



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

**Μέθοδοι Βιοπληροφορικής, Υπολογιστικής
Συστημικής Βιολογίας και Μηχανικής Μάθησης,
Εφαρμοσμένες στην In Silico Ογκολογία**

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Ελευθέριος, Ν. Ουζούνογλου

Αθήνα, Ιούνιος 2018



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

**Μέθοδοι Βιοπληροφορικής, Υπολογιστικής
Συστημικής Βιολογίας και Μηχανικής Μάθησης,
Εφαρμοσμένες στην In Silico Ογκολογία**

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Ελευθέριος, Ν. Ουζούνογλου

**Συμβουλευτική Επιτροπή : Ανδρέας-Γεώργιος, Ν, Σταφυλοπάτης, Καθηγητής
ΣΗΜΜΥ-ΕΜΠ (Επιβλέπων)**

**Γεώργιος, Σ., Σταματάκος, Ερευνητής Α΄
ΕΠΙΣΕΥ/ΕΜΠ, Διευθυντής Ερευνών, Επισκέπτης
Καθηγητής ΣΗΜΜΥ-ΕΜΠ**

**Δημήτριος-Διονύσιος, Γ, Κουτσούρης, Καθηγητής
ΣΗΜΜΥ-ΕΜΠ**

Εγκρίθηκε από την επταμελή εξεταστική επιτροπή την 12^η Ιουνίου 2018.

.....
Ανδρέας-Γεώργιος Σταφυλοπάτης
Καθηγητής ΣΗΜΜΥ-ΕΜΠ

.....
Γεώργιος Σταματάκος
Ερευνητής Α΄ ΕΠΙΣΕΥ/ΕΜΠ

.....
Δημήτριος-Διονύσιος Κουτσούρης
Καθηγητής ΣΗΜΜΥ-ΕΜΠ

.....
Ηλίας Μανωλάκος
Καθηγητής ΕΚΠΑ,
Τμήμα Πληροφορικής &
Τηλ/νιών

.....
Λεωνίδας Αλεξόπουλος
Επ.Καθηγητής ΣΗΜΜΥ-ΕΜΠ

.....
Μαρία Μοσχόβη
Επ.Καθηγήτρια ΕΚΠΑ,
Ιατρική Σχολή

.....
Βασιλική Παπαεωαγγέλου
Καθηγήτρια ΕΚΠΑ,
Ιατρική Σχολή

Αθήνα, Ιούνιος 2018

.....
Ελευθέριος, Ν. Ουζούνογλου
Διδάκτωρ Ε.Μ.Π.

Copyright © Ελευθέριος, Ν. Ουζούνογλου, 2018.
Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Στον πατέρα μου Νίκο και στον γιό μου Νικόλα

Στην μητέρα μου Έλσα και στη γυναίκα μου Τίνα

Στα παιδιά που παλεύουν με τον καρκίνο και σε αυτούς που τα φροντίζουν

Περίληψη

Στα πλαίσια της παρούσας διατριβής μελετήθηκαν οι επιστημονικοί τομείς της Βιοπληροφορικής, της Υπολογιστικής Συστημικής Βιολογίας και της Μηχανικής Μάθησης με στόχο την αξιοποίησή τους σε προβλήματα του έντονα αναπτυσσόμενου και εξελισσόμενου χώρου της *In Silico* Ογκολογίας. Ο χώρος αυτός θέτει ως βασικό στόχο την ανάπτυξη υπολογιστικών μοντέλων ικανών να προσομοιώσουν την ανάπτυξη αλλά και την απόκριση στη θεραπεία καρκινικών όγκων. Κομβική συμμετοχή σε αυτές τις προσπάθειες έχει η Ομάδα για την *In Silico* Ογκολογία και την *In Silico* Ιατρική του Ερευνητικού Πανεπιστημιακού Ινστιτούτου Συστημάτων Επικοινωνιών και Υπολογιστών του Εθνικού Μετσόβιου Πολυτεχνείου (ΕΠΙΣΕΥ-ΕΜΠ) μέσω των πολυετών προσπαθειών ανάπτυξης μιας οικογένειας μοντέλων, των Ογκοπροσομοιωτών. Οι συγκεκριμένοι Ογκοπροσομοιωτές εστιάζουν κυρίως στην προσομοίωση των φαινομένων στο κυτταρικό και σε ανώτερα αυτού επίπεδα. Στην παρούσα διατριβή αναζητήθηκαν τα σημεία εκείνα όπου μέθοδοι της Βιοπληροφορικής, της Υπολογιστικής Συστημικής Βιολογίας και της Μηχανικής μάθησης μπορούν να συμβάλλουν στην εξειδίκευση των Ογκοπροσομοιωτών και στον μοριακό χώρο αλλά και στην αξιοποίησή τους για προβλεπτικές διαδικασίες σχετικά με την εξατομικευμένη απόκριση καρκινικών όγκων σε θεραπεία. Προκειμένου να παρουσιαστούν οι δυνατότητες αυτές, επιλέχθηκαν συγκεκριμένες εφαρμογές και δόθηκε εστίαση στην περίπτωση της Οξείας Λεμφοβλαστικής Λευχαιμίας μέσω της χρήσης και μελέτης της επεκτασιμότητας του Ογκοπροσομοιωτή Λευχαιμίας. Συγκεκριμένα, αναπτύχθηκαν και εκπαιδεύτηκαν μέσω μεθόδων βελτιστοποίησης ένα μοντέλο Συστημικής Βιολογίας για την προσομοίωση της βιοχημικής ρύθμισης του κυτταρικού κύκλου στην Οξεία Λεμφοβλαστική Λευχαιμίας και ένα μοντέλο προσομοίωσης της Φαρμακοκινητικής για το φάρμακο Πρεδνιζόνη, ικανά να τροφοδοτήσουν με πληροφορία τον Ογκοπροσομοιωτή. Επίσης μέσω μεθόδων Βιοπληροφορικής και Μηχανικής Μάθησης αναπτύχθηκε το Υβριδικό Σύστημα Ογκοπροσομοιωτή Λευχαιμίας, δηλαδή διαδικασίες για την αυτοματοποιημένη προσαρμογή του Ογκοπροσομοιωτή σε δεδομένα ασθενών αλλά και την πρόβλεψη της απόκρισης νέων ασθενών σε θεραπεία μέσω εκπαίδευσης επιμέρους μοντέλων μηχανικής μάθησης. Συγκεκριμένα, επιχειρήθηκε η πρόβλεψη της απόκρισης στην Πρεδνιζόνη, ασθενών με Οξεία Λεμφοβλαστική Λευχαιμίας, παιδικής ηλικίας, η οποία αποτελεί βασική πληροφορία για την διαστρωμάτωση των ασθενών σε ομάδες κινδύνου. Η πρωτότυπη προσπάθεια και η επαρκής επιτυχία πρόβλεψης μέσω του

Υβριδικού Ογκοπροσομοιωτή αποτελούν σημαντικό βήμα για την περαιτέρω ανάπτυξη του χώρου της In Silico Ογκολογίας.

Λέξεις Κλειδιά

In Silico Ογκολογίας, Μηχανική Μάθηση, Συστημική Βιολογίας, Βιοπληροφορική, Υπολογιστική Βελτιστοποίηση, Υβριδικός Ογκοπροσομοιωτής, Οξεία Λεμφοβλαστική Λευχαιμία, Καρκίνος

Abstract

In the context of the present thesis, the scientific disciplines of Bioinformatics, Computational Systems Biology and Machine Learning have been studied, aiming at exploiting them in the framework of the constantly growing and evolving field of *In Silico Oncology*. This field has as its central objective the development of computational models capable to simulate the growth of cancerous tumors as well as their response to therapy. The In Silico Oncology and In Silico Medicine group of the Institute of Communication and Computer Systems of the National Technical University of Athens (ICCS-NTUA) has nodal involvement in this through its efforts on developing a family of models, the Oncosimulators. These specific Oncosimulators, are focusing mainly on the simulation of phenomena of the cellular and higher levels. In the present thesis, the aspects in which methods of the fields of Bioinformatics, Computational Systems Biology and Machine Learning may contribute in the extension of the Oncosimulators to the molecular space as well as in the exploitation of them for predictive procedures regarding the personalized response of tumors to therapy. In order for these to be presented, specific applications has been selected focusing on the case of Acute Lymphoblastic Leukemia through the usage and the study of the extendibility of the Leukemia Oncosimulator. Specifically, a Systems Biology oriented model for the simulation of the biochemical regulation of the cell cycle in Acute Lymphoblastic Leukemia as well as a Pharmacokinetic model for the Drug Prednisone, capable to provide input to the Oncosimulator, have been developed and trained through computational optimization methods. Moreover, through the usage of Bioinformatics and Machine Learning methods, a Hybrid Leukemia Oncosimulator System has been developed, which consists of procedures for automated adaptation of the Oncosimulator on patients data as well as the prediction of the response of patients to therapy through training machine learning models. More specifically, the prediction of the response to Prednisone of Acute Lymphoblastic Leukemia pediatric patients, which has a central role in the patient stratification, has been attempted. The original effort and the adequate success in the prediction by the Hybrid Oncosimulator consist of a crucial step for the further development of the field of In Silico Oncology.

Keywords

In Silico Oncology, Machine Learning, Systems Biology, Bioinformatics, Computational Optimization, Hybrid Oncosimulator, Acute Lymphoblastic Leukemia, Cancer

ΠΕΡΙΕΧΟΜΕΝΑ

ΥΠΟΤΙΤΛΟΙ ΕΙΚΟΝΩΝ	15
ΥΠΕΡΤΙΤΛΟΙ ΠΙΝΑΚΩΝ	18
ΠΡΟΛΟΓΟΣ – ΕΥΧΑΡΙΣΤΙΕΣ	19
1. Καρκίνος, Βασικοί βιολογικοί μηχανισμοί, Αιμοποίηση και Οξεία Λεμφοβλαστική Λευχαιμία	21
1.1 Εισαγωγή.....	21
1.2 Ο Καρκίνος	21
1.3 Βασικοί βιολογικοί μηχανισμοί σχετικοί με τον καρκίνο	24
1.3.1 Κυτταρική Κινητική και Κυτταρικός Κύκλος.....	24
1.3.2 Κυτταρικός Θάνατος.....	27
1.4 Αιμοποιητικό Σύστημα και Οξεία Λεμφοβλαστική Λευχαιμία	29
1.4.1 Αιμοποίηση	29
1.4.1.1 Παραγωγή των κυττάρων του αίματος.....	30
1.4.1.1.1 Μη αναγνωρίσιμα κύτταρα	30
1.4.1.1.1.1 Ρύθμιση της παραγωγής	31
1.4.1.1.2 Αναγνωρίσιμα Κύτταρα και Ρύθμιση της διαφοροποίησης και της παραγωγής των αναγνωρίσιμων κυττάρων του μυελού	32
1.4.1.1.2.1 Ερυθρά σειρά.....	32
1.4.1.1.2.2 Κοκκίωδης (μυελική) σειρά	33
1.4.1.1.2.3 Μονοκύτταρα	33
1.4.1.1.2.4 Αιμοπετάλια. Μεγακαρυοκυτταρική σειρά.....	33
1.4.1.1.2.5 Λεμφοκύτταρα.....	33
1.4.2 Η ασθένεια.....	34
1.4.2.1 Μορφολογική κατάταξη ή ταξινόμηση κατά FAB	36
1.4.2.2 Κατάταξη με βάση την προέλευση και τον βαθμό διαφοροποίησης μέσω ανοσοφαινοτυπικής ανάλυσης κυτταρικών δεικτών επιφάνειας.....	36
1.4.2.3 Κατάταξη με βάση τον καρύοτυπο και τις κύριες μοριακές διαταραχές	38
1.4.3 Θεραπευτική Αντιμετώπιση.....	39
1.4.3.1 Η ομάδα εργασίας BFM και η κλινική δοκιμή ALL-BFM 2000	40
1.4.3.1.1 Διαστρωμάτωση κινδύνου.....	41
1.4.3.1.2 Στάδια θεραπείας.....	43
2. Μηχανική Μάθηση, Βιοπληροφορική, Συστημική Βιολογία και In Silico Ογκολογία.....	47
2.1 Εισαγωγή.....	47

2.2	Μηχανική Μάθηση	47
2.2.1	Η ανάγκη για Εξόρυξη Δεδομένων και Γνώσης	48
2.2.2	Η διαδικασία Εξόρυξης Δεδομένων ως μέρος της διαδικασίας Εξόρυξης Γνώσης	48
2.2.3	Η εξόρυξη Δεδομένων	50
2.2.3.1	Κατηγοριοποίηση-Ταξινόμηση	51
2.2.3.2	Παλινδρόμηση.....	52
2.2.3.3	Ομαδοποίηση	52
2.2.3.4	Κανόνες Συσχέτισης.....	53
2.2.3.5	Πρότυπα Ακολουθιών	53
2.2.3.6	Ομοιότητα Χρονολογικών Σειρών	53
2.2.3.7	Απεικόνιση και Μείωση Διαστάσεων	54
2.2.3.8	Μάθηση με ή χωρίς επίβλεψη και Αντεπικύρωση	54
2.3	Βιοπληροφορική	55
2.3.1	Η βιολογική έρευνα μέχρι και την αποκωδικοποίηση του γονιδιώματος	55
2.3.2	Η συμβολή της Πληροφορικής στην αποκωδικοποίηση και στην ανάλυση του γονιδιώματος. Η γέννηση της Βιοπληροφορικής	59
2.3.3	Από την Βιοπληροφορική στη Συστημική Βιολογία.....	62
2.4	Υπολογιστική Συστημική Βιολογία.....	63
2.4.1	Συστημική Βιολογία και Δίκτυα Βιοχημικών Αντιδράσεων.....	64
2.4.1.1	Θεωρητικό Πλαίσιο Δικτύων Βιοχημικών Αντιδράσεων	65
2.4.1.2	Μέθοδοι Προσομοίωσης Δικτύων Βιοχημικών Αντιδράσεων.....	72
2.4.1.2.1	Ντετερμινιστική Προσομοίωση	76
2.4.1.2.2	Στοχαστική Προσομοίωση	78
2.4.1.2.2.1	Ο Αλγόριθμος SSA	81
2.4.1.2.2.2	Η μέθοδος της Επόμενης Αντίδρασης.....	82
2.4.1.3	Εναλλακτικές Μέθοδοι Προσομοίωσης.....	84
2.4.1.3.1	Προσομοίωση με άλματα	84
2.4.1.3.2	Προσομοίωση μέσω Στοχαστικών Διαφορικών Εξισώσεων	85
2.4.1.3.3	Υβριδική Προσομοίωση.....	87
2.4.1.4	Ανάλυση και Εκπαίδευση Δικτύων Βιοχημικών Αντιδράσεων	88
2.4.1.4.1	Μέθοδοι Ανάλυσης Βιοχημικών Δικτύων	88
2.4.1.4.1.1	Ανάλυση Ευαισθησίας	88
2.4.1.4.1.2	Άλλες μέθοδοι Ανάλυσης.....	92
2.4.1.4.2	Εκπαίδευση Βιοχημικών Δικτύων-Εκτίμηση Παραμέτρων.....	93
2.4.1.4.2.1	Εκτίμηση Παραμέτρων μέσω μεθόδων Βελτιστοποίησης	94

2.4.1.4.2.1.1	Θεωρητικό Πλαίσιο Εκτίμησης Παραμέτρων μέσω Μεθόδων Βελτιστοποίησης	94
2.4.1.4.2.1.2	Καταλληλότητα Εφαρμογής μεθόδων Βελτιστοποίησης και Κυριότερες Μέθοδοι.....	96
2.4.1.4.2.1.3	Συμπεράσματα από τη Βιβλιογραφία, Πλεονεκτήματα και Μειονεκτήματα των μεθόδων Βελτιστοποίησης.....	100
2.4.1.4.2.2	Εναλλακτικές μέθοδοι Εκτίμησης Παραμέτρων	101
2.4.1.4.2.3	Συμπεράσματα ανασκόπησης μεθόδων Εκτίμησης Παραμέτρων.....	102
2.5	In Silico Ογκολογία και Ογκοπροσομοιωτής.....	104
2.5.1	Ογκοπροσομοιωτής χωρικής προσέγγισης	106
2.5.2	Ογκοπροσομοιωτής μη χωρικής προσέγγισης.....	113
2.6	Φιλοσοφία εφαρμογής μεθόδων.....	114
3.	Μοντελοποίηση της βιοχημικής ρύθμισης του κυτταρικού κύκλου στην Οξεία Λεμφοβλαστική Λευχαιμία	115
3.1	Εισαγωγή.....	115
3.2	Η βιοχημική ρύθμιση του κυτταρικού κύκλου.....	116
3.2.1	Ο ρόλος των κυκλινών και των εξαρτώμενων από κυκλίνες κινασών ...	116
3.2.2	Αναστολείς των συμπλεγμάτων κυκλινών και εξαρτώμενων από κυκλίνες κινασών	118
3.2.3	Σημεία Ελέγχου του Κυτταρικού Κύκλου	119
3.3	Ειδικά χαρακτηριστικά της βιοχημικής ρύθμισης του κυτταρικού κύκλου στην Οξεία Λεμφοβλαστική Λευχαιμία.....	122
3.4	Βιβλιογραφική ανασκόπηση της μοντελοποίησης της βιοχημικής ρύθμισης του κυτταρικού κύκλου και επιλογή μοντέλου αναφοράς	128
3.5	Δημιουργία και υλοποίηση προτεινόμενου μοντέλου	132
3.6	Στρατηγική προσαρμογής του προτεινόμενου μοντέλου και υλοποίησή της.....	142
3.6.1	Ποιοτική παρουσίαση κριτηρίων προσαρμογής.....	145
3.6.2	Λεπτομερής παρουσίαση κριτηρίων προσαρμογής.....	147
3.7	Προσαρμογή του προτεινόμενου μοντέλου.....	150
3.8	Αποτελέσματα προσαρμογής, δοκιμές εκτέλεσης και αξιολόγησης του μοντέλου ...	156
3.9	Δοκιμή υποθέσεων.....	159
3.10	Σύνοψη, Συμπεράσματα και Μελλοντικοί στόχοι	165
4.	Μοντελοποίηση Φαρμακοκινητικής της από στόματος χρήσης του φαρμάκου Πρεδνιζόνη σε παιδιά με Οξεία Λεμφοβλαστική Λευχαιμία .	167
4.1	Εισαγωγή.....	167
4.2	Βιβλιογραφική ανασκόπηση της μοντελοποίησης και της προσομοίωσης της φαρμακοκινητικής του φαρμάκου Πρεδνιζόνη	168

4.2.1	Φαρμακοκινητικό μοντέλο για την χορήγηση Πρεδνιζολόνης σε παιδιά που πάσχουν από Οξεία Λεμφοβλαστική Λευχαιμία.....	168
4.2.2	Φαρμακοκινητικό μοντέλο για την από στόματος χορήγηση του φαρμάκου Πρεδνιζόνη	170
4.3	Ανάπτυξη εκπαίδευση και δοκιμή προτεινόμενου μοντέλου	173
4.3.1	Ορισμός προβλήματος.....	173
4.3.2	Υλοποίηση προ-υπαρχόντων μοντέλων και έλεγχος αυτής.....	173
4.3.3	Δημιουργία της δομής του προτεινόμενου μοντέλου.....	176
4.3.4	Εκπαίδευση και αξιολόγηση του προτεινόμενου μοντέλου	178
5.	Προσαρμογή Ογκοπροσομοιωτή και Πρόβλεψη τιμών παραμέτρων του με τη χρήση δεδομένων έκφρασης γονιδίων	183
5.1	Εισαγωγή.....	183
5.2	Πρόβλεψη τιμών Ογκοπροσομοιωτή και εξατομίκευση προσομοίωσης-προβλέψεων	183
5.3	Διαθέσιμα δεδομένα και προετοιμασία τους.....	186
5.4	Ο Ογκοπροσομοιωτής Λευχαιμίας.....	187
5.5	Υπό-μονάδες μηχανικής μάθησης του Υβριδικού Ογκοπροσομοιωτή.....	191
5.5.1	Εκτίμηση παραμέτρων μηχανιστικού μοντέλου.....	192
5.5.2	Δόμηση, υλοποίηση και εκπαίδευση συστήματος πρόβλεψης αποτελέσματος θεραπείας.....	195
5.6	Αποτελέσματα.....	211
5.6.1	Αποτελέσματα εκτίμησης παραμέτρων.....	211
5.6.2	Αποτελέσματα πρόβλεψης κατηγορίας απόκρισης στην Πρεδνιζόνη	213
5.7	Συμπεράσματα, γενική αξιολόγηση και μελλοντικοί στόχοι	218
6.	Γενικές παρατηρήσεις, συμπεράσματα και συνολικοί μελλοντικοί στόχοι.....	223
	ΒΙΒΛΙΟΓΡΑΦΙΑ	227
	Δημοσιεύσεις στα πλαίσια της διατριβής.....	241

ΥΠΟΤΙΤΛΟΙ ΕΙΚΟΝΩΝ

Εικόνα 1.1 Οι φάσεις και τα σημεία ελέγχου του κυτταρικού κύκλου	25
Εικόνα 1.2 Κύτταρα του αίματος. Λευκά αιμοσφαίρια ανάμεσα σε ερυθροκύτταρα και αιμοπετάλια	30
Εικόνα 1.3 Κυτταρική διαφοροποίηση της ερυθράς σειράς. Χαρακτηρίζεται από προοδευτική σμίκρυνση του κυττάρου, μεταβολή του χρώματος του κυτταροπλάσματος, σμίκρυνση και πύκνωση του πυρήνα, εξαφάνιση των πυρηνίων και εκβολή του πυρήνα.	32
Εικόνα 1.4 Εξέλιξη της λεμφικής σειράς. Λεμφοβλάστης-Προλεμφοκύτταρα (επάνω). Λεμφοκύτταρα (κάτω).	34
Εικόνα 1.5 Λεμφοβλάστες στην ΟΛΛ σε αντιπαράθεση με μυελοβλάστες σε ΟΜΛ.	36
Εικόνα 1.6 Τα πρώτα στάδια θεραπείας με βάση την κλινική δοκιμή ALL-BFM 2000	44
Εικόνα 2.1 Βήματα διαδικασίας Εξόρυξης ή Ανακάλυψης Γνώσης από Βάσεις Δεδομένων (KDD) [39]	50
Εικόνα 2.2 Ο κύκλος της Συστημικής Βιολογίας [60]	64
Εικόνα 2.3 Παράδειγμα της γλώσσας Περιγραφής Διαδικασίας της SBGN	70
Εικόνα 2.4 Παράδειγμα της γλώσσας Σχέσης Οντοτήτων της SBGN	70
Εικόνα 2.5 Παράδειγμα της γλώσσας Ροής Δραστηριότητας της SBGN	70
Εικόνα 2.6 Χάρτης των κανόνων της SBGN	71
Εικόνα 2.7 Αναπαράσταση του Μοντέλου του Schlögl σε SBGN	71
Εικόνα 2.8 Ο Στοιχειομετρικός πίνακας του Μοντέλου του Schlögl	72
Εικόνα 2.9 Παράδειγμα ντετερμινιστικής προσομοίωσης	74
Εικόνα 2.10 Παράδειγμα Στοχαστικής Προσομοίωσης (1η εκτέλεση)	75
Εικόνα 2.11 Παράδειγμα Στοχαστικής Προσομοίωσης (2η εκτέλεση)	75
Εικόνα 2.12 Η φιλοσοφία του αλγορίθμου Tau-learning	85
Εικόνα 2.13 Η σχέση μεταξύ των διαφόρων μεθόδων προσομοίωσης [79]	87
Εικόνα 2.14 Η φιλοσοφία των υβριδικών αλγορίθμων προσομοίωσης [79]	88
Εικόνα 2.15 Παράδειγμα εκτέλεσης του αλγορίθμου της προσομοιωμένης απόπτωσης (simulated annealing)	97
Εικόνα 2.16 Η φιλοσοφία του αλγορίθμου του εξελικτικού προγραμματισμού (evolutionary programming) [94]	98
Εικόνα 2.17 Η φιλοσοφία του αλγορίθμου Genetic Algorithm [94]	99
Εικόνα 2.18 Αναπαράσταση της κεντρικής ιδέας του αλγορίθμου Particle Swarm [94]	99
Εικόνα 2.19 Τρισδιάστατο πλέγμα διακριτοποίησης που καλύπτει την περιοχή του όγκου (ΓΚ: Γεωμετρική Κυψελίδα).	107
Εικόνα 2.20 Κάθε Γεωμετρική Κυψελίδα (ΓΚ) μπορεί να περιέχει κύτταρα από πέντε κλάσεις κυττάρων. Οι φάσεις του κυτταρικού κύκλου (G1, S, G2, M) και η φάση ανάπαυσης (G0) αποτελούν υποκλάσεις των στελεχειαίων και LIMP κυττάρων.	107
Εικόνα 2.21 Γενικευμένο κυτταροκινητικό μοντέλο που μπορεί να προσαρμοστεί ανάλογα με το εκάστοτε θεραπευτικό πρωτόκολλο. Τα βέλη αναπαριστούν πιθανές μεταβάσεις μεταξύ κυτταρικών καταστάσεων και αντιστοιχούν σε παραμέτρους του μοντέλου. STEM CELLS: στελεχειαία καρκινικά κύτταρα. LIMP CELLS: κύτταρα περιορισμένου μιτωτικού δυναμικού. G1: Gap 1 φάση, S: Φάση σύνθεσης DNA, G2: Gap 2 φάση, M: μίτωση, G0: φάση ανάπαυσης.hit: κύτταρα που έχουν χτυπηθεί θανάσιμα από την ακτινοβολή ή τη χημειοθεραπεία. CCNS φάρμακα: Cell Cycle Non Specific φάρμακα.	109
Εικόνα 3.1 Κύρια μέρη του βιολογικού μονοπατιού της βιοχημικής ρύθμισης του G1/S «Σημείου Περιορισμού» (Restriction Point) στον φυσιολογικό (μη λευχαιμικό, Normal) κυτταρικό κύκλο.	121

Εικόνα 3.2 Τροποποιήσεις και απορυθμίσεις του G1/S «Σημείου Περιορισμού» (Restriction Point) στην BCP-ALL _____	126
Εικόνα 3.3 Αποτελέσματα προσομοίωσης του μοντέλου αναφοράς [139] για 900 <i>min</i> (15 <i>h</i>) _____	131
Εικόνα 3.4 Μέρος της δομής (νέες αντιδράσεις και οντότητες) του προτεινόμενου μοντέλου _____	139
Εικόνα 3.5 (Ημί-)ποσοτικά/ποιοτικά κριτήρια που ορίστηκαν για την κατάσταση και τα επίπεδα κεντρικών οντοτήτων του μοντέλου που χρησιμοποιήθηκαν για την προσαρμογή του προτεινόμενου μοντέλου μέσω βελτιστοποίησης _____	143
Εικόνα 3.6 Αποτελέσματα εκτέλεσης του προτεινόμενου μοντέλου μετά την εκτίμηση των παραμέτρων του και για την μέση περίπτωση _____	157
Εικόνα 3.7 Αποτελέσματα σάρωσης παραμέτρων _____	160
Εικόνα 3.8 Σάρωση παραμέτρων και αποτελέσματα προσομοίωσης για το σενάριο χορήγησης φαρμάκου. _____	163
Εικόνα 4.1 Δομή του μοντέλου που προτείνεται στην [196] (Μοντέλο 1) για την προσομοίωση της φαρμακοκινητικής του Prednisolone (Central Compartment: Κεντρικό διαμέρισμα, Peripheral Compartment: Περιφερικό διαμέρισμα, Oral administration: Από στόματος χορήγηση) _____	169
Εικόνα 4.2 Δομή του μοντέλου που προτείνεται στην [199] για την προσομοίωση της φαρμακοκινητικής του Prednisolone (Depot compartment: αποθηκευτικό διαμέρισμα, Dose: Δόση) _____	171
Εικόνα 4.3 Γραφική αναπαράσταση του μοντέλου για την από στόματος χορήγηση PREDNL όπως υλοποιήθηκε στο SimBiology [194] (Central: Κεντρικό διαμέρισμα, Peripheral: Περιφερικό διαμέρισμα, Dose: Δόση, Absorption: Απορρόφηση, Transport: Μεταφορά, Elimination: Αποβολή) _____	173
Εικόνα 4.4 Τα αποτελέσματα προσομοίωσης του φαρμακοκινητικού μοντέλου όπως δίνονται στην [196] για τον ασθενή που επιλέχθηκε προκειμένου να επιβεβαιωθεί η ορθή υλοποίηση (συμπαγής γραμμή: προσωποποιημένη πρόβλεψη, διακεκομμένη γραμμή: πληθυσμιακή πρόβλεψη, κύκλοι: δείγματα συγκέντρωσης μη δεσμευμένου PREDNL στο πλάσμα του ασθενούς). _____	174
Εικόνα 4.5 Αποτελέσματα προσομοίωσης της φαρμακοκινητικής του PREDNL για το πλήρες σχήμα θεραπείας που ακολουθείται στην [196] μέσω της νέας υλοποίησης του μοντέλου. _____	175
Εικόνα 4.6 Αποτελέσματα προσομοίωσης με την νέα υλοποίηση του μοντέλου για τις ώρες 144 έως 152 σε σύγκριση με τα δεδομένα που λήφθηκαν από την Εικόνα 4.4 _____	176
Εικόνα 4.7 Δομή του προτεινόμενου μοντέλου για την προσομοίωση της φαρμακοκινητικής του PRED (Central Compartment: Κεντρικό διαμέρισμα, Peripheral Compartment: Περιφερικό διαμέρισμα) _____	177
Εικόνα 4.8 Δομή του προτεινόμενου μοντέλου στο περιβάλλον SimBiology (Central Compartment: Κεντρικό διαμέρισμα, Peripheral Compartment: Περιφερικό διαμέρισμα, Depot compartment: Αποθηκευτικό διαμέρισμα) _____	178
Εικόνα 4.9 Δεδομένα που δίνονται στην [201] για τα επίπεδα (μέσες τιμές) της ελεύθερης Πρεδνιζολόνης μετά από χορήγηση 55mg Πρεδνιζόνης σε παιδιά με νεφροτικό σύνδρομο σε φάση ύφεσης. _____	179
Εικόνα 4.10 Αποτελέσματα εκτέλεσης του προτεινόμενου φαρμακοκινητικού μοντέλου, συγκρινόμενα με τα διαθέσιμα δεδομένα _____	181
Εικόνα 5.1 Η περιληπτική δομή του Υβριδικού Ογκοπροσομοιωτή Λευχαιμίας και τα επίπεδα διεπαφής με το χρήστη-προσκήνιο (front-end) και παρασκήνιο (back-end). _____	185
Εικόνα 5.2 Γενικό κυτταροκινητικό μοντέλο του Ογκοπροσομοιωτή Λευχαιμίας για τα διαμερίσματα του μυελού των οστών (bone marrow) και του περιφερικού αίματος (peripheral blood) [130]. STEM: στελεχειαίο κύτταρο/stem cell. LIMP: Limited Mitotic Potential cancer cell. Prolif: Πολλαπλασιαζόμενα κύτταρα/Proliferating cell. G0: Κύτταρα σε ύπωση/dormant cell. Chemo: Χημειοθεραπεία. HIT: Κύτταρα στα οποία έχει εκκινήσει η διαδικασία κυτταρικού θανάτου λόγω της θεραπείας _____	188

