπα διαστήματα.

Σύμφωνα λοιπόν με κάποια παραδείγματα που δημοσίευσαν οι ίδιοι από την εν λόγω προσέγγιση βρέθηκε ότι οι εκτιμήσεις των συνοπτικών αυτών στατιστικών ήταν παρόμοια με αυτή της Guyot. Δηλαδή ενώ οι εκτιμήσεις των διάρκειων επιβίωσης ήταν πολύ κοντά με αυτές που υπήρχαν στις αρχικές δημοσιεύσεις, η προσέγγιση του λόγου κινδύνου ήταν κοντά αλλά δεν ταυτίζονταν. Αυτή η απόκλιση ενδέχεται να οφείλεται στη μη ακριβή εκτίμηση του αριθμού και της θέσης των γεγονότων από τον αλγόριθμο. Μία τελευταία παρατήρηση που έγινε είναι ότι η ανακατασκευή των καμπυλών KM καθώς και η εκτίμηση του λόγου κινδύνου (hazard ratio) είχαν αρκετά καλή ταύτιση με αυτές που μπορούμε να πάρουμε από τα αρχικά δεδομένα. Το γεγονός αυτό υποδηλώνει ότι η χρήση μη αναλογικών κινδύνων δεν επηρεάζει σε σημαντικό βαθμό την ακρίβεια της μεθόδου, εντούτοις χρήζει περαιτέρω διερεύνησης.

II) Αναπροσαρμογή αλγορίθμου Guyot από τους Liu, Zhou και Lee

Έχοντας ως γνώμονα την αναγκαιότητα ανάκτησης των μεμονωμένων δεδομένων των ασθενών από δημοσιευμένες καμπύλες Kaplan Meier και την κλασική μέθοδο που χρησιμοποιείται, η οποία αποτελεί έναν επαναληπτικό αλγόριθμο που βασίζεται στη μέθοδο εκτίμησης των KM δεδομένων, συνοψίζονται διάφοροι περιορισμοί της που μπορούν να προκαλέσουν σοβαρές αποκλίσεις στα εξαγόμενα δεδομένα.

Αρχικά πριν από τη χρήση του αλγορίθμου, για την εξαγωγή των συντεταγμένων των εκάστοτε καμπυλών απαιτείται εξωτερικό λογισμικό (συνήθως DigitizeIt) το οποίο μέσω κλικ του ποντικιού

συλλέγει τα απαιτούμενα δεδομένα. Η διαδικασία αυτή ωστόσο είναι ασταθής και η κατάσταση χειροτερεύει όταν στο ίδιο γράφημα διαπλέκονται αρκετές καμπύλες. Για τα σημεία τα οποία λαμβάνονται θα πρέπει επίσης οι χρήστες να τα ταξινομήσουν έτσι ώστε η πιθανότητα επιβίωσης να μειώνεται με την πάροδο του χρόνου, σε καθορισμένα χρονικά διαστήματα στα οποία αναφέρεται ο αριθμός των ατόμων που βρίσκονται σε κίνδυνο. Επιπλέον λόγω του μη περιορισμού των επιτρεπόμενων λογοκριμένων παρατηρήσεων η επαναληπτική διαδικασία μπορεί να οδηγηθεί σε ατέρμονους βρόγχους καθώς και να υποεκτιμήσει τον αριθμό των συμβάντων σε ένα βήμα αν υπάρχουν πολλά διαθέσιμα εξαγόμενα σημεία σε αυτό. Τέλος δεν παρέχει καμία άμεση αξιολόγηση της ακρίβειας των ανακατασκευασμένων δεδομένων όπως επίσης δεν καθίσταται μία εύκολη διαδικασία για μη εξοικειωμένους χρήστες.

Με αφορμή όλα τα παραπάνω, οι Liu, Zhou και Lee με ορισμένες παρεμβάσεις στον αλγόριθμο της Guyot έδωσαν μία νέα προσέγγιση δύο σταδίων έχοντας λιγότερες απαιτήσεις κατά την εισαγωγή των δεδομένων και έχοντας ενισχύσει την ευρωστία και τη σταθερότητα της αρχικής μεθόδου (Liu, Zhou, & Lee, 2021). Οι βελτιώσεις αυτές προήλθαν κατά κύριο λόγο από την προσαρμογή της μονοτονίας των δεδομένων καθώς και την οριοθέτηση των λογοκρισιών. Πιο αναλυτικά στο πρώτο στάδιο του τροποποιημένου αλγορίθμου γίνεται εξαγωγή των συντεταγμένων από τις καμπύλες KM και στο δεύτερο επιτυγχάνεται η επεξεργασία των δεδομένων και η ανακατασκευή της IPD.

Για τη διεξαγωγή του πρώτου σταδίου δόθηκαν οι απαραίτητες συστάσεις έτσι ώστε να διασφαλιστεί όσο το δυνατόν περισσότερο η ποιότητα των δεδομένων. Αυτές αναφέρουν ότι είναι επιθυμητή η χρήση του Adobe Illustrator προκειμένου να γίνει διαχωρισμός των καμπυλών στις περιπτώσεις όπου ήταν μπερδεμένες, η εξαγωγή όσο το δυνατόν περισσότερων σημείων ομοιόμορφα κατανομημένα στις καμπύλες KM καθώς και η λήψη μόνο των σημείων που αντιστοιχούν στο πάνω και στο κάτω άκρο του κατακόρυφου τμήματος, στο οποίο μειώνεται η πιθανότητα επιβίωσης με την έλευση κάποιου γεγονότος (οι αντίστοιχες πιθανότητες επιβίωσης σε αυτά τα δύο σημεία συμβολίζονται με \hat{S}_k και $\hat{S}_{last(k)}^{KM}$). Το χρονικό αυτό σημείο αναφέρεται ως jumping step και στην ουσία προτείνεται η διασφάλιση μόνο των δύο αυτών σημείων καθώς όταν υπάρχουν πολλά jumping points ο αλγόριθμος υποεκτιμά τον αριθμό των ατόμων που βρίσκονται σε κίνδυνο. Η κατάσταση αυτή είναι γνωστή ως διαδικασία ελέγχου βήματος (step control procedure).

Ακολούθως κατά το δεύτερο στάδιο γίνεται αρχικά η επεξεργασία των ακατέργαστων δεδομένων. Η διαδικασία αυτή περιλαμβάνει την ταξινόμηση των δεδομένων με βάση την πάροδο του χρόνου (T_k , όπου $k=1, \dots, N$; N ο αριθμός των εξαγόμενων σημείων), την εισαγωγή του φράκτη Tukey (Tukey's fence)* για εντοπισμό και αφαίρεση των ακραίων παρατηρήσεων καθώς και έναν έλεγχο τήρησης της μονοτονίας μεταξύ των πιθανοτήτων επιβίωσης και του χρόνου. Πιο συγκεκριμένα η χρήση του ορίου Tukey για την

αυτόματη αφαίρεση των ακραίων τιμών, εξασφαλίζει μία πιο ακριβή εκτίμηση (μικρό το αντίτιμο της αφαίρεσης των ακραίων τιμών στην εκτίμηση) των παραμέτρων και αυξάνει την ευρωστία του αλγορίθμου. Στη συνέχεια χρησιμοποιείται η “force monotonicity” για τη διασφάλιση της μονοτονίας των πιθανοτήτων επιβίωσης με το χρόνο. Εδώ αξίζει να αναφερθεί ότι η «δύναμη μονοτονίας» είναι στις περισσότερες περιπτώσεις καλύτερη από την ισοτονική παλινδρόμηση που εφάρμοσαν οι Wei και Royston, καθώς αυτή χρησιμοποιείται όταν παραβιάζεται η υπόθεση μονοτονίας.

Έπειτα γίνεται η ανακατασκευή των μεμονωμένων δεδομένων των ασθενών (IPD) χρησιμοποιώντας τις εκτιμώμενες παραμέτρους KM για κάθε συντεταγμένη από τον επαναληπτικό αλγόριθμο, έχοντας ωστόσο προηγηθεί ρύθμιση των ορίων των επιτρεπόμενων λογοκριμένων παρατηρήσεων έτσι ώστε να αποφευχθούν οι άπειροι βρόγχοι. Η διαφορά έγκειται δηλαδή στο πέμπτο βήμα του αλγορίθμου όπου η επανεκτέλεση των βημάτων ένα έως τέσσερα, γίνεται τώρα εφόσον ισχύει μία από τις δύο ακόλουθες συνθήκες

$$\hat{n}_{u_{i+1}} > n_{riski+1} \quad \text{και} \quad \hat{n}_{censori} < n_{riski} - n_{riski+1} \quad (3.6.1)$$

$$\hat{n}_{u_{i+1}} < n_{riski+1} \quad \text{και} \quad \hat{n}_{censori} > 0 \quad (3.6.2)$$

Η συνθήκη (3.6.1) υποδηλώνει ότι ο αριθμός των ασθενών σε κίνδυνο για το διάστημα i είναι μεγαλύτερος από την αναφερόμενη τιμή και επομένως η επανάληψη θα πρέπει να συνεχίσει να αυξάνει τον αριθμό των λογοκριμένων ασθενών σε αυτό το διάστημα. Ωστόσο προστίθενται το ανώτερο όριο ($n_{riski} - n_{riski+1}$) στον αριθμό των λογοκριμένων ασθενών ($\hat{n}_{censori}$) για να διασφαλιστεί ότι ο αριθμός των λογοκριμένων ασθενών για κάθε διάστημα θα βρίσκεται εντός του εύρους $[0, n_{riski} - n_{riski+1}]$. Αντίθετα από την άλλη η (3.6.2) δείχνει ότι ο αριθμός των ασθενών σε κίνδυνο στο τέλος του διαστήματος i είναι μικρότερος από την αναφερόμενη τιμή και συνεπώς η επανάληψη θα συνεχίσει να μειώνει την εκτίμηση του αριθμού των λογοκριμένων ασθενών στο διάστημα. Αυτή η οριακή συνθήκη χρειάζεται για να αποφευχθεί ο άπειρος αριθμός επαναλήψεων, κάτι που δεν ελήφθη υπόψη στον αρχικό αλγόριθμο.

**Tukey's fence (φράκτης Tukey): πρόκειται για μία παραμετρική μέθοδο ανίχνευσης ακραίων τιμών. Υπολογίζεται δημιουργώντας ένα όριο «φράκτη» χρησιμοποιώντας το $TF = [Q_1 - k(Q_3 - Q_1), Q_3 + k(Q_3 - Q_1)]$ όπου Q_1, Q_3 είναι το πρώτο και τρίτο τεταρτημόριο των σημείων και το k είναι μία σταθερά που παίρνει τιμές 1.5 και 3. Στην συγκεκριμένη περίπτωση χρησιμοποιείται για $k=3$ ώστε να μην διαγραφούν χρήσιμα σημεία.*

Τέλος για την αξιολόγηση της ακρίβειας της εν λόγω μεθόδου δίνεται η δυνατότητα κατασκευής γραφημάτων απεικόνισης των αρχικών έναντι των ανακατασκευασμένων δεδομένων όπως η εκτιμώμενη πιθανότητα επιβίωσης σε κάθε χρονικό διάστημα και ο εκτιμώμενος αριθμός των ασθενών σε κίνδυνο. Ακόμη είναι εφικτή η παροχή συνοπτικών στατιστικών μέτρων για την αξιολόγηση της ακρίβειας όπως η τετραγωνική ρίζα του μέσου τετραγωνικού σφάλματος (RMSE-Root Mean Square Error) που μετρά τη διαφορά στις ανακατασκευασμένες πιθανότητες επιβίωσης σε σχέση με τις αρχικές, το μέσο απόλυτο σφάλμα (Mean Absolute Error) καθώς και το μέγιστο απόλυτο σφάλμα (Max Absolute Error). Οι προτεινόμενες τιμές των μέτρων αυτών (οι οποίες μπορεί να επιδέχονται βελτιώσεις καθώς αναφέρεται ότι στηρίζονται σε μικρό αριθμό παραδειγμάτων) για την αξιολόγηση της καταλληλότητας των ανακατασκευασμένων δεδομένων είναι $RMSE \leq 0.05$, $MAE \leq 0.02$ και $Max Absolute Error \leq 0.05$. Μία τελευταία αξιολόγηση που παρέχεται γίνεται με τη βοήθεια του Kolmogorov-Smirnov τεστ όπου συγκρίνεται η κατανομή των αρχικών και των εκτιμώμενων καμπυλών επιβίωσης. Γενικά εδώ επιθυμείτε ένα μεγάλο p-value που συνεπάγεται μικρή απόκλιση μεταξύ των καμπυλών.

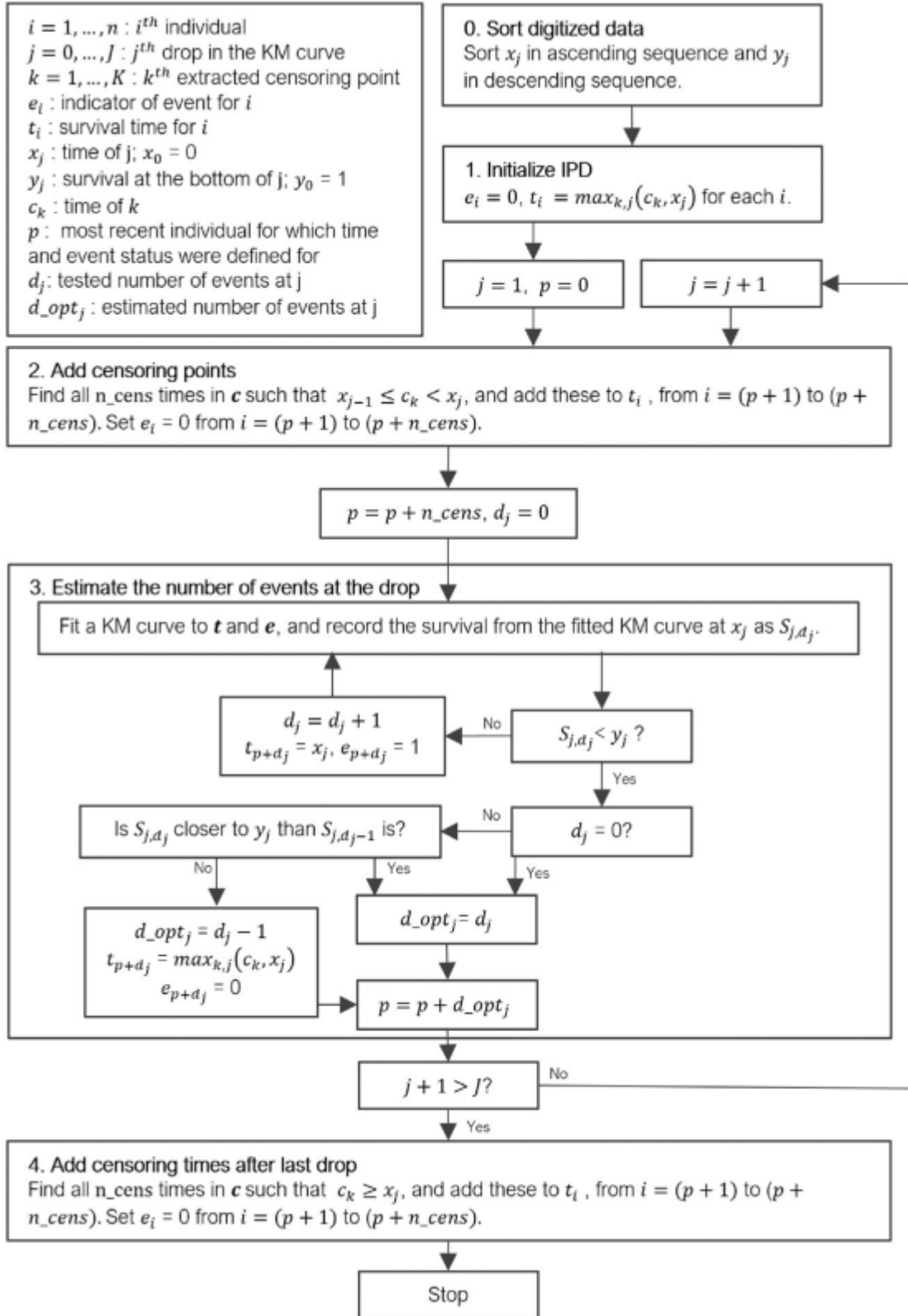
3.7 Ανακατασκευή δεδομένων επιβίωσης ενσωματώνοντας τους σημειωμένους χρόνους λογοκρισίας (marked censoring times)