Εικόνα 5.3 Λεπτομερής δομή του Υβριδικού Ογκοπροσομοιωτή Λευχαιμίας (Hybrid ALL Oncosimulator) και βήματα για την αξιολόγηση του μέσω αντεπικύρωσης (cross-validation). __	196
Εικόνα 5.4 Παράδειγμα ταξινόμησης μέσω του k-NN για k=3 και για k=5 _____	199
Εικόνα 5.5 Παράδειγμα δέντρου παλινδρόμησης με κατηγορικές ανεξάρτητες μεταβλητές. Η εξαρτημένη μεταβλητή περιέχει τις ώρες (σε εβδομαδιαία βάση) δραστηριότητας στην ύπαιθρο ανάλογα με τις μεταβλητές του καιρού. _____	201
Εικόνα 5.6 Παράδειγμα δημιουργίας υποσυνόλων Bootstrap _____	202
Εικόνα 5.7 Δημιουργία και σύνθεση πολλαπλών μοντέλων μάθησης για διαφορετικά υποσύνολα bootstrap _____	203
Εικόνα 5.8 Τα βήματα εκτέλεσης του αλγορίθμου GSVA _____	206
Εικόνα 5.9 Διάγραμμα διασποράς (scatter plot) της μέσης εκτιμηθείσας CKR _{PRED} (6διαφορετικά σενάρια χρόνου διπλασιασμού) για τους ασθενείς που εντάχθηκαν στην μελέτη. Οι ασθενείς που κατατάσσονται στην κατηγορία κακής απόκρισης στην Πρεδνιζόνη τείνουν να εμφανίζουν χαμηλότερη τιμή χημειοευαισθησίας σε σχέση με τους ασθενείς που κατατάσσονται στην κατηγορία καλής απόκρισης στην Πρεδνιζόνη _____	212
Εικόνα 5.10 Ιστογράμματα πολλαπλών εκτελέσεων/προσομοιώσεων του Υβριδικού Ογκοπροσομοιωτή Λευχαιμίας για δύο ενδεικτικές περιπτώσεις ασθενών. Και οι δύο περιπτώσεις ταξινομήθηκαν σωστά από το σύστημα. Το κατώφλι Καλής/Κακής απόκρισης στην Πρεδνιζόνη (Good/Poor response threshold) σημειώνεται με κόκκινη γραμμή ενώ ο αριθμός των λεμφοβλαστών την 8 ^η ημέρα της θεραπείας του κάθε ασθενή με μωβ γραμμή. (a) Οι πλειοψηφία των προσομοιώσεων καταλήγουν σε πρόβλεψη για τον αριθμό των λεμφοβλαστών την 8 ^η ημέρα χαμηλότερο του κατωφλίου για ασθενή που κατατάσσεται στην κατηγορία Καλής απόκρισης. (b) Οι προσομοιώσεις επιστρέφουν συχνότερα τιμές υψηλότερες του κατωφλίου για ασθενή που κατατάσσεται στην κατηγορία Κακής απόκρισης. _____	217

ΥΠΕΡΤΙΤΛΟΙ ΠΙΝΑΚΩΝ

Πίνακας 2.1 Παράμετροι του Ογκοπροσομοιωτή που σχετίζονται με τους υπό μοντελοποίηση βιολογικούς μηχανισμούς	110
Πίνακας 3.1 Κριτήρια επιλογής του μοντέλου αναφοράς και αξιολόγηση υποψήφιων μοντέλων	128
Πίνακας 3.2 Βιοχημικές αντιδράσεις που μοντελοποιούνται στο προτεινόμενο μοντέλο και οι σχετικές κινητικές παράμετροι	132
Πίνακας 3.3 Αρχικά επίπεδα βιοχημικών οντοτήτων που απαρτίζουν το προτεινόμενο μοντέλο	137
Πίνακας 3.4 Μαθηματικός ορισμός των παραμέτρων-ρυθμών αντιδράσεων των νέο-εισαχθέντων αντιδράσεων	141
Πίνακας 3.5 (Ημί-)Ποσοτική πληροφορία εξαχθείσα από την βιβλιογραφία που χρησιμοποιήθηκε για την προσαρμογή του μοντέλου	143
Πίνακας 3.6 Ορισμός των σημαφόρων προσαρμογής	151
Πίνακας 3.7 Ορισμός τους κόστους προσαρμογής για τα επίπεδα της Κυκλίνης A και της Κυκλίνης D	151
Πίνακας 3.8 Ορισμός τους κόστους προσαρμογής για τα επίπεδα της pRb	152
Πίνακας 3.9 Παράμετροι του μοντέλου των οποίων οι τιμές εκτιμήθηκαν και διαστήματα τιμών αυτών	154
Πίνακας 3.10 Αποτελέσματα εκτίμησης παραμέτρων	156
Πίνακας 4.1 Μοντέλο παλινδρόμησης για τις παραμέτρου του Μοντέλου 1 όπως δίδεται στην [196]	170
Πίνακας 4.2 Εκτιμήσεις τιμών παραμέτρων για το φαρμακοκινητικό μοντέλο που προτείνεται στην [196]	170
Πίνακας 4.3 Τιμές παραμέτρων όπως εκτιμήθηκαν στην [199] για το Μοντέλο 2	172
Πίνακας 4.4 Χαρακτηριστικά ασθενούς που επιλέχθηκε από την εργασία [196] για την επιβεβαίωση της ορθής υλοποίησης του μοντέλου	174
Πίνακας 4.5 Οι φαρμακοκινητικές παράμετροι που χρησιμοποιήθηκαν για να προσομοιωθεί η περίπτωση του επιλεγόμενου από την [196] ασθενούς.	175
Πίνακας 4.6 Αρχικές εκτιμήσεις για τις τιμές των παραμέτρων του προτεινόμενου μοντέλου με βάση το μοντέλο παλινδρόμησης που δίνεται στην [196]	180
Πίνακας 4.7 Τελικές εκτιμήσεις παραμέτρων φαρμακοκινητικό μοντέλου	180
Πίνακας 5.1 Στατιστικές πληροφορίες των διαθέσιμων και αξιοποιήσιμων δεδομένων	187
Πίνακας 5.2 Παράμετροι εισόδου του Ογκοπροσομοιωτή Λευχαιμίας (μηχανιστικό μέρος) και εύρος τιμών τους (κατά την εκτίμηση παραμέτρων)	190
Πίνακας 5.3 Περιορισμοί που τέθηκαν για τις ιδιότητες αύξησης της ALL κατά την διαδικασία εκτίμησης παραμέτρων	194
Πίνακας 5.4 Παράμετροι της διαδικασίας προ-επεξεργασίας και των μοντέλων παλινδρόμησης που βελτιστοποιήθηκαν μέσω εσωτερικής αντεπικύρωσης (internal cross-validation)	208
Πίνακας 5.5 Αποτελέσματα εκτίμησης παραμέτρων για τις σχετικές με την ελεύθερη ανάπτυξη του όγκου του μηχανιστικού μέρους του Ογκοπροσομοιωτή Λευχαιμίας για 6 διαφορετικά σενάρια χρόνου διπλασιασμού (doubling times (Td)).	211
Πίνακας 5.6 Αποτελέσματα αξιολόγησης του Υβριδικού Ογκοπροσομοιωτή Λευχαιμίας, μέσω εξωτερικής αντεπικύρωσης (cross-validation, CV) στο να ταξινομεί ασθενείς μεταξύ των κατηγοριών Καλής και Κακής απόκρισης στην Πρεδνιζόνη	213

ΠΡΟΛΟΓΟΣ – ΕΥΧΑΡΙΣΤΙΕΣ

Η ολοκλήρωση της διδακτορικής διατριβής μου αποτελεί ιδιαίτερα σημαντική και συγκινητική στιγμή για την ερευνητική πλευρά αλλά συνολικά και για όλες τις πτυχές της ζωής μου. Όπως συνέβη με όλες τις υπόλοιπες βαθμίδες εκπαίδευσης, κατά την έναρξη τους, η ολοκλήρωση της φάνταζε μακρινή και κατά καιρούς ακόμα και άπιαστη. Σε αντίθεση όμως με τις προηγούμενες, η περίοδος του διδακτορικού αποτέλεσε ιδιαίτερη πρόκληση. Από το να λάβεις και να κατακτήσεις γνώση, καλείσαι πλέον να γεννήσεις και καινούργια. Και παρά τις κατά καιρούς δυσκολίες, τους προβληματισμούς και τις απογοητεύσεις, το τελικό αποτέλεσμα και το αίσθημα ότι μπόρεσες να συμβάλεις έστω και στο ελάχιστο στην επιστήμη σε αποζημιώνει. Ιδιαίτερα όμως, το ανθρωποκεντρικό και ευαίσθητο αντικείμενο της διατριβής αποτέλεσε σημαντικότερο κίνητρο. Εύχομαι προσπάθειες σαν και αυτή να καταφέρουν να συμβάλλουν έστω και ελάχιστα στην υποστήριξη του έργου των κλινικών και στη βελτίωση των εργαλείων που έχουν στα χέρια τους ώστε να συνεχίζουν να βοηθούν τους ασθενείς στην αντιμετώπιση ασθενειών όπως αυτή του καρκίνου.

Οφείλω ένα πολύ μεγάλο ευχαριστώ, αρχικά στους γονείς μου, για την διαχρονική (από γεννήσεως) υποστήριξη που μου προσέφεραν και για τα κίνητρα που μου έδωσαν ώστε να φτάσω μέχρι αυτόν τον τόσο σημαντικό στόχο. Ο πατέρας μου ειδικότερα, λόγω και της ιδιότητάς του, εκτός από τον ρόλο του γονέα, είχε ανέκαθεν και το ρόλο του δασκάλου και του προτύπου, όσο και αν δεν το παραδέχομαι ευθέως σε αυτόν.

Παράλληλα με την εκπόνηση της διατριβής μου, δημιούργησα την δικιά μου οικογένεια. Παντρεύτηκα την Τίνα και αποκτήσαμε τον μικρό Νικόλα. Έχοντας νοσηματοδοτήσει την ζωή μου, αποτέλεσαν και οι δύο το απόλυτο κίνητρο για να συνεχίσω τις προσπάθειες και με υποστήριξαν στο έπακρο με την αγάπη τους. Κάποιες στιγμές έμεινε στην άκρη το διδακτορικό για χάρη τους (όπως και για χάρη της πατρίδας την οποία υπηρέτησα όντας υποψήφιος διδάκτορας) και κάποιες φορές έπρεπε να τους αφήσω για χάρη του διδακτορικού. Τους ευχαριστώ και τους δύο, και θα τους ευχαριστώ για πάντα, για την κατανόησή τους.

Ιδιαίτερες ευχαριστίες θέλω να εκφράσω στους συνεπιβλέποντες της διατριβής μου, τον καθηγητή ΣΗΜΜΥ. Ανδρέα-Γεώργιο Σταφυλοπάτη και τον Ερευνητή Α' ΕΠΙΣΕΥ και επισκέπτη καθηγητή ΣΗΜΜΥ Γιώργο Σταματάκο. Ο Γιώργος Σταματάκος, όντας επικεφαλής της ομάδας για την In Silico Ογκολογία και την In Silico Ιατρική του ΕΠΙΣΕΥ-ΕΜΠ (της οποίας είμαι μέλος), υπήρξε ο βασικός καθοδηγητής στις προσπάθειες μου. Με

πίστη, κατανόηση, επιμονή και υπομονή μου μετέδωσε σημαντικές γνώσεις και εμπειρία ώστε να γίνει αυτό το παραπάνω βήμα στις προσπάθειες της ομάδας ενώ μου έδωσε αρκετή ελευθερία ώστε να ερευνήσω τομείς οι οποίοι με προσέλκυαν. Επίσης η συμμετοχή μου στα ερευνητικά προγράμματα της ομάδας ήταν κομβική για την υποστήριξη της ίδιας της έρευνας αλλά και για την εξασφάλιση των προς το ζην. Ευχαριστώ πολύ τον καθ. Στέφανο Κόλλια ο οποίος επέβλεψε την διατριβή μου κατά τα πρώτα χρόνια της εκπόνησής της.

Ευχαριστώ θερμά την Δρ. Ελένη Κολοκοτρώνη, μέλος της ομάδας για την In Silico Ογκολογία και την In Silico Ιατρική με την οποία η συνεργασία ήταν ιδιαίτερα στενή και μεγάλο μέρος της δικιάς της έρευνας τροφοδότησε το αντικείμενο της παρούσας εργασίας. Τίποτα από όσα παρουσιάζονται σε αυτή ωστόσο δεν θα είχαν πραγματοποιηθεί χωρίς την φιλία, την συνεργασία και την υποστήριξη των υπολοίπων μελών της ομάδας και συγκεκριμένα των Δρ. Δημήτρα Διονυσίου, Δρ. Ελένης Γεωργιάδη, Δρ. Σταυρούλας Γιατίλη, Κατερίνας Αργύρη, Χρήστου Κηρούδη, Μάρκου Αντωνόπουλου, Φαίης Μισιχρόνη, Νίκου Τούσερτ και Νίκου Χριστοδούλου. Μεταξύ μας αναπτύχθηκαν ισχυροί δεσμοί που έδωσαν σημαντική ώθηση στις καθημερινές προσπάθειες για την περάτωση της διατριβής.

Θέλω επίσης να ευχαριστήσω ιδιαίτερα όλα τα μέλη της ομάδας I-Sense του ΕΠΙΣΕΥ-ΕΜΠ της οποίας πλέον αποτελώ μέλος, για την υποστήριξη (και την ανοχή) τους στα τελευταία βήματα ολοκλήρωσης των πρακτικών ζητημάτων συγγραφής και παρουσίασης της διατριβής.

Τέλος, ευχαριστώ θερμά τα υπόλοιπα μέλη της 7μελούς επιτροπής, τον καθ. Δημήτρη Κουτσούρη, τον καθ. Ηλία Μανωλάκο, τον επ. καθ. Λεωνίδα Αλεξόπουλο, την επ. καθ. Μαρία Μοσχόβη και την καθ. Βασιλική Παπαευαγγέλου για την ιδιαίτερα τιμητική αποδοχή της συμμετοχής τους στην εξεταστική επιτροπή. Οφείλω να σημειώσω πως ο καθ. Μανωλάκος ήταν αυτός που με εισήγαγε στην Βιοπληροφορική και στην Συστημική Βιολογία αλλά και μου μετέδωσε σημαντικές αρετές όσον αφορά την μεθοδολογία της έρευνας και έτσι θα κατέχει για πάντα ιδιαίτερη θέση στην «ερευνητική μου καρδιά».

1. Καρκίνος, Βασικοί βιολογικοί μηχανισμοί, Αιμοποίηση και Οξεία Λεμφοβλαστική Λευχαιμία

1.1 Εισαγωγή

Στο παρόν κεφάλαιο γίνεται μία συνοπτική παρουσίαση των χαρακτηριστικών της πολυσύνθετης και πολυποίκιλης νόσου του Καρκίνου και των βασικών βιολογικών μηχανισμών των οποίων η εποπτεία απαιτείται για την καλύτερη κατανόηση αυτής. Ιδιαίτερη έμφαση δίνεται στην περίπτωση της Οξείας Λεμφοβλαστικής Λευχαιμίας (ΟΛΛ, με εστίαση στην παιδιατρική περίπτωση) η οποία αποτέλεσε βασικό πεδίο εφαρμογής των υπό μελέτη και υπό ανάπτυξη μεθόδων και της θεραπευτικής αντιμετώπισης της. Στις ακόλουθες ενότητες παρουσιάζονται τα κυριότερα χαρακτηριστικά του καρκίνου ενώ έμφαση δίνεται στην κυτταρική κινητική και στον κυτταρικό θάνατο. Εν συνεχεία δίνεται μία σύντομη εισαγωγή στα χαρακτηριστικά του αιμοποιητικού συστήματος (ασθένεια του οποίου αποτελεί η Οξεία Λεμφοβλαστική Λευχαιμία) ενώ ακολουθούν επιδημιολογικά στοιχεία και οι ιστολογικοί τύποι της ΟΛΛ. Για την παρουσίαση αυτή βάση αποτέλεσε η [1]. Σε ακόλουθη ενότητα παρουσιάζονται οι βασικές προσεγγίσεις θεραπείας της ΟΛΛ, ενώ ιδιαίτερη εστίαση δίνεται στο θεραπευτικό πρωτόκολλο που ακολουθείται στην κλινική δοκιμή ALL-BFM 2000 από την οποία προήλθαν τα δεδομένα που χρησιμοποιήθηκαν στις μελέτες του Κεφαλαίου 0.

1.2 Ο Καρκίνος

Ο καρκίνος είναι η δεύτερη σε συχνότητα αιτία μη φυσιολογικού θανάτου στις αναπτυγμένες χώρες μετά τη στεφανιαία νόσο [2]. Ο όρος «καρκίνος» αναφέρεται σε μεγάλο αριθμό τύπων καρκινικών όγκων (σε περισσότερους από 100), οι οποίοι εμφανίζουν μεταξύ τους σημαντική ετερογένεια στα ιδιαίτερα βιολογικά χαρακτηριστικά [3]. Η γέννηση αλλά και η συμπεριφορά των διάφορων τύπων καρκίνου επηρεάζεται από μεγάλο αριθμό διαφορετικών παραγόντων, μεταξύ άλλων η ηλικία, το φύλο, η γενετική προδιάθεση και η έκθεση σε στοιχεία του περιβάλλοντος [4]. Ανεξάρτητα της σημαντικής αυτής ετερογένειας, είναι δυνατόν να εντοπιστούν κάποια βασικά γνωρίσματα της νόσου του καρκίνου.

Ο πολλαπλασιασμός των κυττάρων υπό φυσιολογικές συνθήκες υπόκειται σε διαρκή έλεγχο, ώστε να ανταποκρίνεται στις εκάστοτε ανάγκες του ανθρώπινου οργανισμού [5]. Για παράδειγμα, το ποσοστό του κυτταρικού πολλαπλασιασμού υπερβαίνει εκείνο της

κυτταρικής διαφοροποίησης και του κυτταρικού θανάτου κατά το εμβρυακό ή βρεφικό στάδιο, κάτι που επιτρέπει την ραγδαία ανάπτυξη του εμβρύου ή του βρέφους. Στους ιστούς των ενήλικων ανθρώπων, από την άλλη, η αναπαραγωγή των κυττάρων βρίσκεται σε ισορροπία με την απώλεια των διαφοροποιημένων ώριμων κυττάρων (πάντα υπό φυσιολογικές συνθήκες).

Οι μηχανισμοί ελέγχου του κυτταρικού πολλαπλασιασμού σε φυσιολογικές συνθήκες αναλαμβάνουν την επιδιόρθωση πιθανών σφαλμάτων που προέκυψαν κατά την προετοιμασία της κυτταρικής διαίρεσης ή το θάνατο του κυττάρου σε περίπτωση που δεν είναι εφικτή η επιδιόρθωση αυτή. Σε περίπτωση όμως που οι μηχανισμοί αυτοί αποτύχουν, η διαδικασία ανάπτυξης και διαίρεσης συγκεκριμένων κυττάρων, που δεν είναι απαραίτητα στον οργανισμό. Αν ένα τέτοιο κύτταρο καταφέρει να επιζήσει και να μεταβιβάσει στους απογόνους του την ικανότητα του ανεξέλεγκτου πολλαπλασιασμού, δημιουργούνται σημαντικές προϋποθέσεις δημιουργίας όγκων. Τα κύτταρα αυτά ονομάζονται στελεχιαία ή βλαστικά καρκινικά κύτταρα (cancer stem cells) και αποτελούν έναν υποπληθυσμό των κυττάρων ενός όγκου. Τα στελεχιαία καρκινικά κύτταρα μπορούν να ανιχνευθούν, *ex vivo*, μέσω της διεξαγωγής εργαστηριακών διαδικασιών που εξετάζουν την ικανότητα των κυττάρων να σχηματίζουν πολυπληθείς αποικίες απογόνων (colonies) σε ελεγχόμενο περιβάλλον ανάπτυξης. Έτσι, ονομάζονται συχνά και κλωνογόνα καρκινικά κύτταρα (clonogenic cancer cells) [6].

Όπως είναι ευρέως γνωστό και τεκμηριωμένο στη βιβλιογραφία, διάφορες χημικές ουσίες όπως για παράδειγμα συστατικά στον καπνό των τσιγάρων έχουν συνδεθεί άμεσα με την εμφάνιση συγκεκριμένων τύπων καρκίνου. Ωστόσο, υπάρχουν ακόμη και μορφές ιών που συσχετίζονται με την ανάπτυξη καρκινικών όγκων. Για παράδειγμα ο ιός HTLV-I συνδέεται με την εμφάνιση της T-λευχαιμίας ενηλίκων [4] ενώ ο ιός Epstein-Barr ενοχοποιείται για το λέμφωμα Burkitt και το καρκίνωμα του ρινοφάρυγγος [2].

Σημαντικό ρόλο στο φαινόμενο της καρκινογένεσης διαδραματίζουν δύο κατηγορίες γονιδίων του κυττάρου: τα πρώτο-ογκογονίδια (π.χ. τα bcl-2 και c-myc), που υποστηρίζουν γενικά τον πολλαπλασιασμό των κυττάρων, και τα ογκοκατασταλτικά γονίδια (π.χ. τα p53, BRCA1 και BRCA2), που κωδικοποιούν πρωτεΐνες που δύνανται να αναστείλουν τον κυτταρικό κύκλο ή να οδηγήσουν το κύτταρο σε κυτταρικό θάνατο (συνήθως σε απόπτωση).

Σε περίπτωση που εμφανιστεί δυσλειτουργία στα πρωτο-ογκογονίδια (π.χ. λόγω κάποιας μετάλλαξης), δύναται να ευνοηθεί ο ανεξέλεγκτος πολλαπλασιασμός του κυττάρου ή άλλα χαρακτηριστικά των κακοήθων όγκων, καθώς κάποια σχετική πρωτεΐνη που κωδικοποιούν (π.χ. ένας αυξητικός παράγοντας) είτε παράγεται σε υπερβολική ποσότητα είτε παραμένει ενεργή για υπερβολικό χρονικό διάστημα.

Αντίθετα, τα ογκοκατασταλτικά γονίδια μπορεί να συμβάλουν στην εμφάνιση καρκίνου, εάν απενεργοποιηθούν. Η επακόλουθη απώλεια ή δυσλειτουργία της αντίστοιχης

πρωτεΐνης συσχετίζεται με την απώλεια της προστασίας των κυττάρων από ανεξέλεγκτο πολλαπλασιασμό [2].

Αποκλίσεις από την φυσιολογική λειτουργία άλλων γονιδίων είναι δυνατό να συνεισφέρουν στην ανάπτυξη καρκινικών όγκων, επιτρέποντας, για παράδειγμα, στα καρκινικά κύτταρα να διαφεύγουν του ανοσοποιητικού συστήματος, να αποκολλώνται από την κυρίως καρκινική μάζα, να διεισδύουν στους παρακείμενους φυσιολογικούς ιστούς, να επιβιώνουν στην κυκλοφορία του αίματος ή της λέμφου, να προσκολλώνται σε αιμοφόρα αγγεία ή και να αναπτύσσουν νέους όγκους σε διαφορετικές περιοχές του ανθρώπινου σώματος [2], [5].

Επίσης, ιδιαίτερα σημαντικό ρόλο στην ανάπτυξη ενός καρκινικού όγκου διαδραματίζει το φαινόμενο της νεοαγγείωσης, της διαδικασίας δηλαδή σχηματισμού νέων αγγείων που θα υποστηρίξουν τις αυξανόμενες ανάγκες τροφοδοσίας του όγκου [7], [8].

Όπως προαναφέρθηκε το φαινόμενο της απόπτωσης παίζει καθοριστικό ρόλο στην ανάπτυξη διαφόρων τύπων καρκινικών όγκων. Η απόπτωση είναι μία μορφή προγραμματισμένου κυτταρικού θανάτου, που ενεργοποιείται στα φυσιολογικά κύτταρα, όταν έχουν υποστεί ανεπανόρθωτες βλάβες ή έχουν καταστεί επικίνδυνα λόγω μεταλλαγών. Με το μηχανισμό αυτό αποφεύγεται η μεταβίβαση μη αναστρέψιμων λαθών ή επικίνδυνων μεταλλαγών στα κύτταρα απογόνους. Έτσι, οι όγκοι που αναπτύσσονται απαρτίζονται από κύτταρα που καταφέρνουν να διαφύγουν και από τον έλεγχο του μηχανισμού της απόπτωσης. Η δράση ποικίλων γονιδίων (π.χ. p53, bcl-2) έχει συνδεθεί με την απόπτωση [3], [5], [6].

Ένας επιπλέον μηχανισμός περιορισμού του πιθανού ανεξέλεγκτου πολλαπλασιασμού συνίσταται στην «απομνημόνευση» του αριθμού των διαιρέσεων που έχει πραγματοποιήσει ένα κύτταρο. Τα υγιή κύτταρα (π.χ. χωρίς μεταλλαγές στη λειτουργικότητα των γονιδίων p53 και retinoblastoma protein-pRb), μετά από έναν αριθμό διαιρέσεων περνούν σε κατάσταση γήρανσης (senescence) και σταματούν να πολλαπλασιάζονται [7]. Υπάρχουν περιπτώσεις ωστόσο όπου κάποια κύτταρα διαφεύγουν της γήρανσης συνεχίζουν να διαιρούνται και καθίστανται εν δυνάμει «αθάνατα», με πρακτικά άπειρο δυναμικό πολλαπλασιασμού.

Η απαρίθμηση των πολλαπλασιασμών (μιτωτικών κύκλων) που έχει εκτελέσει ένα κύτταρο στηρίζεται τελομερή [9], εξαμερείς επαναλαμβανόμενες αλληλουχίες στα άκρα των χρωμοσωμάτων των οποίων το μήκος μειώνεται κάθε φορά που τα χρωμοσώματα διπλασιάζονται. Όταν φθάσει σε ένα οριακό επίπεδο, σηματοδοτείται η γήρανση του κυττάρου. Σε περίπτωση που παρακαμφθεί ο μηχανισμός της γήρανσης, η περαιτέρω μείωση του μήκους των τελομερών οδηγεί σε τελομερική κρίση. Σε αυτή την περίπτωση τα χρωμοσώματα του κυττάρου αποσταθεροποιούνται.

Στα καρκινικά κύτταρα, ο συγκεκριμένος μηχανισμός θα μπορούσε να αποτρέψει τον ανεξέλεγκτο πολλαπλασιασμό. Ωστόσο μελέτες έχουν δείξει ότι τα συγκεκριμένα κύτταρα

διαθέτουν έναν μηχανισμό ανασύνθεσης των τμημάτων των τελομερών που χάνονται σε κάθε κυτταρικό κύκλο μέσω της ενεργοποίησης μιας ειδικής DNA πολυμεράσης, της τελομεράσης (telomerase) η οποία είναι ανενεργή εν γένει στους φυσιολογικούς ιστούς [10]. Κατά συνέπεια, το μήκος των τελομερών παραμένει ικανό να υποστηρίξει την επέκταση της πολλαπλασιαστικής ικανότητας των καρκινικών κυττάρων και με αυτόν τον τρόπο οι καρκινικοί όγκοι μπορούν να αποκτήσουν μεγάλες διαστάσεις.

Συνοψίζοντας, ο καρκίνος μπορεί να οριστεί ως μία θεμελιώδης διαταραχή στη συμπεριφορά των κυττάρων, με κυρίαρχη χαρακτηριστικά τον ανεξέλεγκτο πολλαπλασιασμό και την αντίσταση στην απόπτωση. Παρότι έχει σημειωθεί σημαντική πρόοδος στην κατανόηση των βιολογικών μηχανισμών που σχετίζονται με την ανάπτυξη και τη συμπεριφορά των καρκινικών όγκων, υπάρχουν ωστόσο πολλά σημεία που παραμένουν ακόμη ασαφή. Η αποσαφήνισή τους αποτελεί μία από τις μεγαλύτερες επιστημονικές προκλήσεις με τους χώρους της Βιοπληροφορικής και της In Silico Ιατρικής/Ογκολογίας (παρουσιάζονται λεπτομερώς στο Κεφάλαιο 2) να αναμένεται να συμβάλουν έτι περαιτέρω.

1.3 Βασικοί βιολογικοί μηχανισμοί σχετικοί με τον καρκίνο

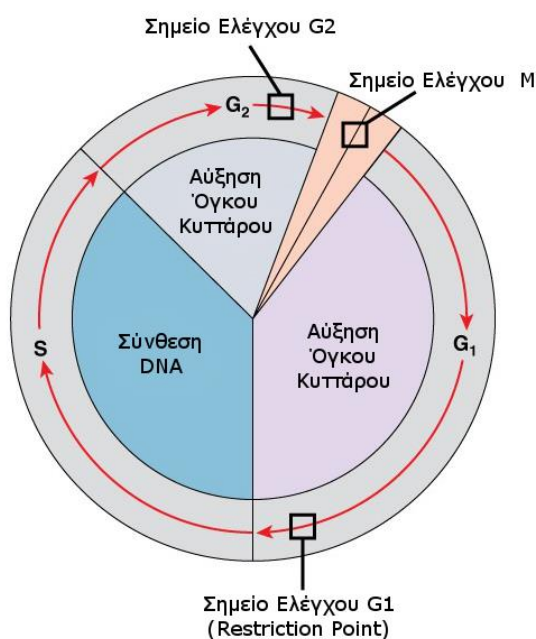
Στην παρούσα υπό-ενότητα παρουσιάζονται τα χαρακτηριστικά δύο βασικών βιολογικών μηχανισμών που εμπλέκονται άμεσα στις σχετικές με την ανάπτυξη αλλά και τη θεραπεία του καρκίνου, αυτών του κυτταρικού κύκλου και του κυτταρικού θανάτου. Θα πρέπει να σημειωθεί ότι στο παρόν κεφάλαιο η εστίαση δίνεται στα χαρακτηριστικά αυτών στο κυτταρικό επίπεδο ενώ λεπτομέρειες που άπτονται του μοριακού επιπέδου παρουσιάζονται λεπτομερώς σε ακόλουθα κεφάλαια της διατριβής.

1.3.1 Κυτταρική Κινητική και Κυτταρικός Κύκλος

Η κινητική των κυττάρων, δηλαδή η αύξηση και ο πολλαπλασιασμός τους, διέπονται από συγκεκριμένες βασικές αρχές οι οποίες επιχειρούνται να παρουσιαστούν στην παρούσα υπο-ενότητα.

Τα υγιή κύτταρα εμφανίζονται διάφορες καταστάσεις διαφοροποίησης, δηλαδή διαφορετικές μορφολογικές και βιοχημικές καταστάσεις ανάλογες των λειτουργιών που καλούνται να επιτελέσουν. Η ικανότητα πολλαπλασιασμού ενός κυττάρου είναι στενά συνδεδεμένη με το βαθμό διαφοροποίησής του. Συγκεκριμένα, όσο περισσότερο διαφοροποιημένα είναι τα κύτταρα ενός ιστού, τόσο λιγότερο τείνουν να βρίσκονται στον ενεργό κύκλο αναπαραγωγής [8] και βρίσκονται συχνότερα σε φάση ανάπαυσης.

Διάφορα ερεθίσματα που σχετίζονται με τις εκάστοτε ανάγκες του οργανισμού λειτουργούν ως ρυθμιστικοί παράγοντες της αναλογίας μεταξύ πολλαπλασιασμού και διαφοροποίησης. Όταν υπάρξει απώλεια κυττάρων λόγω εσωτερικών ή εξωτερικών αιτιών, ενεργοποιούνται τα κατάλληλα σήματα προώθησης του κυτταρικού πολλαπλασιασμού.. Το ερέθισμα για την επανένταξη των αναπαυόμενων κυττάρων στον αναπαραγωγικό κύκλο που συνήθως αναφέρεται ως Κυτταρικός Κύκλος μπορεί να είναι κάποια τοπική μεταβολή στην πυκνότητα των κυττάρων ή η παρουσία μιτογόνων (mitogens) ή αυξητικών παραγόντων (growth factors). Ένα σχηματικό διάγραμμα του κυτταρικού κύκλου δίνεται στην Εικόνα 1.1.



Εικόνα 1.1 Οι φάσεις και τα σημεία ελέγχου του κυτταρικού κύκλου

Ο κυτταρικός κύκλος, όπως συμφωνείται στην βιβλιογραφία, στα ευκαρυωτικά κύτταρα αποτελείται από τις εξής φάσεις: G₁ (Gap 1, ενδιάμεση φάση 1), S (Synthesis, σύνθεση DNA), G₂ (Gap 2, ενδιάμεση φάση 2) και M (Mitosis, μίτωση: διαίρεση κυττάρου σε δύο νέα κύτταρα) [5], [6], [11]. Τα κύτταρα που βρίσκονται στις φάσεις G₁, S, G₂, M ονομάζονται πολλαπλασιαζόμενα (proliferating). Οι φάσεις G₁ και G₂ είναι προπαρασκευαστικές φάσεις, κατά τις οποίες πραγματοποιείται η σύνθεση των απαραίτητων κυτταρικών συστατικών στοιχείων που θα χρειαστούν στις επόμενες φάσεις S και M αντίστοιχα. Συγκεκριμένα κατά τη φάση G₁ το κύτταρο αυξάνει σε μέγεθος ενώ γίνονται οι απαραίτητοι έλεγχοι προκειμένου να διαπιστωθεί ότι το κύτταρο είναι ικανό να συνθέσει το απαραίτητο για την κυτταρική διαίρεση αντίγραφο του DNA. Ο έλεγχος αυτός γίνεται κατά το σημείο ελέγχου (checkpoint) της G₁ ή «Σημείο Περιορισμού» (Restriction Point, RP). Λεπτομερής παρουσίαση του συγκεκριμένου σημείου ελέγχου γίνεται στο

Κεφάλαιο 3 του οποίου αποτελεί κεντρικό αντικείμενο μελέτης. Ακολούθως στη φάση S πραγματοποιείται ο διπλασιασμός του DNA. Κατά τη φάση G2, το κύτταρο συνεχίζει να μεγαλώνει σε μέγεθος ενώ κατά το σημείο ελέγχου G2 (G2 checkpoint) μηχανισμοί ελέγχου επιβεβαιώνουν ότι το κύτταρο μπορεί να διαιρεθεί. Τέλος στην M φάση σταματάει η μεγέθυνση του κυττάρου και πλέον η κυτταρική ενέργεια εστιάζεται στην διαδικασία διαίρεσης του κυττάρου σε δύο θυγατρικά κύτταρα.

Η φάση G0 (dormant, resting ή quiescent phase) είναι η φάση ανάπαυσης ή ηρεμίας, στην οποία μπορεί να περιέλθουν τα φυσιολογικά κύτταρα είτε λόγω διαφοροποίησης, ώστε να εκτελέσουν κάποια εξειδικευμένη λειτουργία, είτε λόγω έλλειψης ανάγκης για πολλαπλασιασμό. Στην περίπτωση αυτή, βρίσκονται εκτός του ενεργού κύκλου αναπαραγωγής, είναι όμως λειτουργικά ενεργά και υπάρχει η δυνατότητα να επανενταχθούν στον κυτταρικό κύκλο, αν οι ανάγκες του οργανισμού το απαιτήσουν. Τα πλήρως διαφοροποιημένα (terminally differentiated, ή στείρα) κύτταρα αποτελούν μία ιδιαίτερη κατηγορία κυττάρων, που έχουν χάσει μόνιμα την ικανότητα αναπαραγωγής τους. Συχνά στη βιβλιογραφία τα στείρα κύτταρα εντάσσονται στη φάση G0.

Σε αντίθεση με τα υγιή κύτταρα, τα καρκινικά κύτταρα συνήθως παρουσιάζουν ανωμαλίες στη διαφοροποίησή τους. Υπάρχει το ενδεχόμενο η διαδικασία διαφοροποίησής τους να πραγματοποιηθεί μερικώς μόνο ή να οδηγήσει στη δημιουργία ακατάλληλα ή εσφαλμένα διαφοροποιημένων κυττάρων. Ο ακριβής μηχανισμός ρύθμισης των καταστάσεων πολλαπλασιασμού και διαφοροποίησης διαταράσσεται, και τα καρκινικά κύτταρα αναπαράγονται σχεδόν ανεξέλεγκτα.

Τα νεοπλασματικά κύτταρα σε έναν καρκινικό όγκο που βρίσκονται στον ενεργό κύκλο αναπαραγωγής ανήκουν στην κατηγορία (ή διαμέρισμα) των πολλαπλασιαζόμενων κυττάρων η οποία συνήθως αναφέρεται ως κλάσμα ανάπτυξης (Growth Fraction, GF). Όλα τα νέα καρκινικά κύτταρα προκύπτουν από τον πολλαπλασιασμό των κυττάρων αυτής της κατηγορίας.

Τα αναπαυόμενα G0 κύτταρα και τα πλήρως διαφοροποιημένα κύτταρα αποτελούν δύο ακόμη κατηγορίες. Η βασική διαφορά μεταξύ τους είναι ότι, σε αντίθεση με τα πλήρως διαφοροποιημένα κύτταρα, τα G0 κύτταρα είναι ικανά να επανενταχθούν στον ενεργό κύκλο αναπαραγωγής. Στα G0 κύτταρα συγκαταλέγονται και κύτταρα που έχουν εισέλθει στη φάση ανάπαυσης όχι λόγω διαφοροποίησης, αλλά λόγω συνθηκών που δεν ευνοούν την παραμονή στον ενεργό κύκλο αναπαραγωγής όπως για παράδειγμα η έλλειψη οξυγόνου ή γλυκόζης.

Τέλος, υπάρχει και η κατηγορία που απαρτίζεται από τα νεκρά κύτταρα. Η έκταση των νεκρωτικών περιοχών (που απαντάται στους συμπαγείς όγκους) ποικίλει ανά περίπτωση.

Μεταξύ των ανωτέρω αναφερόμενων κατηγοριών πραγματοποιείται διαρκής μετακίνηση κυττάρων. Η μετατόπιση κυττάρων από το διαμέρισμα των αναπαυόμενων κυττάρων στο διαμέρισμα των πολλαπλασιαζόμενων κυττάρων ονομάζεται και στρατολόγηση (recruitment), ενώ η αντίστροφη διαδικασία παρατηρείται στην περίπτωση ανεπαρκούς τροφοδοσίας με οξυγόνο ή θρεπτικά συστατικά. Επίσης κάποια κύτταρα μπορεί να εισέλθουν στο διαμέρισμα των διαφοροποιημένων κυττάρων λόγω φυσιολογικών μηχανισμών διαφοροποίησης. Τέλος, υπάρχουν κύτταρα που εγκαταλείπουν την περιοχή του καρκινικού όγκου, είτε λόγω κυτταρικού θανάτου είτε επειδή αποκολλώνται από την κυρίως καρκινική μάζα και μεταναστεύουν σε άλλες περιοχές.

Τα στελεχειαία καρκινικά κύτταρα (cancer stem cells) αποτελούν έναν υποπληθυσμό των κυττάρων ενός καρκινικού όγκου που χαρακτηρίζεται από την ικανότητα πραγματοποίησης πολλαπλών κύκλων μιτώσεων. Συνήθως αποτελούν ένα μικρό ποσοστό του συνόλου των κυττάρων ενός όγκου.

Τέλος, ο ρυθμός ανάπτυξης ενός καρκινικού όγκου εξαρτάται από τις ακόλουθες βασικές παραμέτρους [6], [7]:

- τη διάρκεια του κυτταρικού κύκλου, TC (ο χρόνος που απαιτείται για την ολοκλήρωση μίας κυτταρικής διαίρεσης και περιλαμβάνει τις φάσεις G1, S, G2, M),
- το κλάσμα ανάπτυξης (Growth Fraction, GF) και τον παράγοντα των κυτταρικών απωλειών (Cell Loss Factor, CLF) ή το ρυθμό των κυτταρικών απωλειών (Cell Loss Rate, CLR).

Οι καρκινικοί όγκοι αναπτύσσονται ταχύτερα όταν έχουν μικρή διάρκεια κυτταρικού κύκλου, υψηλό ποσοστό πολλαπλασιαζόμενων κυττάρων και μικρές κυτταρικές απώλειες. Οι παράμετροι αυτές εξαρτώνται από τα ιστοπαθολογικά χαρακτηριστικά του όγκου, τα οποία με τη σειρά τους καθορίζονται κατά κύριο λόγο από τα γενετικά χαρακτηριστικά των καρκινικών κυττάρων (τις μεταβολές που προκλήθηκαν στο κύτταρο κατά την καρκινογένεση) και την ιστολογική προέλευσή τους.

1.3.2 Κυτταρικός Θάνατος

Οι μορφές κυτταρικού θανάτου που έχουν μελετηθεί εκτενέστερα στη βιβλιογραφία είναι αυτές της νέκρωσης της απόπτωσης.

Η απόπτωση αποτελεί μια μορφή εν γένει ελεγχόμενου προγραμματισμένου κυτταρικού θανάτου. Ενίοτε αναφέρεται και ως «αυτοκτονία» του κυττάρου. Έχουν μελετηθεί και επαρκώς χαρακτηριστεί δύο βασικοί οδοί για την ενεργοποίηση της απόπτωσης, η ενδογενής (intrinsic) κατά την οποία η έναρξη του βιολογικού μονοπατιού πραγματοποιείται από γεγονότα εντός του κυττάρου και η εξωγενής (extrinsic) κατά την οποία εξωγενή βιοχημικά σήματα (τα οποία λαμβάνονται από υποδοχείς του κυττάρου) εκκινούν την διαδικασία. Χαρακτηρίζεται συνήθως από συμπύκνωση και αποδόμηση της πυρηνικής χρωματίνης σε μικρά τμήματα, χαρακτηριστικού μεγέθους. Το κύτταρο συρρικνώνεται, ενώ τα κυτταρικά οργανίδια παραμένουν συνήθως άθικτα. Είναι πιθανό να δημιουργηθούν τα λεγόμενα ψευδοπόδια (pseudopodes-membrane budding), που συνήθως αποκολλώνται από το κύτταρο, το οποίο αποικοδομείται έτσι στα λεγόμενα αποπτωτικά σωματίδια (apoptotic bodies). Αυτά καταστρέφονται από φαγοκύτταρα (μακροφάγα ή άλλα γειτονικά κύτταρα [12]), χωρίς να προκαλείται εκτεταμένη φλεγμονώδης αντίδραση, γιατί η κυτταροπλασματική μεμβράνη διαφυλάσσεται και έτσι το ενδοκυτταρικό περιεχόμενο δε διαχέεται στον περιβάλλοντα ιστό. Σημαντικό ρόλο στο φαινόμενο της απόπτωσης διαδραματίζουν οι κασπάσες (caspases), μια οικογένεια πρωτεολυτικών ενζύμων της κυστεΐνης [13], ενώ έχει βρεθεί ότι η παρουσία τριφωσφορικής αδενοσίνης (adenosine triphosphate, ATP), δηλαδή διαθέσιμης ενέργειας, είναι καθοριστικής σημασίας [14]. Ο αποπτωτικός θάνατος επομένως είναι μία διαδικασία ελεγχόμενης καταστροφής του κυττάρου. Όπως προαναφέρθηκε μπορεί να προκύψει είτε ενδογενώς είτε ως αποτέλεσμα ποικίλων ερεθισμάτων που εξαρτώνται και από τον τύπο του κυττάρου [13], [15].

Αντίθετα, η νέκρωση συχνά αναφέρεται ως μία μορφή βίαιου, ανεξέλεγκτου κυτταρικού θανάτου που αποτελεί αποτέλεσμα κυτταρικού τραυματισμού (cell injury). Κατά τη νέκρωση προκαλείται εκτενής καταστροφή των μιτοχονδρίων και ραγδαία ενεργειακή πτώση, που οδηγεί σε διαταραχή της ενδοκυτταρικής ιοντικής ομοιόστασης, με αποτέλεσμα το σχηματισμό φυσαλίδων (membrane blebbing) και την καταστροφή της κυτταρικής μεμβράνης. Τα κύτταρα και τα κυτταρικά οργανίδια διογκώνονται και λύνονται απελευθερώνοντας ενδοκυτταρικό υλικό που πυροδοτεί φλεγμονώδεις αντιδράσεις.

Από τις σύγχρονες έρευνες προκύπτουν πολλαπλές ενδείξεις ότι η απόπτωση και η νέκρωση αποτελούν στενά συνδεδεμένα φαινόμενα, καθώς όχι μόνο προκαλούνται συχνά από τους ίδιους παράγοντες, αλλά πολλές φορές χαρακτηρίζονται και από κοινά τμήματα στους μηχανισμούς σηματοδότησης (signaling pathways) [13]. Σε in vitro πειράματα, έχει βρεθεί ότι κάποια ερεθίσματα που θεωρούνται κατά κανόνα αποπτωτικά, επιφέρουν την νέκρωση, όταν είναι ιδιαίτερα ισχυρά [16]–[18]. Επιπλέον, παράγοντες όπως η έλλειψη

καλής οξυγόνωσης, η ακτινοβολία, διάφοροι ιοί κ.α. μπορούν να προκαλέσουν τόσο απόπτωση όσο και νέκρωση.

1.4 Αιμοποιητικό Σύστημα και Οξεία Λεμφοβλαστική

Λευχαιμία

Στην παρούσα ενότητα επιχειρείται μία συνοπτική εισαγωγή στην βιολογία του αιμοποιητικού συστήματος αλλά και στα χαρακτηριστικά της Οξείας Λεμφοβλαστικής Λευχαιμίας η οποία αποτέλεσε βασικό πεδίο εφαρμογής της σχετικής με την παρούσα διατριβή έρευνα. Ως βάση για τα όσα παρουσιάζονται στις ακόλουθες υποενότητες, μεταξύ άλλων, χρησιμοποιήθηκε η [1].

1.4.1 Αιμοποίηση

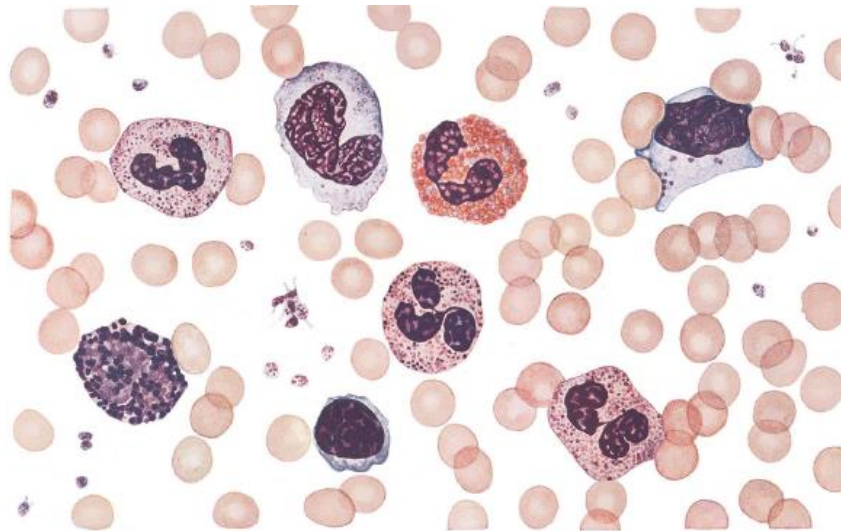
Η λειτουργία της αιμοποίηση εξασφαλίζει την κανονική αναπλήρωση των κυττάρων του αίματος και ρυθμίζεται μέσω ενός πλήθους μορίων, τα οποία εκκρίνονται από τα κύτταρα του υποστρώματος, από άλλα ειδικά κύτταρα αλλά και από τα ίδια τα αιμοποιητικά κύτταρα. Η αιμοποίηση επιτελείται στον μυελό των οστών (από το τέλος της εμβρυϊκής ζωής). Αρχίζει από κύτταρα που δεν μπορούν να αναγνωρισθούν μορφολογικά τα οποία έχουν την ιδιότητα να πολλαπλασιάζονται και να διαφοροποιούνται εξασφαλίζοντας και τη δική τους αυτό-ανανέωση. Σε δεύτερη φάση, τα αρχέγονα αυτά κύτταρα δεσμεύονται προς εξέλιξη προς μία από τις τέσσερις αιμοποιητικές σειρές (ερυθρά, κοκκιώδης-μονοκυτταρική, μεγακαρυοκυτταρική και λεμφική). Τα δεσμευμένα προγονικά κύτταρα, τα οποία πλέον είναι μορφολογικά αναγνωρίσιμα, συνεχίζουν να πολλαπλασιάζονται και να διαφοροποιούνται μέχρι να καταλήξουν ώριμα κύτταρα των αιμοποιητικών σειρών. Τα κύτταρα αυτά πλέον που εξέρχονται στην κυκλοφορία του αίματος προκειμένου να επιτελέσουν τις λειτουργίες τους.

Στα ώριμα κύτταρα, διακρίνονται τέσσερις κύριες ομάδες (Εικόνα 1.2):

- Τα **ερυθροκύτταρα** τα οποία ομοιόμορφα, ισόχρωμα και ισομεγέθη, δισκοειδή, απύρρηνα κύτταρα γεμάτα με αιμοσφαιρίνη για την μεταφορά οξυγόνου
- Τα κοκκιοκύτταρα και τα μονοκύτταρα με πολύλοβους πυρήνες και κυτταρόπλασμα γεμάτο από ηωσινόφιλα, πολυχρωματόφιλα ή βασεόφιλα κοκκία (συμμετέχουν στην άμεση άμυνα του οργανισμού)
- τα **λεμφοκύτταρα** (συμμετέχουν στην άμεση και έμμεση άμυνα και στην ανοσία).

- τα **αιμοπετάλια**, τα οποία είναι απύρηντα θραύσματα μεγαλύτερων κυττάρων του μυελού απαραίτητα για την αιμόσταση.

Η παραγωγή των κυττάρων αυτών γίνεται στον μυελό των οστών με πολλαπλασιασμό και διαφοροποίηση των αντίστοιχων στελεχιαίων/προγονικών κυττάρων και ρυθμίζεται με ακρίβεια με ειδικά «σήματα», εξωκυττάρια (αυξητικοί παράγοντες, ερυθροποιητίνη, θρομβοποιητίνη, κ.α. κυτταροκίνες) και ενδοκυττάρια (κινάσες και μεταγραφικοί παράγοντες), ως απάντηση έναντι συγκεκριμένων αναγκών.



Εικόνα 1.2 Κύτταρα του αίματος. Λευκά αιμοσφαίρια ανάμεσα σε ερυθροκύτταρα και αιμοπετάλια

1.4.1.1 Παραγωγή των κυττάρων του αίματος

Στις ακόλουθες υπο-ενότητες δίνεται μία συνοπτική παρουσίαση της διαδικασίας παραγωγής των διαφόρων τύπων κυττάρων του αίματος.

1.4.1.1.1 Μη αναγνωρίσιμα κύτταρα

Το πρώτο βήμα για την εξέλιξη της αιμοποίησης του ενήλικου ανθρώπου αρχίζει την 4η εβδομάδα της κύησης με την αυτόνομη εμφάνιση στελεχιαίων αιμοποιητικών κυττάρων (HSC, human stem cells), που προοδευτικά αποικίζουν το ήπαρ, τον σπλήνα και τον μυελό των οστών.

Στον ενήλικο άνθρωπο η παραγωγή των κυττάρων του αίματος γίνεται στον μυελό των οστών αρχίζοντας από το μη δεσμευμένο πολυδύναμο στελεχιαίο κύτταρο. Το κύτταρο αυτό δεν αναγνωρίζεται μορφολογικά μέσα στο πλήθος των κυττάρων του μυελού (η αναλογία του είναι της τάξεως του 1 στα 20 εκατομμύρια κύτταρα). Το στελεχιαίο πολυδύναμο κύτταρο αναγνωρίζεται από τις πρωτεΐνες-αντιγόνα που αναγνωρίζονται με ειδικούς αντιορούς στην επιφάνειά του (CD41, CD46, c-kit και sca 1). Επιπλέον, στα αρχικά στάδια της εμβρυϊκής ζωής φέρει και το αντιγόνο CD34, ενώ αυτό δεν διαπιστώνεται μετά την 19η εβδομάδα (CD34 αρνητικό). Τα αντιγόνα αυτά μπορούν να

ανιχνευθούν με κυτταρομετρία ροής και είναι ιδιαίτερα χρήσιμα για την μέτρηση του αριθμού των στελεχιαίων κυττάρων.

Το στελεχιαίο αιμοποιητικό κύτταρο έχει τη δυνατότητα να αυτοανανεώνεται και να διαφοροποιείται, δηλαδή, κατά την κυτταρική διαίρεση το ένα θυγατρικό κύτταρο παραμένει ως στελεχιαίο, ενώ το άλλο δεσμεύεται για εξέλιξη, δηλαδή διαφοροποιείται για να παραγάγει ερυθρά ή λευκά αιμοσφαίρια (τα λευκά αιμοσφαίρια διακρίνονται σε Ουδετερόφιλα, Λεμφοκύτταρα, Μονοκύτταρα, Ηωσινόφιλα και Βασεόφιλα) ή αιμοπετάλια. Υπολογίζεται ότι ένα τέτοιο κύτταρο μπορεί να δώσει 106 ώριμα κύτταρα μετά από 20 διαιρέσεις.

Κατά κανόνα, τα στελεχιαία αιμοποιητικά κύτταρα παραμένουν στον μυελό σε φάση ηρεμίας (G0) και κατορθώνουν να εξασφαλίσουν την ισορροπία της αιμοποίησης με πολλαπλασιασμό τους που γίνεται πολύ σπάνια, μία φορά κάθε 21 εβδομάδες. Με τον πολύ βραδύ ρυθμό πολλαπλασιασμού το στελεχιαίο αιμοποιητικό κύτταρο δεν εξαντλείται σύντομα, αν και η ικανότητα αυτοανανέωσης προοδευτικά μειώνεται λόγω μεταλλάξεων, επιγενετικών βλαβών και μείωσης του μήκους των τελομεριδίων.

Στη συνέχεια, το πολυδύναμο στελεχιαίο κύτταρο αποδίδει πολυδύναμα προγονικά κύτταρα, τα οποία διαφοροποιούνται ολοένα και περισσότερο προς μονοδύναμα προγονικά κύτταρα, που συνεχίζουν να είναι ελάχιστα χωρίς να μπορούν να ανιχνευθούν μορφολογικά, αλλά αναγνωρίζονται χάρη στην ιδιότητα να σχηματίζουν αποικίες (ομάδες ομοειδών θυγατρικών κυττάρων). Όλες οι παραπάνω διεργασίες ρυθμίζονται από ένα πλήθος εξωγενών παραγόντων, οι υποδοχείς των οποίων εκφράζονται κατά κύματα στην μεμβράνη των εξελισσόμενων κυτταρικών σειρών.

1.4.1.1.1 Ρύθμιση της παραγωγής

Οι διαδικασίες παραγωγής κυττάρων του αίματος ρυθμίζονται με μία σειρά ενδοκυττάρων μεταγραφικών παραγόντων που ενεργοποιούνται διαδοχικά, αυτόνομα ή κάτω από την επίδραση διαφόρων εξωγενών παραγόντων, που επίσης εμφανίζονται διαδοχικά ανάλογα με το στάδιο διαφοροποίησης (κυτταροκίνες, ιντερλευκίνες κ.α.).

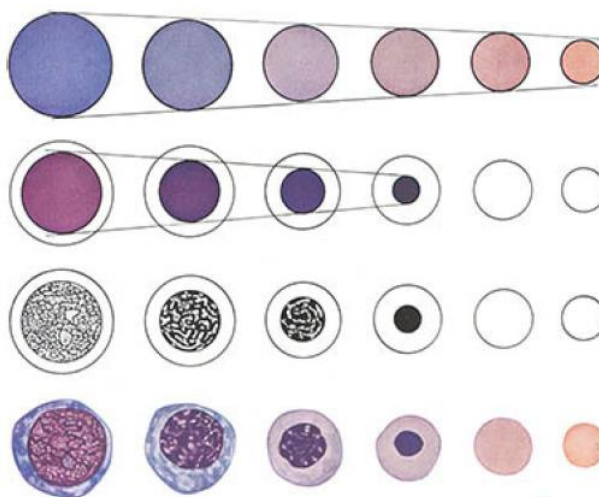
Ο γενικός μηχανισμός αυτών των αλληλεπιδράσεων αρχίζει με την σύνδεση των κυτταροκινών με υποδοχείς που αναπτύσσονται, επίσης διαδοχικά, στην επιφάνεια των εξελισσόμενων προγονικών κυττάρων. Η σύνδεση αυτή ενεργοποιεί την ενδομεμβρανική πλευρά του υποδοχέα (κατά κανόνα οι υποδοχείς είναι διαμεμβρανικές πρωτεΐνες) και αυτή, με τη σειρά της, μεταφέρει το σήμα και ενεργοποιεί μία ακολουθία άλλων ενδοκυττάρων πρωτεϊνών (μεταγωγή σήματος). Έπειτα, οι τελικοί ενεργοποιημένοι διαβιβαστές εισέρχονται στον πυρήνα, συνδέονται με το σύμπλεγμα της πολυμεράσης επάνω στις

κατάλληλες νουκλεοτιδικές ακολουθίες και ξεκινούν τη διαδικασία σύνθεσης πρωτεϊνών, με χρήση των οποίων το κύτταρο θα επιτελέσει την λειτουργία του (στην προκείμενη περίπτωση τον πολλαπλασιασμό και την διαφοροποίηση).

1.4.1.1.2 Αναγνωρίσιμα Κύτταρα και Ρύθμιση της διαφοροποίησης και της παραγωγής των αναγνωρίσιμων κυττάρων του μυελού

Με την δέσμευση των μονοδύναμων κυττάρων στον μυελό αρχίζει η τελική τους διαφοροποίηση προς κύτταρα της ερυθράς, της κοκκιώδους (μυελικής), και της μεγακαρυοκυτταρικής σειράς. Τα λεμφοειδή κύτταρα ακολουθούν άλλη οδό διαφοροποίησης, που επιτελείται κυρίως στον θύμο αδένα και τους λεμφαδένες.

Κατά την διαφοροποίηση, ο γενικός κανόνας για κάθε σειρά είναι η προοδευτική σμίκρυνση του κυττάρου και η πύκνωση του πυρήνα (Εικόνα 1.3).



Εικόνα 1.3 Κυτταρική διαφοροποίηση της ερυθράς σειράς. Χαρακτηρίζεται από προοδευτική σμίκρυνση του κυττάρου, μεταβολή του χρώματος του κυτταροπλάσματος, σμίκρυνση και πύκνωση του πυρήνα, εξαφάνιση των πυρηνίων και εκβολή του πυρήνα.

1.4.1.1.2.1 Ερυθρά σειρά

Η εξέλιξη της ερυθράς σειράς αρχίζει με την Προερυθροβλάστη. Ο προερυθροβλάστης πολλαπλασιάζεται και διαφοροποιείται προς βασεόφιλη ερυθροβλάστη, κύτταρο μικρότερο, στρογγυλό, με πυκνότερο πυρήνα, η χρωματίνη του οποίου παίρνει τροχοειδή διάταξη. Ακολουθεί η πολυχρωματόφιλη και έπειτα η οξύφιλη ερυθροβλάστη, όπου το κυτταρόπλασμα γίνεται προοδευτικά οξύφιλο και ροδίζει (πλήρωση με αιμοσφαιρίνη), ενώ ο πυρήνας πυκνώνει περισσότερο και γίνεται ένα βαθύχρωμο σφαιροειδές μόρφωμα. Στο τέλος, ο πυρήνας εκβάλλεται και το άωρο ερυθροκύτταρο, που ονομάζεται δικτυοερυθροκύτταρο αποδίδεται στην κυκλοφορία. Τα φυσιολογικά ερυθροκύτταρα έχουν σχήμα αμφίκιουλου δίσκου (η αιμοσφαιρίνη διατάσσεται περιμετρικά) με διάμετρο περί τα 7 μ, και είναι όλα ισομεγέθη και ισόχρωμα.

1.4.1.1.2.2 Κοκκιώδης (μυελική) σειρά

Το πρώτο αναγνωρίσιμο κύτταρο της κοκκιώδους σειράς είναι η μυελοβλάστη. Πρόκειται για μεγάλο κύτταρο με γαλάζιο πρωτόπλασμα χωρίς κοκκία και μεγάλο πυρήνα. Η μυελοβλάστη εξελίσσεται προς προμυελοκύτταρο, που είναι μεγαλύτερο, έχει πυκνότερο πυρήνα, και το πρωτόπλασμά του γεμίζει με αδρά βαθυκύανα κοκκία. Έπειτα, το προμυελοκύτταρο διαφοροποιείται προς μυελοκύτταρο, όπου ο πυρήνας γίνεται περισσότερο πυκνωτικός, ενώ το πρωτόπλασμα γεμίζει με κοκκία, που καθορίζουν και την περαιτέρω εξέλιξή του σε ουδετερόφιλο μεταμυελοκύτταρο, ηωσινόφιλο μεταμυελοκύτταρο ή βασεόφιλο μεταμυελοκύτταρο. Προοδευτικά, ο πυρήνας των μυελοκυττάρων γίνεται επιμήκης και καταλήγει με την μορφή ραβδίου στο ραβδοπύρηνο κύτταρο. Τελικά, ο πυρήνας των ραβδοπύρηνων κυττάρων κατατέμεται σε δύο ή περισσότερους λοβούς και τα παράγωγα πολυμορφοπύρηνα κύτταρα αποδίδονται στην κυκλοφορία.

1.4.1.1.2.3 Μονοκύτταρα

Η εξέλιξη της σειράς των μονοκυττάρων είναι παρόμοια με προηγουμένως. Το άωρο αναγνωρίσιμο κύτταρο είναι η μονοβλάστη, η οποία διαφοροποιείται προς μονοκύτταρο. Το μονοκύτταρο είναι μεγαλύτερο από τα πολυμορφοπύρηνα (14-20 μm), έχει ανώμαλο, ελαφρά ιώδες πρωτόπλασμα με λεπτά κοκκία και ψευδοπόδια και πολύμορφο, αλλά μονήρη πυρήνα.

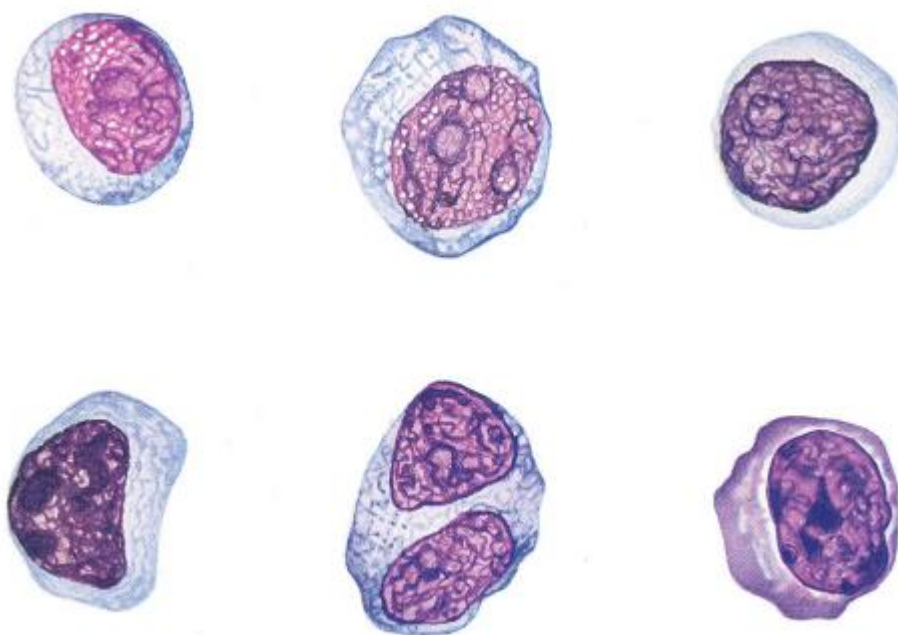
1.4.1.1.2.4 Αιμοπετάλια. Μεγακαρυοκυτταρική σειρά

Τα αιμοπετάλια αποτελούν τμήματα του πρωτοπλάσματος των μεγακαρυοκυττάρων, τα οποία δεν εξέρχονται στην κυκλοφορία. Η προγονική μορφή τους είναι και εδώ η μεγακαρυοβλάστη. Η μεγακαρυοβλάστη εξελίσσεται προς άωρο μεγακαρυοκύτταρο, που είναι πολύ μεγαλύτερο από όλα τα άλλα κύτταρα του μυελού, έχει πλούσιο κυτταρόπλασμα, σαφές περίγραμμα και πυρήνα, μονήρη και στρογγυλό στην αρχή, που διαιρείται μέσα στο κύτταρο χωρίς αυτό να προχωρεί σε μίτωση και καταλήγει στο ώριμο μεγακαρυοκύτταρο, κύτταρο πολύ μεγάλο και πολυπύρηνο με πρωτόπλασμα, το οποίο κατατέμεται άτακτα αποδίδοντας μικρά θραύσματα που αποτελούν τα αιμοπετάλια.