Στην ανάλυση ευαισθησίας η λογοκρισία συμβαίνει όταν για έναν ασθενή δεν υπάρχουν προσβάσιμες πληροφορίες κατά την περίοδο παρακολούθησης ή στο τέλος της περιόδου αυτής (γι αυτούς που δεν βίωσαν το συμβάν). Για παράδειγμα σε μία μελέτη όπου το υπό μελέτη γεγονός είναι η εμφάνιση εμφράγματος σε μία ομάδα ατόμων για ένα χρονικό διάστημα ενός έτους, εάν ένας ασθενής χανόταν από την παρακολούθηση κάποια χρονική στιγμή εντός του έτους τότε η στιγμή της λογοκρισίας θα ήταν η στιγμή της τελευταίας παρακολούθησης. Αν όμως δεν πάθαινε έμφραγμα μέχρι το τέλος της μελέτης, τότε λογοκρίνεται στο τέλος της δοκιμής. Κατά την ανακατασκευή των μεμονωμένων δεδομένων ασθενών (IPD) είναι σημαντική η χρήση των πληροφοριών λογοκρισίας καθώς η ενσωμάτωσή τους μεγιστοποιεί την ακρίβεια των δεδομένων και αποφεύγεται η μεροληψία. Ωστόσο σε καμία από τις υπάρχουσες μεθόδους για την ανακατασκευή IPD από καμπύλες Kaplan Meier δεν χρησιμοποιούνται οι σημειωμένοι χρόνοι λογοκρισίας οι οποίοι συνήθως υποδεικνύονται σε καμπύλες KM με σύμβολα όπως σταυρούς κ.ά.. Αντ' αυτού χρησιμοποιούνται ο αριθμός των ατόμων σε κίνδυνο και ο αριθμός των συμβάντων για την εκτίμηση του μοτίβου λογοκρισίας. Το πρόβλημα εδώ είναι ότι αυτές οι πληροφορίες δεν αναφέρονται πάντα και όταν είναι διαθέσιμες είναι σε μικρότερη ακρίβεια από τους μεμονωμένους δείκτες λογοκρισίας. Συνεπώς παρόλο που σε κάποιες περιπτώσεις υπάρχουν διαθέσιμα τα δεδομένα στις καμπύλες και πάλι δεν είναι εφικτός ο ακριβής υπολογισμός των χρόνων λογοκρισίας.

Μία προσπάθεια είναι καταγεγραμμένη από τον Irvine και τους συνεργάτες για την εκτίμηση της λογοκρισίας από τις καμπύλες KM με σκοπό τον υπολογισμό συνοπτικών στατιστικών στοιχείων

συμπεριλαμβανομένου του λογαρίθμου του λόγου κινδύνου καθώς και της διακύμανσής του, απαιτώντας ωστόσο τη διαθεσιμότητα της τιμής p ή της χ^2 δοκιμής (Irvine, et al., 2020). Σε συνέχεια της συγκεκριμένης διαδικασίας η Rogula και οι συνεργάτες της παρουσίασαν μία μέθοδο η οποία χρησιμοποιεί παρόμοιες εισροές αλλά εστιάζει στην ανακατασκευή του IPD και επομένως μπορεί να χρησιμοποιηθεί και σε περιπτώσεις όπου ένα στατιστικό τεστ δεν είναι διαθέσιμο (Rogula, et al., 2022). Ο σκοπός της εν λόγω μεθόδου ήταν η βελτίωση της ποιότητας της IPD εφόσον εάν τα σημεία λογοκρισίας είναι επισημασμένα και εύκολα διακριτά δύναται η εκτίμηση της IPD να γίνει πιο προσεκτικά, ενσωματώνοντας τους ακριβείς χρόνους λογοκρισίας όπως ακριβώς αυτοί διαφαίνονται στο εκάστοτε σχήμα, αποφεύγοντας τις υποθέσεις σχετικά με το χρόνο λογοκρισίας.

Η μέθοδος ακολουθεί έναν άπληστο (greedy) αλγόριθμο, δηλαδή έναν αλγόριθμο όπου στοχεύει στην εύρεση της βέλτιστης συνολικής λύσης επιλέγοντας την τοπικά βέλτιστη λύση κάθε φορά. Συγκεκριμένα ο αλγόριθμος διατρέχει κάθε διαδοχική πτώση στην καμπύλη και υπολογίζει τον αριθμό των γεγονότων σε κάθε πτώση. Αν υπάρχουν σημεία λογοκρισίας στο διάστημα πριν την πτώση αυτά προστίθενται στο ανακατασκευασμένο IPD πριν την εκτίμηση των συμβάντων. Πιο αναλυτικά για την υλοποίηση του αλγορίθμου απαιτείται η εισαγωγή του συνολικού αριθμού των ατόμων που βρίσκονται σε κίνδυνο τη χρονική στιγμή 0, οι εξαγόμενες συντεταγμένες στον άξονα x και y καθώς και ένα διάνυσμα με τους χρόνους λογοκρισίας $c = (c_1, c_2, \dots, c_k)$. Στη συνέχεια περιγράφονται αναλυτικά τα βήματα που απεικονίζονται στο διάγραμμα ροής (Εικόνα 3.7.1):



Εικόνα 4: Διάγραμμα ροής αλγορίθμου Rogula (Rogula, et al., 2022)

ΒΗΜΑ 0

Ταξινόμηση των x_j - συντεταγμένων σε αύξουσα σειρά ή ισοδύναμα των y_j σε φθίνουσα σειρά.

ΒΗΜΑ 1

Αρχικοποίηση των ανακατασκευασμένων ατομικών δεδομένων ασθενών (IPD), ώστε όλα τα άτομα να λογοκρίνονται στο μεγαλύτερο χρόνο παρακολούθησης. Στη συνέχεια ο αλγόριθμος ξεκινάει από την πρώτη πτώση στην επιβίωση, έστω στο x_1 και με προς τα εμπρός κίνηση φτάνει στην x_j , επισκέπτοντας το κάθε σημείο μόνο μία φορά.

ΒΗΜΑ 2

Σε κάθε πτώση x_j ελέγχεται ο αριθμός των ατόμων που λογοκρίνονται στο διάστημα c για το χρονικό διάστημα $[x_{j-1}, x_j]$ και στη συνέχεια προστίθενται στο ανακατασκευασμένο σύνολο δεδομένων.

ΒΗΜΑ 3

Εκτίμηση του αριθμού των γεγονότων που πρέπει να προστεθούν στο x_j . Βρίσκεται μέσω του αριθμού των γεγονότων των οποίων η πιθανότητα επιβίωσης βρίσκεται μέσα στο x_j από μια KM ανάλυση προσαρμοσμένη στο ανακατασκευασμένο IPD που βρίσκεται πλησιέστερα στην ψηφιοποιημένη τιμή y_j . Ένας τρόπος υλοποίησης που διαφαίνεται και στο διάγραμμα ροής (Εικόνα 4) είναι η αρχική θεώρηση μηδενικών συμβάντων και στη συνέχεια η αύξησή τους έως ότου η προσαρμοσμένη πιθανότητα επιβίωσης βρεθεί κάτω από την y_j . Στη συνέχεια ελέγχεται αν αυτή η πιθανότητα ή η προηγούμενη της είναι πιο κοντά στην ψηφιοποιημένη τιμή. Σε κάποιες περιπτώσεις ενδέχεται να μην προστεθούν και συμβάντα. Ωστόσο αφού καθοριστεί ο αριθμός των συμβάντων κατά την πτώση, προστίθενται στο ανακατασκευασμένο IPD πριν γίνει η μετάβαση στο επόμενο σκαλοπάτι.

ΒΗΜΑ 4

Η ανακατασκευή της IPD ολοκληρώνεται με την επίσκεψη όλων των πτώσεων και των σημείων λογοκρισίας (συμπεριλαμβανομένων και κάποιων μετά την τελευταία πτώση)

Η αξιοπιστία του αλγορίθμου αξιολογήθηκε μέσω προσομοίωσης. Πιο αναλυτικά έγινε προσομοίωση των καμπυλών επιβίωσης για δύο υποθετικές θεραπείες χρησιμοποιώντας τρεις κατανομές (τη λογαριθμική κανονική, τη λογαριθμική λογιστική και την εκθετική) κάνοντας αρκετούς συνδυασμούς στις παραμέτρους και θεωρώντας τέσσερα διαφορετικά μεγέθη δείγματος (20, 50, 100 και 200). Τα αποτελέσματα μέσω οπτικής επιθεώρησης έδειξαν ότι οι καμπύλες από το ανακατασκευασμένο IPD

παρείχαν στενή προσαρμογή στις αρχικές καμπύλες KM. Επίσης δεν βρέθηκαν συστηματικές διαφορές στους λόγους κινδύνου μεταξύ των δύο θεραπειών για κανένα από τα υλοποιημένα σενάρια.

Παρόλα αυτά η εν λόγω μέθοδος υπόκειται σε πολλούς περιορισμούς. Αρχικά οι χρόνοι λογοκρισίας για να μπορέσουν να εξαχθούν με σαφήνεια θα πρέπει να υπάρχουν σημειωμένοι και να είναι ευκολοδιάκριτοι, γεγονός που δεν ισχύει πάντα. Ειδικά αν ο αριθμός των δειγμάτων είναι μεγάλος η κατάσταση γίνεται όλο και πιο δύσκολη. Γι' αυτό και στην προσομοίωση το μεγαλύτερο μέγεθος δείγματος ήταν τα 200 άτομα. Γενικά η χρήση του αλγορίθμου προτείνεται για μικρά μεγέθη δειγμάτων εν τούτοις σημαντικό ρόλο παίζει ο ρυθμός λογοκρισίας και η ποιότητα της δημοσιευμένης εικόνας. Βέβαια ακόμα και να είναι εύκολη η διάκριση μπορεί να υπάρξει αλληλοκάλυψη ή μη καταγραφή όταν συμπίπτει με τη διενέργεια ενός γεγονότος, τα οποία γίνονται διακριτά μόνο μετά την ανακατασκευή των δεδομένων, όπου η καμπύλη δεν έχει την απαραίτητη προσαρμογή. Επιπρόσθετα ανάλογα με την ποιότητα της εικόνας πολλές φορές δεν είναι εύκολη η παρατήρηση των πτώσεων της καμπύλης γεγονός που επηρεάζει την ανακατασκευή των δεδομένων. Τέλος και αυτή η μέθοδος πάσχει από τον κοινό περιορισμό των μεθόδων ανακατασκευής δεδομένων από δημοσιευμένες καμπύλες, όπου η αξιοπιστία των δεδομένων επηρεάζεται από το σφάλμα που εισάγεται από το χρήστη κατά τη διαδικασία ψηφιοποίησης.

Κεφάλαιο 4^ο-Μελέτη Προσομοίωσης

4.1 Περιγραφή Προσομοίωσης

Έπειτα από την περιγραφή όλων των διαθέσιμων μεθόδων στο Κεφάλαιο 3 και θέλοντας να αξιολογήσουμε την ακρίβεια των εκτιμήσεων που παράγονται εκτελέσαμε μία μελέτη προσομοίωσης 50 δοκιμών χρησιμοποιώντας τη νέα προσέγγιση του αλγορίθμου της Guyot δύο σταδίων (Liu, et al., 2014). Η χρήση μόνο 50 καμπυλών έγινε καθώς η ανακατασκευή των δεδομένων από τα οποία έχουν παρέλθει εκτελείται με χειροκίνητο τρόπο και αποτελεί μία αρκετά χρονοβόρα διαδικασία. Η επιλογή της εν λόγω μεθόδου έγινε μιας και από τις μελέτες επικύρωσης ανάμεσα στις μεθόδους του Parmar, Williamson, Hoyle και Henley και της Guyot, η τελευταία αναμφίβολα δίνει τις πιο κοντινές εκτιμήσεις στα πρωτότυπα δεδομένα καθώς και ότι οι νέες βελτιώσεις μοιάζουν αρκετά υποσχόμενες. Όσο αφορά τη μέθοδο της Rogula παρ' ότι είναι η πιο πρόσφατη, δεν θα εφαρμοστεί στη στην εν λόγω εργασία διότι έχει αρκετούς εγγενείς περιορισμούς και απευθύνεται σε περιορισμένο μέγεθος δείγματος και αριθμό λογοκρισιών.

Γενικά πολυάριθμη είναι η φύση των δεδομένων και οι κατανομές που μπορούν να χρησιμοποιηθούν για να αναπαραστήσουν το χρόνο επιβίωσης, καθώς αρκετές είναι και οι επιλογές του μεγέθους δείγματος των δοκιμών, εφόσον ανάλογα με την εκάστοτε μελέτη και το υπο διερεύνηση γεγονός ενδιαφέροντος, ο απαιτούμενος αριθμός ατόμων εισαγωγής στη μελέτη ποικίλλει αρκετά. Ωστόσο η συγκεκριμένη διατριβή θα επικεντρωθεί στην ανάλυση των δύο δειγμάτων μιας και αυτή είναι η συνηθέστερη κατάσταση στην ανάλυση επιβίωσης την οποία μελετάμε. Έτσι κάθε μία από τις 50 δοκιμές που προσομοιώθηκαν αποτελούνται από μία ομάδα ελέγχου (control arm) και μία ομάδα θεραπείας (treatment arm). Το μέγεθος δείγματος για κάθε μία από αυτές ήταν $n_j = 400, j = 1, \dots, 50$ με τους μισούς ασθενείς να περιλαμβάνονται στην πρώτη ομάδα και οι υπόλοιποι στη δεύτερη. Στη συνέχεια ο υποκείμενος χρόνος επιβίωσης (survival time) $t_i, i = 1, \dots, 200$ για τα άτομα κάθε ομάδας δημιουργήθηκε χρησιμοποιώντας την κατανομή Weibull (Kim & Yum, 2008) με συνάρτηση πιθανότητας