1.4.1.1.2.5 Λεμφοκύτταρα

Μητρική μορφή των λεμφοκυττάρων του μυελού είναι η λεμφοβλάστη. Πρόκειται για ευμέγεθες σχετικά στρογγυλό κύτταρο με βαθυκύανο πρωτόπλασμα και μεγάλο πυρήνα που συχνά περιορίζει το πρωτόπλασμα σε μια στεφάνη. Ο πυρήνας μπορεί να έχει ένα ή δύο πυρήνια, λιγότερο ευκρινή από εκείνα της μυελοβλάστης. Η λεμφοβλάστη εξελίσσεται προς προλεμφοκύτταρο και κατόπιν σε λεμφοκύτταρο. Κατά γενικό κανόνα τα λεμφοκύτταρα

είναι μικρά και στρογγυλά κύτταρα με πυκνωτικό πυρήνα που περιβάλλεται από πρωτόπλασμα.



Εικόνα 1.4 Εξέλιξη της λεμφικής σειράς. Λεμφοβλάστης-Προλεμφοκύτταρα (επάνω). Λεμφοκύτταρα (κάτω).

Η κινητική των λεμφοκυττάρων είναι περίπλοκη γιατί, παρά την μορφολογική τους ομοιογένεια, τα κύτταρα αυτά κατανέμονται σε δύο κύριους διακριτούς πληθυσμούς: τα Β- και τα Τ- λεμφοκύτταρα. Η εξέλιξη της Β-σειράς αρχίζει στον μυελό από όπου τα Β-λεμφοκύτταρα μεταναστεύουν στον σπλήνα και τους λεμφαδένες για να μετάσχουν, μετά από ειδική διαφοροποίηση, στην ανοσολογική λειτουργία. Τα Β-λεμφοκύτταρα αποτελούν το 5-15% των λεμφοκυττάρων του περιφερικού αίματος. Τα Τ-λεμφοκύτταρα, με αφετηρία και πάλι τον μυελό, μεταναστεύουν στον θύμο αδένα, όπου πολλαπλασιάζονται και ευαισθητοποιούνται, αποκτούν δηλαδή την ιδιότητα να απαντούν σε ορισμένα ειδικά αντιγονικά ερεθίσματα. Έπειτα, μεταναστεύουν και αυτά σε ειδικές περιοχές των λεμφαδένων και του σπλήνα, όπου μετέχουν στη ανοσολογική λειτουργία του οργανισμού. Τα Τ-λεμφοκύτταρα κατανέμονται περαιτέρω στις υποομάδες των βοηθητικών, των κατασταλτικών και των κυττάρων φυσικών φονέων.

1.4.2 Η ασθένεια

Η Οξεία Λεμφοβλαστική Λευχαιμία χαρακτηρίζεται από τον αναίτιο, άσκοπο και ανεξέλεγκτο πολλαπλασιασμό των κυττάρων της λεμφικής σειράς με ελάχιστη ή καμία

διαφοροποίηση. Αποτελεί το 80% των λευχαιμιών της παιδικής ηλικίας, είναι λιγότερο συχνή στους νέους και ξαναγίνεται συχνή στους ενήλικους μετά την ηλικία των 60 ετών. Επίσης αποτελεί μία από τις συχνότερες μορφή καρκίνου στην παιδική ηλικία, αγγίζοντας το 25% των παιδικών καρκίνων και το 80% των παιδικών λευχαιμιών [19], [20]. Προσβάλλει συχνότερα τους λευκούς της Ευρώπης και βορείων ΗΠΑ και είναι σχετικά σπάνια στους Ασιάτες, τους Αφρικανούς και τους κατοίκους της Νότιας Αμερικής.

Η παιδική λευχαιμία προσβάλλει συχνότερα παιδιά που έχουν σύνδρομο Down σύνδρομο Bloom's, αναιμία Fanconi ή άλλες γενετικές ανωμαλίες [21]–[23] και μπορεί να εμφανίζεται συχνότερα μετά από έκθεση των μητέρων (ή και των παιδιών) σε ακτινοβολία, εντομοκτόνα, κάπνισμα ή διάφορα φάρμακα. Με το δεδομένο ότι η εκδήλωση της ΟΛΛ αποτελεί το τελικό γεγονός μιας σειράς γενετικών διαταραχών, είναι πιθανό ότι σε μερικές περιπτώσεις ΟΛΛ, η λευχαιμογένεση αρχίζει ήδη στην εμβρυϊκή ζωή [24].

Υπάρχουν πολλοί τύποι ΟΛΛ με κοινό χαρακτηριστικό την παρουσία βλαστικών κυττάρων στον μυελό και το περιφερικό αίμα. Η αναγνώριση των ειδικών αυτών τύπων έχει πολύ σημαντική αξία για την πρόγνωση και τον καθορισμό της θεραπευτικής αγωγής. Τα κριτήρια αναγνώρισης είναι (α) η μορφολογία των βλαστών, (β) η προέλευση και ο βαθμός διαφοροποίησης, (γ) ο ανοσοφαινότυπος, (δ) οι χαρακτηριστικές («όχι τυχαίες») καρυοτυπικές ανωμαλίες και (ε) οι χαρακτηριστικές μοριακές μεταβολές.

Η μορφολογία εξακολουθεί να παραμένει η μέθοδος που αρχικά γίνεται η διάγνωση της Οξείας Λευχαιμίας και μαζί με τις κυτταροχημικές χρώσεις αποτελεί τον βασικό τρόπο διαχωρισμού της Οξείας Μυελογενούς Λευχαιμίας από την ΟΛΛ [25].

Για την πιο ακριβή ταξινόμηση της ΟΛΛ σε Β ή Τ κυτταρική σειρά, χρησιμοποιούνται ανοσολογικές τεχνικές, προκειμένου να ανιχνευθούν ειδικά αντιγόνα που χαρακτηρίζουν κάθε κυτταρική σειρά.

Η κυτταρογενετική ανάλυση εξακολουθεί να αποτελεί σημαντικό μέρος της διάγνωσης της ΟΛΛ, καθώς έχει μεγάλη προγνωστική αξία. Παρόλα αυτά μοριακές τεχνικές για την ανίχνευση κάποιων υποομάδων της ΟΛΛ (π.χ. BCR-ABL + ΟΛΛ) αρχίζουν να αποκτούν όλο και μεγαλύτερο ενδιαφέρον. Μοριακοί ιχνηθέτες (markers), όπως ανακατατάξεις των γονιδίων του υποδοχέα του Τ-κυττάρου (TCR:T-cell receptor) και της αλυσίδας της ανοσοσφαιρίνης (IgH) χρησιμοποιούνται σήμερα για να εκτιμήσουν το θεραπευτικό αποτέλεσμα σε συγκεκριμένους ασθενείς με την ανίχνευση ελάχιστης υπολειμματικής νόσου (MRD-Minimal Residual Disease) [26].

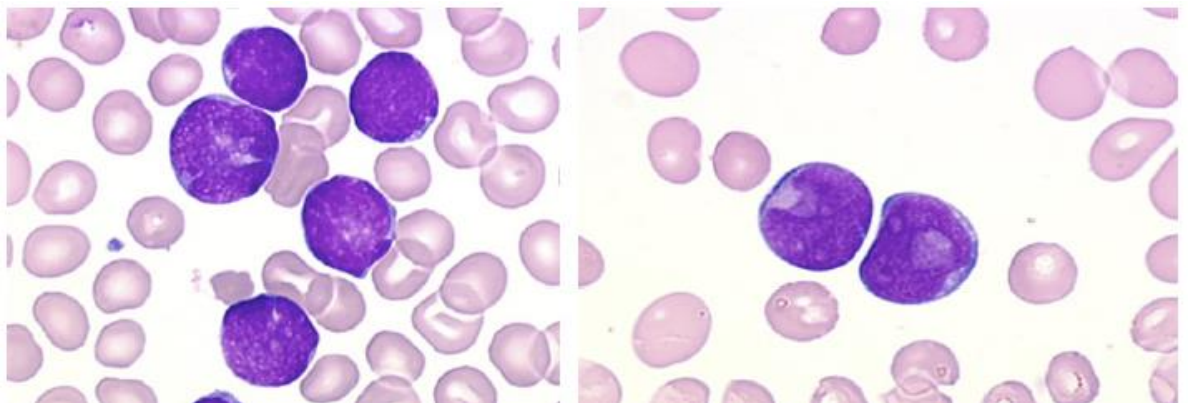
Με βάση τα ανωτέρω οι ΟΛΛ ταξινομούνται με βάση τα συστήματα που παρουσιάζονται στις ακόλουθες υποενότητες.

1.4.2.1 Μορφολογική κατάταξη ή ταξινόμηση κατά FAB

Οι λεμφοβλάστες μπορεί να χαρακτηριστούν σύμφωνα με την FAB (French-American-British) ταξινόμηση ως

- L1: σχετικά μικροί μονόμορφοι,
- L2: μεγαλύτεροι και πιο πλειομορφικοί και
- L3: με βασεόφιλο πρωτόπλασμα και με κενोटόπια.

Ο διαχωρισμός στους υπότυπους L1-L2 είναι ελάχιστονος σημασίας για την πρόγνωση του αποτελέσματος της θεραπείας. Ο μορφολογικός υπότυπος L3 πρέπει να αναγνωρίζεται, καθώς είναι ενδεικτικός ώριμου B-ΟΛΛ φαινότυπου και απαιτεί διαφορετικό είδος θεραπείας. Όπως ωστόσο σημειώνεται στην σχετική βιβλιογραφία, η διάγνωση θα πρέπει πάντα να επιβεβαιώνεται από την ανοσοφαινοτυπική εξέταση δεικτών επιφάνειας [1]. Ως παράδειγμα της διαδικασίας, στην Εικόνα 1.5 παρουσιάζονται λεμφοβλάστες σε σύγκριση με μυελοβλάστες. Στην ΟΛΛ ο πυρήνας είναι σχεδόν μεγαλύτερος από το κύτταρο, αναδιπλώνεται και δεν φαίνεται να περιέχει πυρήνια. Τα χαρακτηριστικά αυτά δεν υπάρχουν στις μυελοβλάστες της ΟΜΛ.



Εικόνα 1.5 Λεμφοβλάστες στην ΟΛΛ σε αντιπαράθεση με μυελοβλάστες σε ΟΜΛ.

1.4.2.2 Κατάταξη με βάση την προέλευση και τον βαθμό διαφοροποίησης μέσω ανοσοφαινοτυπικής ανάλυσης κυτταρικών δεικτών επιφάνειας

Η ανθρώπινη ΟΛΛ προέρχεται από ένα μοναδικό πρόδρομο κύτταρο που έχει υποστεί γενετική βλάβη και οδηγήθηκε σε απορρύθμιση της αύξησης και αναστολή της διαφοροποίησής του [27]. Ωστόσο, σημαντικό εύρημα αποτελεί το ότι οι κακοήθεις λεμφοβλάστες μοιράζονται πολλά κοινά χαρακτηριστικά με τα φυσιολογικά πρόδρομα λεμφικά κύτταρα [27]. Έτσι, οι βλάστες της ΟΛΛ αναδιατάσσουν τα γονίδια της βαρείας αλυσίδας της ανοσοσφαιρίνης (IgH) και υποδοχέα του T λεμφοκυττάρου (TCR) και εκφράζουν μέρη από μόρια που αποτελούν αντιγονικούς υποδοχείς ή άλλες γλυκοπρωτεΐνες

που σχετίζονται με τη διαφοροποίηση, με τρόπο ώστε να υπάρχει αντιστοιχία με χαρακτηριστικά που ανήκουν σε φυσιολογικά στάδια ανάπτυξης των Β και των Τ λεμφοκυττάρων [28]. Στις περισσότερες περιπτώσεις οι βλάστες της ΟΛΛ μοιάζουν να αντιπροσωπεύουν τον πολλαπλασιασμό προδρόμων της λεμφικής σειράς που έχουν παραμείνει σε πρώιμα στάδια της Τ ή Β διαφοροποίησης[29]. Ωστόσο, οι βλάστες της ΟΛΛ μπορούν να παρουσιάζουν ασύγχρονη έκφραση γονιδίων ή μικρές διαφοροποιήσεις του ανοσοφαινότυπού τους (19).

Με βάση τα παραπάνω η ΟΛΛ ταξινομείται σύμφωνα με το αντίστοιχο «φυσιολογικό» στάδιο ανάπτυξης για μία λεμφοβλάστη [30]. Η Ευρωπαϊκή ομάδα για τον ανοσολογικό χαρακτηρισμό της Οξείας Λευχαιμίας (EGIL-European Group for the Immunological Characterization of Acute Leukemia) προτείνει μία κοινή ανοσοφαινοτυπική ταξινόμηση των ΟΛΛ η οποία δίνεται ακολούθως.

- **ΟΛΛ Β-κυτταρικής σειράς:**

1. Προ-B-ΟΛΛ (Pro-B-ALL): είναι CD10 αρνητική ΟΛΛ και δεν χαρακτηρίζεται από ειδικά αντιγόνα επιφάνειας Β ή Τ κυτταρικής σειράς, αλλά εκφράζει το μείζον αντιγόνο ιστοσυμβατότητας HLA-DR, την Tdt και το CD19. Επιπλέον παρουσιάζει αναδιάταξη των γονιδίων της IgH.
2. Κοινή Β-ΟΛΛ (Common B-ALL): Χαρακτηρίζεται από την εμφάνιση του CD10. Τα βλαστικά κύτταρα δεν φέρουν αντιγόνα που χαρακτηρίζουν σχετικά ώριμα Β-λεμφοκύτταρα, όπως ανοσοσφαιρίνη κυτταροπλασματική ή επιφανείας και είναι θετικά για την Tdt και το CD19.
3. Πρώιμος Β-ΟΛΛ (Pre-B-ALL ή Precursor-B ALL): Τα ανοσοφαινοτυπικά χαρακτηριστικά της είναι όμοια με της κοινής Β-ΟΛΛ ωστόσο χαρακτηρίζεται από την παρουσία κυτταροπλασματικής ανοσοσφαιρίνης.
4. Ωριμος Β-ΟΛΛ(mature B-ALL): αποτελεί περίπου το 5% των περιπτώσεων των ΟΛΛ των ενηλίκων. Τα βλαστικά κύτταρα εκφράζουν αντιγόνα επιφάνειας των ώριμων Β-κυττάρων, μεταξύ των οποίων επιφανειακή ανοσοσφαιρίνη.

- **ΟΛΛ Τ-κυτταρικής σειράς:**

Όλες οι περιπτώσεις εκφράζουν το Τ-κυτταρικό αντιγόνο CD7 και μπορεί να εκφράζουν ανάλογα με το βαθμό διαφοροποίησης άλλα αντιγόνα (π.χ. το CD2-δηλαδή τον υποδοχέα των Τ λεμφοκυττάρων στον οποίο συνδέονται τα ερυθρά αιμοσφαίρια, το CD1 φλοιϊκό θυμοκυτταρικό αντιγόνο και άλλα). Μία μειονότητα Τ-ΟΛΛ μπορεί να

εκφράζει το CD10 μαζί με T-κυτταρικά αντιγόνα. Στις περισσότερες περιπτώσεις των T-ΟΛΛ ένα ή περισσότερα γονίδια του T-κυτταρικού υποδοχέα μπορεί να είναι αναδιαταγμένα. Η ανοσοφαινοτυπική τυποποίηση επιτρέπει την ταξινόμηση της T-ΟΛΛ ανάλογα με το στάδιο διαφοροποίησης σε:

1. Πρώιμες-προδρομικές T-ΟΛΛ (Early T-ALL) : Έχουν χαρακτηριστικά T κυττάρων, π.χ. cyCD3 και CD7, αλλά όχι περαιτέρω διαφοροποίηση.
2. Θυμικές-φλοιϊκές T-ΟΛΛ (cortical T-ALL ή Thy ALL): Χαρακτηρίζονται κυρίως από την ύπαρξη του αντιγόνου CD1a. Επιπλέον μπορεί να υπάρχει το επιφανειακό CD3 (sCD3).
3. Ωριμες T-ΟΛΛ (mature T-ALL): Δεν εκφράζουν το CD1a, αλλά εκφράζουν το sCD3.

1.4.2.3 Κατάταξη με βάση τον καρυότυπο και τις κύριες μοριακές διαταραχές

Σύμφωνα με την τελευταία ταξινόμηση του παγκόσμιου οργανισμού υγείας (World Health Organization, WHO) του 2008 και η οποία βασίζεται σε πληροφορίες από τον καρυότυπο και την μοριακή ανάλυση, τα νεοπλάσματα από προγονικά κύτταρα της λεμφικής σειράς ταξινομούνται σε [31]:

- **B λεμφοβλαστική λευχαιμία/λέμφωμα**

1. B λεμφοβλαστική λευχαιμία/λέμφωμα μη αλλιώς ταξινομούμενη, χωρίς ειδικά γενετικά χαρακτηριστικά (not other way specified-NOS)
2. B λεμφοβλαστική λευχαιμία/λέμφωμα με επαναλαμβανόμενες γενετικές ανωμαλίες:
 - B λεμφοβλαστική λευχαιμία/λέμφωμα με t(9;22) (q34;q11.2);BCR-ABL1
 - B λεμφοβλαστική λευχαιμία/λέμφωμα με t(v;11q23); MLL αντιμεταθέσεις
 - B λεμφοβλαστική λευχαιμία/λέμφωμα με t(12;21)(p13;q22); TEL-AML1 (ETV6-RUNX1)
 - B λεμφοβλαστική λευχαιμία/λέμφωμα με υπερδιπλοειδία
 - B λεμφοβλαστική λευχαιμία/λέμφωμα με υποδιπλοειδία
 - B λεμφοβλαστική λευχαιμία/λέμφωμα με t(5;14)(q31;q32);IL3-IGH
 - B λεμφοβλαστική λευχαιμία/λέμφωμα με t(1;19)(q23;p13.3); E2A-PBX1 (TCF3-PBX1)

- **Τ λεμφοβλαστική λευχαιμία/λέμφωμα**

1.4.3 Θεραπευτική Αντιμετώπιση.

Η θεραπεία της παιδικής ΟΛΛ αποτελεί μία από τις σημαντικότερες επιτυχίες της κλινικής ογκολογίας με το τρέχον ποσοστό ίασης στις ανεπτυγμένες χώρες να αγγίζει το 80% [19]. Τα αποτελέσματα αυτά έχουν επιτευχθεί μέσω της εφαρμογής εντατικής και πολυδιάστατης χημειοθεραπευτικής θεραπείας αλλά και ακτινοθεραπείας ή/και μεταμόσχευσης αιμοποιητικών βλαστικών κυττάρων σε συγκεκριμένες υποκατηγορίες ασθενών.

Αν και η στρατηγική της θεραπείας είναι όμοια στα παιδιά και τους ενηλίκους, η ανταπόκριση και η τελική πρόγνωση είναι πολύ διαφορετικές, με αποτέλεσμα η ΟΛΛ να είναι «ιάσιμη» σε μεγάλο ποσοστό παιδιών, αλλά να αποτελεί δύσκολο πρόβλημα με κακή πρόγνωση στους ενηλίκους. Οι γενικές αρχές της θεραπευτικής προσπέλασης βασίζονται στο γεγονός ότι όταν η λευχαιμία εκδηλώνεται κλινικά τα λευχαιμικά κύτταρα είναι ήδη πάρα πολλά και πλέον δεν πολλαπλασιάζονται γρήγορα, όπως θα αναμενόταν (περαιτέρω συζήτηση επί του θέματος γίνεται στο Κεφάλαιο 3). Έτσι στην πρώτη φάση της αντιμετώπισης χρησιμοποιούνται κυρίως φάρμακα που δεν εστιάζουν στην καταστολή του πολλαπλασιασμού των λευχαιμικών κυττάρων όπως αλκυλιούντες παράγοντες, κορτικοειδή [32], ανθρακυκλίνες και ασπαραγινάση. Ο στόχος αυτής της φάσης είναι η μείωση του φορτίου της νόσου με αποτέλεσμα τα κύτταρα να επανέλθουν σε φάση γρήγορου πολλαπλασιασμού. Όταν αυτό επιτευχθεί, χρησιμοποιούνται κυρίως κυτταροστατικά φάρμακα (αναστέλλουν τον κυτταρικό πολλαπλασιασμό) ειδικά για τις φάσεις S (σύνθεση DNA) και M (μίτωση) όπως μεθοτρεξάτη, μερκαπτοπουρίνη, υδροξυκαρβαμίδη και αραβινοσίδη της κυτοσίνης και αλκαλοειδή της vinca.

Οι σύγχρονες θεραπευτικές αγωγές αποτελούνται τουλάχιστον από 4 φάσεις:

1. Περίοδος θεραπείας εφόδου (induction period) που στοχεύει στην αρχική όσο το δυνατόν μεγαλύτερη μείωση της ασθένειας με διάρκεια 4 έως 6 εβδομάδων και με τη χρήση πολλαπλών χημειοθεραπευτικών φαρμάκων
2. Περίοδος ενίσχυσης της θεραπείας, εδραίωση του αποτελέσματος και επανάληψη εφόδου (consolidation/intensification and reinduction segments) με στόχο την εκρίζωση εναπομεινάντων λευχαιμικών κυττάρων από ασθενείς που βρίσκονται σε ύφεση (με βάση μορφολογικά κριτήρια).
3. Εξωδιαμερισματική θεραπεία (extracompartment therapy) όπως για παράδειγμα προστατευτική θεραπεία για το Κεντρικό Νευρικό Σύστημα-

ΚΝΣ (central nervous system (CNS) preventive therapy), προκειμένου να αποφευχθεί η προσβολή αυτού από λευχαιμικά κύτταρα.

4. Περίοδος συντήρησης (maintenance period) για την περαιτέρω σταθεροποίηση της ύφεσης μέσω της καταπίεσης της επανεμφάνισης πιθανών ανθεκτικών στη θεραπεία κλώνων [33], [34]

Θα πρέπει να σημειωθεί ωστόσο ότι με δεδομένο πως συγκεκριμένοι ασθενείς με διαφορετικό κλινικό και βιολογικό υπόβαθρο μπορεί να εμφανίσουν σημαντικά μειωμένη απόκριση στην τυπική θεραπεία της ALL, δημιουργήθηκαν και κλινικά πρωτόκολλα θεραπειών που εστιάζουν ειδικά στις ανάγκες των υπο-ομάδων αυτών. Για παράδειγμα έχουν δημιουργηθεί υβριδικά πρωτόκολλα για νεογνά (infants) ή ειδικά πρωτόκολλα για ασθενείς με BCR/ABL-θετική ALL [35], [36] .

1.4.3.1 Η ομάδα εργασίας BFM και η κλινική δοκιμή ALL-BFM 2000

Η κλινική δοκιμή ALL-BFM 2000 αποτέλεσε τη συνέχεια (<http://www.bfm-international.org/>) σημαντικών κλινικών δοκιμών που οργανώνονται από την διεθνή πλέον ομάδα εργασίας BFM (Berlin, Frankfurt, Münster).

Η ομάδα εργασίας BFM ιδρύθηκε το 1975 στη Γερμανία όταν οι Hansjörg Riehm στο Βερολίνο (B), Bernhard Kornhuber στη Φρανκφούρτη (F) και Günther Schellong στο Μύνστερ (M) ξεκίνησαν την πρώτη πολυκεντρική (multicentre) δοκιμή. Η φιλοσοφία θεραπειών BFM, η οποία βασίζεται στην ιδιαίτερα εντατική χημειοθεραπευτική προσέγγιση με τη χρήση 8 διαφορετικών φαρμάκων, οδήγησε σε επαναστατική αύξηση της επιβίωσης παιδιών και εφήβων με Οξεία Λεμφοβλαστική Λευχαιμία. Από τότε, ο επονομαζόμενος «σκελετός θεραπειών BFM» παραμένει σημαντικό τμήμα πολλών πρωτοκόλλων θεραπείας της ΟΛΛ σε παιδιά και ενήλικες. Οι προσέγγιση υιοθετήθηκε από συνεχώς αυξανόμενες εθνικές ομάδες εργασίας παγκοσμίως, οδηγώντας σε καρποφόρες συνεργασίες για την εξέλιξη της θεραπείας.

Σήμερα, πολλές από αυτές τις ομάδες εργάζονται μαζί ώστε να βελτιωθούν περαιτέρω οι θεραπείες για τη Λευχαιμία και το Λέμφωμα υπό τη σκέπη της I-BFM (International BFM) ομάδας εργασίας η οποία δημιουργήθηκε το 1987 αρχικά από ομάδες εργασίας στην Αυστρία (BFM), Γαλλία/Βέλγιο (EORTC-CLG), Γερμανία (BFM), Ιταλία (AIEOP) και Ολλανδία (DCLSG). Η I-BFM εξελίχθηκε σε σημαντική διεθνή πλατφόρμα για την προώθηση της έρευνας και της κλινικής φροντίδας παιδιών και εφήβων με Λευχαιμία και Λέμφωμα όπου σήμερα περιλαμβάνει 28 ομάδες εργασίας από 34 χώρες.

Το 1991, το I-BFM ξεκίνησε μία προοπτική μελέτη (prospective study) με στόχο την αξιολόγηση της μοριακής παρακολούθησης της πρώιμης απόκρισης στη θεραπεία της παιδικής ΟΛΛ (Ελάχιστη Υπολειπόμενη Ασθένεια, Minimal Residual Disease-MRD). Η επονομαζόμενη MRD-αξιολόγηση επιτρέπει την υπο-μικροσκοπική ανίχνευση των σχετικών με λευχαιμικούς κλώνους αιμοσφαιρίνης και/ή ανακατατάξεις του γονιδίου του υποδοχέα των T-κυττάρων μέσω της διαδικασίας της αλυσιδωτής αντίδρασης πολυμεράσης (polymerase chain reaction-PCR) και είναι περίπου 1000 με 10000 φορές πιο ευαίσθητη σε σχέση με μεθόδους που βασίζονται στην ανίχνευση μέσω μορφολογίας. Τα αποτελέσματα της διεθνούς πολυκεντρικής δοκιμής έδειξαν ότι η απόκριση του κάθε ασθενούς ξεχωριστά μετρούμενη με την MRD διαδικασία ήταν καλύτερη από την μορφολογική αξιολόγηση και ο σημαντικότερος προβλέπτης της τελικής απόκρισης. Για πρώτη φορά ήταν εφικτή με πολύ μικρή πιθανότητα υποτροπής (μη μετρήσιμη MRD μετά από 5 όπως και μετά από 11 εβδομάδες θεραπείας) αλλά και να εντοπιστούν ασθενείς με περισσότερο από 70% πιθανότητα υποτροπής ($MRD \geq 10^{-3}$ μετά από 11 εβδομάδες θεραπείας). Οι υπόλοιποι ασθενείς μέσου ρίσκου εντοπίζονται με ανιχνεύσιμα αλλά μειούμενα επίπεδα MRD και βρίσκονται να έχουν πρόγνωση με περίπου 75% ελεύθερη υποτροπής επιβίωση. Παρόμοια αποτελέσματα επιδείχθηκαν και από άλλες διεθνείς ομάδες εργασίας. Τα αποτελέσματα αυτά οδήγησαν στο να καταστεί η MRD ανάλυση ως βασικό εργαλείο διαστρωμάτωσης ασθενών ξεκινώντας από την δοκιμή ALL-BFM 2000 το 2000.

Στις ακόλουθες ενότητες παρουσιάζονται βασικά χαρακτηριστικά της δοκιμής, από την οποία προήλθαν τα δεδομένα που χρησιμοποιήθηκαν για τις σχετικές με το Κεφάλαιο 0 μελέτες.

1.4.3.1.1 Διαστρωμάτωση κινδύνου

Η διαστρωμάτωση των ασθενών στην κλινική δοκιμή ALL-BFM 2000 βασίζεται σχεδόν αποκλειστικά στα συμπεράσματα από την απόκριση στη χημειοθεραπεία στις ακόλουθες φάσεις τις θεραπείας:

- Απόκριση στην Πρεδνιζόνη (Prednisone Response): Μέτρηση των βλαστών στο περιφερικό αίμα την 8 ημέρα της θεραπείας έπειτα από 7 ημέρες θεραπείας προ-φάσης (prephase) με Πρεδνιζόνη και μία δόσης ενδοραχιαίας (intrathecal) Μεθοτρεξάτης (Methotrexate) την πρώτη ημέρα. Οι ασθενείς που εμφανίζουν περισσότερες από 1000 βλάστες ανά μλ αίματος κατατάσσονται στην κατηγορία κακής απόκρισης στην Πρεδνιζόνη (Prednisone Poor Responders), εντάσσονται στην ομάδα Υψηλού Κινδύνου

(High Risk) και ακολουθούν το πρωτόκολλο θεραπείας I (protocol I) που αφορά την ομάδα αυτή.

- Ύφεση της ασθένειας στο μυελό των οστών (Bone Marrow-Remission) μετά από 33 ημέρες θεραπείας εφόδου: Αν μετά από 33 ημέρες θεραπείας του πρωτοκόλλου I, το ποσοστό των λευχαιμικών κυττάρων στο μυελό δεν είναι κάτω από 5%, ο ασθενής κατατάσσεται στην ομάδα υψηλού κινδύνου.
- Δυναμική της απόκρισης της λευχαιμίας στη θεραπεία σε υπο-μικροσκοπικό επίπεδο: Μέτρηση της υπολειπόμενης ασθένειας μικροσκοπικά την ημέρα 33 και πριν το πρωτόκολλο M.

Εξαιρέσεις:

- Ασθενείς με Ph+ ΟΛΛ (BCR/ABL ή t(9;22) ή με μεταθέσεις t(4;11)(q11;q23)-MLL/AF4 εντάσσονται στην ομάδα Υψηλού Κινδύνου ασχέτως της απόκρισης τους.
- Ασθενείς των οποίων το MRD δεν μπορεί να προσδιοριστεί, εντάσσονται είτε στην ομάδα Μέσου Κινδύνου (MR) είτε στην ομάδα Υψηλού Κινδύνου ανάλογα με την απόκρισή τους στην Πρεδνιζόνη.

Υπάρχουν τρεις διαφορετικές διαστρωματώσεις κινδύνου για τις οποίες τα κριτήρια ένταξης ενός ασθενούς δίνονται ακολούθως

1. Βασικός Κίνδυνος (Standard Risk Group-SR)

- Επαρκής απόκριση στην 7 ημερών θεραπεία προ-φάσης με Πρεδνιζόνη και μία δόση Μεθοτρεξάτης ενδοραχιαία την 1^η ημέρα. **Καλή απόκριση στην Πρεδνιζόνη (Prednisone Good Response, PRG): < 1000 λευχαιμικά κύτταρα ανά μl αίματος την ημέρα 8.**
- Πλήρης ύφεση την 33^η ημέρα σε μορφολογικό επίπεδο
- Μη ανίχνευση μεταθέσεων (translocation) t(9;22)- BCR/ABL ή t(4;11)- MLL/AF4
- Αρνητικό MRD την 33^η ημέρα και πριν το πρωτόκολλο M.