$$S(t) = e^{-(\lambda t)^\gamma},$$

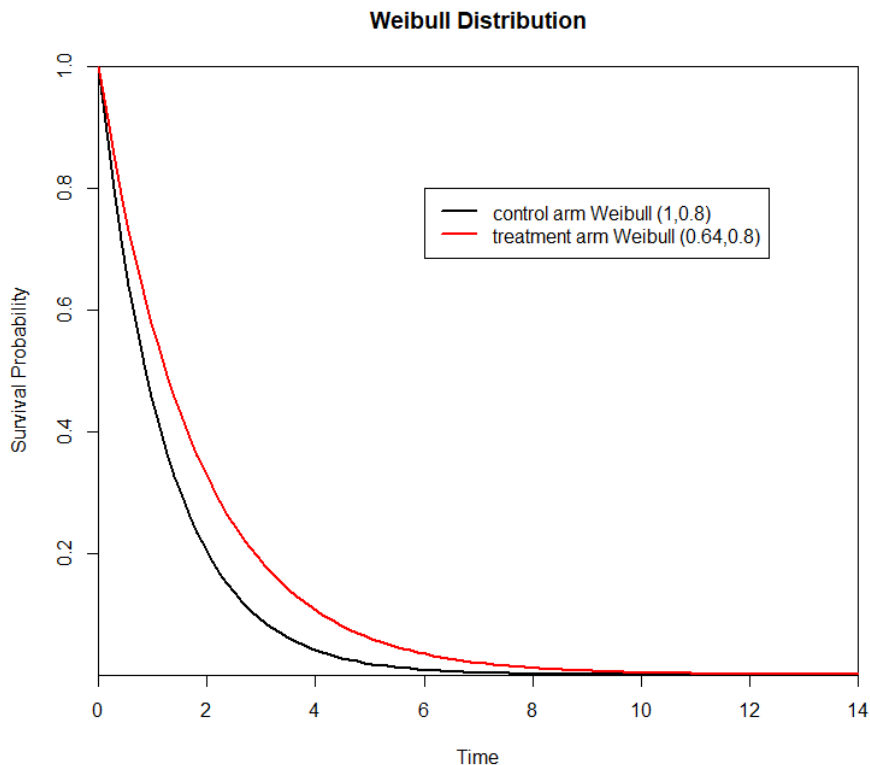
όπου γ και λ είναι η παράμετρος σχήματος (shape parameter) και η παράμετρος κλίμακας (scale parameter) αντίστοιχα. Πιο συγκεκριμένα για την εν λόγω προσομοίωση χρησιμοποιήθηκαν για την ομάδα ελέγχου παράμετρος σχήματος ίση με 0.8 και παράμετρος κλίμακας ίση με τη μονάδα, με τη συνάρτηση επιβίωσης να γίνεται:

$$S(t) = e^{-(1t)^{0.8}} = e^{-t^{0.8}}.$$

Ακόμη, προκειμένου ο λόγος κινδύνου να παραμείνει σταθερός με το χρόνο χρησιμοποιήσαμε την ίδια παράμετρο σχήματος για την ομάδα θεραπείας αλλά επιλέξαμε παράμετρο κλίμακας τέτοια ώστε

$$HR = \frac{h_1(t)}{h_0(t)} = \frac{\lambda_1^\gamma \gamma t^{\gamma-1}}{\lambda_0^\gamma \gamma t^{\gamma-1}} = \left(\frac{\lambda_1}{\lambda_0}\right)^\gamma = 0.7.$$

Όπως γίνεται αισθητό και στο παρακάτω γράφημα (Γράφημα 4), εφόσον η παράμετρος σχήματος είναι μικρότερη της μονάδας η καμπύλη επιβίωσης αναπαριστά μειούμενο κίνδυνο.



Γράφημα 4: Κατανομές Weibull με παραμέτρους κλίμακας $\lambda_0=1$ και $\lambda_1=0.64$ και κοινή παράμετρο σχήματος $\gamma=0.8$ για την ομάδα ελέγχου και την ομάδα θεραπείας αντίστοιχα

Ακόμη έχοντας ως στόχο τα προσομοιωμένα δεδομένα να αναπαριστούν όσο πιο κοντά γίνεται κάποιο πραγματικό σενάριο θεωρήσαμε ότι υπάρχει στα δεδομένα ένα ποσοστό τυχαίας λογοκρισίας. Με άλλα λόγια ο χρόνος λογοκρισίας για κάθε άτομο θεωρείται ανεξάρτητος από την ομάδα στην οποία ανήκει και ακολουθεί μία Weibull κατανομή με ρυθμό που θα έχει ως αποτέλεσμα το 10% των λογοκριμένων παρατηρήσεων, ελλείψει σταθερής λογοκρισίας.

Στη συνέχεια προχωρήσαμε στην εξαγωγή των δεδομένων και στη δημιουργία των καμπυλών Kaplan Meier για κάθε ξεχωριστή δοκιμή με τη χρήση της R. Όλα τα γραφήματα αποθηκεύτηκαν ως bitmap αρχεία και πιο συγκεκριμένα ως εικόνες .png. Ακόμη για όλες τις μελέτες υπολογίστηκαν τα εξής

συνοπτικά μέτρα: ο λόγος κινδύνου (HR) και ο περιορισμένος μέσος χρόνος επιβίωσης (RMST) καθώς και ο μέσος όρος τους (mean) και η τυπική απόκλιση (sd). Επιπλέον για κάθε μία δοκιμή παρουσιάζονταν τρία ενδεχόμενα εξαγωγής ατόμων σε κίνδυνο (number at risk). Την πρώτη φορά θεωρήσαμε γνωστό τον αριθμό των ατόμων σε κίνδυνο σε 15 χρονικά σημεία (times at risk) κατά τη διάρκεια της δοκιμής και πιο συγκεκριμένα πάρθηκαν ανά μισό έτος μέχρι τα πρώτα 5 έτη παρακολούθησης και μετά ανά έτος μέχρι τα εννιά έτη. Στο δεύτερο σενάριο θεωρήσαμε ότι είναι διαθέσιμοι οι αριθμοί των ατόμων σε κίνδυνο σε 7 χρονικά σημεία και ειδικά ανά έτος για τα πρώτα 6 έτη της εκάστοτε μελέτης ενώ στο τρίτο υποθέσαμε ότι δεν είναι γνωστός ο αριθμός των ατόμων σε κίνδυνο σε κανένα χρονικό σημείο και το μόνο που γνωρίζουμε είναι ο συνολικός αριθμός ασθενών σε κάθε ομάδα, που στην περίπτωσή μας ήταν 200 άτομα. Τέλος για κάθε μια δοκιμή υπολογίστηκε ο συνολικός αριθμός των συμβάντων (total events), ο οποίος χρησιμοποιήθηκε μόνο στα δύο πρώτα σενάρια.

Έχοντας λοιπόν αυτές τις πληροφορίες, που πρακτικά είναι οι συνήθεις πληροφορίες που υπάρχουν σε δημοσιευμένες μελέτες είτε στην ολότητά τους είτε μέχρι σε ένα βαθμό, προχωρήσαμε στο πρώτο στάδιο του τροποποιημένου αλγορίθμου που έγινε η εξαγωγή των συντεταγμένων. Η διαδικασία αυτή πραγματοποιήθηκε στο στατιστικό πρόγραμμα R (version 4.1.3) με χρήση της βιβλιοθήκης IPDfromKM. Εκεί σε αρχικό επίπεδο επιλέχθηκε η συλλογή σημείων που αντιστοιχούσαν στο μισό πλήθος του συνολικού αριθμού των ασθενών, με τη βοήθεια των κλικ του ποντικιού πάνω στην εκάστοτε καμπύλη (mouse-clicks). Με άλλα λόγια έγινε εξαγωγή 100 σημείων για κάθε καμπύλη. Ωστόσο η επιλογή αυτών των συντεταγμένων δεν έγινε με τυχαίο τρόπο αλλά με τον ενδεδειγμένο. Πιο αναλυτικά τα σημεία που εξήχθησαν ήταν όσο το δυνατό πιο ομοιόμορφα κατανεμειμένα στις καμπύλες Kaplan Meier, καθώς επίσης στα σημεία που ήταν διακριτή η μείωση της πιθανότητας επιβίωσης γινόταν προσπάθεια λήψης των δύο ακρινών σημείων του κατακόρυφου τμήματος μόνο και όχι των ενδιάμεσων που δημιουργούν κάποια σύγχυση στον αλγόριθμο. Εντούτοις δεν έγινε χρήση του Adobe Illustrator καθώς εφόσον θεωρήσαμε ότι ισχύει η αναλογικότητα των κινδύνων στα δεδομένα που προσομοιώσαμε, οι καμπύλες επιβίωσης ήταν κάπως παράλληλες και δεν κρίθηκε σκόπιμη η χρήση του. Στη συνέχεια δημιουργήθηκαν δύο υποσύνολα με το $\frac{1}{2}$ και το $\frac{1}{4}$ του πλήθους των συντεταγμένων επιλέγοντας ανά δύο και ανά τέσσερα σημεία αντίστοιχα, από τα ήδη εξαγόμενα στοιχεία των καμπυλών. Η επιλογή των σημείων έγινε με αυτό τον τρόπο ώστε να συνάδει με την ομοιόμορφα κατανεμημένη εξαγωγή των σημείων που χρειάζεται για την ποιοτική εξαγωγή των συντεταγμένων.

Στη συνέχεια προχωρήσαμε στο επόμενο στάδιο όπου έγινε η προεπεξεργασία των συντεταγμένων των δεδομένων όπως αναφέρεται στο 3^ο κεφάλαιο. Πραγματοποιήθηκε δηλαδή η ταξινόμηση των δεδομένων κατά το χρόνο επιβίωσης t_i , έγινε προσαρμογή της μονοτονίας των εξαγόμενων συντεταγμένων σύμφωνα με το φράκτη Tukey αφαιρώντας τις παράλογες εισόδους και τέλος εκτελέστηκε

έλεγχος βημάτων προκειμένου να γίνει βελτίωση της ακρίβειας των εκτιμήσεων. Έπειτα εκτελέστηκε η ανακατασκευή των ατομικών δεδομένων των ασθενών με χρήση του επαναληπτικού προσαρμοσμένου αλγορίθμου έχοντας εκτιμηθεί ο αριθμός των ασθενών που είναι σε κίνδυνο σε διάφορες χρονικές στιγμές σύμφωνα με τα προαναφερθέντα σενάρια, ο αριθμός των συμβάντων καθώς και ο αριθμός των ασθενών που λογοκρίθηκαν, έχοντας εφαρμόσει ένα όριο στις λογοκριμένες παρατηρήσεις για την αποφυγή των ατέρμονων βρόγχων.

Λαμβάνοντας λοιπόν τα ανακατασκευασμένα IPD και θέλοντας να αξιολογήσουμε την ακρίβειά τους δημιουργήσαμε για κάθε δοκιμή ένα γράφημα το οποίο απεικονίζει τόσο την εκτιμώμενη πιθανότητα επιβίωσης για κάθε χρονικό διάστημα από τα ανακατασκευασμένα δεδομένα όσο και από τα αρχικά. Τα γραφήματα αυτά αναπαριστούν τις καμπύλες Kaplan Meier και προσφέρουν μία άμεση οπτική εκτίμηση του βαθμού συμφωνίας των πιθανοτήτων επιβίωσης. Επιπρόσθετα εξήχθησαν συνοπτικά στατιστικά στοιχεία, όπως η τετραγωνική ρίζα του μέσου τετραγωνικού σφάλματος (root mean square error- RMSE) η οποία μετρά τη διαφορά στις πιθανότητες επιβίωσης που υπολογίζονται χρησιμοποιώντας τα ανακατασκευασμένα και τα πρωτότυπα στοιχεία. Έστω s_{e_i} η αναμενόμενη τιμή της πιθανότητας επιβίωσης, s_{o_i} η ανακατασκευασμένη τιμή και N το συνολικό μέγεθος δείγματος της δοκιμής, τότε η

τετραγωνική ρίζα του μέσου τετραγωνικού σφάλματος θα υπολογίζεται ως εξής: $RMSE = \sqrt{\frac{\sum_{i=1}^N (s_{e_i} - s_{o_i})^2}{N}}$.

Ακόμη για τον έλεγχο της ακρίβειας υπολογίστηκαν το μέσο απόλυτο σφάλμα (mean absolute error- $MAE = \frac{\sum_{i=1}^N |s_{e_i} - s_{o_i}|}{N}$) και το μέγιστο απόλυτο σφάλμα (max absolute error= actual value-measured value).

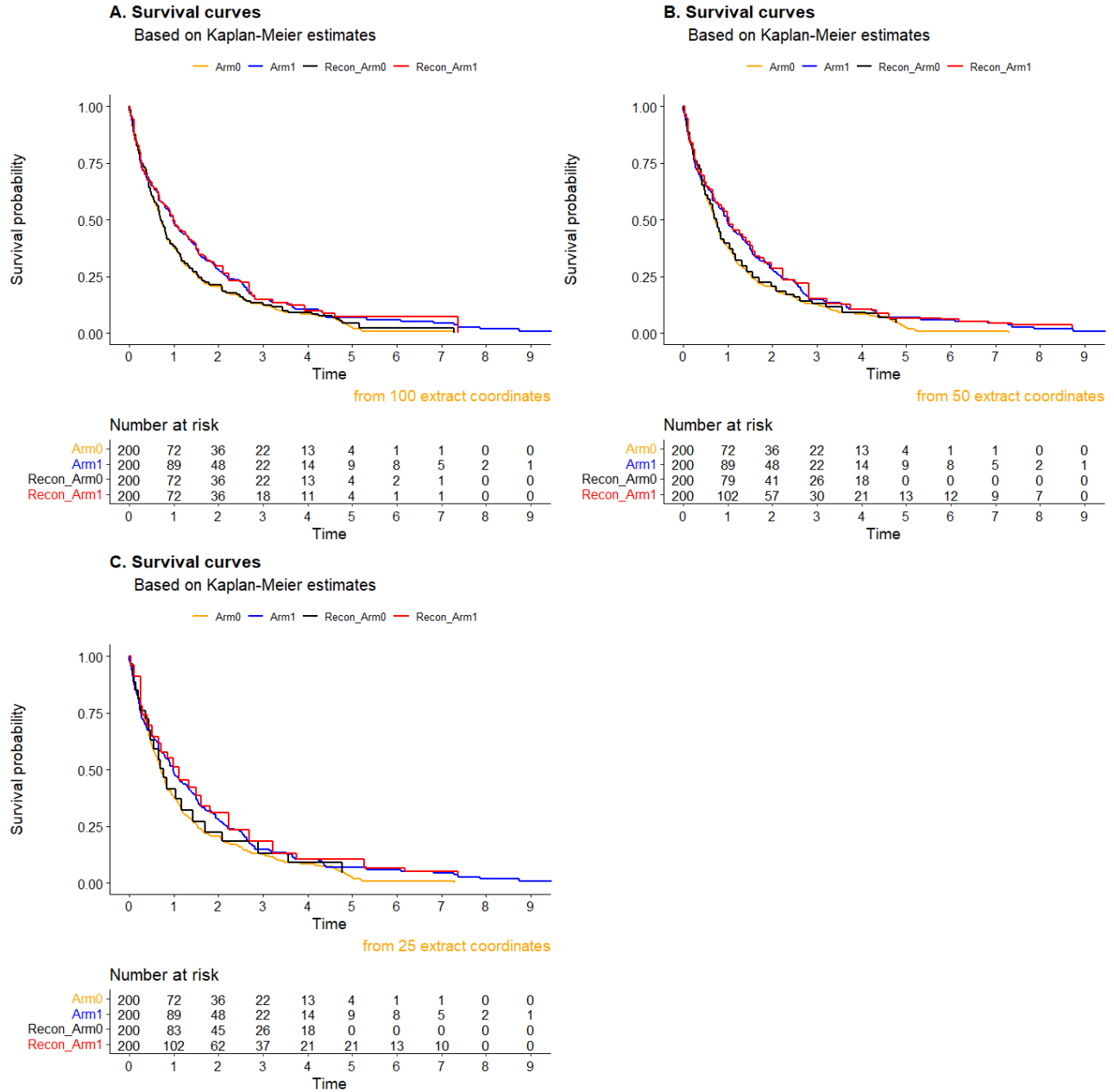
Τέλος διεκπεραιώσαμε το Kolmogorov-Smirnov τεστ (δύο δειγμάτων) προκειμένου να συγκρίνουμε τις κατανομές των καμπυλών ανάγνωσης και εκτιμώμενης επιβίωσης και να αποφανθούμε αν προέρχονται από την ίδια κατανομή.

Στη συνέχεια προχωρήσαμε στον υπολογισμό των συνοπτικών μέτρων για κάθε μία από τις 50 δοκιμές σε κάθε σενάριο που εφαρμόστηκε για την ανακατασκευή, που είχαν υπολογιστεί με τα αρχικά δεδομένα, είναι χρήσιμα στην ανάλυση επιβίωσης και ενδείκνυται η αναφορά τους. Τα μέτρα αυτά αντανακλούν το λόγο κινδύνου (HR) και τον περιορισμένο χρόνο επιβίωσης (RMST). Για την αξιολόγηση της συσχέτισης (correlation) των εν λόγω εκτιμήσεων αρχικά κατασκευάσαμε κάποια scatterplots μεταξύ των πρωτότυπων και των ανακατασκευασμένων τιμών και έπειτα για κάθε σενάριο υπολογίσαμε τη μέση τιμή τους (mean) καθώς και την τυπική απόκλιση (sd).

4.2 Αποτελέσματα

Για τα διάφορα σενάρια που έχουν ληφθεί υπόψη τόσο για τον αριθμό των χρονικών σημείων στα οποία αναφέρεται ο αριθμός των ατόμων σε κίνδυνο, όσο και για τον αριθμό των εξαγόμενων συντεταγμένων από τα αρχικά γραφήματα Kaplan Meier, δημιουργήθηκαν οι ανακατασκευασμένες καμπύλες επιβίωσης χρησιμοποιώντας τα ανακατασκευασμένα δεδομένα των ασθενών. Κάθε μία από αυτές συνδυάστηκε με την αντίστοιχή της, που προέκυψε κάτω από το ίδιο σενάριο χρησιμοποιώντας τα προσομοιωμένα δεδομένα με στόχο την αξιολόγηση της απόδοσης της μεθόδου.

Η πρώτη περίπτωση αφορά το ενδεχόμενο να είναι γνωστά 15 χρονικά σημεία στα οποία είναι διαθέσιμος ο αριθμός των ατόμων σε κίνδυνο. Πιο συγκεκριμένα αυτά τα σημεία τοποθετήθηκαν ανά μισό έτος έως τα πρώτα πέντε έτη και από εκεί ανά έτος μέχρι τα εννιά έτη. Λαμβάνοντας λοιπόν αυτή την πληροφορία πραγματοποιήθηκε η ανακατασκευή της IPD και κατ'επέκταση των καμπυλών Kaplan Meier για τα εξής σενάρια: Α. Έχοντας εξάγει 100 σημεία από την αρχική καμπύλη KM Β. Έχοντας εξάγει 50 σημεία από την αρχική καμπύλη KM και C. Έχοντας εξάγει 25 σημεία από την αρχική καμπύλη KM. Στα παρακάτω γραφήματα (Εικόνα 5) απεικονίζονται οι αρχικές και οι ανακατασκευασμένες καμπύλες KM μιας τυχαίας δοκιμής, από τις 50 οι οποίες κατασκευάστηκαν, για όλες τις προαναφερθείσες περιπτώσεις.

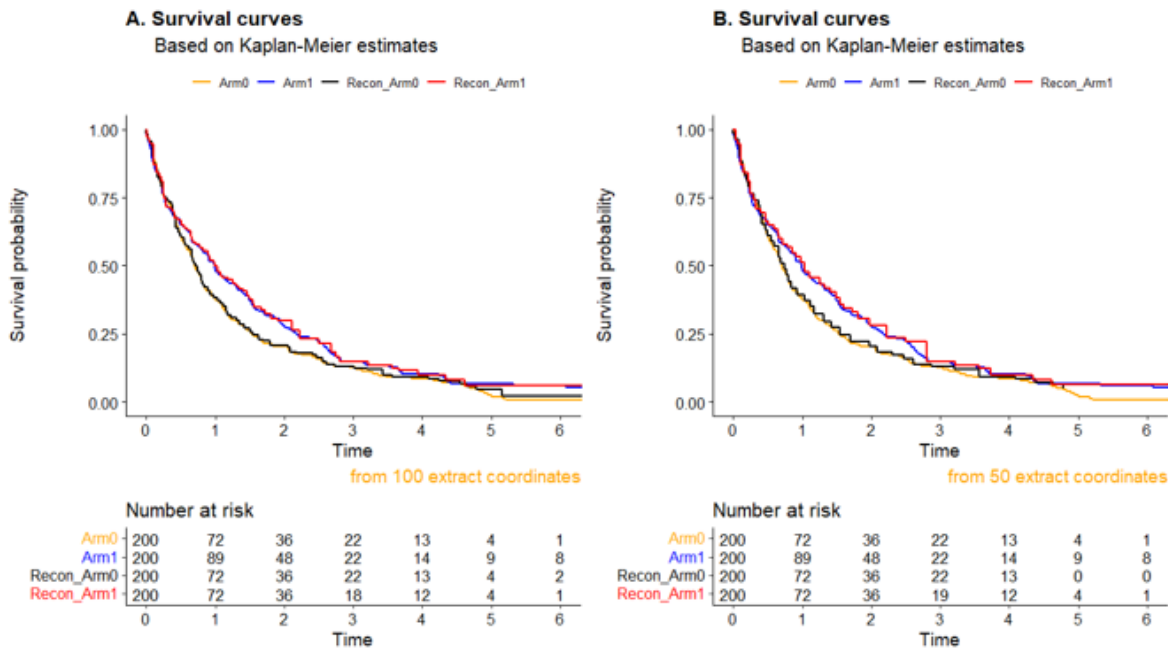


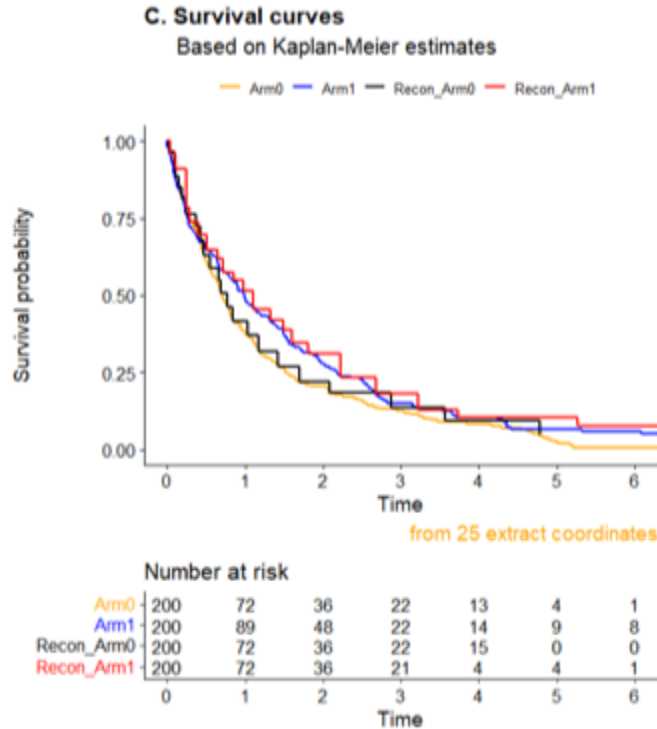
Εικόνα 5: Αρχικές και ανακατασκευασμένες καμπύλες Kaplan Meier μαζί με τον πίνακα του πλήθους των ατόμων σε κίνδυνο, έχοντας χρησιμοποιήσει για την ανακατασκευή 15 σημεία αναφέροντας τον αριθμό των ατόμων σε κίνδυνο. **A.** Έχοντας γίνει εξαγωγή 100 σημείων από την αρχική καμπύλη KM. **B.** Έχοντας γίνει εξαγωγή 50 σημείων από την αρχική καμπύλη KM. **C.** Έχοντας γίνει εξαγωγή 25 σημείων από την αρχική καμπύλη

Όπως φαίνεται λοιπόν στις δύο πρώτες περιπτώσεις όπου εισάγαμε στον αλγόριθμο περισσότερη πληροφορία σχετικά με τα αρχικά σημεία επιβίωσης, οι εκτιμήσεις των ανακατασκευασμένων πιθανοτήτων επιβίωσης προσεγγίζουν με αρκετά μεγάλη ακρίβεια τις αρχικές. Κάτι αντίστοιχο δεν διαφαίνεται στην τελευταία περίπτωση όπου έχει γίνει χρήση μόνο 25 εξαγόμενων συντεταγμένων. Εδώ

βλέπουμε μία πιο τραχεία απεικόνιση των βημάτων της καμπύλης επιβίωσης με αρκετές αποκλίσεις από τα μέσα της καμπύλης και έπειτα. Η διαφορά αυτή αποτυπώνεται και στις εκτιμήσεις του πλήθους των ατόμων που βρίσκονται σε κίνδυνο στις διάφορες χρονικές στιγμές (Εικόνα 5). Ακόμη πρέπει να αναφερθεί ότι σε όλες τις περιπτώσεις οι ανακατασκευασμένες καμπύλες σταματούν λίγο νωρίτερα από τις αρχικές. Το γεγονός αυτό μπορεί να οφείλεται σε αδυναμία του αλγορίθμου, μιας και οι ασθενείς σε κίνδυνο στο τελευταίο χρονικό διάστημα είναι αρκετά περιορισμένοι ή στον τρόπο εξαγωγής των συντεταγμένων (παρ'όλο που έχει γίνει όσο πιο ομοιόμορφα δύναται). Γενικά όμως μπορούμε να πούμε ότι στις δύο πρώτες περιπτώσεις οι προσεγγίσεις είναι αρκετά κοντά στις αρχικές εκτιμήσεις.

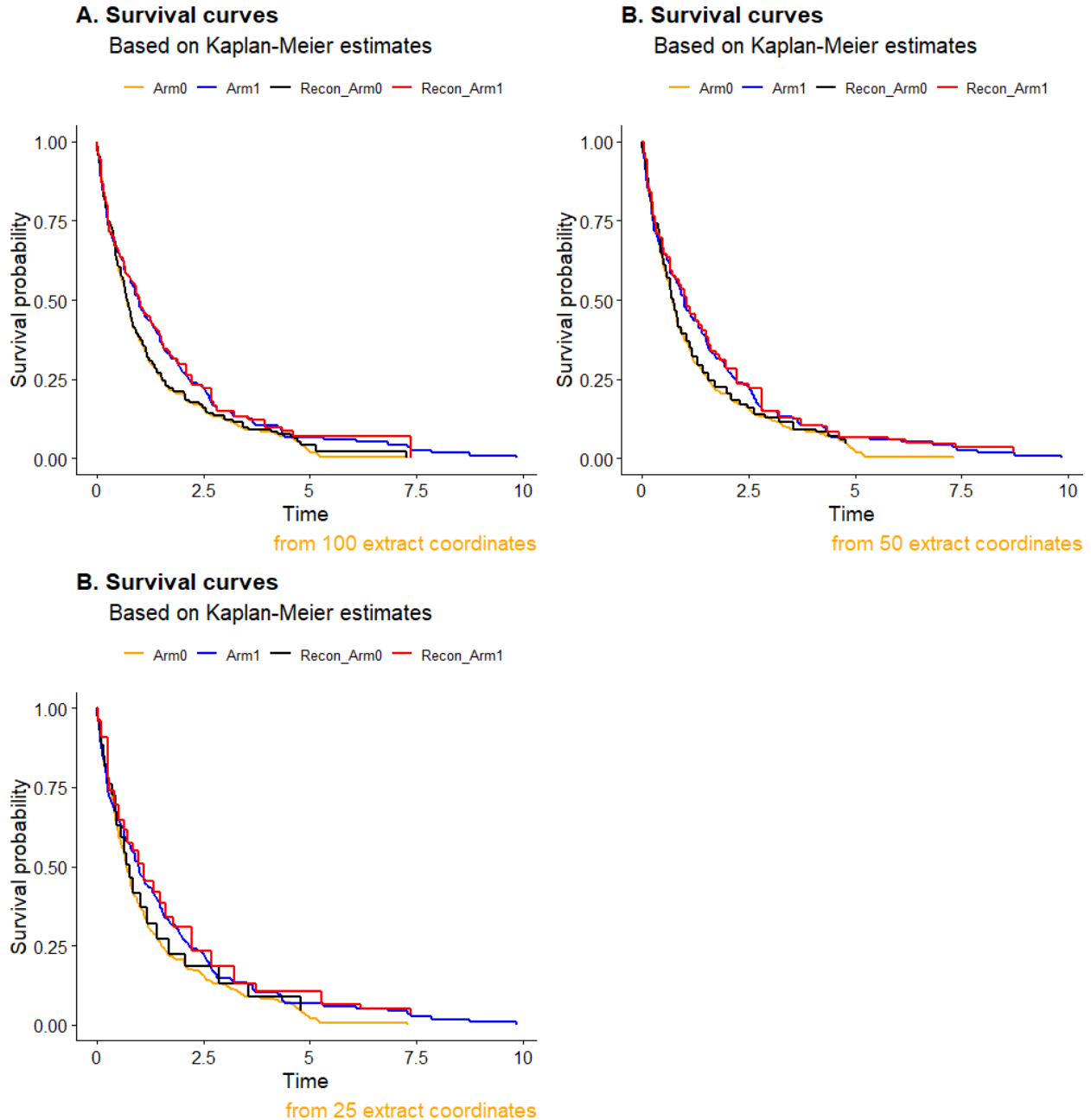
Κάτι αντίστοιχο παρατηρείται και στο δεύτερο σενάριο (Εικόνα 6) όπου έχει θεωρηθεί ότι είναι διαθέσιμοι οι αριθμοί των ατόμων σε κίνδυνο σε επτά χρονικά σημεία και πιο συγκεκριμένα για τα πρώτα 6 χρόνια. Και εδώ λοιπόν παρατηρούμε ότι στο πρώτο ενδεχόμενο όπου έχουν χρησιμοποιηθεί 100 συντεταγμένες από την αρχική απεικόνιση, οι ανακατασκευασμένες καμπύλες Kaplan Meier σχεδόν συμπίπτουν με τις αρχικές. Το ίδιο συμβαίνει και με τη χρήση 50 συντεταγμένων. Ωστόσο στην τελευταία περίπτωση διαφαίνονται ξανά κάποιες αποκλίσεις με πιο αδρές εκτιμήσεις καθώς έχουμε επίσης το φαινόμενο του πρώιμου τερματισμού των καμπυλών.





Εικόνα 6: Αρχικές και ανακατασκευασμένες καμπύλες Kaplan Meier μαζί με τον πίνακα του πλήθους των ατόμων σε κίνδυνο, έχοντας χρησιμοποιήσει για την ανακατασκευή 7 σημεία αναφέροντας τον αριθμό των ατόμων σε κίνδυνο A. Έχοντας γίνει εξαγωγή 100 σημείων από την αρχική καμπύλη KM B. Έχοντας γίνει εξαγωγή 50 σημείων από την αρχική καμπύλη KM C. Έχοντας γίνει εξαγωγή 25 σημείων από την αρχική καμπύλη KM

Τέλος, στο τρίτο σενάριο όπου δεν θεωρείται γνωστός ο αριθμός των ατόμων που βρίσκονται σε κίνδυνο σε κανένα χρονικό σημείο και εισάγεται μόνο ο συνολικός αριθμός των ατόμων σε κάθε ομάδα, παρατηρείται μία κατάσταση η οποία δεν διαφέρει σε μεγάλο βαθμό από τις προηγούμενες, ωστόσο οι αποκλίσεις μεγαλώνουν προς το τέλος των καμπυλών (Εικόνα 7). Επιπρόσθετα παρατηρείται ξανά ο πρώιμος τερματισμός των ανακατασκευασμένων καμπυλών Kaplan Meier. Στη συγκεκριμένη περίπτωση θα περιμέναμε οι ανακατασκευασμένες καμπύλες να μην εφαρμόζουν τόσο καλά στις αρχικές λόγω της λιγότερης πληροφορίας που εισάγαμε στον αλγόριθμο. Εν τούτοις επειδή η ποιότητα της εικόνας με βάση της οποίας έγινε η ανακατασκευή καθώς και η συλλογή των σημείων που πραγματοποιήθηκε με τη δέουσα προσοχή τηρώντας τις προτεινόμενες οδηγίες συντελούν στην εξαγωγή ποιοτικών δεδομένων δικαιολογείται αυτή η καλή προσαρμογή των δύο καμπυλών.