Και τα 4 κριτήρια θα πρέπει να καλύπτονται

2. Μέσος Κίνδυνος (Medium Risk Group-MR)

- Επαρκής απόκριση στην 7 ημερών θεραπεία προ-φάσης με Πρεδνιζόνη και μία δόση Μεθοτρεξάτης ενδοραχιαία την 1^η ημέρα.
- Πλήρης ύφεση την 33^η ημέρα σε μορφολογικό επίπεδο

- Μη ανίχνευση μεταθέσεων (translocation) t(9;22)- BCR/ABL ή t(4;11)- MLL/AF4.
- Δεν πληρούνται τα κριτήρια SR-MRD
- Δεν πληρούνται τα κριτήρια HR-MRD

Όλα τα κριτήρια θα πρέπει να πληρούνται προκειμένου να εισαχθεί ο ασθενής στο σύνολο Μέσου Κινδύνου

3. Υψηλός Κίνδυνος (High Risk Group-MR)

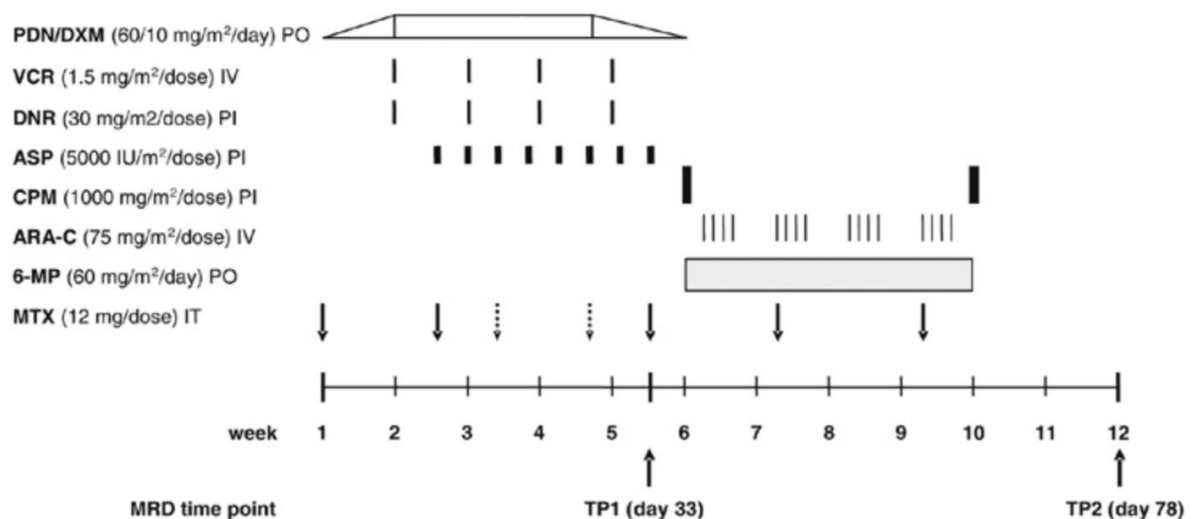
- Μη επαρκής απόκριση στην 7 ημερών θεραπεία προ-φάσης με Πρεδνιζόνη και μία δόση Μεθοτρεξάτης ενδοραχιαία την 1η ημέρα.
Κακή απόκριση στην Πρεδνιζόνη (Prednisone POOR Response, PRG): > 1000 λευχαιμικά κύτταρα ανά μl αίματος την ημέρα 8.
- Μη πλήρης ύφεση την 33η ημέρα σε μορφολογικό επίπεδο
- Ανίχνευση μεταθέσεων (translocation) t(9;22)- BCR/ABL ή t(4;11)- MLL/AF4
- Θετική MRD εξέταση πριν το πρωτόκολλο M ($\geq 10^{-3}$, υψηλό ποσοστό υπολιπούσας ασθένειας)

Η πλήρωση ενός κριτηρίου αρκεί για την κατάταξη του ασθενούς στην ομάδα Υψηλού Κινδύνου.

Όπως είναι εμφανές από τα παραπάνω, αλλά όπως σημειώνεται και στη σχετική βιβλιογραφία ([32], [37]) , η απόκριση στην Πρεδνιζόνη έχει ιδιαίτερη αξία για την διαστρωμάτωση Κινδύνου (αρκεί η κακή απόκριση σε αυτή για να χαρακτηριστεί η περίπτωση ασθενούς ως Υψηλού Κινδύνου) κάτι που είναι ιδιαίτερα σημαντικό για υποστήριξη της αξίας της έρευνας που παρουσιάζεται στο Κεφάλαιο 5.

1.4.3.1.2 Στάδια θεραπείας

Στην παρούσα υπο-ενότητα παρουσιάζεται εν συντομία το πρόγραμμα θεραπείας που εφαρμόζεται στην κλινική δοκιμή ALL-BFM 200. Στην ακόλουθη εικόνα δίνεται μία σύνοψη της θεραπείας εφόδου ενώ ακολουθεί η λεπτομερής παρουσίαση των φάσεων θεραπείας.



Εικόνα 1.6 Τα πρώτα στάδια θεραπείας με βάση την κλινική δοκιμή ALL-BFM 2000

1. **Θεραπεία Προ-φάσης:** Οι ασθενείς λαμβάνουν Πρεδνιζόνη (Prednisone,PRED) από το στόμα τις ημέρες 1-7 και μία δόση Μεθοτρεξάτης (Methotrexate,MTX) ενδοραχιαίως (intrathecally, IT) την ημέρα 1.
2. **Θεραπεία Εφόδου και εδραίωση του αποτελέσματος:** Οι ασθενείς χωρίζονται τυχαία σε δύο κλάδους θεραπείας (arms)
 - a. **Κλάδος I:** Οι ασθενείς λαμβάνουν Πρεδνιζόνη τις ημέρες 8-28
 - b. **Κλάδος II:** Οι ασθενείς λαμβάνουν Δεξαμεθασόνη (Dexamethasone, DEXA) τις ημέρες 8-28

Οι ασθενείς και στους δύο κλάδους λαμβάνουν επίσης Βινκριστίνη (Vincristine,VCR) και Υδροχλωρική Δαουνορουβικίνη (Daunorubicin hydrochloride, DNR) εβδομαδιαίως στις εβδομάδες 2-5, Ασπαραγινάση (Asparaginase, ASP) τις ημέρες 12-33, Κυκλοφοσφαμίδη (Cyclophosphamide, CPM) τις ημέρες 36 και 64, Κυταραβίνη (Cytarabine, ARA-C) στις εβδομάδες 6-9, Μερκαπτοπουρίνη (Mercaptopurine,MP) τις ημέρες 36-63 και Μεθοτρεξάτη (ενδοραχιαίως) τις ημέρες 12,33,45 και 59 (ασθενείς των οποίων το Κεντρικό Νευρικό Σύστημα έχει προσβληθεί λαμβάνουν επίσης τις ημέρες 18 και 27)
3. **Εξωδιαμερισματική θεραπεία, πρωτόκολλο M:** Οι ασθενείς λαμβάνουν MP τις ημέρες 1-56 και MTX τις ημέρες 8,22,36 και 50
4. **Θεραπεία επανεφόδου/επαναεντατικοποίησης:**

- a. **Κλάδος I (βασική θεραπεία επανεφόδου, πρωτόκολλο II):** οι ασθενείς που κατατάσσονται στις ομάδες κινδύνου SR και MR λαμβάνουν DEXA τις ημέρες 1-22, VCR και Υδροχλωρική Δοξορουβικίνη (Doxorubicin hydrochloride, DOX) στις εβδομάδες 2-5, ASP τις ημέρες 8,11,15 και 18, CPM την 36^η ημέρα, ARA-C και Θειογουανίνη (Thioguanine, TG) τις ημέρες 36-49 και MTX IT τις ημέρες 38 και 45 (ασθενείς των οποίων το Κεντρικό Νευρικό Σύστημα έχει προσβληθεί λαμβάνουν επίσης τις ημέρες 1 και 18). Οι ασθενείς ακολούθως λαμβάνουν θεραπεία συντήρησης.
- b. **Κλάδος II (θεραπεία επανεφόδου μειωμένης έντασης, πρωτόκολλο III):** οι ασθενείς που κατατάσσονται στην ομάδα SR λαμβάνουν DEXA τις ημέρες 1-15, VCR και DOX τις ημέρες 1 και 8, ASP τις ημέρες 1,4,8 και 11, CPM την ημέρα 15, ARA-C και TG τις ημέρες 15-28 και MTX IT τις ημέρες 17 και 24 (ασθενείς των οποίων το Κεντρικό Νευρικό Σύστημα έχει προσβληθεί λαμβάνουν επίσης την ημέρα 1). Οι ασθενείς ακολούθως λαμβάνουν θεραπεία συντήρησης.
- c. **Κλάδος III (θεραπεία επανεφόδου μειωμένης έντασης, δεύτερη καθυστερημένη θεραπεία επανεφόδου/διπλή θεραπεία επαναεντατικοποίησης):** οι ασθενείς που κατατάσσονται στην ομάδα MR λαμβάνουν θεραπεία όπως στον Κλάδο II. Μετά από 10 ημέρες ενδιάμεσης φάσης συντήρησης, επαναλαμβάνεται η θεραπεία. Οι ασθενείς ακολούθως λαμβάνουν θεραπεία συντήρησης.
- d. **Κλάδος IV (βασική θεραπεία επαναεντατικοποίησης):** οι ασθενείς που κατατάσσονται στην ομάδα HR λαμβάνουν δύο ακολούθιες τμημάτων θεραπείας (π.χ. 1,2,3,1,2,3) από το ακόλουθα:
- i. **Τμήμα HR-1:** Οι ασθενείς λαμβάνουν DEXA τις ημέρες 1-5, VCR τις ημέρες 1 και 6, ARA-C εις διπλούν την ημέρα 5, MTX και CPM κάθε 12 ώρες τις ημέρες 2-4 (5 δόσεις), ASP τις ημέρες 6 και 11 και MTX/ARA-C/PRED IT την ημέρα 1.
 - ii. **Τμήμα HR-2:** Οι ασθενείς λαμβάνουν DEXA τις ημέρες 1-5, Βινδεσίνη (vindesine) τις ημέρες 1 και 6, DNR την ημέρα 5, MTX και Ιφωσφαμίδη (ifosfamide) κάθε 12 ώρες τις ημέρες 2-4 (5 δόσεις), ASP τις ημέρες 6 και 11 και MTX/ARA-C/PRED IT την ημέρα 1 (ασθενείς HR των οποίων το Κεντρικό Νευρικό

Σύστημα έχει προσβληθεί λαμβάνουν επίσης IT θεραπεία την ημέρα 5).

- iii. Τμήμα HR-3:** Οι ασθενείς λαμβάνουν DEXA τις ημέρες 1-5, ARA-C κάθε 12 ώρες τις ημέρες 1-2 (4 δόσεις), Ετοποσίδη (etoposide) 5 φορές ημερησίως τις ημέρες 3-5, ASP την ημέρα 6 και 11 και MTX/ARA-C/PRED IT την ημέρα 1.

Στη συνέχεια οι ασθενείς ακολουθούν την θεραπεία του Κλάδου I και τέλος λαμβάνουν θεραπεία συντήρησης.

- e. Κλάδος V (εκτεταμένη θεραπεία επαναινατικοποίησης/τριπλό πρωτόκολλο III):** Οι HR ασθενείς λαμβάνουν τα τμήματα HR3,2,1 όπως στον κλάδο IV και ακολούθως τη θεραπεία επαναινατικοποίησης του κλάδου II με επανάληψη του τμήματος θεραπείας εις διπλούν με μεσολάβηση ενδιάμεσης θεραπείας συντήρησης 4 εβδομάδων. Οι ασθενείς ακολούθως λαμβάνουν θεραπεία συντήρησης.
- f. Θεραπεία ενδιάμεσης συντήρησης/θεραπεία συντήρησης:** Οι ασθενείς λαμβάνουν MTX εβδομαδιαίως και MP ημερησίως μέχρι την εβδομάδα 104.
- g. Ραδιοθεραπεία:** Ασθενείς HR ή ασθενείς με T-λευχαιμία ή ασθενείς των οποίων το Κεντρικό Νευρικό Σύστημα έχει προσβληθεί λαμβάνουν Ραδιοθεραπεία του Κεντρικού Νευρικού Συστήματος.

2. Μηχανική Μάθηση, Βιοπληροφορική, Συστημική Βιολογία και In Silico Ογκολογία

2.1 Εισαγωγή

Στο παρόν Κεφάλαιο παρουσιάζονται οι βασικές αρχές των επιστημονικών χώρων της Μηχανικής Μάθησης της Βιοπληροφορικής και της Συστημικής Βιολογίας όπως επίσης και αυτός της In Silico Ογκολογίας ο οποίος αποτελεί των τομέα εφαρμογής των πρώτων στα πλαίσια της παρούσας διατριβής, Θα πρέπει να σημειωθεί ότι η παρουσίαση που γίνεται στο παρόν Κεφάλαιο είναι όσο το δυνατόν πιο γενική, ενώ οποιεσδήποτε λεπτομέρειες των διαφόρων μεθόδων και τις εφαρμογής τους δίνονται σε αντίστοιχα Κεφάλαια της διατριβής.

2.2 Μηχανική Μάθηση

Η Μηχανική Μάθηση (Machine Learning) αποτελεί τομέα της Επιστήμης των Υπολογιστών και της Πληροφορικής και εντάσσεται στην ευρύτερη επιστημονική περιοχή της Τεχνητής Νοημοσύνης (Artificial Intelligence).

Οι Stuart Russell και Peter Norvig στο βιβλίο τους «Τεχνητή Νοημοσύνη, μια σύγχρονη προσέγγιση» [38] συμπυκνώνουν τον ορισμό της Τεχνητής Νοημοσύνης σε τέσσερις βασικούς άξονες:

1. Στα Συστήματα που σκέπτονται σαν τον άνθρωπο,
2. Στα Συστήματα που ενεργούν σαν τον άνθρωπο,
3. Στα Συστήματα που σκέπτονται ορθολογικά και
4. Στα Συστήματα που ενεργούν ορθολογικά.

Όπως αναφέρουν, ένα σύστημα θεωρείται ορθολογικό αν κάνει «το σωστό», με δεδομένα όσα γνωρίζει.

Βασικές διεργασίες που εκτελούνται στα πλαίσια της μηχανικής μάθησης είναι αυτές της Εξόρυξης Δεδομένων και Γνώσης (Data and Knowledge Mining). Σημαντική συμβολή αυτών είναι η όσο το δυνατόν καλύτερη, αποδοτικότερη και καταλληλότερη συλλογή και επεξεργασία των αναγκαίων δεδομένων που θα συμβάλουν στην αύξηση της ορθολογικότητας των παραπάνω συστημάτων.

2.2.1 Η ανάγκη για Εξόρυξη Δεδομένων και Γνώσης

Στον σημερινό κόσμο οι βάσεις δεδομένων έχουν εξελιχθεί ως ένα από τα σημαντικότερα εργαλεία που υποστηρίζουν την ανάπτυξη του ανθρωπίνου πολιτισμού. Καθημερινά συλλέγονται τεράστιες ποσότητες δεδομένων για ερευνητικούς, ιατρικούς, στατιστικούς και οικονομικούς σκοπούς. Η ύπαρξη αυτού του τεράστιου όγκου δεδομένων και πληροφοριών δημιουργεί την ανάγκη ανάλυσης και ερμηνείας των συνόλων δεδομένων με στόχο την εξαγωγή γνώσης από αυτά. Το πραγματικά πολύ μεγάλο φάσμα επιστημών που ωφελούνται από αυτή την απόκτηση γνώσης, όπως η Τεχνητή Νοημοσύνη, η Στατιστική, οι Βάσεις Δεδομένων, ο Αυτοματισμός, η Ιατρική, η Βιολογία και πολλές άλλες, οδήγησε στην δημιουργία ενός από τους πιο σύγχρονους τομείς της Επιστήμης της Πληροφορικής, αυτόν της Εξόρυξης Γνώσης και Δεδομένων.

2.2.2 Η διαδικασία Εξόρυξης Δεδομένων ως μέρος της διαδικασίας Εξόρυξης Γνώσης

Αν και ο όρος Εξόρυξη Δεδομένων έχει επικρατήσει ως αντιπροσωπευτικός της διαδικασίας ανάλυσης και ερμηνείας δεδομένων, πολλοί ερευνητές καταδεικνύουν την ανάγκη ύπαρξης μιας γενικότερης έννοιας, αυτής της Εξόρυξης ή Ανακάλυψης Γνώσης από Βάσεις Δεδομένων (Knowledge Discovery in Databases – KDD). Έτσι, ο όρος KDD χρησιμοποιείται για την περιγραφή ολόκληρης της διαδικασίας ανάλυσης ενός συνόλου δεδομένων ενώ ο όρος Εξόρυξη Δεδομένων αναφέρεται στις μεθόδους και τις τεχνικές που χρησιμοποιούνται στη διαδικασία της ανάλυσης.

Για τον όρο KDD δίνεται ο εξής ορισμός [39]:

«KDD είναι η ντετερμινιστική διαδικασία αναγνώρισης έγκυρων, καινοτόμων, ενδεχομένως χρήσιμων και εν τέλει κατανοητών προτύπων στα δεδομένα»

Η διαδικασία KDD είναι επαναληπτική και αποτελείται από τα ακόλουθα βήματα τα οποία παρουσιάζονται και συγκεντρωτικά στην Εικόνα 2.1:

1. Την ανάπτυξη και κατανόηση της περιοχής της εφαρμογής, της σχετικά προγενέστερης γνώσης του προς εξέταση τομέα και τους στόχους του τελικού χρήστη.
2. Την ολοκλήρωση των δεδομένων. Υπάρχουν διαφορετικά είδη αποθηκών πληροφοριών που μπορούν να χρησιμοποιηθούν στη διαδικασία Εξόρυξης Γνώσης. Κατά

συνέπεια, οι πολλαπλές πηγές δεδομένων μπορούν να συνδυαστούν καθορίζοντας το σύνολο στο οποίο τελικά η διαδικασία Εξόρυξης Γνώσης πρόκειται να εφαρμοστεί

3. Τη δημιουργία του στόχου-συνόλου δεδομένων. Επιλογή του συνόλου δεδομένων (δηλαδή μεταβλητές, δείγματα δεδομένων) στο οποίο η διαδικασία εξόρυξης πρόκειται να εκτελεστεί.

4. Τον καθορισμό και την προ-επεξεργασία δεδομένων. Αυτό το βήμα περιλαμβάνει βασικές διαδικασίες όπως η αφαίρεση του θορύβου ή των δεδομένων που παρουσιάζονται αναπάντεχα διαφορετικά από τα υπόλοιπα (outliers), η συλλογή των απαραίτητων πληροφοριών για τη διαμόρφωση ή τη μέτρηση του θορύβου, η απόφαση σχετικά με τις στρατηγικές διαχείρισης των ελλειπόντων πεδίων δεδομένων.

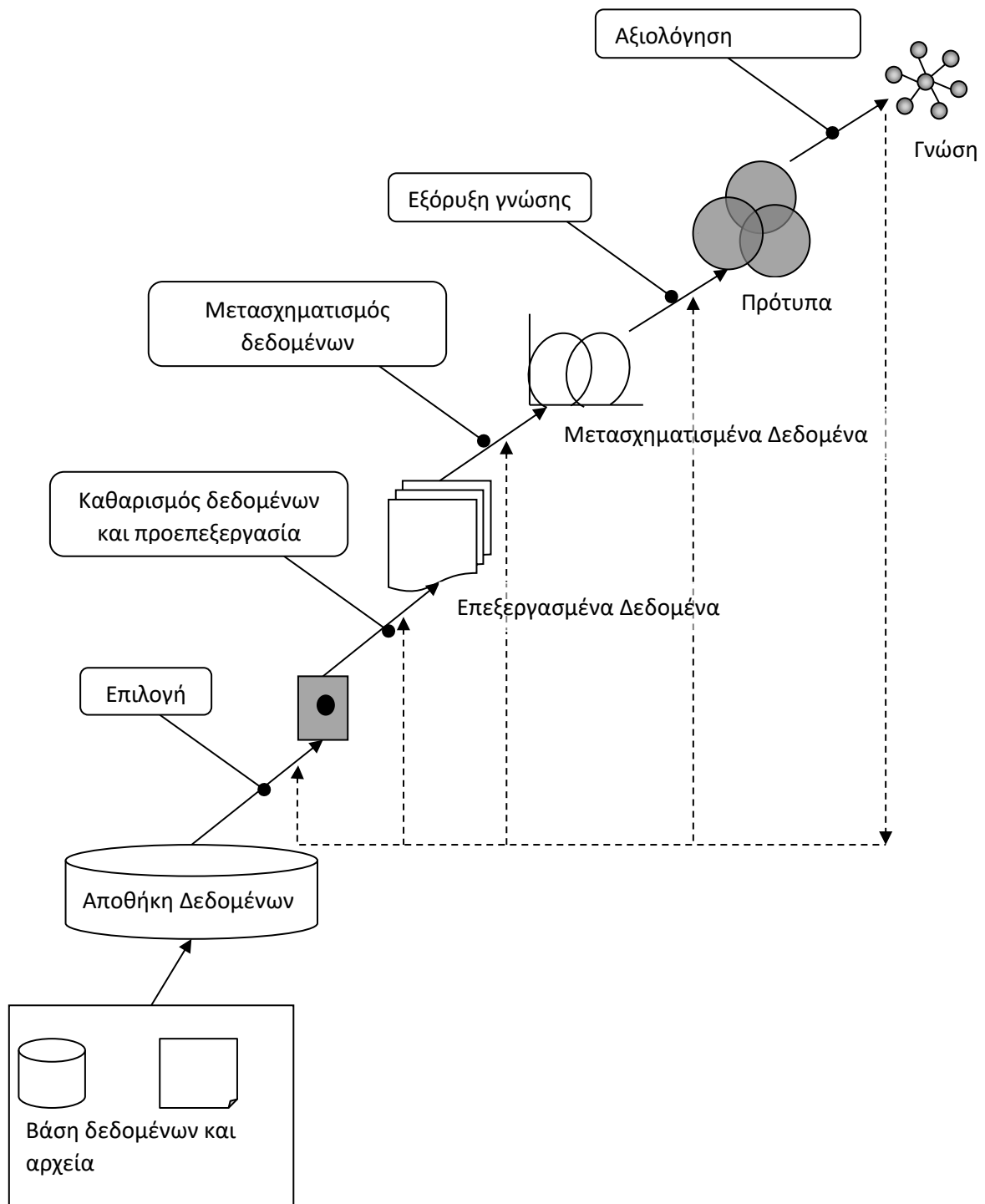
5. Το μετασχηματισμό των δεδομένων. Τα δεδομένα μετασχηματίζονται ή παγιώνονται σε μορφές κατάλληλες για εξόρυξη. Χρήση των μεθόδων μείωσης διαστάσεων ή μετασχηματισμού για τη μείωση του αριθμού των υπό εξέταση μεταβλητών ή την εύρεση κατάλληλης αντιπροσώπευσης των δεδομένων.

6. Την επιλογή των στόχων και των αλγορίθμων Εξόρυξης Δεδομένων. Σε αυτό το βήμα αποφασίζονται οι στόχοι της διαδικασίας KDD, επιλέγοντας τους στόχους Εξόρυξης Δεδομένων προς επίτευξη. Επίσης, επιλέγονται οι μέθοδοι που θα χρησιμοποιηθούν. Αυτό περιλαμβάνει την επιλογή του κατάλληλου μοντέλου και παραμέτρων (π.χ. κατηγορικό ή αριθμητικό μοντέλο δεδομένων). Επίσης η μέθοδος Εξόρυξης Δεδομένων πρέπει να αντιστοιχηθεί με τις απαιτήσεις και τα γενικά κριτήρια της διαδικασίας KDD.

7. Την Εξόρυξη Δεδομένων. Εφαρμόζοντας ευφυείς μεθόδους, αναζητούνται ενδιαφέροντα πρότυπα γνώσης. Τα πρότυπα θα μπορούσαν να είναι μιας συγκεκριμένης αντιπροσωπευτικής μορφής ή ενός συνόλου τέτοιων αντιπροσωπεύσεων, όπως κανόνες Κατηγοριοποίησης-Ταξινόμησης, Δέντρα, Συσταδοποίηση κ.λπ. Η απόδοση και τα αποτελέσματα της μεθόδου Εξόρυξης Δεδομένων εξαρτώνται από τα προηγούμενα βήματα.

8. Την αξιολόγηση των προτύπων. Τα εξαγόμενα πρότυπα αξιολογούνται με κάποια μέτρα, προκειμένου να προσδιοριστούν τα πρότυπα τα οποία αντιπροσωπεύουν τη γνώση, δηλαδή τα αληθινά ενδιαφέροντα πρότυπα.

9. Τη σταθεροποίηση και παρουσίαση της γνώσης. Σε αυτό το βήμα, η εξορυγμένη γνώση ενσωματώνεται στο σύστημα και τεχνικές αντιπροσώπευσης γνώσης χρησιμοποιούνται για να παρουσιάσουν της εξορυγμένη γνώση στο χρήστη. Επίσης ελέγχονται πιθανές συγκρούσεις με προηγούμενη εξορυγμένη γνώση.



Εικόνα 2.1 Βήματα διαδικασίας Εξόρυξης ή Ανακάλυψης Γνώσης από Βάσεις Δεδομένων (KDD) [39]

2.2.3 Η εξόρυξη Δεδομένων

Η διαδικασία της Εξόρυξης Δεδομένων καλείται να επιτελέσει δύο διαφορετικές εργασίες σε μεγάλα σύνολα δεδομένων οι οποίες διαφοροποιούνται ανάλογα με την εφαρμογή των αποτελεσμάτων της. Η εργασία της περιγραφής εντοπίζεται στην εύρεση

προτύπων που παρουσιάζονται στα δεδομένα και μέσω αυτής στην αναπαράσταση των δεδομένων σε μία πιο άμεση και κατανοητή από τον άνθρωπο μορφή. Η εργασία της πρόβλεψης από την άλλη, επικεντρώνεται στην πρόβλεψη της συμπεριφοράς κάποιων μεταβλητών και στην εύρεση της εξάρτησης τους από άλλες. Η δύο αυτές εργασίες αξιοποιούνται από διάφορους τομείς της επιστήμης που αποτελούν τον γενικότερο επιστημονικό χώρο της Τεχνητής Νοημοσύνης. Η περιγραφή κρίνεται ιδιαίτερα σημαντική για την Εξόρυξη Γνώσης ενώ η πρόβλεψη αποτελεί τον κορμό της Αναγνώρισης Προτύπων (Pattern Recognition).

Ένα παράδειγμα που θα μπορούσε να αποσαφηνίσει τη διαφορά μεταξύ πρόβλεψης και περιγραφής είναι το ακόλουθο: έχοντας ένα μεγάλο σύνολο δεδομένων με στοιχεία και χαρακτηριστικά κάποιων ανθρώπων, μέσω της περιγραφής προσπαθούμε να εντοπίσουμε κάποια κοινά χαρακτηριστικά τους (όπως συνήθειες, χαρακτηριστικά σώματος, φύλο κτλ.) και έτσι να αποκτήσουμε επιπλέον γνώση για τις επικρατέστερες περιπτώσεις σε αυτά, π.χ πιο είναι το πιο διαδεδομένο ύψος. Από την άλλη μέσω της πρόβλεψης, μπορούμε να δημιουργήσουμε κάποια προηγούμενη γνώση για την εμφάνιση αυτών των χαρακτηριστικών και να γνωρίζουμε έτσι τις πιθανότητες εμφάνισης κάποιων τιμών για αυτά τα χαρακτηριστικά για έναν νέο άνθρωπο έξω από το υπάρχον σύνολο δεδομένων.

Οι εργασίες που η Εξόρυξη Δεδομένων καλείται να επιτελέσει είναι πολυποίκιλες και έτσι ένας πολύ μεγάλος αριθμός μεθόδων έχει προταθεί για την εξυπηρέτηση των αναγκών των διαφόρων εφαρμογών. Στις ακόλουθες ενότητες γίνεται μία σύντομη περιγραφή των κύριων ομάδων μεθόδων που ακολουθούνται στην Εξόρυξη Δεδομένων.

2.2.3.1 Κατηγοριοποίηση-Ταξινόμηση

Η βασική έννοια που διέπει την εργασία της Κατηγοριοποίησης ή Ταξινόμησης (Classification) είναι αυτή της Κατηγορίας ή Κλάσης (Class). Η διαδικασία της Κατηγοριοποίησης βασίζεται στην ύπαρξη δύο συνόλων αντικειμένων, αυτό της Εκπαίδευσης (Training Set) του οποίου οι κατηγορίες των αντικειμένων είναι από πριν γνωστές και αυτό του οποίου τα αντικείμενα επιθυμούμε να κατηγοριοποιήσουμε. Για λόγους αποδοτικής δημιουργίας του μοντέλου κατηγοριοποίησης χρησιμοποιείται και ένα τρίτο σύνολο αντικειμένων Δοκιμής (Test Set) παρόμοιο με αυτό της εκπαίδευσης.

Τα αντικείμενα αποτελούνται από εγγραφές χαρακτηριστικών, π.χ. για έναν άνθρωπο το ύψος, το βάρος και την ηλικία. Για τα αντικείμενα του συνόλου Εκπαίδευσης και του συνόλου Δοκιμής παρουσιάζεται και ένα ακόμα χαρακτηριστικό αυτό της Κατηγορίας. Για την περίπτωση του ανθρώπου θα μπορούσε να είναι ο χαρακτηρισμός του

δείκτη μάζας σώματος του π.χ. λιποβαρής, κανονικός, υπέρβαρος. Ο στόχος της Κατηγοριοποίησης είναι να δημιουργηθεί βάσει του συνόλου Εκπαίδευσης και μέσω της επανεξέτασής του μέσω του συνόλου Δοκιμής ένα μοντέλο κατηγοριοποίησης τέτοιο ώστε να είναι δυνατή η Κατηγοριοποίηση ενός νέου αντικειμένου, στο παράδειγμα ενός ανθρώπου, για τον οποίο δεν γνωρίζουμε την κατηγορία που ανήκει. Ωστόσο, οι πιθανές κατηγορίες που μπορεί ένα νέο αντικείμενο να ενταχθεί είναι ορισμένες από πριν.

2.2.3.2 Παλινδρόμηση

Η Παλινδρόμηση (Regression) αποτελεί το αντίστοιχο της ταξινόμησης όταν η μεταβλητή προς πρόβλεψη δεν είναι διακριτή (π.χ. οι διάφορες κλάσεις) αλλά συνεχής. Αποτελεί τεχνική για την έρευνα της συσχέτισης μίας εξαρτώμενης μεταβλητής (προς πρόβλεψη) και μίας ή περισσότερων ανεξάρτητων μεταβλητών. Στα πλαίσια της εξόρυξης δεδομένων, μέσω της Παλινδρόμησης οδηγούμαστε στη δημιουργία ενός μοντέλου που αποσκοπεί στην πρόβλεψη της τιμής ενός χαρακτηριστικού λαμβάνοντας υπόψη τις τιμές των υπόλοιπων χαρακτηριστικών. Τα μοντέλα παλινδρόμησης περιλαμβάνουν τις ακόλουθες μεταβλητές:

- Οι άγνωστες παράμετροι συσχέτισης β
- Οι ανεξάρτητες μεταβλητές X .
- Η εξαρτημένη μεταβλητή Y

Απλούστερο παράδειγμα αποτελεί η γραμμική παλινδρόμηση όπου για μία ανεξάρτητη μεταβλητής x_i το μοντέλο λαμβάνει την ακόλουθη μορφή:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

όπου ϵ_i το σφάλμα της πρόβλεψης.

Στόχος της παλινδρόμησης αποτελεί η αναζήτηση των κατάλληλων τιμών για τις παραμέτρους β_0, β_1 κ.ο.κ.