Εικόνα 7: Αρχικές και ανακατασκευασμένες καμπύλες Kaplan Meier έχοντας χρησιμοποιήσει μόνο το συνολικό αριθμό των ασθενών **A.** Έχοντας γίνει εξαγωγή 100 σημείων από την αρχική καμπύλη KM **B.** Έχοντας γίνει εξαγωγή 50 σημείων από την αρχική καμπύλη KM **C.** Έχοντας γίνει εξαγωγή 25 σημείων από την αρχική καμπύλη

Έπειτα λοιπόν από αυτή την γραφική αξιολόγηση, για όλες τις δοκιμές υπολογίστηκαν συνοπτικά στοιχεία για την εύρεση της ακρίβειας των ανακατασκευασμένων τιμών. Πιο αναλυτικά για τη μέτρηση της διαφοράς στις πιθανότητες επιβίωσης που υπολογίζονται χρησιμοποιώντας ανακατασκευασμένα και πρωτότυπα δεδομένα υπολογίστηκε η ρίζα του μέσου τετραγωνικού σφάλματος (RMSE) όπως επίσης το

μέσο απόλυτο σφάλμα (Mean Absolute Error) και το μέγιστο απόλυτο σφάλμα (Max Absolute Error). Τα προτεινόμενα όρια που προτείνονται για τα μέτρα αυτά ώστε να θεωρήσουμε επαρκής την ανακατασκευή αυτών των δεδομένων είναι το $RMSE \leq 0.05$, το μέσο απόλυτο σφάλμα μικρότερο ή ίσο του 0.02 και το μέγιστο απόλυτο σφάλμα μικρότερο ή ίσο του 0.05.

Τα αποτελέσματα έδειξαν ότι για όλα τα σενάρια το RMSE ήταν πάντα μικρότερο του 0.05 όπως επίσης και το μέσο απόλυτο σφάλμα ήταν πάντα μικρότερο του 0.02, ωστόσο υπήρχαν περιπτώσεις όπου το μέγιστο απόλυτο σφάλμα ξεπερνούσε το 0.05 φτάνοντας μέχρι το 0.16. Το γεγονός αυτό χρήζει μία περαιτέρω διερεύνηση, όμως επειδή τα όρια αυτά έχουν βασιστεί σε μικρό αριθμό παραδειγμάτων και οι ίδιοι ερευνητές αναφέρουν ότι δέχονται τροποποιήσεις, καθώς και ότι οι περιπτώσεις αυτές ήταν μικρότερες του 6% μπορούμε να θεωρήσουμε με αρκετά μεγάλη βεβαιότητα ότι τα δεδομένα που ανακατασκευάσαμε είναι αξιόπιστα για μεταγενέστερες αναλύσεις.

Στη συνέχεια προκειμένου να συγκρίνουμε τις κατανομές των καμπυλών ανάγνωσης και εκτιμώμενης επιβίωσης εφαρμόσαμε το Kolmogorov-Smirnov test το οποίο για όλες τις δοκιμές μας έδωσε πολύ υψηλές τιμές σημαντικότητας (p-value) με τις πλείστες των περιπτώσεων να είναι κοντά στη μονάδα. Το γεγονός αυτό συνεπάγεται ότι ελλείπει στατιστικών στοιχείων που να δείχνουν τη διαφορά, η απόκλιση μεταξύ της καμπύλης ανάγνωσης και της εκτιμώμενης καμπύλης επιβίωσης είναι μικρή.

Σύμφωνα λοιπόν με τα παραπάνω η ανακατασκευή των ατομικών δεδομένων των ασθενών έχει επιτευχθεί με μεγάλη επιτυχία σε ένα αρκετά μεγάλο ποσοστό. Ωστόσο στις εκδοχές όπου οι πληροφορίες εισαγωγής στον αλγόριθμο ήταν ελλειπείς παρατηρούνται αρκετές αποκλίσεις, όπως άλλωστε περιμέναμε.

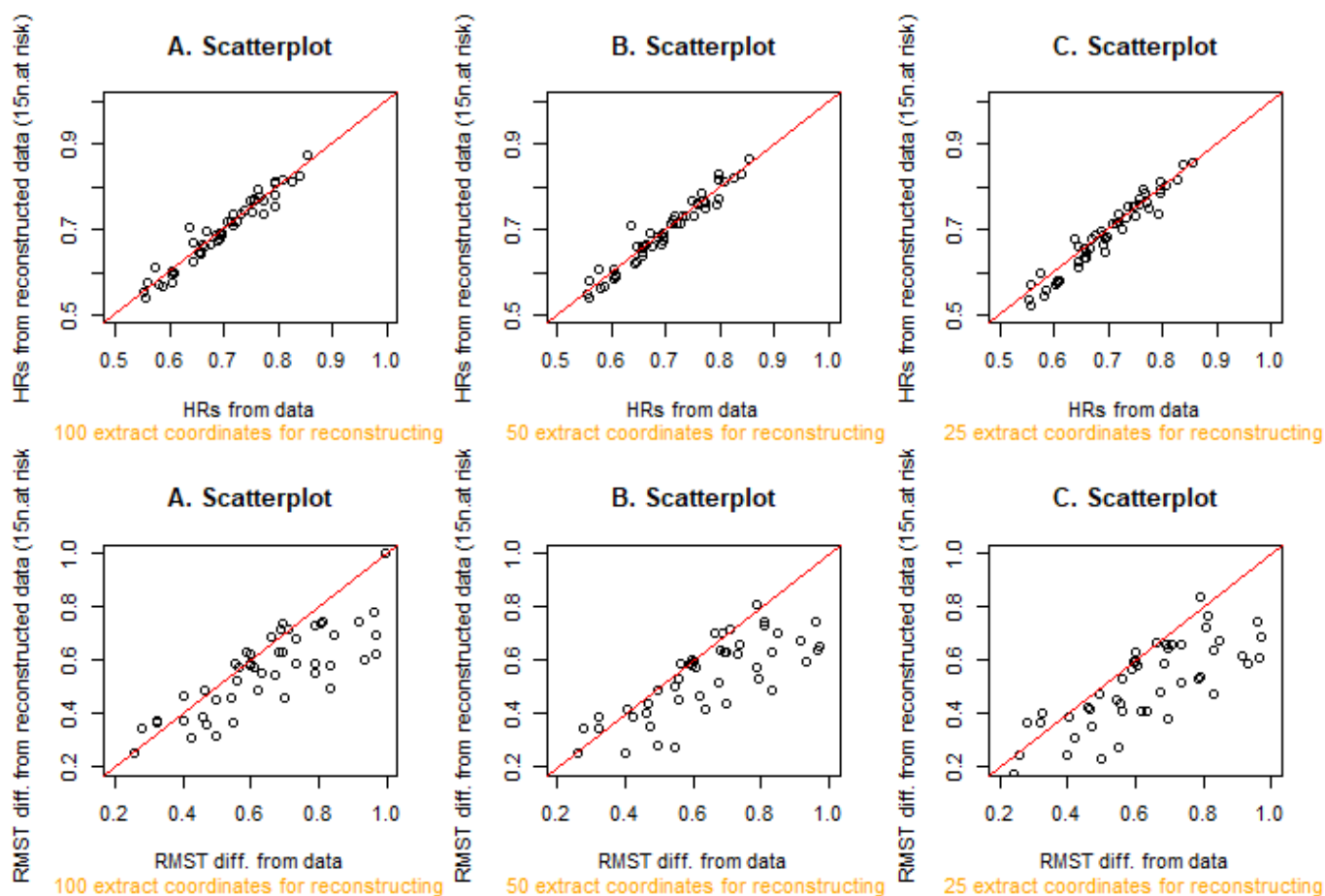
Έπειτα έχοντας το λόγο κινδύνου και τον περιορισμένο μέσο χρόνο επιβίωσης για τα αρχικά δεδομένα θελήσαμε να κατασκευάσουμε αυτά τα συνοπτικά μέτρα χρησιμοποιώντας τα ανακατασκευασμένα δεδομένα, ώστε να ελέγξουμε το βαθμό συμφωνίας τους.

Εδώ κρίνεται σκόπιμο να αναφερθεί ότι ο περιορισμένος μέσος χρόνος επιβίωσης υπολογίστηκε σε όλες τις περιπτώσεις επιλέγοντας ως καθορισμένο σημείο τ , το μέγιστο παρατηρούμενο χρόνο κάθε δοκιμής. Στη συνέχεια και έχοντας κατά νου ότι εμάς μας αφορά η σύγκριση μεταξύ των ομάδων (ομάδα ελέγχου και ομάδα θεραπείας) υπολογίσαμε τη διαφορά των περιορισμένων μέσων χρόνων επιβίωσης (RMST difference), ξανά στο ίδιο χρονικό σημείο. Με άλλα λόγια η διαφορά αυτή αντανάκλα τη σύγκριση των περιοχών κάτω από τις δύο καμπύλες επιβίωσης και όσο μεγαλύτερη είναι τόσο πιο πιθανό να απορριφθεί η μηδενική υπόθεση της ισότητας των περιορισμένων μέσων χρόνων επιβίωσης.

Έχοντας υπολογίσει και ανακατασκευάσει αυτά τα συνοπτικά μέτρα για όλες τις δοκιμές και για όλα τα προαναφερθέντα σενάρια προβήκαμε στην οπτική τους αναπαράσταση μέσω διαγραμμμάτων

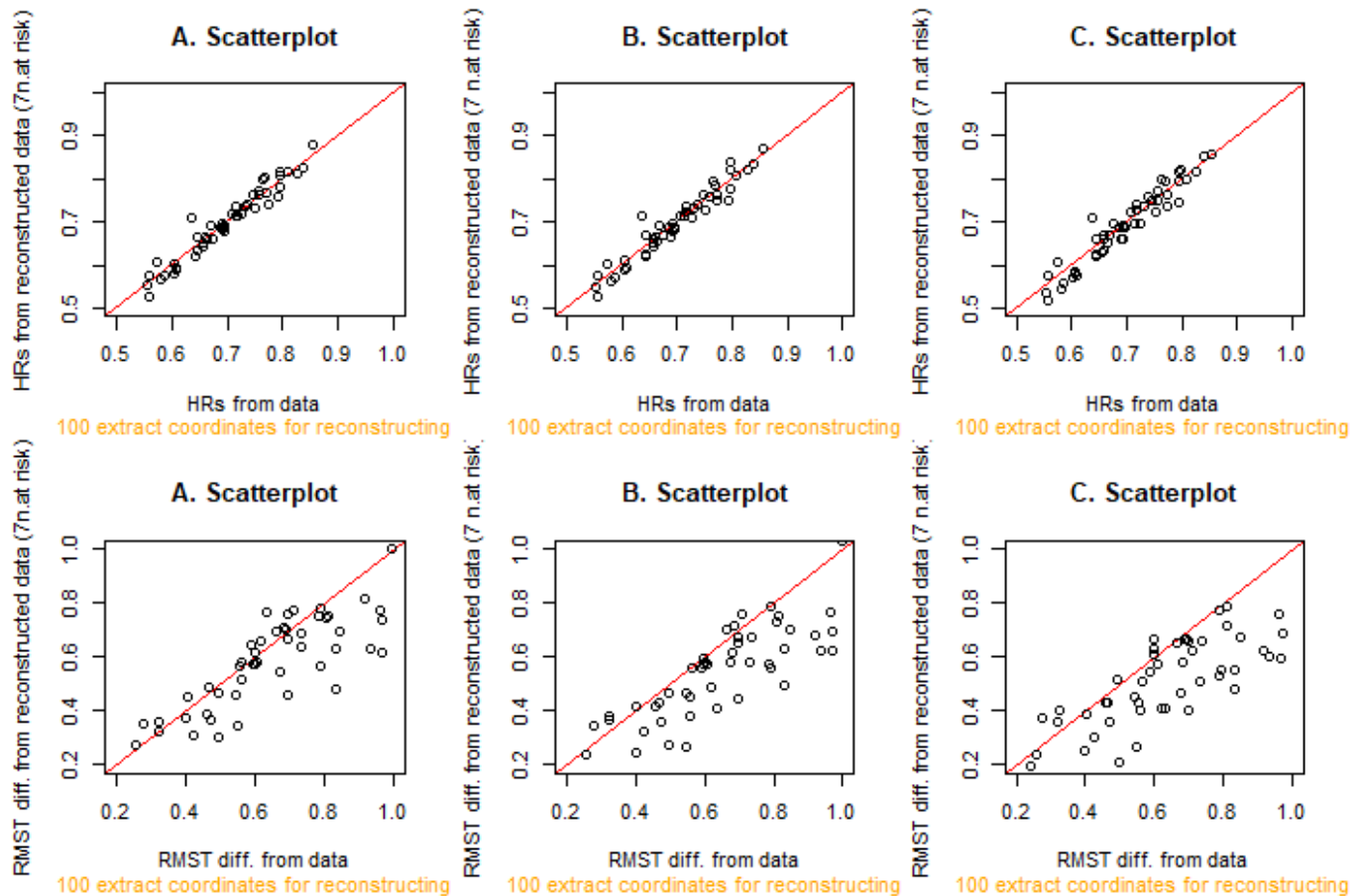
δασποράς (scatter plots) ώστε να ληφθεί μία πρώτη ένδειξη της τάσης που υπάρχει μεταξύ τους. Με τον όρο τάση εννοούμε τον τρόπο με τον οποίο αλλάζουν οι τιμές της μίας μεταβλητής καθώς μεταβάλλονται οι τιμές της άλλης μεταβλητής. Στην περίπτωση μας θα περιμέναμε μία γραμμική τάση (αντανακλάται από την ευθεία $y=x$) καθώς θέλουμε οι ανακατασκευασμένες τιμές να συμπίπτουν με τις αρχικές. Στα εν λόγω διαγράμματα η αρχικές τιμές των παραμέτρων απεικονίζονται στον άξονα x'x και οι ανακατασκευασμένες στον y'y.

Όπως λοιπόν διαφαίνεται στα παρακάτω διαγράμματα (Εικόνα 8), που αναφέρονται στην περίπτωση που έχουν χρησιμοποιηθεί 15 σημεία αναφοράς του αριθμού των ατομών σε κίνδυνο για την ανακατασκευή, έχουμε μία πολύ καλή γραμμική συσχέτιση μεταξύ των αρχικών και των ανακατασκευασμένων λόγων κινδύνου ανεξάρτητα από το πλήθος των εξαγόμενων συντεταγμένων στο αρχικό γράφημα.



Εικόνα 8: Διαγράμματα διασποράς μεταξύ των αρχικών και των ανακατασκευασμένων συνοπτικών μέτρων HR και διαφορά RMSTs κάτω από τα διαφορετικά σενάρια εξαγωγής συντεταγμένων και έχοντας χρησιμοποιήσει 15 σημεία αναφοράς των ατόμων σε κίνδυνο.