2.2.3.3 Ομαδοποίηση

Σε αντίθεση με την Κατηγοριοποίηση στην εργασία της Ομαδοποίησης (Clustering) δεν υπάρχουν γνωστές κατηγορίες ούτε γενικότερα για το σύνολο των αντικειμένων ούτε για το κάθε ένα ξεχωριστά. Απεναντίας ο στόχος της Ομαδοποίησης είναι να εντοπίσει τις διάφορες κατηγορίες που μπορεί να παρουσιάζονται μεταξύ των αντικειμένων τα οποία υπόκεινται σε Ομαδοποίηση με βάση την ομοιότητα που παρουσιάζουν μεταξύ τους στα διάφορα χαρακτηριστικά τους.

2.2.3.4 Κανόνες Συσχέτισης

Οι Κανόνες Συσχέτισης (Association Rules) ως αρχική και κύρια εφαρμογή έχουν το χώρο της αγοράς και της διαχείρισης του εμπορίου. Για την δημιουργία των Κανόνων Συσχέτισης ορίζεται ένα σύνολο $I = \{i_1, \dots, i_n\}$ κατηγορημάτων και ένα σύνολο δοσοληψιών D , στο οποίο κάθε δοσοληψία T αποτελεί ένα σύνολο από κατηγορήματα ($T \subseteq I$). Μία δοσοληψία περιέχει ένα σύνολο αντικειμένων X αν ισχύει $X \subseteq T$.

Βάσει των παραπάνω, ένας Κανόνας Συσχέτισης είναι μία συσχέτιση της μορφής $X \rightarrow Y$, όπου $X \subseteq I, Y \subseteq I$ και $X \cap Y = \emptyset$. Επίσης, ορίζεται ότι ένας κανόνας συσχέτισης $X \rightarrow Y$ ισχύει σε ένα σύνολο δοσοληψιών D με εμπιστοσύνη c , αν το $c\%$ των δοσοληψιών που περιέχουν το X περιέχουν επίσης και το Y . Τέλος ένας κανόνας ορίζεται ότι έχει υποστήριξη s , αν το $s\%$ των δοσοληψιών στο D περιέχουν το $X \cup Y$.

Ανακαλύπτοντας από τα δεδομένα τέτοιες μορφές συσχέτισης, μία επιχείρηση για παράδειγμα που θα εντοπίσει μία συσχέτιση *προϊόν1* \rightarrow *προϊόν2* με εμπιστοσύνη 60% μπορεί να τοποθετήσει το *προϊόν2* σε μία θέση του καταστήματος πιο κοντά στο *προϊόν1* ώστε να αυξήσει τις πωλήσεις του.

2.2.3.5 Πρότυπα Ακολουθιών

Μέσω της εργασίας Εξόρυξης Προτύπων Ακολουθιών, αναζητούνται τα συχνά εμφανιζόμενα πρότυπα που υπάρχουν στα δεδομένα τα οποία σχετίζονται με το χρόνο ή άλλες ακολουθίες. Διατηρώντας τους ορισμούς για την εμπιστοσύνη και την υποστήριξη από την αμέσως προηγούμενη ενότητα, για την διαδικασία εύρεσης προτύπων μπορεί να δοθεί η εξής περιγραφή [39]:

Λαμβάνοντας υπόψη ένα ενδεχομένως μεγάλο πρότυπο (συμβολοσειρά) S , ενδιαφερόμαστε για τα πρότυπα ακολουθιών της μορφής $a \rightarrow b$, όπου τα a, b, ab είναι υποσυμβολοσειρές μέσα στο S , τέτοιες ώστε η συχνότητα του ab να μην είναι μικρότερη από κάποια ελάχιστη υποστήριξη και η πιθανότητα ότι το a ακολουθείται αμέσως από το b να μην είναι μικρότερη από την ελάχιστη εμπιστοσύνη.

Τα πρότυπα ακολουθιών βρίσκουν πολλές σημαντικές εφαρμογές εκ των οποίων οι σημαντικότερες είναι η ανάλυση κειμένων, δεδομένων καιρού, ροών δεδομένων αλλά και η ανάλυση των ακολουθιών του DNA και των αρχείων ιατρικού ιστορικού.

2.2.3.6 Ομοιότητα Χρονολογικών Σειρών

Ως Χρονολογική Σειρά, ορίζεται μία ακολουθία αριθμών όπου κάθε ένας από αυτούς χαρακτηρίζεται από μία ετικέτα χρόνου. Παραδείγματα Χρονολογικών Σειρών συναντώνται

καθημερινά σε διάφορους τομείς. Το ζητούμενο για τις Χρονολογικές Σειρές μέσω της Εξόρυξης Δεδομένων είναι η δημιουργία τεχνικών και μοντέλων που επιτρέπουν την εύρεση της ομοιότερης Χρονολογικής Σειράς που υπάρχει σε μία βάση δεδομένων, δοθείσας μίας νέας Χρονολογικής Σειράς.

2.2.3.7 Απεικόνιση και Μείωση Διαστάσεων

Μέσω των τεχνικών Απεικόνισης και Μείωσης Διαστάσεων επιχειρείται η προβολή των δεδομένων n διαστάσεων (χαρακτηριστικών, μεταβλητών) σε k διαστάσεις όπου $k \ll n$ με κύρια προϋπόθεση την όσο το δυνατόν μεγαλύτερη διατήρηση των αποστάσεων-διαφορών μεταξύ των διαφόρων αντικειμένων (ή εγγραφών) που υπάρχουν στο αρχικό σύνολο δεδομένων.

Μέσω της συγκεκριμένης διαδικασίας επιτυγχάνεται η συμπίεση των δεδομένων ενώ επίσης διευκολύνεται η κατανόηση και την απεικόνιση των πολυδιάστατων δεδομένων. Ακόμα, πολλές Τεχνικές Εξόρυξης Δεδομένων μπορούν να επωφεληθούν από τη μείωση διαστάσεων δεχόμενες ένα απλοποιημένο σύνολο δεδομένων που απορρίπτει όσο το δυνατόν λιγότερες πληροφορίες. Τέλος, οι Τεχνικές Μείωσης Διαστάσεων συμβάλουν σημαντικά στην βελτίωση της απόδοσης διαφόρων τεχνικών ευρετηρίασης.

2.2.3.8 Μάθηση με ή χωρίς επίβλεψη και Αντεπικύρωση

Συνοψίζοντας τα όσα παρουσιάστηκαν προηγουμένως, ο κυριότερος διαχωρισμός των μεθόδων που παρουσιάστηκαν στην προηγούμενη ενότητα γίνεται μεταξύ των μεθόδων μάθησης με επίβλεψη (supervised learning) και μάθησης χωρίς επίβλεψη (unsupervised learning).

Στην πρώτη περίπτωση, επιχειρείται η εξαγωγή ή μάθηση μια σχέσης ή συνάρτησης που μπορεί να περιγράψει τον μηχανισμό με τον οποίο παράγονται τα παρατηρούμενα δεδομένα. Η συνάρτηση αυτή χρησιμοποιείται για την πρόβλεψη της τιμής μιας εξαρτημένης μεταβλητής δοθέντων των τιμών ενός συνόλου ανεξάρτητων μεταβλητών ή χαρακτηριστικών. Τα δύο κύρια προβλήματα της μάθησης με επίβλεψη είναι αυτό της ταξινόμησης (Classification), με στόχο την δημιουργία συναρτήσεων ικανών να ταξινομήσουν δείγματα σε διάφορες, ήδη καθορισμένες, κατηγορίες ή κλάσεις, παρατηρώντας ένα σύνολο χαρακτηριστικών τους και αυτό της παλινδρόμησης ή παρεμβολής (regression) με στόχο την δημιουργία συναρτήσεων που μπορούν να προβλέψουν την αριθμητική τιμή μιας εξαρτημένης μεταβλητής δοθέντων των τιμών των χαρακτηριστικών (ανεξάρτητων μεταβλητών).

Όσον αφορά τη δεύτερη περίπτωση, αυτή της μάθησης χωρίς επίβλεψη, το σύστημα εδώ καλείται, παρατηρώντας μόνο τις τιμές των χαρακτηριστικών, να εξάγει/ανακαλύψει ομάδες/κατηγορίες στο σύνολο δεδομένων, χωρίς να είναι γνωστή η ύπαρξη αλλά και ο αριθμός τους. Κυριότερη κατηγορία μεθόδων της μάθησης χωρίς επίβλεψη αποτελούν οι μέθοδοι ομαδοποίησης (clustering).

Μια ευρέως διαδεδομένη μεθοδολογία ανάλυσης της αποδοτικότητας των διαφόρων μεθόδων στο να εξάγουν την εν λόγω συνάρτηση, δοθέντος ενός συνόλου δεδομένων με γνωστές τιμές για τα χαρακτηριστικά και την εξαρτημένη μεταβλητή είναι αυτή της Αντεπικύρωσης ή Διασταυρούμενης Επιβεβαίωσης (Cross Validation) η οποία χρησιμοποιείται συχνά και στα πλαίσια της παρούσας διατριβής (και παρουσιάζεται λεπτομερώς στα επιμέρους Κεφάλαια). Μέσω της συγκεκριμένης μεθοδολογίας, το διαθέσιμο σύνολο δεδομένων διαιρείται σε δύο τμήματα, σε αυτά του Συνόλου Εκπαίδευσης (Training Set) και του Συνόλου Δοκιμής (Test Set). Το πρώτο σύνολο χρησιμοποιείται για την εκπαίδευση του συστήματος (δηλαδή για την εξαγωγή της συνάρτησης), παρατηρώντας τις τιμές των χαρακτηριστικών αλλά και της εξαρτημένης μεταβλητής, ενώ το δεύτερο σετ χρησιμοποιείται για την αξιολόγηση της απόδοσης του συστήματος στο να προβλέπει την κλάση ή την τιμή της εξαρτημένης μεταβλητής για νέα δείγματα για τα οποία η τιμή της αποκρύπτεται μέχρι να εξαχθεί η σχετική πρόβλεψη.

2.3 Βιοπληροφορική

Η Βιοπληροφορική είναι ο επιστημονικός τομέας εκείνος ο οποίος έχει ως κύριο στόχο την αξιοποίηση υπαρχόντων και την ανάπτυξη νέων μεθόδων Πληροφορικής, με στόχο την εξαγωγή, την διαχείριση και την ανάλυση Βιολογικής Πληροφορίας. Στις ακόλουθες ενότητες παρουσιάζεται αρχικά μία ιστορική αναδρομή στην βιολογική έρευνα της οποίας οι εξελίξεις οδήγησαν στην δημιουργία του τομέα της Βιοπληροφορικής.

2.3.1 Η βιολογική έρευνα μέχρι και την αποκωδικοποίηση του γονιδιώματος

Αν και η μελέτη της έμβιας φύσης, των κανόνων αυτής αλλά και των διαφόρων ειδών αναφέρεται από αρχαιότατων χρόνων στην ανθρώπινη ιστορία (με σημαντικές αναφορές στην Αιγυπτιακή και την Κινέζικη Ιατρική, τις μελέτες του Αριστοτέλη για την ταξινόμηση των ειδών, την συνεισφορά του Ιπποκράτη και του Γαληνού στην Ιατρική και τη Φαρμακολογία αλλά τα σημαντικά επιτεύγματα των Μουσουλμάνων Φυσιολόγων [40]), είναι κοινώς αποδεκτό ότι από τον 19ο αιώνα η επιστήμη της βιολογίας, που πλέον ορίζεται

ως διακριτός επιστημονικός χώρος, έχει να επιδείξει ιδιαίτερα σημαντικά επιτεύγματα τα οποία ραγδαία άλλαξαν τον τρόπο με τον οποίο η ανθρωπότητα αντιλαμβάνεται τη φύση αλλά και τον ίδιο της τον εαυτό. Στις ακόλουθες παραγράφους επιχειρείται να δοθούν τα σημαντικότερα από αυτά τα επιτεύγματα τα οποία καθόρισαν την εξέλιξη της βιολογικής έρευνας στον 20ο και στον 21ο αιώνα, όπως παρουσιάζονται στην [41].

Στην διάρκεια του 19ου αιώνα πραγματοποιήθηκαν τα βασικά βήματα έτσι ώστε να καθοριστεί η οπτική με την οποία αντιλαμβανόμαστε σήμερα την βιολογία. Το 1859 ο Charles Darwin δημοσιεύει το βιβλίο του «Η Καταγωγή των Ειδών» (On the Origin of Species) [42]. Το έργο αυτό του Δαρβίνου εισήγαγε την θεωρία ότι οι πληθυσμοί εξελίσσονται από γενιά σε γενιά με τη διαδικασία της φυσικής επιλογής. Ουσιαστικά πρότεινε τη θεωρία της εξέλιξης υποστηρίζοντας ότι κάτι τέτοιο ισχύει και για τον άνθρωπο. Μία άλλη σημαντική συνεισφορά έγινε από τον Gregor Mendel ο οποίος ουσιαστικά όρισε τον επιστημονικό τομέα της γενετικής (genetics) [43], αποδεικνύοντας μέσω των πειραμάτων του σε φασόλια, ότι η κληρονομικότητα ακολουθεί συγκεκριμένα μοτίβα, τα οποία πλέον αποτελούν τους νόμους για την κληρονομικότητα του Mendel [44]. Το 1869 ο Friedrich Miescher απομόνωσε για πρώτη φορά υλικό DNA, στην προσπάθειά του να μελετήσει την χημεία των κυττάρων. Το υλικό που απομόνωσε ήταν πλούσιο σε φώσφορο και του έδωσε το όνομα νουκλεΐνη (nuclein). Στις αρχές του 20ου αιώνα τον ακολούθησαν πολύ ερευνητές, ανακαλύπτοντας τις χημικές ιδιότητες του DNA σε βάθος. Τέλος, μία ιδιαίτερα σημαντική βιολογική ανακάλυψη του 19ου αιώνα ήταν η περιγραφή της μίτωσης κατά τον πολλαπλασιασμό των κυττάρων, το 1879 από τον Walter Flemming. Όπως είναι ξεκάθαρο, ήδη από τον 19ο αιώνα η βιολογική μελέτη είχε εστιαστεί στην μελέτη του υπο-κυττάρου (sub-cellular) χώρου.

Ο 20ος αιώνας, τώρα, έχει να επιδείξει σημαντικές εξελίξεις στο χώρο της Βιολογίας οι οποίες σε καμία περίπτωση δεν μπορούν να συμπεκνωθούν σε ένα εισαγωγικό κείμενο, όπως το παρόν. Ωστόσο επιχειρείται να δοθούν οι κυριότερες από αυτές. Το 1902, ο Βρετανός φυσιολόγος Archibald Garrod, υποστήριξε ότι η ασθένεια Αλκαπτονουρία (alkaptonuria) κληρονομείται και μάλιστα η κληρονομικότητα αυτή ακολουθεί τους νόμους του Mendel. Η συγκεκριμένη ασθένεια ήταν από τις πρώτες που περιγράφηκε να έχει γενετικής φύσεως αιτίες (κάτι που επιβεβαιώθηκε πλήρως την δεκαετία του 1990). Την ίδια χρονιά επίσης ο Walter Sutton, διαπίστωσε ότι κατά την μείωση, τα προκύπτοντα κύτταρα (σπερματοζωάρια και ωάρια) λαμβάνουν μόνο αντίγραφο του κάθε τύπου χρωμοσώματος, ενώ το πρότυπο διαχωρισμού αυτών βρέθηκε να συμφωνεί με τα αντίστοιχα πρότυπα του Mendel. Εξέχουσα σημασία συμπέρασμα εξήχθη το 1911 από τον Thomas Hunt Morgan

και την ομάδα του, οι οποίοι απέδειξαν ότι τα γονίδια, τα οποία αποτελούν τμήματα των χρωμοσωμάτων, είναι οι υπεύθυνες μονάδες για την κληρονομικότητα, ορίζοντας πλέον την σύγχρονη βάση της γενετικής (για την οποία το 1933 τιμήθηκε με το βραβείο Nobel).

Περνώντας στη δεκαετία του 1940, τα δύο σημαντικότερα βιολογικά ευρήματα έχουν να κάνουν πρώτον με την πρώτη απεικόνιση του DNA, το 1943, μέσω X-Ray από τον William Astbury και με την ανακάλυψη της Barbara McClintock το 1944 ότι τα γονίδια μπορούν να τροποποιήσουν τη θέση τους στα χρωμοσώματα, αποδεικνύοντας τη δυναμική φύση του γονιδιώματος.

Η δεκαετία του 1950 ήταν περίοδος με ραγδαίες εξελίξεις. Το 1953 οι Francis Crick και James Watson περιέγραψαν την δομή του DNA ως διπλή έλικα ανοίγοντας νέους δρόμους στη βιολογική έρευνα ενώ το 1955 ο Joe Hin Tjio απέδειξε τον ακριβή αριθμό των ανθρώπινων χρωμοσωμάτων, που είναι 46. Την ίδια χρονιά ο Arthur Kornberg και η ομάδα του κατάφεραν να απομονώσουν την DNA πολυμεράση, ένα ένζυμο που κατέχει κεντρικό ρόλο στην αντιγραφή του DNA και είναι απαραίτητο για την επιτυχή δημιουργία ανασυνδυασμένου DNA αλλά και την αποκωδικοποίηση του γονιδιώματος.

Περνώντας στη δεκαετία του 1960, το 1961 οι Sydney Brenner, Francois Jacob, και Matthew Meselson ανακάλυψαν ότι το mRNA είναι το μόριο που μεταφέρει την πληροφορία που βρίσκεται στο DNA από τον πυρήνα στους μηχανισμούς παραγωγής πρωτεϊνών ενώ το 1966 ο μηχανισμός της μετάφρασης του αλφαβήτου των 4 γραμμάτων των DNA-RNA στα 20 διαφορετικά αμινοξέα ανακαλύφθηκε από τους Marshall Nirenberg, Har Khorana και Severo Ochoa.

Η δεκαετία του 1970, μεταξύ άλλων είχε να επιδείξει δύο ιδιαίτερα σημαντικά επιτεύγματα. Τις χρονιές 1975 και 1977 ο Sanger και οι συνεργάτες του αλλά και οι Maxam και Gilbert εξέλιξαν μεθόδους αποδοτικής και ταχείας αλληλούχισης του DNA. Επίσης το 1977 στα εργαστήρια των Richard Roberts και Phil Sharp ανακαλύφθηκε η ύπαρξη των εσονίων στα γονίδια των ευκαρυωτικών οργανισμών, δηλαδή των περιοχών των γονιδίων που δεν συμμετέχουν στον σχηματισμό πρωτεϊνών και αποκόπτονται κατά τη διαδικασία της μεταγραφής.

Περνώντας στην δεκαετία 1980, συνέβησαν κομβικές εξελίξεις που καθόρισαν τις μεθόδους βιοτεχνολογίας των επομένων ετών αλλά και εισήγαγαν τη χρήση των υπολογιστών και της επιστήμης της Πληροφορικής ως όλο και περισσότερο αναγκαίο εργαλείο για την διαχείριση της πολυπλοκότητας των βιολογικών προβλημάτων. Το 1981-82 επετεύχθη η δημιουργία των πρώτων διαγονιδιακών θηλαστικών (ποντίκι) και εντόμων (μύγα), εξέλιξη η οποία άλλαξε σε πολύ μεγάλο βαθμό την βιολογική έρευνα. Το 1982,

δημιουργήθηκε για πρώτη φορά μία βάση δεδομένων με δεδομένα ακολουθίας DNA, η GenBank, στην οποία κατατίθεντο ακολουθίες από διάφορους μηχανισμούς. Έτσι, ξεκίνησε η συμβολή της πληροφορικής στη διαχείριση της βιολογικής πληροφορίας η οποία επέτρεψε την διαχείρισή της και την εξαγωγή γνώσης από αυτή. Το 1983, προετοιμάστηκε το έδαφος για την ιδιαίτερα συνηθισμένη στις μέρες μας συσχέτιση ασθενειών με γονίδια, με την πρώτη σύνδεση της ασθένειας Huntington με γενετικό βιοδείκτη στο χρωμόσωμα 4. Την ίδια χρονιά ανακαλύφθηκε η τεχνική PCR (Polymerase Chain Reaction) [45], η οποία επιτρέπει την δημιουργία πολλαπλών αντιγράφων ενός τμήματος DNA, κάτι που άλλαξε ριζικά την βιοτεχνολογική έρευνα. Συνεχίζοντας, το 1986, εντοπίστηκε το πρώτο γονίδιο με τη χρήση της μεθόδου Κλωνοποίηση Θέσης (Positional cloning) [46], μέσω της οποίας μπορεί να εντοπιστεί ένα γονίδιο χωρίς καμία γνώση για την πρωτεΐνη που αυτό κωδικοποιεί. Η συγκεκριμένη διαδικασία, άλλαξε πλήρως μετά την αποκωδικοποίηση του DNA και πλέον πραγματοποιείται απευθείας στα διαθέσιμα δεδομένα των βάσεων δεδομένων. Τέλος, την χρονιά 1987, δημιουργήθηκε ο πρώτος Γενετικός Χάρτης του ανθρώπου [47], αναδεικνύοντας τις περιοχές των χρωμοσωμάτων που έχουν την τάση να κληρονομούνται μαζί.

Από την δεκαετία του 1990 και εντεύθεν, παρουσιάστηκαν καταγιστικές εξελίξεις στην βιολογική και βιοτεχνολογική έρευνα οι οποίες λίγο έως πολύ καθορίστηκαν από την έναρξη του Προγράμματος Αποκωδικοποίησης του ανθρωπίνου γονιδιώματος (Human Genome Project) το 1990. Οι κεντρικοί στόχοι του προγράμματος ήταν η επιτυχής ανάγνωση των 3,2 δισεκατομμυρίων βάσεων στο DNA και η κατάλληλη χαρτογράφηση του ανθρωπίνου γονιδιώματος. Μαζί με αυτούς τους πρωτεύοντες στόχους, τέθηκαν οι επίσης σημαντικοί και αναγκαίοι στόχοι, της ανάπτυξης της κατάλληλης τεχνολογίας για την ανάλυση του DNA αλλά και η μελέτη των ηθικών, νομικών και κοινωνικών ζητημάτων που αφορούν την γενωμική έρευνα. Στα επόμενα χρόνια υπήρξαν συνεχείς παρουσιάσεις ανάγνωσης γονιδιωμάτων ολόκληρων οργανισμών αλλά και συνεχής βελτίωση των γενετικών χαρτών του ανθρώπου. Σημαντική στιγμή για την αναγκαία συνεργασία της Πληροφορικής και της Βιολογίας-Βιοτεχνολογίας, αποτελεί ο καθορισμός των βασικών οδηγιών για την κατάλληλη αποθήκευση και δημοσίευση των αποτελεσμάτων αποκωδικοποίησης του DNA κατά το Human Genome Project, το 1997, ανοίγοντας έτσι το δρόμο για το πολύτιμο εργαλείο των βάσεων δεδομένων με βιολογική-γενωμική πληροφορία η οποία το 2000 αποφασίστηκε να παραμείνει ελεύθερη προς πρόσβαση για τον οποιοδήποτε.

Ιδιαίτερα σημαντική στιγμή για την νέα εποχή της γενωμικής, αποτέλεσε η αποδοχή το 2001 από τον FDA (Food and Drug Administration), φαρμάκου που βασιζόμενο σε γενετική πληροφορία μπορεί να στοχεύει μόνο τα καρκινικά κύτταρα στην χρόνια μυελώδη λευχαιμία, ενώ το 2002 ανακαλύφθηκε γενετική συσχέτιση για τον καρκίνο του προστάτη.

Με την ολοκλήρωση του Human Genome Project το 2003, πλέον στα χέρια της επιστημονικής κοινότητας διατίθετο μία άνευ προηγουμένου πληθώρα πληροφοριών, σε συνδυασμό με όλα τα ευρήματα των περασμένων δεκαετιών. Ωστόσο, η ποσότητα της πληροφορία αυτή απαιτήσε και απαιτεί κατάλληλες μεθόδους διαχείρισης, επεξεργασίας και ανάλυσής της. Έτσι στα χρόνια που ακολούθησαν την ολοκλήρωση του Human Genome Project η εφαρμογή μεθόδων Πληροφορικής στον χώρο της Βιολογίας, γιγαντώθηκε και αποτέλεσε αναγκαιότητα για την αξιοποίηση της διαθέσιμης πληροφορίας. Η συμβολή της Πληροφορικής στην μελέτη της Βιολογίας και των φαινομένων που συμβαίνουν σε αυτή αναλύεται στις ακόλουθες ενότητες.

2.3.2 Η συμβολή της Πληροφορικής στην αποκωδικοποίηση και στην ανάλυση του γονιδιώματος. Η γέννηση της Βιοπληροφορικής

Η αποκωδικοποίηση ενός γονιδιώματος, όπως του ανθρώπινου που έγινε μέσω του Human Genome Project, όπως είναι αναμενόμενο, είναι μία ιδιαίτερα σύνθετη και δύσκολη διαδικασία η οποία απαιτεί σημαντικές προσπάθειες, όχι μόνο στην ανάγνωση της ακολουθίας των βάσεων αλλά και στη μετέπειτα ανάλυση αυτών. Πέραν του κεντρικού ρόλου που έπαιξε η εξέλιξη της βιοτεχνολογίας στον δρόμο προς την επίτευξη του παραπάνω στόχου, η επιστήμη που θα μπορούσε να συμβάλει αποφασιστικά ήταν η αυτή της Πληροφορικής έχοντας ήδη αναπτύξει μεθόδους διαχείρισης και ανάλυσης δεδομένων αλλά και εξαγωγής γνώσης και πληροφορίας από διάφορου είδους σήματα όπως οι εικόνες και πολλά άλλα ήδη μετρήσεων. Έτσι, ταυτόχρονα με την επανάσταση της γενωμικής, γεννήθηκε ο επιστημονικός τομέας της Βιοπληροφορικής (Bioinformatics) [48].

Ένας από τους πρώτους και βασικούς ρόλους των μεθόδων Βιοπληροφορικής στην επίτευξη του στόχου της αποκωδικοποίησης του γονιδιώματος, ήταν η κατάλληλη διαχείριση του σήματος που εξάγεται από τις διάφορες βιοτεχνολογικές μεθόδους που έχουν ως στόχο την ανάγνωση της ακολουθίας του DNA (όπως περιγράφονται στην [49]). Χωρίς την συμβολή των μεθόδων Πληροφορικής, θα ήταν αδύνατο όλη αυτή η τεράστια σε ποσότητα και πολυπλοκότητα πληροφορία να διαχειριστεί και να εξαχθεί η ζητούμενη

γνώση από αυτή. Ωστόσο, η Βιοπληροφορική έπαιξε ακόμα σημαντικότερο ρόλο στην εξακρίβωση του περιεχομένου της αναγνωσμένης πληροφορίας.

Ένας σύντομος ορισμός για τον επιστημονικό τομέα της Βιοπληροφορικής μπορεί να δοθεί ως ακολούθως:

Η Βιοπληροφορική είναι ο επιστημονικός τομέας εκείνος ο οποίος έχει ως κύριο στόχο την αξιοποίηση υπαρχόντων και την ανάπτυξη νέων μεθόδων Πληροφορικής, με στόχο την εξαγωγή, την διαχείριση και την ανάλυση Βιολογικής Πληροφορίας.

Οι κυριότερες περιοχές της Πληροφορικής οι οποίες συνέβαλλαν και συμβάλουν στην δημιουργία των μεθόδων Βιοπληροφορικής είναι αυτές της Αναγνώρισης Προτύπων (Pattern Recognition) [50], της Εξαγωγής Γνώση από Δεδομένα (Data Mining) [51], της Μηχανικής Μάθησης (Machine Learning) [52] και της Οπτικοποίησης [53] Δεδομένων (Visualization). Ωστόσο, σημαντική είναι και η συνεισφορά άλλων τομέων όπως των γλωσσών προγραμματισμού (με αρκετές γλώσσες να αναπτύσσονται σχεδόν κατά αποκλειστικότητα για την Βιοπληροφορική Ανάλυση), των βάσεων δεδομένων (για την αποθήκευση δεδομένων και αναλύσεων αυτών), της ασφάλειας (για την ασφαλή φύλαξη και μεταφορά των δεδομένων) και πολλές άλλες.

Οι σημαντικότερες μέθοδοι Βιοπληροφορικής οι οποίες συνέβαλαν σημαντικά στην ολοκλήρωση του Human Genome Project και συνεχίζουν να υποστηρίζουν την βιολογική έρευνα συνοψίζονται ακολούθως:

- Διάταξη Γονιδιώματος (Genome Assembly) [54]: Συνδυασμός μικρών ακολουθιών DNA ώστε να ανασυντεθεί η σειρά με την οποία βρίσκονται στα χρωμοσώματα από τα οποία και προέρχονται.
- Στοιχίση ακολουθιών (Sequence Alignment) [55]: Κατάλληλη διάταξη δύο ή περισσότερων ακολουθιών DNA, RNA ή αμινοξέων πρωτεΐνης ώστε να εντοπιστούν περιοχές ομοιότητας που μπορεί να οφείλονται σε λειτουργικές, δομικές ή εξελικτικές σχέσεις μεταξύ των ακολουθιών.
- Χαρακτηρισμός Γονιδιώματος (Genome Annotation) [48]: Εύρεση των περιοχών στο DNA που αποτελούν γονίδια ή έχουν άλλα βιολογικά χαρακτηριστικά.
- Στοιχίση Δομών Πρωτεϊνών (Protein Structure Alignment) [56]: Εύρεση πιθανής ομολογίας μεταξύ πρωτεϊνών και RNA πραγματοποιώντας συγκρίσεις σε επίπεδο τριδιάστατης δομής και όχι σε επίπεδο ακολουθίας αμινοξέων.

- Πρόβλεψη Δομών Πρωτεϊνών (Protein Structure Prediction) [57]: Πρόβλεψη της τρισδιάστατης δομής μίας πρωτεΐνης μέσω της ακολουθίας των αμινοξέων από τα οποία απαρτίζεται.
- Μέτρηση, Ανάλυση και Πρόβλεψη της έκφρασης γονιδίων (Gene Expression Measure, Analysis and Prediction)
 - Ανάλυση των ρυθμιστικών λειτουργιών [48]: Ανάλυση όλων εικόνων των φαινομένων που μπορούν να οδηγήσουν στην ρύθμιση της έκφρασης γονιδίων σε ένα κύτταρο. Παράδειγμα αποτελεί η ανάλυση υποκινητών, δηλαδή ο εντοπισμός μοτίβων ακολουθιών στην γειτονική περιοχή των γονιδίων που ρυθμίζουν την έκφραση αυτού.
 - Εύρεση των αλληλεπιδράσεων μεταξύ Πρωτεϊνών (Protein-Protein Interactions) [58]: Εντοπισμός των πρωτεϊνών που πρέπει να δημιουργήσουν σύμπλοκο μεταξύ τους ώστε να πραγματοποιήσουν τη βιολογική τους λειτουργία.
 - Μελέτες Συσχέτισης σε κλίμακα Γονιδιώματος (Genome Wide association studies) [59]: Μελέτη του γονιδιώματος διαφορετικών ατόμων του ίδιου είδους με στόχο την εξακρίβωση των μεταξύ τους διαφορών αλλά και τη συσχέτιση αυτών με διάφορα γνωρίσματα, όπως οι ασθένειες.