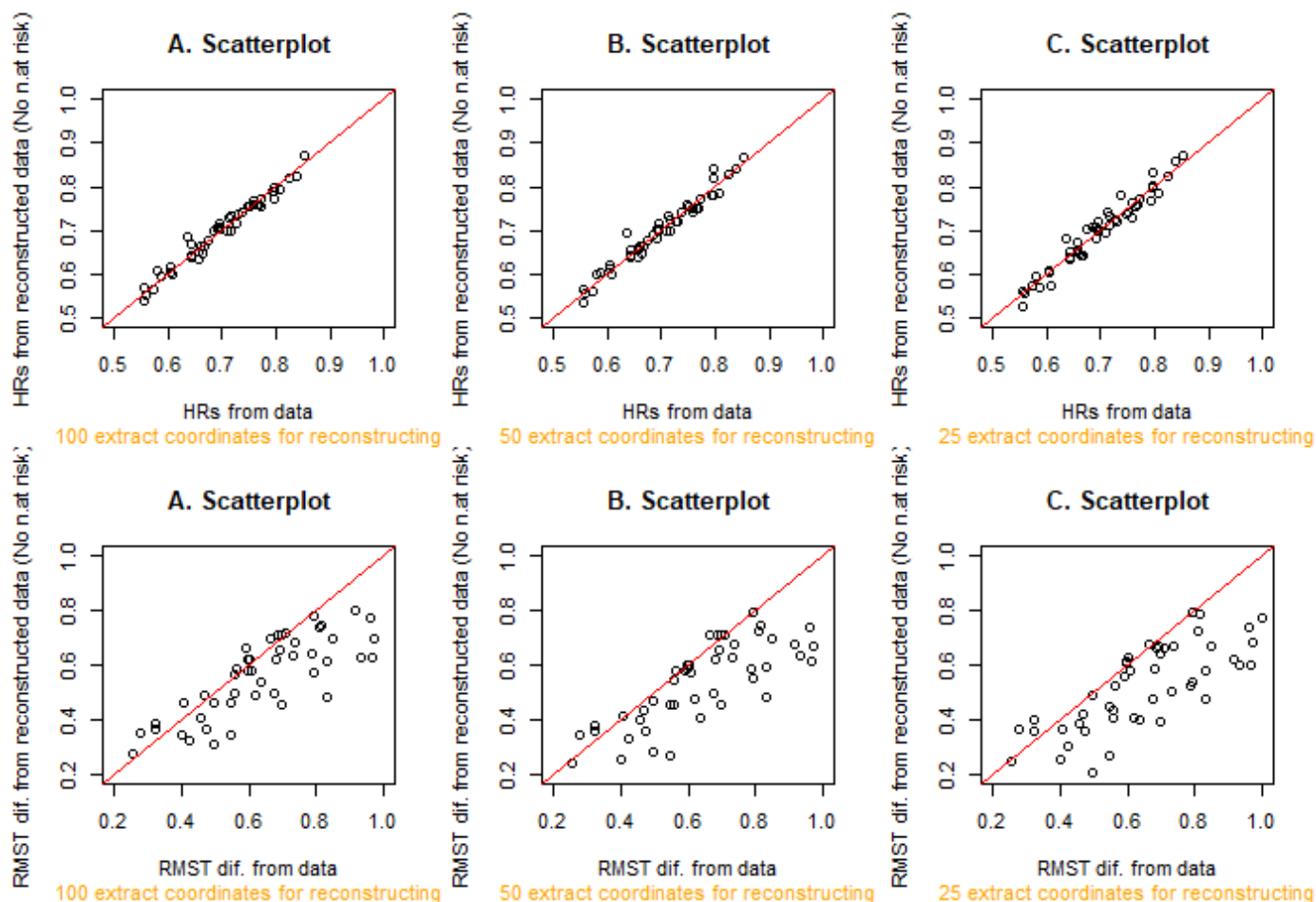
Αυτό προκύπτει από το γεγονός ότι όλες οι δοκιμές που αντιπροσωπεύονται από τα κυκλάκια είναι κοντά στην ευθεία $y=x$, με μόνο μία μελέτη να φαίνεται ότι έχει υπερεκτιμηθεί από την ανακατασκευή του λόγου κινδύνου. Κάτι αντίστοιχο δεν συμβαίνει στην ανακατασκευή της διαφοράς των περιορισμένων μέσων χρόνων επιβίωσης, καθώς παρατηρούμε ότι η πλειοψηφία των δοκιμών βρίσκεται μακριά από την ευθεία $y=x$ για όλες τις εκδοχές της εξαγωγής συντεταγμένων. Πιο συγκεκριμένα φαίνεται οι ανακατασκευασμένες τιμές να υποεκτιμούν τη διαφορά των περιορισμένων μέσων χρόνων επιβίωσης. Το γεγονός αυτό πιθανώς οφείλεται στην επιλογή του προκαθορισμένου σημείου τ . Όπως αναφέραμε και προηγουμένως το σημείο τ ορίστηκε ως ο μέγιστος παρατηρούμενος χρόνος σε κάθε δοκιμή, ωστόσο όπως παρατηρήθηκε και από τις καμπύλες επιβίωσης (Εικόνα 5,6 & 7) η ανακατασκευή των ατομικών δεδομένων των ασθενών προκαλεί έναν πρόωμο τερματισμό της μελέτης, το οποίο συνεπάγεται μικρότερο παρατηρούμενο χρόνο. Εφόσον λοιπόν για τον υπολογισμό της διαφοράς των περιορισμένων μέσων χρόνων επιβίωσης το σημείο τ που χρησιμοποιείται είναι μικρότερο από ότι στην πραγματικότητα εξηγεί το φαινόμενο της υποεκτίμησης που παρατηρείται στο παραπάνω γράφημα.



Εικόνα 9: Διαγράμματα διασποράς μεταξύ των αρχικών και των ανακατασκευασμένων συνοπτικών μέτρων HR και διαφορά RMSTs κάτω από τα διαφορετικά σενάρια εξαγωγής συντεταγμένων και έχοντας χρησιμοποιήσει 7 σημεία αναφοράς των ατόμων σε κίνδυνο.

Τα παραπάνω γραφήματα (Εικόνα 9) αντανακλούν τα διαγράμματα διασποράς για τις αρχικές και τις ανακατασκευασμένες τιμές των συνοπτικών μέτρων για όλα τα σενάρια εξαγωγής συντεταγμένων έχοντας χρησιμοποιήσει 7 σημεία αναφοράς του αριθμού των ατόμων σε κίνδυνο. Και σε αυτή την περίπτωση ωστόσο παρατηρούμε τα ίδια. Μία πολύ καλή συμφωνία μεταξύ των αρχικών και των ανακατασκευασμένων λόγων κινδύνου και μία συνεχής υποεκτίμηση του περιορισμένου μέσου χρόνου επιβίωσης από τα ανακατασκευασμένα δεδομένα που πιθανώς οφείλεται στο λόγο που αναφέραμε προηγουμένως.

Τέλος και στην περίπτωση όπου δεν έχουν εισαχθεί οι αριθμοί των ατόμων σε κίνδυνο αλλά μόνο ο συνολικός αριθμός των ατόμων σε κάθε ομάδα, τα διαγράμματα διασποράς για όλα τα σενάρια εξαγωγής συντεταγμένων παρουσιάζουν μία εντελώς αντίστοιχη εικόνα (Εικόνα 10).



Εικόνα 10: Διαγράμματα διασποράς μεταξύ των αρχικών και των ανακατασκευασμένων συνοπτικών μέτρων HR και διαφορά RMSTs κάτω από τα διαφορετικά σενάρια εξαγωγής συντεταγμένων και έχοντας χρησιμοποιήσει μόνο το συνολικό αριθμό των ασθενών κάθε ομάδας.

Ολοκληρώνοντας θέλοντας να βγάλουμε πιο συγκεντρωτικά συμπεράσματα για το σύνολο της διαδικασίας υπολογίσαμε ένα μέτρο θέσης, το μέσο όρο και ένα μέτρο διασποράς, την τυπική απόκλιση των συνοπτικών μέτρων (HR και RMST dif) για κάθε ένα σενάριο. Ο μέσος όρος των συνοπτικών μέτρων θα μας υποδείξει τη «μέση απόσταση» των μέτρων ανάμεσα σε αυτές τις 50 δοκιμές που προσομοιώθηκαν και η τυπική απόκλιση το ποσό της διασποράς των εκτιμώμενων τιμών.

Στον παρακάτω πίνακα (Πίνακας 2) παρουσιάζονται όλοι οι υπολογισμοί για τα εκάστοτε σενάρια.

Data	Measures	Mean		Standard deviation (sd)			
Original	HR	0.6967		0.0790			
	RMST dif	0.6468		0.2056			
		15 number at risk		7 number at risk		No number at risk	
		Mean	sd	Mean	sd	Mean	sd
Using 100 coordinates	HR	0.6965	0.0815	0.6970	0.0833	0.6988	0.0769
	RMST dif	0.5595	0.1642	0.5781	0.1773	0.5676	0.1745
Using 50 coordinates	HR	0.6950	0.0818	0.6960	0.0830	0.6990	0.0790
	RMST dif	0.5477	0.1769	0.5427	0.1764	0.5461	0.1758
Using 25 coordinates	HR	0.6901	0.0862	0.6923	0.0872	0.6982	0.0827
	RMST dif	0.5291	0.1772	0.5281	0.1749	0.5190	0.1599

Πίνακας 2: Πίνακας του μέσου όρου και της τυπικής απόκλισης των συνοπτικών μέτρων HR (λόγος κινδύνου) και RMST (περιορισμένος μέσος χρόνος επιβίωσης) για όλα τα σενάρια που πραγματοποιήθηκαν όσο αφορά τον αριθμό των εξαγόμενων μεταβλητών από το αρχικό γράφημα και το πλήθος των αριθμών σε κίνδυνο που εισήχθησαν στον αλγόριθμο ανακατασκευής

Παρατηρείται ότι ο μέσος όρος των λόγων κινδύνων χρησιμοποιώντας τα πραγματικά δεδομένα και τα ανακατασκευασμένα δεν έχουν σοβαρές αποκλίσεις και πιο συγκεκριμένα οι διαφορές τους αναφέρονται στο τρίτο δεκαδικό ψηφίο. Ακόμη για όλες τις τιμές, η τυπική απόκλιση είναι μικρή, της τάξης του 0.08, γεγονός που υποδηλώνει ότι οι μέσες τιμές των λόγων κινδύνων για όλες τις δοκιμές είναι συγκεντρωμένες γύρω από τη μέση τιμή. Συνεπώς μπορούμε να εξάγουμε το συμπέρασμα ότι οι εκτιμήσεις του λόγου κινδύνου των δοκιμών από τα ανακατασκευασμένα δεδομένα επιτυγχάνεται με μεγάλη ακρίβεια και οι τιμές αυτές είναι αξιόπιστες για μετέπειτα αναλύσεις. Το γεγονός αυτό έρχεται σε πλήρη συμφωνία και με τη γραφική αξιολόγηση που πραγματοποιήθηκε προηγουμένως. Εν τούτοις είναι προτιμητέα η εισαγωγή τουλάχιστον 7 σημείων που να αναφέρεται ο αριθμός των ατόμων σε κίνδυνο όπως επίσης συνίσταται η εξαγωγή συντεταγμένων από τις αρχικές καμπύλες Kaplan Meier όπου το πλήθος τους θα είναι μεγαλύτερο από το $\frac{1}{4}$ των ατόμων που περιέχει κάθε ομάδα.

Από την άλλη μεριά όσο αφορά τη μέση διαφορά του περιορισμένου μέσου χρόνου επιβίωσης (RMST) παρατηρείται ένα αρκετά διαφορετικό μοτίβο. Ενώ η μέση εκτίμηση της διαφοράς του

περιορισμένου μέσου χρόνου επιβίωσης των δύο ομάδων από τα αρχικά δεδομένα υπολογίστηκε στο 0.64 (με τυπική απόκλιση 0.21, γεγονός που προσδίδει μεγάλη μεταβλητότητα στις τιμές του εν λόγω μέτρου) οι μέσες εκτιμήσεις από τα ανακατασκευασμένα δεδομένα όταν προσφέρονταν σε 15 σημεία οι αριθμοί των ατόμων σε κίνδυνο και είχαν εξαχθεί 100 συντεταγμένες από τις αρχικές καμπύλες, ήταν 0.56 (με τυπική απόκλιση 0.16). Παρατηρούμε δηλαδή μία υποεκτίμηση περίπου μισής τυπικής απόκλισης, όπως άλλωστε είχαμε εντοπίσει και από την γραφική αξιολόγηση, η οποία αυξάνεται καθώς η εισαγωγή πληροφοριών στον αλγόριθμο ανακατασκευής των ατομικών δεδομένων (Πίνακας 2) γίνεται πιο ελλιπής. Το γεγονός αυτό πιθανώς να οφείλεται στην επιλογή του καθορισμένου σημείου τ (στην περίπτωση μας μέγιστη παρατηρούμενη χρονική στιγμή) κατά τον υπολογισμό του περιορισμένου μέσου χρόνου επιβίωσης όπου όπως είδαμε και προηγουμένως διαφέρει αρκετά ανάμεσα στα αρχικά και τα ανακατασκευασμένα δεδομένα. Συνεπώς ο υπολογισμός του μέτρου αυτού από τα ανακατασκευασμένα δεδομένα χρήζει περαιτέρω διερεύνηση καθώς η αβεβαιότητα που παρουσιάζεται είναι μεγάλη.

Κεφάλαιο 5^ο-Συζήτηση και περαιτέρω έρευνα

Στην παρούσα διπλωματική εργασία διενεργήθηκε αρχικά μία ανασκόπηση και ανάλυση της υπάρχουσας μεθοδολογίας ανακατασκευής δεδομένων επιβίωσης, χρησιμοποιώντας τις καμπύλες Kaplan Meier καθώς και άλλες διαθέσιμες πληροφορίες, με σκοπό την ενίσχυση της ακρίβειας των αποτελεσμάτων. Στη συνέχεια επιλέγοντας τον πιο εύρωστο αλγόριθμο, αυτόν της Guyot και των συνεργατών της, εισάγοντας ωστόσο κάποιες διαθέσιμες βελτιώσεις, πραγματοποιήθηκε μία μελέτη προσομοίωσης 50 δοκιμών από την Weibull κατανομή. Οι δοκιμές αυτές ανακατασκευάτηκαν μέσω του βελτιωμένου αλγορίθμου για διάφορα σενάρια εξαγωγής συντεταγμένων από τις αρχικές καμπύλες KM και αναφοράς των ατόμων σε κίνδυνο, η κάθε μία από τις οποίες αποτελούνταν από μία ομάδα ελέγχου και μία ομάδα θεραπείας, υποθέτοντας την ισχύ της αναλογικότητας των κινδύνων. Με την εν λόγω προσομοίωση διερευνήθηκε ο ρόλος του αριθμού των σημείων που εξάγονται από τις αρχικές καμπύλες KM για την ανακατασκευή των δεδομένων καθώς και του αριθμού των χρονικών σημείων για τα οποία αναφέρεται ο αριθμός των ατόμων σε κίνδυνο. Επιπρόσθετα εξετάστηκε ο βαθμός συμφωνίας των αρχικών συνοπτικών μέτρων του λόγου κινδύνου (HR) και της διαφοράς του περιορισμένου μέσου χρόνου επιβίωσης (RMST difference) για τις δύο ομάδες θεραπείας με τα ανακατασκευασμένα συνοπτικά μέτρα χρησιμοποιώντας την ανακατασκευασμένη IPD.