Μέσω των παραπάνω μεθόδων, οι οποίες αποτελούν ένα υποσύνολο των μεθόδων Βιοπληροφορικής, τις τελευταίες δύο δεκαετίες αποκαλύφθηκαν πολλές ιδιότητες του γονιδιώματος αλλά και γενικότερα της βιολογίας με τον εντοπισμό διαφόρων βιομοριών. Ωστόσο, οι μορφές αναλύσεων που περιγράφηκαν αλλά και τα δεδομένα που χρησιμοποιούν αυτές, δεν παύουν να αποτελούν στιγμιότυπα των μελετώμενων βιολογικών διεργασιών οι οποίες είναι δυναμικές διαδικασίες. Επίσης σε πολλές περιπτώσεις, οι αποκτούμενες πληροφορίες έχουν αποσπασματικό χαρακτήρα και δεν είναι συσχετισμένες με άλλες βιολογικές διεργασίες και τα σχετικά εμπλεκόμενα βιομόρια. Έτσι, διαμορφώθηκε και συνεχίζει να υφίσταται η ανάγκη ανάπτυξης μεθοδολογιών που θα επιτρέπουν την μελέτη των βιολογικών διαδικασιών κατά την δυναμική τους εξέλιξη στο χρόνο αλλά και θα συνδυάζουν όλες τις διαθέσιμες πληροφορίες. Για την κάλυψη της ανάγκης αυτής, διαμορφώθηκε και συνεχώς εξελίσσεται, ο επιστημονικός χώρος της Συστημικής Βιολογίας (Systems Biology), ο οποίος αποτελεί αντικείμενο των επομένων δύο ενοτήτων αλλά και μεγάλος μέρος της παρούσας διατριβής.

2.3.3 Από την Βιοπληροφορική στη Συστημική Βιολογία

Όπως αναφέρθηκε στην προηγούμενη ενότητα, παρά την σημαντική συνεισφορά των Μεθόδων Βιοπληροφορικής στην ανακάλυψη όλων των σημαντικών πληροφοριών που παρουσιάστηκαν προηγουμένως, τα ευρήματα αυτά δεν παύουν να αποτελούν αποσπασματικές πληροφορίες από το σύνολο των βιολογικών συστημάτων. Όπως αναφέρεται στην [60], ο εντοπισμός όλων των γονιδίων, των πρωτεϊνών και των άλλων βιομορίων ενός οργανισμού ισοδυναμεί με την τοποθέτηση σε μία λίστα όλων των εξαρτημάτων ενός μηχανήματος. Παρότι αυτή η λίστα μπορεί να προσφέρει σημαντική πληροφόρηση, δεν είναι αρκετή ώστε να κατανοηθεί σε όλη του την πολυπλοκότητα ο τρόπος λειτουργίας του μηχανήματος και κάτι τέτοιο απαιτεί την γνώση του τρόπου με τον οποίο τα διάφορα εξαρτήματα συνδέονται μεταξύ τους ώστε να δημιουργήσουν την τελική δομή και λειτουργικότητα του μηχανήματος. Κατά ανάλογο τρόπο, η εξαντλητική δημιουργία δομών βιολογικής πληροφορίας, όπως τα δίκτυα ρύθμισης γονιδίων (gene regulatory networks) αλλά και η γνώση των διαφόρων βιοχημικών αλληλεπιδράσεων που συμβαίνουν, δεν είναι αρκετά ώστε να εξαχθεί γνώση για την επιρροή που έχει το ένα τμήμα του συστήματος στο άλλο. Όπως προτείνεται, λοιπόν, στην [60], θα πρέπει η βιολογική έρευνα να κατευθυνθεί προς την κατεύθυνση μελέτης Συστημάτων και η σχετική γνώση για αυτά μπορεί να προκύψει από τους ακόλουθους τομείς:

1. Δομές Συστημάτων: Δίκτυα αλληλεπίδρασης γονιδίων και βιοχημικά βιολογικά μονοπάτια, όπως επίσης και οι μηχανισμοί που ορίζουν τις φυσικές ιδιότητες των δομών σε επίπεδο κυττάρου και σε επίπεδο πληθυσμών
2. Δυναμική Συστημάτων: Ο τρόπος με τον οποίο ένα σύστημα συμπεριφέρεται στη διάρκεια του χρόνου και σε διαφορετικές συνθήκες.
3. Έλεγχος: Μηχανισμοί που ελέγχουν την κατάσταση του κυττάρου οι οποίοι μπορούν να τροποποιηθούν ώστε να ελαχιστοποιηθούν οι δυσλειτουργίες και να προκύψουν πιθανοί θεραπευτικοί στόχοι .
4. Σχεδιασμός: Μέθοδοι τροποποίησης και δημιουργίας βιολογικών συστημάτων, που διαθέτουν επιθυμητές ιδιότητες. Τα συγκεκριμένα συστήματα μπορούν να βασιστούν σε καθορισμένες αρχές σχεδιασμού και σε προσομοιώσεις, αποφεύγοντας τη διαδικασία δοκιμής και απόρριψης (trial and error)

Η ανάγκη, λοιπόν, της μελέτης της βιολογίας σε επίπεδο δυναμικών συστημάτων, οδήγησε στην δημιουργία του χώρου της Συστημικής Βιολογίας. Αν και οι προτεινόμενες από αυτόν τον χώρο προσεγγίσεις αποτελούν κεντρικό ζήτημα της παρούσας εργασίας,

στην ακόλουθη ενότητα δίνονται εν συντομία οι βασικές αρχές που τον διέπουν, ενώ στα ακόλουθα Κεφάλαια, δίνονται οι λεπτομέρειες για την πρακτική εφαρμογή των προτεινόμενων προσεγγίσεων.

2.4 Υπολογιστική Συστημική Βιολογία

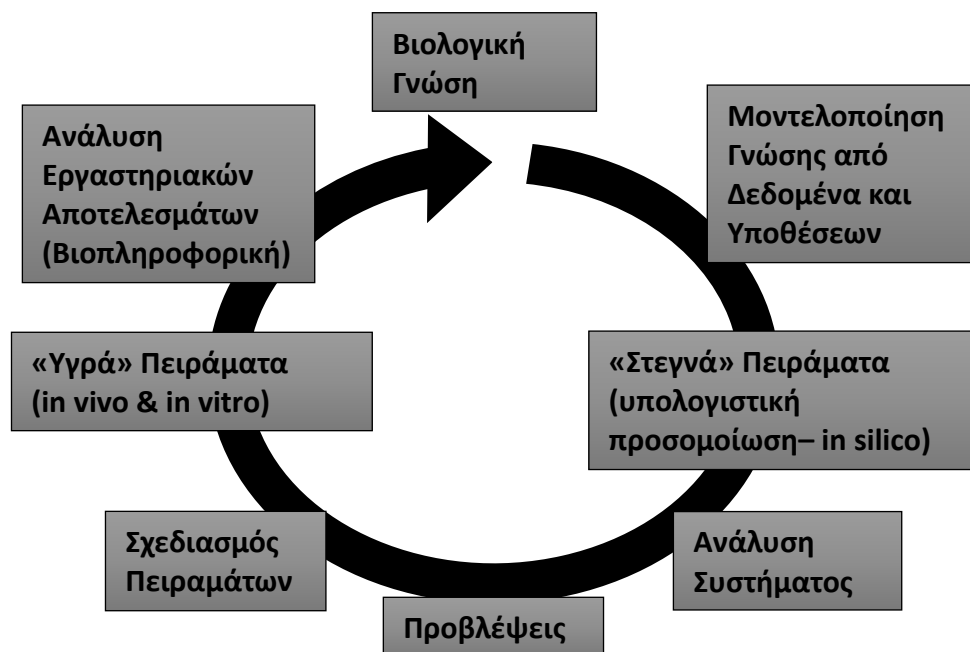
Όπως, λοιπόν, σκιαγραφήθηκε στην προηγούμενη ενότητα, η Συστημική Βιολογία αποτελεί και προβλέπεται να αποτελέσει ακόμα περισσότερο στο μέλλον, μία νέα εποχή στην Βιολογική Έρευνα. Αν και ο συγκεκριμένος χώρος συνεχίζει να εξελίσσεται και να προτείνει νέες μεθόδους για την ανάλυση και την απόκτηση βιολογικής γνώσης, μπορούν να του αποδοθούν κάποια συγκεκριμένα χαρακτηριστικά.

Η Συστημική Βιολογία, αποτελεί τον επιστημονικό χώρο που έχει ως στόχο την αξιοποίηση:

- της συνεχούς αυξανόμενης ποιοτικής αλλά και ποσοτικές βιολογικής γνώσης,
- της πληθώρας –ομικών (γενωμικών, πρωτεωμικών, μεταβολομικών κ.ο.κ.) δεδομένων που παράγονται στις μέρες μας μέσω των ανεπτυγμένων μεθόδων βιοτεχνολογίας
- των ανεπτυγμένων μαθηματικών-πληροφορικών μεθόδων που εφαρμόζονται στα –ομικά δεδομένα (μέθοδοι Βιοπληροφορικής)
με στόχο:
- τον συνδυασμό σε κοινές δομές γνώσεις (συστήματα/μοντέλα) των παραπάνω γνώσεων
- την αποχώρηση από τη μελέτη συγκεκριμένων βιομορίων ή την μελέτη των αλληλεπιδράσεων μεταξύ συγκεκριμένων μόνο βιομορίων
- τον ορισμό ολιστικών μοντέλων για τις διάφορες βιολογικές-βιοχημικές διαδικασίες-φαινόμενα που συμβαίνουν φυσιολογικώς ή κατά την κατάσταση ασθένειας στο επίπεδο εντός του κυττάρου αλλά και στο επίπεδο πληθυσμών κυττάρων
- την υποστήριξη της βιολογικής έρευνας μέσω της δημιουργίας προβλέψεων για την δυναμική συμπεριφορά των βιολογικών συστημάτων στο χρόνο αλλά και κατά την τροποποίηση των συνθηκών που βρίσκονται αυτά ή την αλλαγή συγκεκριμένων παραμέτρων-χαρακτηριστικών τους.

Η Συστημική Βιολογία, βασίζεται στις Υπολογιστικές Μεθόδους έτσι ώστε να είναι εφικτή η διαχείριση της πολύ μεγάλης διαθέσιμης πληροφορίας αλλά και η υψηλή πολυπλοκότητα των βιολογικών συστημάτων. Επίσης, η Συστημική Βιολογία, πολλές φορές καλείται και Υπολογιστική Βιολογία ή και Εκτελέσιμη (Executable) Βιολογία δεδομένου ότι στα πλαίσια της δημιουργούνται μοντέλα των βιολογικών διαδικασιών-φαινομένων που μπορούν να προσομοιωθούν υπολογιστικά. Επίσης δίνεται η δυνατότητα επέμβασης στις διαδικασίες αυτές τροποποιώντας κατάλληλα την δομή και τις παραμέτρους των δημιουργηθέντων μοντέλων.

Η γενικότερη φιλοσοφία και ο τρόπος με τον οποίο η Συστημική Βιολογία επιχειρεί να αποτελέσει σημαντικό κλάδο της βιολογικής έρευνας μπορεί να δοθεί από το ακόλουθο σχήμα, στο οποίο εμφανίζονται τα βήματα κατά τα οποία μπορεί να συμβάλει αλλά και να ορίσει η Συστημική Βιολογία



Εικόνα 2.2 Ο κύκλος της Συστημικής Βιολογίας [60]

Λεπτομέρειες για τα παραπάνω αναφερόμενα βήματα δίνονται στις ακόλουθες υπό-ενότητες της εργασίας, ενώ παράδειγμα χρήσης της γενικότερης προσέγγισης της Συστημικής Βιολογίας αποτελεί το σύνολο της παρούσας διατριβής.

2.4.1 Συστημική Βιολογία και Δίκτυα Βιοχημικών Αντιδράσεων

Κατά τις δύο τελευταίες δεκαετίες στις οποίες αναπτύχθηκε ουσιαστικά ο χώρος της Συστημικής Βιολογίας, έχει προταθεί μία πληθώρα τρόπων Μοντελοποίησης των υπό μελέτη Βιολογικών Συστημάτων και αναπαράστασης των μοντέλων αυτών. Ωστόσο, όσο η πρόσβαση σε λεπτομερή δεδομένα σε μοριακό/βιοχημικό επίπεδο, γίνεται ευκολότερη και

κυρίως οικονομικότερη, η κατηγορία των μοντέλων που σταδιακά επικρατεί των άλλων είναι αυτή των Δικτύων Βιοχημικών Αντιδράσεων. Το σημαντικό πλεονέκτημα των συγκεκριμένων δικτύων είναι ότι βασικό εργαλείο μοντελοποίησης σε αυτά αποτελούν οι Βιοχημικές Αντιδράσεις (χημικές αντιδράσεις που συμβαίνουν στα κύτταρα των ζώντων οργανισμών), Έτσι, η μοντελοποίηση των βιολογικών συστημάτων γίνεται με όρους ιδιαίτερα κοντά στους φυσικούς/χημικούς νόμους που διέπουν το σύστημα, σε αντίθεση με τις έννοιες της Καταπίεσης ή της Ενεργοποίησης οι οποίες αποτελούν υψηλότερου επιπέδου περιγραφές της δυναμικής του συστήματος και συναντώνται στα δίκτυα Αλληλεπιδράσεων. Ένα ακόμα πλεονέκτημα των Δικτύων Βιοχημικών Αντιδράσεων είναι ότι οι βασικές αρχές μοντελοποίησης που ορίζονται από αυτά, μπορούν να γίνουν εύκολα αντιληπτές από ένα ευρύ φάσμα επιστημόνων συμπεριλαμβανομένων των βιοχημικών, των φυσικών, των βιολόγων αλλά και των επιστημόνων πληροφορικής.

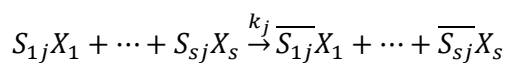
Στις ακόλουθες ενότητες παρουσιάζεται το θεωρητικό πλαίσιο το οποίο διέπει τα συγκεκριμένα δίκτυα (η παρουσίαση βασίζεται στην αντίστοιχη παρουσίαση που δίνεται στην [61]) ενώ παρουσιάζονται οι κυριότερες μέθοδοι προσομοίωσης, ανάλυσης και εκπαίδευσης τους.

2.4.1.1 Θεωρητικό Πλαίσιο Δικτύων Βιοχημικών Αντιδράσεων

Για την παρουσίαση του θεωρητικού πλαισίου των Δικτύων Βιοχημικών Αντιδράσεων, θεωρούμε τα μόρια από s διαφορετικές χημικές οντότητες (chemical species), τα οποία είναι ομοιογενώς κατανεμημένα σε μία περιοχή/διαμέρισμα σταθερού όγκου V . Επίσης θεωρούμε ότι το συγκεκριμένα σύστημα βρίσκεται σε θερμική ισορροπία ενώ τα διάφορα μόρια αλληλεπιδρούν μεταξύ τους μέσω κάποιων r μη αναστρέψιμων/μονής κατεύθυνσης καναλιών αντιδράσεων (reaction channels).

Ένα κανάλι αντίδρασης καλείται στοιχειώδες (elementary) όταν μέσω αυτού πραγματοποιείται ένα απλό βήμα, ωστόσο μπορεί να αναπαριστά το σύνολο ενός αριθμού θεμελιωδών βημάτων σε ένα απλό βήμα. Επίσης, οι αντιστρέψιμες (reversible) αντιδράσεις μπορούν να οριστούν ως δύο μη αντιστρέψιμες αντιδράσεις.

Συμβολίζοντας τώρα την i -οστή οντότητα με X_i και την j -οστή αντίδραση με R_j , το κανάλι αντίδρασης R_j αναπαριστάται από το παρακάτω γενικό σχήμα:



όπου:

- $X_i(t) = \frac{N_i(t)}{\Omega}$, ως μέτρηση των επιπέδων της X_i στο σύστημα την χρονική στιγμή t , με $N_i(t)$ να ορίζεται ως η πολλαπλότητα της συγκεκριμένης οντότητας και Ω να αφορά στο μέγεθος του συστήματος το οποίο:
 - ο Για μοριακές συγκεντρώσεις (molar concentrations) το μέγεθος του συστήματος δίνεται ως $\Omega = N_A V$, όπου N_A ο αριθμός-σταθερά Avogadro [62].
 - ο Για σχετικές συγκεντρώσεις ως Ω τίθεται ένα σταθερός αριθμός.
- S_{ij} , οι στοιχειομετρικοί συντελεστές (stoichiometric coefficients) της αντίδρασης, όπου ο \underline{S}_{ij} αφορά στη συμμετοχή του X_i ως αντιδρών και ο \overline{S}_{ij} , αντιστοίχως ως προϊόν.
- k_j , η σταθερά ρυθμού της αντίδρασης

Η δραστηριότητα του καναλιού αντίδρασης R_j , ποσοτικοποιείται μέσω της μέτρησης της τιμής $Z_j(t)$, του αριθμού των πραγματοποιήσεων της αντίδρασης R_j στο χρονικό διάστημα $[0, t]$.

Εφόσον, τώρα, κάθε πραγματοποίηση της αντίδρασης R_j , αλλάζει την πολλαπλότητα της X_i κατά $S_{ij} = \overline{S}_{ij} - \underline{S}_{ij}$, τότε η συνολική αλλαγή στην πολλαπλότητα της συγκεκριμένης οντότητας στο χρονικό διάστημα $[0, t]$, καθορίζεται από τις στοιχειομετρικές και από τον αριθμό των πραγματοποιήσεων των αντιδράσεων και έτσι δίνεται από:

$$N_i(t) = N_i(0) + \sum_{j=1}^r S_{ij} Z_j(t)$$

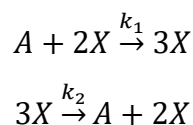
και σε μορφή διανύσματος ως

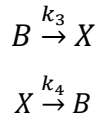
$$N(t) = N(0) + SZ(t)$$

ενώ διαιρώντας με την Ω λαμβάνουμε τη σχέση με όρους συγκεντρώσεων

$$X(t) = X(0) + \frac{SZ(t)}{\Omega}$$

Για την ευκολότερη κατανόηση του παραπάνω πλαισίου, μπορεί να δοθεί ένα απλό παράδειγμα Δικτύου Βιοχημικών Αντιδράσεων. Ένα τέτοιο μοντέλο είναι αυτό του Schlögl που παρουσιάστηκε στην [63] και περιγράφεται από τις ακόλουθες αντιδράσεις:





Δεδομένου, τώρα, ότι ο αριθμός των προτεινόμενων μοντέλων για τα διάφορα βιολογικά συστήματα που μελετώνται και σε υπολογιστικά πλαίσια, αυξάνεται με ραγδαίους ρυθμούς αλλά και ότι υπάρχει η ανάγκη ευκολότερης και αποδοτικότερης δημοσίευσης και ανταλλαγής μοντέλων, δημιουργήθηκε ο στόχος ανάπτυξης οντολογιών και προτύπων περιγραφής αλλά και γραφικής αναπαράστασης των μοντέλων. Όσον αφορά τις οντολογίες, αυτή που πλέον έχει επικρατήσει είναι η SBML (Systems Biology Markup Language) [64]–[66] ενώ όσον αφορά τα πρότυπα γραφικής αναπαράστασης έχει επικρατήσει αυτό της SBGN (Systems Biology Graphical Notation) [67].

Η SBML, αποτελεί μία μορφή αναπαράστασης, που βασίζεται στην XML [68], και χρησιμοποιείται για την δημοσίευση και την αποθήκευση υπολογιστικών μοντέλων βιολογικών διαδικασιών. Επίσης, η χρήση της επιτρέπει την ανταλλαγή των μοντέλων αυτών και την επαναχρησιμοποίηση τους μεταξύ των διαφόρων εργαλείων που υποστηρίζουν της SBML (τα οποία την περίοδο συγγραφής της εργασίας αγγίζουν τις 290 [66]) χωρίς την ανάγκη μετατροπής αυτών. Μέσω της SBML μπορούν να περιγραφούν διάφορες κλάσεις βιολογικών φαινομένων, συμπεριλαμβανομένων μεταβολικών δικτύων (metabolic networks), βιολογικών μονοπατιών σηματοδότησης (signaling pathways) και ρυθμιστικών δικτύων (regulatory networks).

Μέσω της SBML, κωδικοποιούνται μοντέλα που απαρτίζονται από οντότητα (Species) οι οποίες δρουν μέσω συγκεκριμένων διαδικασιών που ονομάζονται στην γλώσσα της SBML Reactions (αντιδράσεις). Θα πρέπει ωστόσο να σημειωθεί ότι η εφαρμογή της SBML δεν περιορίζεται μόνο στα Δίκτυα Βιοχημικών Αντιδράσεων, και παρά την φιλοσοφία της αναπαράστασης διαδικασιών τύπου χημικών αντιδράσεων, ο ίδιος φορμαλισμός μπορεί να χρησιμοποιηθεί για την αναπαράσταση πολλών άλλων τύπων διαδικασιών. Επίσης η SBML εμπεριέχει διάφορα άλλα χαρακτηριστικά γλώσσας που επιτρέπουν τον άμεσο ορισμό μαθηματικών εκφράσεων και ασυνεχών γεγονότων, ανεξάρτητων των διαδικασιών αντίδρασης, επιτρέποντας έτσι την περιγραφή πολλών περισσότερων φαινομένων εκτός των βιοχημικών αντιδράσεων. Ακόμα, σημαντικό χαρακτηριστικό της SBML αποτελεί το ότι η αναπαράσταση ενός μοντέλου μέσω αυτής δεν απαιτεί τον ορισμό των μαθηματικών εργαλείων με τα οποία θα περιγραφεί η δυναμική του συστήματος, όπως οι Διαφορικές εξισώσεις ή άλλες μορφές ερμηνείας (λεπτομέρειες για

τον μαθηματικό ορισμό των μοντέλων δίνονται σε ακόλουθες ενότητες). Αυτό το χαρακτηριστικό δίνει σημαντική ευελιξία στις εφαρμογές που υποστηρίζουν την SBML να ορίζουν αυτές τον τρόπο με τον οποίο θα περιγραφεί η δυναμική του συστήματος χωρίς να χρειάζεται να γίνει κάποια σχετική μετατροπή.

Τα βασικά μέρη ενός αρχείου SBML(για τα οποία δεν χρειάζεται να οριστούν κατ' ανάγκη όλα) είναι τα ακόλουθα [66]:

- **Ορισμός Συνάρτησης (Function Definition):** Μία μαθηματική συνάρτηση που μπορεί να χρησιμοποιηθεί σε όλη την έκταση του μοντέλου
- **Ορισμός Μονάδων (Unit Definition):** Ορισμός νέων μονάδων μέτρησης ή επανορισμός των ενσωματωμένων στην SBML μονάδων. Οι μονάδες αυτές χρησιμεύουν στην αναπαράσταση των ποσοτικών πληροφοριών ενός μοντέλου.
- **Τύπος Διαμερίσματος (Compartment):** Τύπος μιας περιοχής του μοντέλου στην οποία βρίσκονται οι αντιδρούσες οντότητες του συστήματος.
- **Τύπος Οντότητας (Species Type):** Τύπος οντότητας που συμμετέχει στις αντιδράσεις. Συνήθεις τύποι οντοτήτων αποτελούν τα Ιόντα (όπως το ασβέστιο), διάφορα μόρια όπως η γλυκόζη και το ATP, σημεία πρόσδεσης επί των πρωτεϊνών κ.α.
- **Διαμέρισμα (Compartment):** Ομοιογενής διακριτός χώρος του συστήματος, με συγκεκριμένο τύπο και ορισμένο μέγεθος, στον οποίο βρίσκονται διάφορες οντότητες και συμβαίνουν συγκεκριμένες αντιδράσεις. Ένα μοντέλο μπορεί να περιέχει πολλά διαμερίσματα ακόμα και του ίδιου τύπου, ενώ κάθε οντότητα του μοντέλου θα πρέπει να βρίσκεται εντός κάποιου διαμερίσματος.
- **Οντότητες (Species):** Ένα σύνολο οντοτήτων του ίδιου τύπου που βρίσκονται σε συγκεκριμένο διαμέρισμα.
- **Παράμετρος (Parameter):** Μία ποσότητα με ορισμένο όνομα. Στην SBML ο όρος παράμετρος έχει γενική έννοια, χαρακτηρίζοντας ποσότητες άσχετα με τον αν αυτές είναι σταθερές ή μεταβλητές του μοντέλου.
- **Αρχική Ανάθεση (Initial Assignment):** Μαθηματική έκφραση που χρησιμοποιείται για τον ορισμό των αρχικών συνθηκών του μοντέλου. Μπορεί να χρησιμοποιηθεί μόνο για τον ορισμό του τρόπου που η τιμή μίας μεταβλητής θα υπολογιστεί από την τιμή άλλων ποσοτήτων του συστήματος στην αρχή της προσομοίωσης.
- **Κανόνας (Rule):** Μαθηματική έκφραση που χρησιμοποιείται σε συνδυασμό με τον μαθηματικό ορισμό του μοντέλου (π.χ. διαφορικές εξισώσεις) έτσι ώστε να ορίσουν

τον τρόπο που η τιμή μίας μεταβλητής μπορεί να υπολογιστεί από τις τιμές άλλων μεταβλητών. Οι κανόνες μπορούν να συνδυαστούν με τις εξισώσεις κινητικών νόμων των μοντέλων έτσι ώστε να καθορίσουν την συμπεριφορά του μοντέλου στο χρόνο και επηρεάζουν το σύστημα για όλη τη διάρκεια της προσομοίωσης.

- **Περιορισμός (Constraint):** Μαθηματική έκφραση μέσω της οποίας ορίζονται περιορισμοί για τις τιμές των μεταβλητών του μοντέλου. Οι περιορισμοί αυτοί εφαρμόζονται σε όλη τη διάρκεια της προσομοίωσης, ενώ δεν θα πρέπει να χρησιμοποιούνται για τον καθορισμό της δυναμικής συμπεριφοράς του συστήματος στο χρόνο.
- **Αντίδραση (Reaction):** Ορισμός που περιγράφει μία τροποποίηση, μετακίνηση ή διαδικασία δέσμευσης μέσω της οποίας μπορούν να αλλάξουν τα επίπεδα μίας ή περισσότερων οντοτήτων.
- **Γεγονός (Event):** Ορισμός που περιγράφει μία στιγμιαία και ασυνεχή αλλαγή σε ένα σύνολο μεταβλητών οποιουδήποτε τύπου (επίπεδα οντοτήτων, μέγεθος διαμερίσματος, τιμές παραμέτρων), η οποία συμβαίνει μόνο όταν μία κατάλληλη συνθήκη πληρείται κατά τη διάρκεια της προσομοίωσης του μοντέλου.

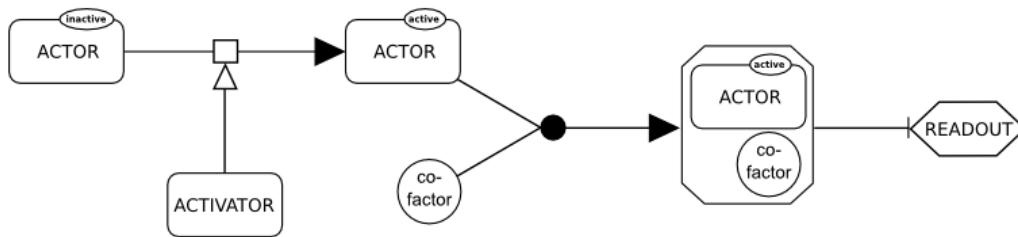
Όσον αφορά την SBGN, τώρα, είναι κοινός αποδεκτό ότι σημαντικός παράγοντας για την διεύρυνση της αποδοχής των Μοντέλων Συστημικής Βιολογίας, αποτελεί η δυνατότητα οπτικοποίησης αυτών, ώστε να δημιουργούνται αλλά και να δημοσιεύονται ευκολότερα. Ωστόσο, στα πρώτα χρόνια ανάπτυξης του χώρου της Συστημικής Βιολογίας, κάθε εφαρμογή μοντελοποίησης υιοθετούσε, παρότι παραπλήσιους, διαφορετικούς τρόπους γραφικής αναπαράστασης των μοντέλων. Λύση στο συγκεκριμένο πρόβλημα έρχεται να δώσει το αντίστοιχο της SBML πρότυπο όσον αφορά τη γραφική αναπαράσταση, το πρότυπο SBGN.

Μέσω της SBGN ορίζονται συγκεκριμένοι κανόνες για τον τρόπο που πρέπει να δομούνται οι γράφοι των μοντέλων (είτε αποτελούν Δίκτυα Βιοχημικών Αντιδράσεων είτε άλλων μορφών Δίκτυα), τη μορφή που πρέπει να έχουν οι κάθε είδους οντότητες αλλά και ο τρόπος που θα πρέπει να αναπαριστούνται οι διάφορες διαδικασίες.

Η SBGN δομείται από τρεις βασικές γλώσσες [69] οι οποίες είναι:

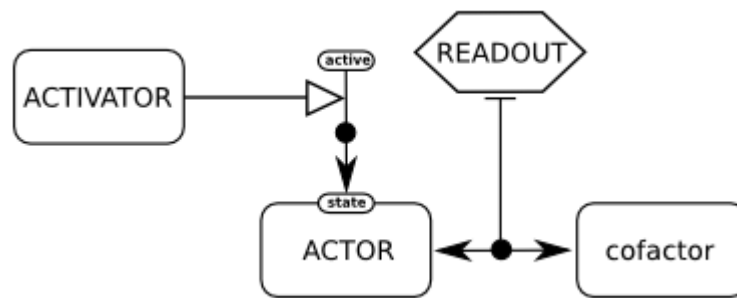
- **SBGN Περιγραφή Διαδικασίας (Process Description)** [70]: Απεικονίζει τη χρονική εξέλιξη των βιοχημικών αλληλεπιδράσεων ή αντιδράσεων στο δίκτυο. Μπορεί να χρησιμοποιηθεί ώστε να εμφανίζει όλες τις αλληλεπιδράσεις/αντιδράσεις μεταξύ των μοριακών οντοτήτων που λαμβάνουν χώρα στο δίκτυο, με την ίδια

οντότητα να μπορεί να εμφανίζεται πολλές φορές στο ίδιο διάγραμμα. Ένα παράδειγμα της SBGN Process Description δίνεται στην ακόλουθη εικόνα



Εικόνα 2.3 Παράδειγμα της γλώσσας Περιγραφής Διαδικασίας της SBGN

- **SBGN Σχέση Οντοτήτων (Entity Relationship) [71]:** Επιτρέπει την επισκόπηση των σχέσεων στις οποίες μία οντότητα συμμετέχει, αγνοώντας τη διάσταση του χρόνου. Οι σχέσεις μπορούν να θεωρηθούν ως κανόνες που περιγράφουν την επιρροή των οντοτήτων (κόμβοι του δικτύου) σε άλλες σχέσεις. Οι λογική της SBGN Entity Relationship δίνεται στην ακόλουθη εικόνα.



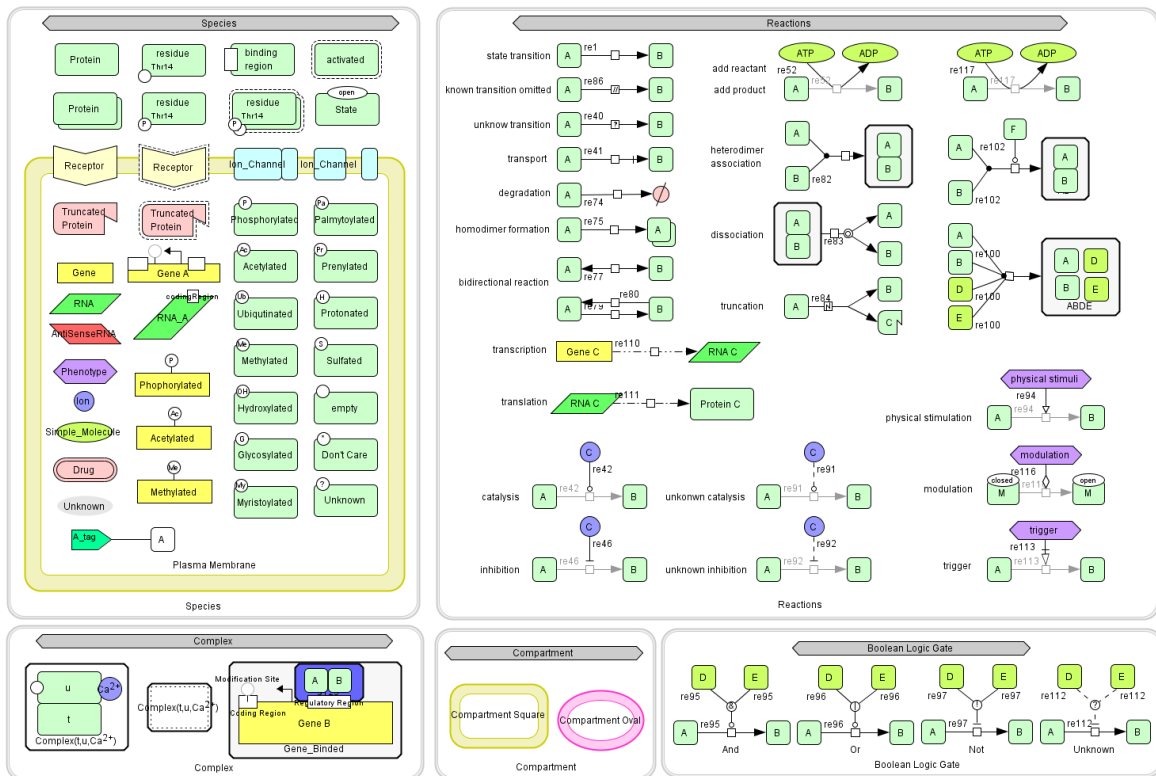
Εικόνα 2.4 Παράδειγμα της γλώσσας Σχέσης Οντοτήτων της SBGN

- **SBGN Ροή Δραστηριότητας (Activity Flow) [72]:** Απεικονίζει την μεταφορά πληροφορίας μεταξύ των βιοχημικών οντοτήτων του δικτύου, αγνοώντας τις αλλαγές κατάστασης αυτών, όπως εμφανίζεται στην ακόλουθη εικόνα.