Τα αποτελέσματα έδειξαν ότι όσο μεγαλύτερος είναι ο αριθμός των εξαγόμενων συντεταγμένων από τις αρχικές καμπύλες KM τόσο πιο κοντά στις πραγματικές εκτιμήσεις είναι οι εκτιμήσεις του αλγορίθμου. Ακόμη έδειξε μία αρκετά καλή εκτίμηση των λόγων κινδύνου η οποία πλησίασε την απόλυτη ταύτιση στα σενάρια όπου υπήρξε ικανοποιητική εξαγωγή συντεταγμένων (ίση με το 50% των συνολικών ατόμων της εκάστοτε ομάδας) και επαρκή σημεία στα οποία αναφερόταν ο αριθμός των ατόμων σε κίνδυνο.

Αυτό δεν συνέβει με τη μέση διαφορά του περιορισμένου μέσου χρόνου επιβίωσης όπου οι διαφορές στις εκτιμήσεις ήταν μεγαλύτερες, παρουσιάζοντας μεγάλη αβεβαιότητα. Το γεγονός αυτό μπορεί να οφείλεται στο ότι ο λόγος κινδύνου υπολογίζεται λαμβάνοντας υπόψη όλο το μήκος της καμπύλης ενώ η διαφορά του περιορισμένου μέσου χρόνου επιβίωσης επηρεάζεται αισθητά από τον πρώιμο τερματισμό που παρατηρήθηκε στα ανακατασκευασμένα δεδομένα, το οποίο χρήζει περαιτέρω διερεύνηση και ίσως κάποια ακόμη βελτίωση στον τρόπο λειτουργίας του αλγορίθμου.

Τέλος δεν παρατηρήθηκε κάποια αξιοσημείωτη διαφορά στις εκτιμήσεις των συνοπτικών μέτρων όσο αφορά το διαθέσιμο πλήθος των σημείων στα οποία διαφαίνεται ο αριθμός των ατόμων σε κίνδυνο. Συμπερασματικά λοιπόν η εν λόγω μέθοδος φαίνεται να ανακατασκευάζει σε ικανοποιητικό βαθμό τα δεδομένα όταν έχει εξαχθεί ένας εύλογος αριθμός συντεταγμένων από τις αρχικές καμπύλες και δίνονται

σε κάποια σημεία ο αριθμός των ατόμων σε κίνδυνο. Ωστόσο χρήζει ιδιαίτερης προσοχής στον υπολογισμό μέτρων όπως ο περιορισμένος μέσος χρόνος επιβίωσης καθώς διαφαίνεται μία συστηματική υποεκτίμηση λόγω του πρώιμου τερματισμού.

Στο σημείο αυτό κρίνεται σκόπιμο να αναφερθεί ότι λόγω της προσομοίωσης και της αποθήκευσης των καμπυλών ως αρχεία .png υπήρξε μια αρκετά καλή ανάλυση κατά την εξαγωγή των συντεταγμένων. Αυτό πιθανώς να έχει επηρεάσει τη διαδικασία ανακατασκευής καθώς υπό πραγματικές συνθήκες οι εικόνες που βρίσκονται δημοσιευμένες συνήθως δεν δύναται να απομονωθούν σε εικόνες υψηλής ανάλυσης. Για παράδειγμα αν μία καμπύλη KM απεικονίζεται σε μία μικρή εικόνα, κατά την προσπάθεια μεγέθυνσης για την εξαγωγή των σημείων, η ανάλυση πιθανώς να χαλάσει αρκετά, με αποτέλεσμα να υπάρξουν μεγάλες αποκλίσεις στα ανακατασκευασμένα δεδομένα. Ακόμη και ο τρόπος εξαγωγής των σημείων από τις καμπύλες, αν δεν είναι ομοιόμορφα κατανεμημένα πάνω σε αυτές, πιθανώς να επιφέρει διαφοροποιήσεις.

Όπως έγινε λοιπόν αντιληπτό η διαδικασία εξαγωγής δεδομένων από καμπύλες επιβίωσης είναι ένα ζήτημα που απασχολεί έντονα την κοινότητα της βιοστατιστικής. Στις αρχές της χιλιετίας πολλοί ήταν οι ερευνητές που εξέφραζαν την επιφυλακτικότητά τους και υποστήριζαν ότι ο αριθμός των γεγονότων δεν θα πρέπει να εκτιμάται από τις καμπύλες Kaplan Meier για μετα-αναλυτικούς σκοπούς (Duchateau, et al., 2000), με εξαίρεση μόνο περιπτώσεις όπου κανένας ασθενής δεν είχε χαθεί στην παρακολούθηση ή λογοκριθεί και υπήρχε ικανοποιητικό πλήθος ατόμων σε κίνδυνο τη στιγμή της έλευσης του γεγονότος. Σήμερα φτάνουμε στο σημείο όπου όχι μόνο έχει υποστηριχθεί η άποψη ότι είναι δυνατή η εξαγωγή δεδομένων από καμπύλες επιβίωσης σε ορισμένες περιπτώσεις και κάτω από προϋποθέσεις (Earle & Wells, 2000; Tierney, et al., 2007; Williamson, et al., 2002) αλλά έχουν αναπτυχθεί μέθοδοι προκειμένου να ξεπεράσουν πολλές πρακτικές δυσκολίες (Guyot, et al., 2012). Ωστόσο ένα σημείο που αναφέρεται συχνά ως ένας μεγάλος περιορισμός στην εξαγωγή ακριβών και αξιόπιστων δεδομένων και παρ' όλα αυτά δεν έχει βρεθεί ακόμα μία πανάκεια αντιμετώπιση, αφορά την ποιότητα της εικόνας και κατ' επέκταση τον τρόπο με τον οποίο μπορεί να γίνει η συλλογή των σημείων.

Γενικά οι δημοσιευμένες εικόνες χωρίζονται στις εικόνες ράστερ (raster images) και στις διανυσματικές εικόνες (vector images). Οι εικόνες ράστερ (ή bitmap) αποτελούνται από διατεταγμένα εικονοστοιχεία (pixels) σε ένα διατεταγμένο πλέγμα που αντιπροσωπεύεται με μία τελεία ή ένα τετράγωνο και έχει τις δικές του συντεταγμένες και χρώμα (Liu, et al., 2014). Κατά τη μεγέθυνση τους γίνονται όλο και πιο κοκκώδης και οι μεμονωμένες κουκίδες που αποτελούν τις γραμμές και τα σύμβολα στο γράφημα γίνονται όλο και πιο εμφανή. Οι καταλήξεις των συνηθέστερων εικόνων ράστερ είναι .jpeg .png .tiff καθώς και .gif. Ακόμη όσο μεγαλύτερος είναι ο αριθμός των pixels τόσο καλύτερα αντικατοπτρίζονται οι αρχικές τιμές και η εξαγωγή των σημείων επιτυγχάνεται με δύο τρόπους. Αρχικά δύναται μέσω προγραμματισμού

(π.χ μέσω της γλώσσας προγραμματισμού C++) να γίνει η εξαγωγή κάποιων σημείων σε πίνακα (τεχνικός τρόπος, μεγάλες προκλήσεις) καθώς και με τη χρήση ψηφιοποιητών γραφημάτων. Σε αυτούς, αφού εισαχθεί η εικόνα και εμφανιστεί στην οθόνη, επιτρέπει στους χρήστες την αναγνώριση οριζόντιων και κάθετων ορόσημων μέσω του δρομέα και έπειτα μέσω κλικ δίνεται η δυνατότητα συλλογής των επιθυμητών σημείων καθώς και η αποθήκευσή τους. Ευρέως χρησιμοποιούμενοι ψηφιοποιητές είναι το DigitizeIt, ScanIt, Engauge Digitizer και το Plot Digitizer ενώ είναι εφικτή η πραγματοποίηση της διαδικασίας αυτής και μέσω της R με τη βοήθεια κάποιων βιβλιοθηκών. Η διαδικασία αυτή χρησιμοποιείται σε όλες τις προαναφερθείσες μεθόδους εξαγωγής ατομικών δεδομένων από τις καμπύλες επιβίωσης. Από την άλλη, οι διανυσματικές εικόνες βασίζονται σε διανύσματα και αποτελούνται από γεωμετρικά πρωτότυπα στοιχεία ή στοιχεία όπως σημεία και γραμμές. Αναγνωρίζονται κυρίως από το γεγονός ότι μπορούν να μεγεθύνονται επ' αορίστου χωρίς απώλεια ποιότητας. Η πιο κοινή γλώσσα για την παραγωγή τους είναι η γλώσσα PostScript (Liu, et al., 2014). Η εξαγωγή των σημείων επιτυγχάνεται μέσω των διαβασμάτων των μονοπατιών (paths) όπου περιέχουν τις εντολές μέσω των οποίων έχουν δημιουργηθεί τα εκάστοτε σημεία. Συνεπώς εδώ παραλείπεται ο ανθρώπινος παράγοντας στην εξαγωγή δεδομένων.

Ορμώμενοι λοιπόν από αυτή την υπεροχή στην ανάλυση των διανυσματικών εικόνων έγινε μία προσπάθεια μετατροπής των εικόνων που απεικονίζουν τις καμπύλες Kaplan Meier και λαμβάνονται από τις δημοσιεύσεις (συνήθως μορφή ράστερ) σε διανυσματικές. Στόχος ήταν η εκμετάλλευση της υψηλότερης αυτής ανάλυσης (PostScript) για εξάλειψη της διακύμανσης που εισάγεται από τους ανθρώπινους ψηφιοποιητές και η επίτευξη υψηλότερης ακρίβειας. Η εν λόγω θα βοηθούσε τόσο στην απόκτηση μεγαλύτερης λεπτομέρειας στον αριθμό των ατόμων που βρίσκονται σε κίνδυνο σε πολλαπλά χρονικά σημεία, όσο και στο χειρισμό καμπυλών επιβίωσης με εκατοντάδες βήματα. Δεν πρέπει να παραλείψουμε την αναφορά στο γεγονός ότι ακόμα και χωρίς σημάδια λογοκρισίας στην καμπύλη (όπως σταυρούς για παράδειγμα), οι διαφορές στο μέγεθος των αλμάτων υποδεικνύουν μερική λογοκρισία (σε αντίθετη περίπτωση όλα τα άλματα θα ήταν ίσου μεγέθους $\frac{1}{n_0}$ ή πολλαπλάσια αυτού). Γενικά καθώς κινούμαστε προς τα δεξιά τα άλματα γίνονται μεγαλύτερα από αυτά που έχουν προηγηθεί (το οποίο γίνεται εύκολα αντιληπτό από τον αλγόριθμο ανακατανομής προς τα δεξιά του Efron (Efron, 1967)). Ωστόσο αυτή η προσπάθεια δεν ευδοκίμησε καθώς δεν κατέστη εφικτή η μετατροπή της εικόνας και η συλλογή των δεδομένων με αυτοματοποιημένο τρόπο χωρίς απαιτούμενα κλικ. Πιο συγκεκριμένα έγιναν διάφορες δοκιμές με πολυάριθμα πακέτα, με επικρατέστερο το grImport όπου εισάγει εικόνες Postscript (Murrell, 2009). Εν τούτοις ούτε με το εν λόγω πακέτο πραγματοποιήθηκε η εύρεση των εντολών που σχεδιάζουν την εικόνα, καθώς απαιτούσε Ghostscript και δε δινόταν να χειριστεί τις πρόσφατες εκδόσεις του Adobe Acrobat. Βέβαια πρέπει να αναφέρουμε ότι ακόμα και η εικόνα της αρχικής ανάλυσης να ήταν Postscript

αρχείο με πολύ μεγαλύτερη ανάλυση και πάλι θα υπήρχε το πρόβλημα της περιορισμένης ανάλυσης της οθόνης, ωστόσο η απόκλιση αυτή θα ήταν φανερά μικρότερη.

Καταλήγοντας θεωρείται επιτακτική η ανάγκη για περαιτέρω έρευνα σε αυτή την κατεύθυνση καθώς λόγω της βελτιωμένης ανάλυσης και της φύσης των εικόνων αυτών θα γίνει εφικτός ένας αυτοματοποιημένος τρόπος εξαγωγής των συντεταγμένων ακυρώνοντας κάθε είδους σφάλματος που προκαλείται από τον ερευνητή λόγω των χειροκίνητων μεθόδων. Επιπλέον κρίνεται απαραίτητη περαιτέρω έρευνα για τη διερεύνηση της ακρίβειας της προτεινόμενης μεθόδου με εναλλακτικές κατανομές επιβίωσης και χωρίς την ισχύ της αναλογικότητας των κινδύνων (εδώ έχει γίνει χρήση μόνο της Weibull κατανομής) καθώς και αλλαγές στον ρυθμό και στον τρόπο λογοκρισίας ώστε να αντικατοπτρίζονται πιο ρεαλιστικές δοκιμές. Για παράδειγμα τόσο στη δικιά μας προσέγγιση όσο και γενικά στο μεγαλύτερο ποσοστό των μελετών γίνεται η υπόθεση ότι η πρόσθετη λογοκρισία δεν είναι ενημερωτική καθώς ο χρόνος λογοκρισίας είναι ανεξάρτητος από το χρόνο εμφάνισης του γεγονότος. Με άλλα λόγια έχουμε θεωρήσει έναν συνεχή κίνδυνο πρόσθετης λογοκρισίας κατά τη διάρκεια παρακολούθησης της μελέτης. Κάποιοι πιο περίπλοκοι μηχανισμοί που επιτρέπουν την αλλαγή του ρυθμού λογοκρισίας στην πάροδο του χρόνου ενσωματώνοντας ενημερωτική λογοκρισία αντιστακλούν περισσότερο τα πραγματικά σκηνικά και χρειάζονται επιπλέον διερεύνηση. Τέλος χρήσιμη θα ήταν η διεξαγωγή έρευνας με σκοπό την εύρεση τρόπων ώστε να είναι εφικτή η δυνατότητα διερεύνησης των αλληλεπιδράσεων μεταξύ των μεταβλητών κατά την εξαγωγή των δεδομένων.

Περίληψη

Ανακατασκευή δεδομένων επιβίωσης από δημοσιευμένες καμπύλες Kaplan-Meier

Η ανάγκη για αξιόπιστες και έγκυρες αποφάσεις σε θέματα που αφορούν τη δημόσια υγεία και την καθημερινή πρακτική καθώς και τη συνεχόμενη αύξηση της παρεχόμενης πληροφορίας για κάθε επιστημονική υπόθεση, οδήγησαν στην επιτακτική διεξαγωγή μετα-αναλύσεων, με στόχο την αξιοποίηση όλου του διαθέσιμου όγκου των δεδομένων και την εξαγωγή έμπιστων συμπερασμάτων. Ωστόσο η ποικιλομορφία των αναφορών στην ανάλυση επιβίωσης δημιούργησε την ανάγκη ανακατασκευής των ατομικών δεδομένων επιβίωσης από τις καμπύλες KM, αφού αποτελεί την πιο κοινή δημοσιευμένη πληροφορία, χρησιμοποιώντας ωστόσο και άλλα διαθέσιμα στοιχεία.

Στην παρούσα διπλωματική εργασία πραγματοποιήθηκε μία ανασκόπηση και ανάλυση των διαθέσιμων μεθόδων ανακατασκευής δεδομένων από καμπύλες Kaplan Meier καθώς και από διαθέσιμα συνοπτικά μέτρα. Οι μέθοδοι αυτοί προτάθηκαν με χρονολογική σειρά από τους Parmar, Williamson, Hoyle και Henley, Guyot και Rogula. Στη συνέχεια επιλέγοντας την πιο εύρωστη από τις πέντε μεθόδους, αυτή της Guyot και των συνεργατών της (2012) και εισάγοντας κάποιες διαθέσιμες βελτιώσεις που άπτονταν θέματα μονοτονίας των εξαγόμενων συντεταγμένων και οριοθέτηση των τιμών προς αποφυγή ατέρμονων βρόγχων, υλοποιήθηκε μία προσομοίωση 50 δοκιμών. Η προσομοίωση διεξήχθη με χρήση της Weibull κατανομής και υποθέτοντας την ισχύ της αναλογικότητας των κινδύνων. Στη συνέχεια διερευνήθηκε η ποιότητα των ανακατασκευασμένων ατομικών δεδομένων επιβίωσης σε εναλλακτικά σενάρια χρησιμοποίησης του αριθμού των σημείων που εξάγονται από τα αρχικά γραφήματα KM και του αριθμού των χρονικών στιγμών στα οποία αναφέρονται το πλήθος των ατόμων σε κίνδυνο. Επίσης μελετήθηκε ο βαθμός συμφωνίας των αρχικών συνοπτικών μέτρων (του λόγου κινδύνου (HR) και της διαφοράς του περιορισμένου μέσου χρόνου επιβίωσης (RMST)) από τα δεδομένα, με τα ανακατασκευασμένα συνοπτικά μέτρα από την ανακατασκευασμένη IPD.

Τα αποτελέσματα έδειξαν ότι όσο μεγαλύτερος είναι ο αριθμός των εξαγόμενων συντεταγμένων από τις αρχικές καμπύλες Kaplan Meier τόσο πιο κοντά στην πραγματική τιμή είναι η εκτίμηση του αλγορίθμου. Δεν παρατηρήθηκε κάποια αξιοσημείωτη διαφορά στις εκτιμήσεις των συνοπτικών μέτρων όσο αφορά το διαθέσιμο πλήθος των σημείων στα οποία διαφαίνεται ο αριθμός των ατόμων σε κίνδυνο. Ωστόσο παρατηρήθηκε ότι τα ανακατασκευασμένα δεδομένα αντανακλούν ένα πρόωρο τερματισμό της μελέτης, γεγονός που γίνεται εύκολα αντιληπτό από τις ανακατασκευασμένες καμπύλες KM. Τέλος, ενώ οι εκτιμήσεις των λόγων κινδύνου σε κάθε ένα από τα σενάρια που εκτελέστηκαν ήταν ικανοποιητικές και στιβαρές, ιδίως στις καταστάσεις όπου προσφέρονταν και επιπλέον πληροφορίες, δεν ίσχυε το ίδιο για τη

διαφορά του περιορισμένου μέσου χρόνου επιβίωσης όπου υπήρχαν αποκλίσεις στις εκτιμήσεις με μεγάλη αβεβαιότητα. Τόσο λοιπόν ο πρόωρος τερματισμός που παρατηρείται από τα ανακατασκευασμένα δεδομένα όσο και οι μη καλές εκτιμήσεις κάποιων συνοπτικών μέτρων απαιτούν επιπλέον διερεύνηση και μεγάλη προσοχή. Ακόμη θεωρείται σκόπιμη η περαιτέρω έρευνα για την ακρίβεια της εν λόγω μεθόδου σε καταστάσεις όπου χρησιμοποιούνται εναλλακτικές κατανομές επιβίωσης (πέραν της Weibull) καθώς και όταν εισάγεται ενημερωτική λογοκρισία, γενικά σε καταστάσεις που αντιπροσωπεύουν καλύτερα ενδεχομένως τα πραγματικά δεδομένα.

Τέλος έγινε μία προσπάθεια μετατροπής των εικόνων ράστερ (raster images) στις οποίες απεικονίζονται οι καμπύλες KM σε διανυσματικές (vector images), με στόχο την επίτευξη ενός αυτοματοποιημένου τρόπου εξαγωγής των συντεταγμένων εξαλείφοντας σφάλματα που προέρχονται από τους ερευνητές λόγω των χειροκίνητων μεθόδων, ωστόσο δεν κατέστη δυνατή η πραγματοποίησή της. Εν τούτοις η ανάλυση της εικόνας καθώς και η ακρίβεια της συλλογής των σημείων από τους ερευνητές αντανακλούν σε συστηματικά παρεχόμενα σφάλματα και είναι επιτακτική η ανάγκη βελτιστοποίησης στην εν λόγω κατεύθυνση.

Abstract

Reconstructing survival data from published Kaplan–Meier curves

The necessity of reliable and valid decision-making regarding public health issues and everyday practice, as well as the continuous accumulation of information in each scientific case, led to the imperative implementation of meta-analyses, aiming at utilizing all the available data and making dependable inferences. Nevertheless, the variety of references in Survival Analysis necessitated the reconstruction of the individual survival data from Kaplan-Meier curves, since they constitute the most commonly published figures, taking into account additional statistical elements too.

In the current thesis, a literature review along with an in-depth explanation of the available methods for reconstructing data from Kaplan-Meier curves and other summary measures at hand have been conducted. In chronological order, these approaches were proposed by Parmar, Williamson, Hoyle & Henley, Guyot, and Rogula. Choosing the most robust method amongst them, i.e, the one suggested by Guyot et al. (2012), and after modifying and enhancing it, a simulation of 50 replications, based on the Weibull distribution and the proportionality assumption, has been carried out. The quality of the reconstructed data has been examined under various alternative scenarios regarding both the number of points extracted from the initial KM graphs and the number of time points at which the size of the risk set is known. Moreover, the consistency between the initial summary measures (HR & RMST) and the ones occurring from the reconstructed data has been investigated.

The results have shown that the larger the number of the extracted coordinates from the initial Kaplan-Meier curves is, the closer to the true value is the algorithm estimate. The discrepancies among the estimates when the number of time points (at which the number of subjects at risk is reported) changes have been negligible. However, it was evident from the Kaplan-Meier curves that the termination of a study of interest occurred earlier in the reconstructed data. Finally, the HR estimates in the applied scenarios were pleasing and robust, especially when additional information was provided. On the contrary, Restricted Mean Survival Time differences estimates exhibited greater variability. Consequently, both the early termination phenomenon in the reconstructed data and the poor estimates of some of the summary measures require further investigation and careful consideration. Additionally, an in-depth examination of the aforementioned method's accuracy is of the utmost importance on occasions where different distributions are used and under the presence of censoring, two scenarios that reflect better real-life problems.

Lastly, an attempt to convert raster pictures of KM curves to vector images has been made, with the intention of automatizing the coordinate extraction procedure, rectifying in this way the issue of additional errors (OR BIAS) due to manual implementation of the methods. Since this was not possible, it should be stressed that the picture quality, as well as the accuracy of point selection from the investigators, introduce systematic errors, and it is necessary to improve these approaches and optimize them with that in mind.

BIBΛΙΟΓΡΑΦΙΑ

- A'hern, R. P., Smith, I. E. & Ebbs, S. R., 1993. Chemotherapy and survival in advanced breast cancer: the inclusion of doxorubicin in Cooper type regimens. *British Journal of Cancer*, pp. 67, 801–805.
- Altman, D., De Stavola, B., Love, S. & Stepniewska, K., 1995. Review of survival analyses published in cancer journals. *British Journal of Cancer*, p. 72:511–518.
- Andersen, P. & Perme, M., 2010. Pseudo-observations in survival analysis. *Statistical Methods in Medical Research*, p. 19(1):71–99.
- Barlow, R. E., Bartholomew, D. J., Bremner, J. M. & Brunk., H. D., 1972. *Statistical Inference Under Order Restrictions: Theory and Application of Isotonic Regression..* New York: Wiley.
- Bland, J. M. & Altman, D. G., 1986. Statistical methods for assessing agreement between two methods of clinical measurement. pp. 1(8476),307–310..
- Bland, J. M. & Altman, D. G., 1999. Measuring agreement in method comparison studies. *Statistical methods in medical research*, pp. 8(2), 135–160..
- Breslow, N., 1974. Covariance analysis of censored survival data. *Biometrics*, p. 89–99.
- Cox, D., 1979. A note on the graphical analysis of survival data.. *Biometrika*, p. 66: 188–190.
- Cox, R. D., 1975. Partial likelihood. *Biometrika*, p. 62(2):269–276.
- Dear, K. B. G., 1994. Iterative generalised least squares for meta-analysis of survival data at multiple times. *Biometrics*, pp. 50, 989—1002.
- Deeks, J. J., Higgins, J. P. T. & Altman, D. G., 2008. Analysing data and undertaking meta-analyses.. *Cochrane Handbook for Systematic Reviews of Interventions*, p. 243–296.
- Duchateau, L., Collette, L., Sylvester, R. & Pignon, J., 2000 . Estimating number of events from the Kaplan-Meier curve for incorporation in a literature-based meta-analysis: what you don't see you can't get!. *Biometrics*, pp. 56 (3): 886-892.
- Earle, C. & Wells, G., 2000. An Assessment of Methods to Combine Published Survival Curves.. *Medical Decision Making*, pp. 20: 104-111.
- Efron, B., 1967. The Two-Sample Problem with Censored Data.. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, pp. Vol. 4, 831-852..
- Efron, B., 1977. The efficiency of cox's likelihood function for censored data.. *Journal of the American statistical Association*, p. 72(359):557–565.
- Guyot, P., Ades, A. E., Ouwens, M. J. N. M. & Welton., N. J., 2012. Enhanced secondary analysis of survival data: Reconstructing the data from published Kaplan– Meier survival curves. *BMC Medical Research Methodology* 12: 9.
- Higgins, J. & Thomas, J., 2022. *Cochrane Handbook for Systematic Reviews of Interventions* Version 6.3.

Hosmer, D. & Lemeshow, S., 1999. *Applied Survival Analysis: Regression Modelling of Time to Event Data*. New York: Wiley.

Hoyle, M. W. & Henley, W., 2011. Improved curve fits to summary survival data: application to economic evaluation of health technologies.. *BMC medical research methodology*, pp. 11, 139.

Hunink, M. G. M. & Wong, J. B., 1994. Meta-analysis of failure-time data with adjustment for covariates. *Medical Decision Making*, pp. 14, 59—70.

Irvine, A. F., Waise, S., Green, E. W. & Stuart, B., 2020. A non-linear optimisation method to extract summary statistics from Kaplan-Meier survival plots using the published P value.. *BMC Medical Research Methodology*, pp. 20(1), 269.

Irwin, J., 1949. The standard error of an estimate of expectation of life, with special reference to expectation of tumourless life in experiments with mice. *Epidemiology & Infection*, p. 47(2):188–189.

Jenkins, S. P., 2005. Survival analysis. Unpublished manuscript, Institute for Social and Economic Research. *University of Essex, Colchester, UK*, , pp. 42, 54-56..

Kaplan, E. & Meier, P., 1958. "Nonparametric estimation from incomplete observations.". *Journal of the American Statistical Association*, p. 53(282):457–81.

Kim, J. S. & Yum, B.-J., 2008. Selection between Weibull and lognormal distributions: A comparative simulation study. *Computational Statistics & Data Analysis*, pp. 477-485.

Kleinbaum, D. G. & Klein, M., 1996. *Survival analysis*. s.l.:Springer,pp 38-42.

Kloecker, D., Davies, M., Khunti, K. & Zaccardi, F., 2020. Uses and Limitations of the Restricted Mean Survival Time: Illustrative Examples From Cardiovascular Outcomes and Mortality Trials in Type 2 Diabetes. *Annals of Internal Medicine*, pp. 172(8):541-552.

Liu, N., Zhou, Y. & Lee, J., 2021. IPDfromKM: Reconstruct Individual Patient Data from Published Kaplan-Meier Survival Curves. *BMC Medical Research Methodology*.

Liu, Z., Rich, B. & Hanley, J., 2014. Recovering the raw data behind a non-parametric survival curve.. *Systematic Reviews*, pp. 3, 151 .

McCaw, Z., Yin, G. & Wei, L. J., 2019. Using the Restricted Mean Survival Time Difference as an Alternative to the Hazard Ratio for Analyzing Clinical Cardiovascular Studies. *Circulation*, pp. 140(17):1366-1368.

Messori, A., Bartoli, L. & Trippoli, S., 2021. The restricted mean survival time as a replacement for the hazard ratio and the number needed to treat in long-term studies. *ESC Heart Fail*, pp. 8(3):2345-2348.

Moher, D., Hopewell, S., Schulz, K. & al., e., 2010. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *British Medical Journal*, p. 340:c869.

Murrell, P., 2009. "Importing Vector Graphics: The grImport Package for R.". *Journal of Statistical Software*, pp. 30(4), 1–37.

Parmar, M., Torri, V. & Stewart, L., 1998. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints.. *Statistics in Medicine*, p. 17:2815 – 2834.

- Parner, E. & Andersen, P., 2010. Regression Analysis of Censored Data Using Pseudo-observations. *The Stata Journal*, pp. 10, 408 - 422.
- Peto, R. και συν., 1977. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples.. *Br J Cancer*, p. 35: 1–39.
- Pigott, T. D. & Polanin, J. R., 2020. Methodological Guidance Paper: High-Quality Meta-Analysis in a Systematic Review. *Review of Educational Research*, pp. 90(1), pp. 24–46.
- Rogula, B., Lozano-Ortega, G. & Johnston, K., 2022. A Method for Reconstructing Individual Patient Data From Kaplan-Meier Survival Curves That Incorporate Marked Censoring Times. *MDM Policy & Practice*, 7(1).
- Royston, P. & Parmar, M., 2013. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. *Medical Research Methodology*, p. 13:152.
- Royston, P. & Parmar, M. K., 2011. The use of restricted mean survival time to estimate the treatment effect in randomized clinical trials when the proportional hazards assumption is in doubt.. *Statistics in medicine*, p. 30(19):2409– 2421.
- Saluja, R., Cheng, S., Delos Santos, K. & Chan, K., 2019. Estimating hazard ratios from published Kaplan-Meier survival curves: A methods validation study.. *Research Synthesis Methods*.
- Simmonds, M. C. και συν., 2005. Meta-analysis of individual patient data from randomized trials: a review of methods used in practice. *Clinical trials (London, England)*, pp. 2(3), 209–217.
- Syn, N. L. και συν., 2020. Survival Advantage of Laparoscopic Versus Open Resection For Colorectal Liver Metastases: A Meta-analysis of Individual Patient Data From Randomized Trials and Propensity-score Matched Studies.. *Annals of surgery*, pp. 272(2), 253–265.
- Tierney, J. et al., 2007. Practical methods for incorporating summary time-to-event data into meta-analysis.. *Trials* 8.
- Uno, H. και συν., 2015. Alternatives to Hazard Ratios for Comparing the Efficacy or Safety of Therapies in Noninferiority Studies.. *Ann Intern Med.*, pp. 163(2):127-34.
- van Putten, W. & Royston, P., 2017. irax: Stata module to perform isotonic regression analysis. Statistical Software Components. *Department of Economics, Boston College* <https://ideas.repec.org/c/boc/bocode/s458406.html>..
- Wei, Y. & Royston, P., 2017. «Reconstructing Time-to-event Data from Published Kaplan–Meier Curves.». *The Stata Journal* , pp. 17(4), 786–802..
- Wei, Y., Royston, P., Tierney, J. F. & Parmar, M. K. B., 2015. Meta-analysis of time- to-event outcomes from randomized trials using restricted mean survival time: application to individual participant data. *Statist. Med.*, p. 34: 2881–2898.
- Williamson, P., Smith, C., Hutton, J. & Marson, A., 2002. Aggregate data metaanalysis with time-to-event outcomes.. *Statistics in Medicine*, pp. 21:3337-3351.
- Πανάγου, Φ., 2014. Περιορισμένος Μέσος Χρόνος Επιβίωσης και Εφαρμογές στη Μετα – Ανάλυση, Διπλωματική εργασία ΜΠΣ. *Αθήνα: Πανεπιστήμιο Αθηνών*.