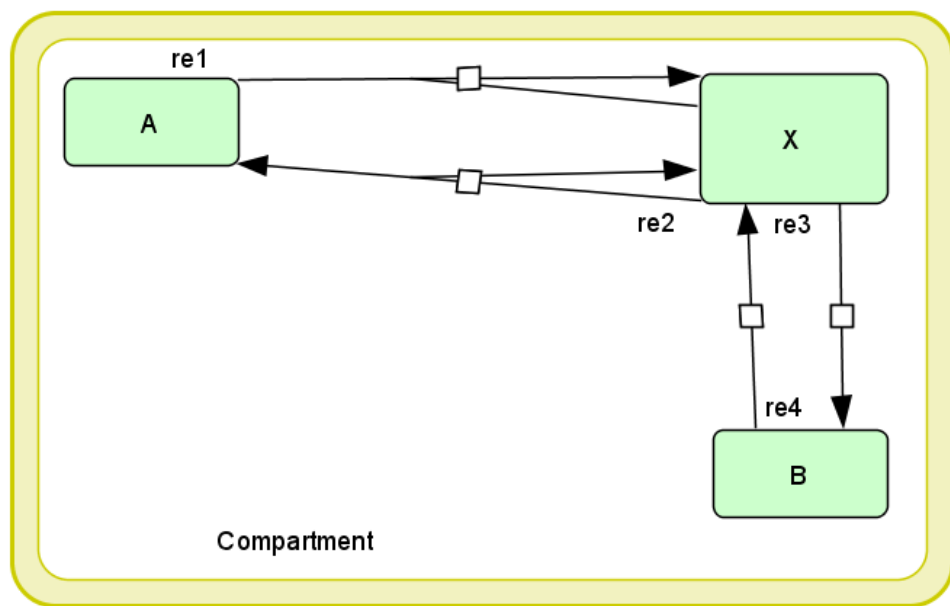
Εικόνα 2.5 Παράδειγμα της γλώσσας Ροής Δραστηριότητας της SBGN

Σημαντικό κομμάτι της SBGN, όπως προαναφέρθηκε, αποτελεί ορισμός της μορφής που πρέπει να παρουσιάζουν οι διάφορες οντότητες αλλά και οι διαδικασίες στις οποίες εμπλέκονται. Ο πλήρης χάρτης των συγκεκριμένων κανόνων δίνεται στην ακόλουθη εικόνα.



Εικόνα 2.6 Χάρτης των κανόνων της SBGN

Ένα απλό παράδειγμα της χρήσης της SBGN μπορεί να δοθεί για το μοντέλο του Schlögl που παρουσιάστηκε προηγουμένως. Η αναπαράσταση του μοντέλου σε SBGN Process Description δίνεται στην ακόλουθη εικόνα.



Εικόνα 2.7 Αναπαράσταση του Μοντέλου του Schlögl σε SBGN

Οι διάφορες αντιδράσεις που απαρτίζουν το μοντέλο, απεικονίζονται ως βέλη που συνδέουν τις οντότητες, ενώ οι κατεύθυνση του βέλους ορίζει τα αντιδρώντα και τα προϊόντα της αντίδρασης (με το βέλος να δείχνει τα προϊόντα). Θα πρέπει να σημειωθεί ιδιαίτερα, ωστόσο, ότι αναπόσπαστο κομμάτι της δημιουργίας/σχεδίασης οποιουδήποτε μοντέλου αποτελεί ο ορισμός των στοιχειομετρικών συντελεστών κάθε αντίδρασης. Έτσι, μπορεί στην ανωτέρω εικόνα η οντότητα X να εμφανίζεται μία φορά, ωστόσο στη στοιχειομετρία της αντίδρασης re1 θα πρέπει να ορίσουμε την συμμετοχή 2 μορίων της X και την απολαβή τελικά 3 μορίων της ίδιας οντότητας. Εναλλακτικά, θα μπορούσε να εισαχθεί η οντότητα X περισσότερες από μία φορές, όπως σημειώνεται στην περιγραφή της SBGN, ώστε η επιθυμητή στοιχειομετρία να δημιουργηθεί άμεσα.

Έχοντας ορίσει τα παραπάνω, προκύπτει ο στοιχειομετρικός πίνακας του δικτύου, ο οποίος δίνεται ακολούθως.

	(re1)	(re2)	(re3)	(re4)
X	1	-1	-1	1
A	-1	1	0	0
B	0	0	1	-1

Εικόνα 2.8 Ο Στοιχειομετρικός πίνακας του Μοντέλου του Schlögl

Οι παραπάνω ορισμοί, παρότι δοσμένοι σε αυστηρό μαθηματικό πλαίσιο, από μόνοι τους δεν είναι αρκετοί για την υπολογιστική προσομοίωση των Βιοχημικών Συστημάτων που περιγράφονται μέσω αυτών. Τα βασικά στοιχεία τα οποία υπολείπονται για να είναι εφικτό κάτι τέτοιο, είναι ο ορισμός της μεθόδου με την οποία θα αποφασίζεται η πραγματοποίηση των αντιδράσεων αλλά και ο τρόπος που θα εμπλέκονται οι σταθερές ρυθμών αντίδρασης σε αυτή αλλά και στη συνολική εξέλιξη της προσομοίωσης.

Οι δύο κύριες προσεγγίσεις για την απαιτούμενη αυτή εξειδίκευση είναι η ντετερμινιστική και η στοχαστική οι οποίες παρουσιάζονται στην ακόλουθη ενότητα.

2.4.1.2 Μέθοδοι Προσομοίωσης Δικτύων Βιοχημικών Αντιδράσεων

Όπως προαναφέρθηκε, οι δύο κύριες κατηγορίες μεθόδων για την Προσομοίωση ενός Δικτύου Βιοχημικών αντιδράσεων, είναι η ντετερμινιστική και η στοχαστική. Κύριο χαρακτηριστικό της Στοχαστικής προσέγγισης είναι ότι κατά την μοντελοποίηση, θεωρούμε τα διάφορα μόρια ως διακριτές οντότητες (που είναι και στην πραγματικότητα) και ότι οι αντιδράσεις μεταξύ τους πραγματοποιούνται μέσω της συνεύρεσης τους η οποία, όπως ορίζουν οι νόμοι της φυσικής, αποτελεί τυχαία διαδικασία. Απεναντίας στην Ντετερμινιστική προσομοίωση δεν υπάρχει κανένας παράγοντας τυχειότητας. Ανάλογα με

την μέθοδο που θα επιλεγεί, απαιτείται και διαφορετική εξειδίκευση των ορισμών που δόθηκαν στην προηγούμενη ενότητα, οι οποίες κύριες μορφές εξειδίκευσης παρουσιάζονται στις υπο-ενότητες της παρούσας ενότητας.

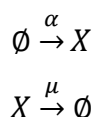
Όσον αφορά την καταλληλότητα εφαρμογής κάθε προσέγγισης ο γενικότερος κανόνας αφορά στα επίπεδα που βρίσκονται οι διάφορες οντότητες του συστήματος. Πιο συγκεκριμένα όπως αναφέρεται στην [73], η ντετερμινιστική προσέγγιση μπορεί να είναι επαρκής ώστε να προσομοιωθούν συστήματα στα οποία τα επίπεδα των οντοτήτων είναι αρκετά υψηλά (όπως αυτά που παρουσιάζονται στην κλίμακα του δοκιμαστικού σωλήνα). Ωστόσο, όταν τα μελετώμενα συστήματα είναι αρκούντως μικρά ώστε οι μοριακοί πληθυσμοί τουλάχιστον κάποιων από τις εμπλεκόμενες οντότητες να μην είναι ιδιαίτερα μεγαλύτεροι της μονάδος, τότε είναι αναγκαία η Στοχαστική προσέγγιση για την ακριβέστερη προσομοίωση των Συστημάτων αυτών. Επίσης, ένα ιδιαίτερα συνηθισμένος τρόπος διάκρισης της καταλληλότητας εφαρμογής της μία ή της άλλης μεθόδου αφορά στον τύπο της βιολογικής διαδικασίας που μοντελοποιείται. Πιο συγκεκριμένα, η ντετερμινιστική Προσομοίωση θεωρείται επαρκής για μοντέλα που αφορούν διαδικασία του Μεταβολισμού (Metabolic Networks), στον οποίο συνήθως εμπλέκονται μεγάλοι αριθμοί μορίων. Απεναντίας, για τα μοντέλα που αναφέρονται στην Σηματοδοσία εντός του κυττάρου και στην Έκφραση Γονιδίων (και στην διαδικασία ρύθμισης αυτής), διαδικασίες στις οποίες και εμπλέκονται μικροί αριθμοί μορίων αλλά και κάθε γεγονός έχει σημαντική αξία για την εξέλιξη του συστήματος, η Στοχαστική Προσομοίωση θεωρείται ότι προσφέρει σημαντικά πλεονεκτήματα.

Ένας ακόμα παράγοντας που υποστηρίζει την χρήση Στοχαστικής Προσομοίωσης είναι αυτός του θορύβου που εμφανίζεται στα διάφορα βιολογικά συστήματα. Με τον όρο Θόρυβος, εννοείται η τυχαιότητα που εμφανίζεται στις διάφορες βιολογικές διαδικασίες. Για παράδειγμα ο θόρυβος στην έκφραση γονιδίων μπορεί να αναφέρεται στην στοχαστική-τυχαία διαφοροποίηση των επιπέδων της ίδιας εκφραζόμενης πρωτεΐνης μεταξύ ομοίων κυττάρων τα οποία έχουν το ίδιο ιστορικό και βρίσκονται υπό τις ίδιες συνθήκες [61].

Οι δύο μορφές θορύβου που επηρεάζουν την συμπεριφορά ενός βιολογικού συστήματος είναι [61], [74] ο εγγενής θόρυβος (intrinsic noise) και ο εξωγενής θόρυβος (extrinsic noise). Οι συνθήκες και ο τρόπος κατά τον οποίο εμφανίζεται ο κάθε θόρυβος περιγράφονται με λεπτομέρεια στις προηγούμενες αναφερόμενες εργασίες. Ωστόσο ένα παράδειγμα διαχωρισμού των δύο θορύβων μπορεί να δοθεί για ένα (όχι κατ' ανάγκη) υποθετικό σύστημα το οποίο περιέχει δύο υποκινητές γονιδίου στο ίδιο κύτταρο. Ποσοτικοποιώντας την έκφραση της αντίστοιχης πρωτεΐνης, μπορούμε να πούμε ότι ο

εγγενής θόρυβος αφορά σε πηγές θορύβου που οδηγούν σε διαφορές στην έκφραση μέσω των δύο υποκινητών στο ίδιο κύτταρο ενώ ο εξωγενής θόρυβος αφορά σε πηγές θορύβου που έχουν την ίδια επιρροή και στους δύο υποκινητές του ίδιου κυττάρου αλλά προκαλούν διαφορές μεταξύ δύο κυττάρων. Τα ζητήματα αυτά είναι ιδιαίτερα σημαντικά μιας και πλέον είναι αποδεκτό ότι μεταξύ ακόμα και ίδιων κυττάρων που βρίσκονται στο ίδιο περιβάλλον εμφανίζεται σε μικρό ή μεγάλο βαθμό ετερογένεια στην συμπεριφορά την οποία μόνο η Στοχαστική Προσομοίωση μπορεί να προβλέψει [74].

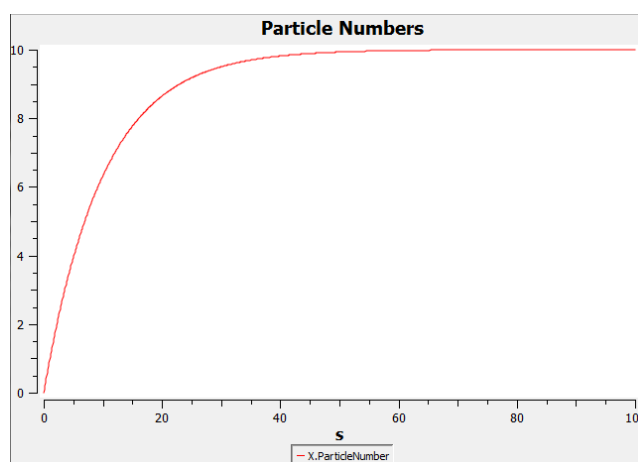
Για την παρουσίαση, τώρα, των διαφορετικών χαρακτηριστικών των αποτελεσμάτων που μπορούν να μας δώσουν οι δύο προσεγγίσεις, αρκεί η χρήση ενός ιδιαίτερα απλού μοντέλου, όπως προτείνεται στην [74]. Το μοντέλο αποτελείται από μία οντότητα, έστω τη X και δύο αντιδράσεις, την παραγωγή της X από μία πηγή και την αποδόμηση της X . Οι αντιδράσεις αυτές μπορούν να γραφούν σύμφωνα με τα όσα παρουσιάστηκαν προηγουμένως, ως εξής:



οι παράμετροι α και μ αποτελούν τους ρυθμούς των αντίστοιχων αντιδράσεων με τιμές 1 και 0.1. αντίστοιχα, ενώ αρχικά τα επίπεδα της X είναι μηδενικά.

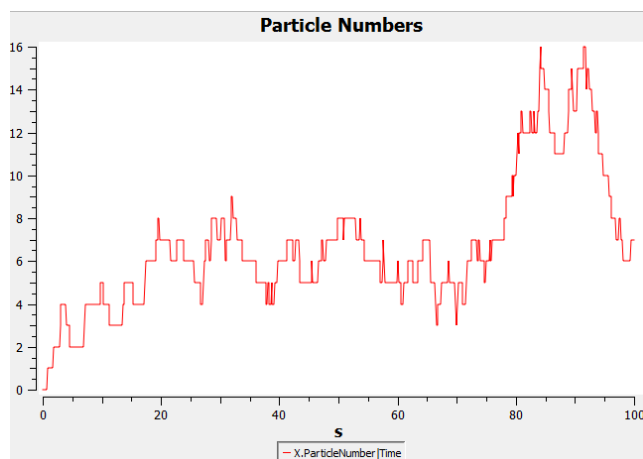
Για τις ανάγκες παρουσίασης, το μοντέλο δημιουργήθηκε στο λογισμικό COPASI και τα αποτελέσματα της προσομοίωσης για 100 δευτερόλεπτα δίνονται στα ακόλουθα σχήματα.

Ντετερμινιστική Προσομοίωση



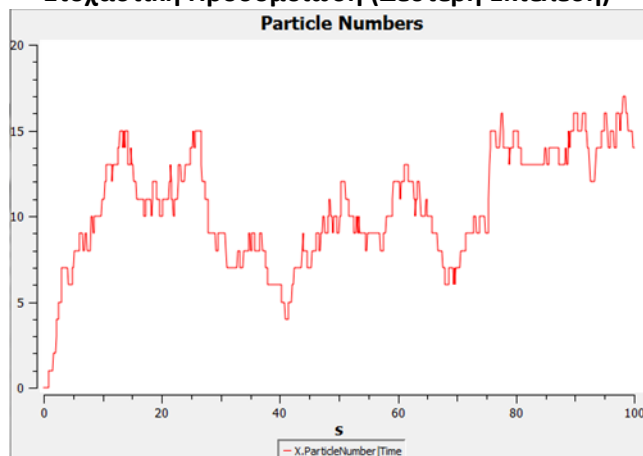
Εικόνα 2.9 Παράδειγμα ντετερμινιστικής προσομοίωσης

Στοχαστική Προσομοίωση (Πρώτη Εκτέλεση)



Εικόνα 2.10 Παράδειγμα Στοχαστικής Προσομοίωσης (1η εκτέλεση)

Στοχαστική Προσομοίωση (Δεύτερη Εκτέλεση)



Εικόνα 2.11 Παράδειγμα Στοχαστικής Προσομοίωσης (2η εκτέλεση)

Όπως είναι εύκολα αντιληπτό, υπάρχουν σημαντικές διαφορές μεταξύ των αποτελεσμάτων που δίνονται από τους δύο τύπους προσομοίωσης. Κύρια διαφορά αποτελεί ότι για την περίπτωση της ντετερμινιστικής προσομοίωσης, από ένα σημείο της προσομοίωσης και μετά, το σύστημα παύει να παρουσιάζει την οποιαδήποτε δυναμική συμπεριφορά και τα επίπεδα της X , παραμένουν σταθερά. Απεναντίας κάτι τέτοιο δεν συμβαίνει για την περίπτωση της στοχαστικής προσομοίωσης, κατά την οποία λόγω της τυχαιότητας στο αν θα συμβεί αντίδραση παραγωγής ή αντίδραση αποδόμησης κάθε στιγμή, τα επίπεδα της X , μεταβάλλονται συνεχώς. Αν η συγκεκριμένη οντότητα, αποτελούσε μέρος ενός μεγαλύτερου συστήματος με σημαντική επιρροή σε αυτό, πολύ πιθανόν αυτές οι διακυμάνσεις να είχαν σημαντικό αντίκτυπο στη συνολική συμπεριφορά του συστήματος, κάτι που δεν θα συνέβαινε κατά τον ίδιο τρόπο με τη χρήση ντετερμινιστικής προσομοίωσης. Ακόμα, ένα πολύ σημαντικό συμπέρασμα που μπορεί να

εξαχθεί από τα δύο σχήματα που αφορούν στην Στοχαστική Προσομοίωση, έχει να κάνει με τα διαφορετικά αποτελέσματα που προκύπτουν από αυτή σε κάθε περίπτωση, πάλι λόγω της τυχαιότητας στην πραγματοποίηση των αντιδράσεων. Η συμπεριφορά αυτή είναι ιδιαίτερα σημαντική για την όσο το δυνατόν ακριβέστερη προσομοίωση των Βιολογικών Συστημάτων, δεδομένης της σημαντικής ετερογένειας που εμφανίζεται στην συμπεριφορά αυτών, όπως σημειώθηκε προηγουμένως.

Ακολούθως παρουσιάζονται τα πλαίσια εξειδίκευσης του ορισμού των Δικτύων Βιοχημικών Αντιδράσεων και οι μέθοδοι που ακολουθούνται κατά περίπτωση για την προσομοίωση αυτών.

2.4.1.2.1 Ντετερμινιστική Προσομοίωση

Η παρουσίαση που ακολουθεί στην συγκεκριμένη ενότητα βασίζεται στην αντίστοιχη παρουσίαση που δίνεται στην [61].

Θεωρώντας ότι οι αντιδράσεις που συμμετέχουν στο σύστημα συμβαίνουν τόσο συχνά ώστε ο αριθμός πραγματοποίησης αντιδράσεων $Z(t)$, μπορεί να προσεγγιστεί από μία συνεχή ποσότητα $z(t)$ και ότι:

- Ένας μεγάλος αριθμός αντιδρώντων μορίων είναι διαθέσιμος σε ένα αρκετά μεγάλο όγκο με δυνατότητα να αντιδράσουν εύκολα
- Η ενέργεια και ο προσανατολισμός των μορίων ευνοούν την πραγματοποίηση της αντίδρασης (κάτι που περιγράφεται από την σταθερά ρυθμού της αντίδρασης) και ότι
- Λόγω του υψηλού αριθμού μορίων η μία πραγματοποίηση μίας αντίδρασης οδηγεί σε μικρής κλίμακας αλλαγή στη συνολική κατάσταση του συστήματος τότε, η πολλαπλότητα των οντοτήτων $N(t)$ αλλά και η συγκέντρωση αυτών $X(t)$, μπορούν και αυτές να προσεγγιστεί από συνεχείς ποσότητες $n(t)$ και $x(t)$.

Με βάση τα παραπάνω, λαμβάνουμε την ντετερμινιστική μορφή των εξισώσεων η οποία είναι

$$n(t) = n(0) + Sz(t)$$

και

$$x(t) = x(0) + \frac{Sz(t)}{\Omega}$$

Παραγωγίζοντας τώρα ως προς το χρόνο, λαμβάνουμε τις Ντετερμινιστικές Χημικές Κινητικές Εξισώσεις:

$$\dot{x}(t) = Sv(x(t))$$

όπου $v = \frac{\dot{z}}{\rho}$, το διάνυσμα ρυθμών αντίδρασης, με v_j να ισούται με το ρυθμό της αντίδρασης R_j . Άρα ο ρυθμός της αντίδρασης είναι ίσος με τον αριθμό των επαναλήψεων της αντίδρασης ανά μονάδα χρόνου, διαιρεμένος με το μέγεθος του συστήματος, ενώ ο συμβολισμός $v(x(t))$ επιλέγεται λόγω της παραδοχής ότι οι συγκεκριμένοι ρυθμοί εξαρτώνται μόνο από τις συγκεντρώσεις των αντιδρώντων. Παρότι αυτή η παραδοχή είναι ρεαλιστική σε πολλές περιπτώσεις, γενικά οι ρυθμοί των αντιδράσεων, όπως είναι αναμενόμενο, μπορούν να εξαρτώνται από την θερμοκρασία, την πίεση αλλά και από τις συγκεντρώσεις άλλων οντοτήτων που εμπλέκονται στο υπό μελέτη σύστημα.

Έχοντας ορίσει τα παραπάνω, η συναρτησιακή μορφή $v_j(\cdot)$ του ρυθμού της R_j καλείται Κινητικός Νόμος και αποτελεί συμπέρασμα των παραδοχών μοντελοποίησης των καναλιών αντιδράσεων. Ο ορισμός του Κινητικού Νόμου για κάθε αντίδραση αποτελεί το ουσιαστικό βήμα εξειδίκευσης των Διαφορικών Εξισώσεων της ανωτέρω εξίσωσης έτσι ώστε να μπορούν να χαρακτηρίσουν ένα Δίκτυο Βιοχημικών Αντιδράσεων και τη συμπεριφορά αυτού, ενώ χωρίς των ορισμό τους αποτελούν ένα σύνολο κανόνων διατήρησης της μάζας μεταξύ αντιδρώντων και προϊόντων.

Μία ιδιαίτερα σημαντική κατηγορία Κινητικών Νόμων είναι αυτή κατά την οποία ο ρυθμός της αντίδρασης είναι ανάλογος της συγκέντρωσης ή των επιπέδων του κάθε αντιδρώντος μορίου, υψωμένων σε κάποια δύναμη. Οι κινητικοί νόμοι αυτής της μορφής ονομάζονται Κινητικοί Νόμοι με οριστική κατάταξη (Definite Orders) και έχουν την ακόλουθη αναπαράσταση:

$$v_j(x) = k_j \prod_{i=1}^s x_i^{g_{ij}}$$

με την σταθερά ρυθμού της αντίδρασης k_j να αντιπροσωπεύει διάφορους παράγοντες, όπως την επάρκεια ενέργειας και τον κατάλληλο προσανατολισμό των αντιδρώντων για να πραγματοποιηθεί η αντίστοιχη αντίδραση.

Αν τώρα οι αντιδράσεις είναι στοιχειώσεις τότε οι g_{ij} ισούνται με τις στοιχειομετρικές των αντιδράσεων S_{ij} (τις σχετικές με τα αντιδρώντα) και έτσι έχουμε:

$$v_j(x) = k_j \prod_{i=1}^s x_i^{S_{ij}}$$

Οι κινητικοί νόμοι που δίδονται σε αυτή τη μορφή αποτελούν την κατηγορία νόμων της Δράσης Μάζας (Mass Action)

Ως ένα πρακτικό παράδειγμα των παραπάνω, στο μοντέλου Schlögl που δόθηκε στην προηγούμενη ενότητα, αν συμβολίσουμε ως x_A, x_B και x τις συγκεντρώσεις των οντοτήτων A, B και X τότε για την περίπτωση του Κινητικού Νόμου Mass action έχουμε:

$$v_1 = k_1 x_A x^2, v_2 = k_2 x^3, v_3 = k_3 x_B, v_4 = k_4 x$$

και τα επίπεδα της X στο χρόνο δίνονται από την ντετερμινιστική συνήθη διαφορική εξίσωση:

$$\frac{dx}{dt} = v_1 - v_2 + v_3 - v_4 = k_1 x_A x^2 - k_2 x^3 + k_3 x_B - k_4 x$$

Όπως εύκολα μπορεί να γίνει αντιληπτό, άρα, για την ντετερμινιστική προσομοίωση ενός δικτύου βιοχημικών αντιδράσεων, έπειτα από τον ορισμό του, όπως αυτός παρουσιάστηκε προηγουμένως, αρκεί η χρήση ενός προγράμματος επίλυσης Συνήθων Διαφορικών Εξισώσεων. Η λεπτομερής περιγραφή των διαφόρων αλγορίθμων ξεφεύγει από τα όρια της σύντομης παρουσίασης που γίνεται στην παρούσα εργασία, ωστόσο, αυτού του τύπου οι αλγόριθμοι προσφέρονται από μία πληθώρα εργαλείων είτε γενικού σκοπού (όπως η MATLAB, η R και η Mathematica) είτε εξειδικευμένης φύσης για προβλήματα Συστημικής Βιολογίας (όπως τα COPASI, Cell Designer, SimBiology Toolbox).

2.4.1.2.2 Στοχαστική Προσομοίωση

Στην παρούσα ενότητα δίνεται το μαθηματικό πλαίσιο της Στοχαστικής Προσομοίωσης και ο τρόπος ορισμού των Βιοχημικών Δικτύων μέσω αυτού. Η παρουσίαση που ακολουθεί βασίζεται στην αντίστοιχη που δίνεται στην [73].

Θεωρώντας ένα ομοιογενές (well-stirred) σύστημα μοριακών οντοτήτων που αντιδρούν μεταξύ τους μέσω βιοχημικών αντιδράσεων, και κάνοντας την παραδοχή ότι το σύστημα βρίσκεται σε σταθερό όγκο V και σε θερμική (αλλά όχι χημική) ισορροπία, στόχος (και) της Στοχαστικής Προσομοίωσης είναι η εκτίμηση του διανύσματος $\mathbf{X}(t) = (X_1(t), \dots, X_N(t))$, δηλαδή των επιπέδων των διαφόρων οντοτήτων, δεδομένου ότι το σύστημα αρχικά βρισκόταν στην κατάσταση $\mathbf{X}(t_0) = \mathbf{x}_0$.

Κεντρική παραδοχή για την προσέγγιση που θα παρουσιαστεί είναι ότι μπορούμε να περιγράψουμε επαρκώς την κατάσταση του συστήματος, καθορίζοντας μόνο τα επίπεδα των διαφόρων μοριακών οντοτήτων, αγνοώντας τις θέσεις τους και τις ταχύτητές τους. Οι λόγοι που καθιστούν κάτι τέτοιο ρεαλιστικό δίνονται με λεπτομέρεια στην [73].

Όπως και στην ντετερμινιστική προσομοίωση, οι αλλαγές των επιπέδων των οντοτήτων είναι αποτέλεσμα της πραγματοποίησης αντιδράσεων. Για την περίπτωση της Στοχαστικής Προσομοίωσης, κάθε κανάλι αντίδρασης χαρακτηρίζεται από δύο ποσότητες:

- Από το διάνυσμα $\mathbf{S}_j = (S_{1j}, \dots, S_{Nj})$, των αλλαγών στα επίπεδα των οντοτήτων κατά την πραγματοποίηση της αντίδρασης R_j , όπως ακριβώς και στην ντετερμινιστική περίπτωση
- Από τη συνάρτηση propensity a_j η οποία ορίζεται ως $a_j(\mathbf{x})dt \triangleq$ η πιθανότητα, δοσμένου ότι $\mathbf{X}(t) = \mathbf{x}$, να πραγματοποιηθεί μία αντίδραση R_j , εντός του όγκου στον οποίο βρίσκεται το σύστημα, στο επόμενο απειροστό χρονικό διάστημα $[t, t + dt)$.

Για την πιθανότητα αυτή, τώρα, ισχύουν τα εξής:

- Αν η αντίδραση R_j είναι μονομοριακή, η κβαντική μηχανική ορίζει ότι υπάρχει κάποια σταθερά c_j , τέτοια ώστε η πιθανότητα ότι κάθε ξεχωριστή οντότητα X_j θα αντιδράσει στο επόμενο απειροστό χρονικό διάστημα dt , ισούται με $c_j dt$. Έτσι από τους νόμους των πιθανοτήτων προκύπτει ότι αν υπάρχει x_1 μόρια της οντότητας X_1 στο σύστημα, η πιθανότητα ότι κάποιο από αυτά θα ακολουθήσει την R_j αντίδραση είναι $x_1 \cdot c_j dt$. Έτσι για αυτή την περίπτωση:

$$a_j(\mathbf{x}) = c_j x_1$$

- Αν τώρα στην R_j εμπλέκονται δύο μόρια, η κινητική θεωρία αλλά και η ομοιογενής (well-stirred) κατάσταση του συστήματος, ορίζουν ότι υπάρχει μία σταθερά c_j , τέτοια ώστε η πιθανότητα ένα τυχαία επιλεγμένα ζευγάρι μορίων X_1 και X_2 θα αντιδράσει μέσω της R_j στο επόμενο απειροστό χρονικό διάστημα dt , να ισούται με $c_j dt$. Έτσι, η πιθανότητα ότι ένα από τα $x_1 x_2$ ζευγάρια οντοτήτων X_1 και X_2 θα αντιδράσει μέσω της R_j εντός του όγκου του συστήματος ισούται με $x_1 x_2 \cdot c_j dt$ και τότε η συνάρτηση propensity δίνεται από την:

$$a_j(\mathbf{x}) = c_j x_1 x_2$$

- Στην περίπτωση τώρα που τα δύο αντιδρώντα μόρια είναι του ίδιου τύπου, η propensity δίνεται από την:

$$a_j(\mathbf{x}) = c_j \frac{1}{2} x_1 (x_1 - 1)$$

Ιδιαίτερα σημαντικό συμπέρασμα, το οποίο περιγράφεται με λεπτομέρεια στην [73], είναι ότι για τις τιμές των c_j ισχύει:

- Για τις μονομοριακές αντιδράσεις η c_j ισούται αριθμητικά με την τιμή της σταθεράς ρυθμού αντίδρασης k_j όπως αυτή ορίστηκε για τη ντετερμινιστική Προσομοίωση
- Για τις διμοριακές αντιδράσεις η c_j ισούται με $\frac{k_j}{V}$ για την περίπτωση διαφορετικών οντοτήτων και με $\frac{2k_j}{V}$ για την περίπτωση ίδιων οντοτήτων.

Παρά τον ορισμό των παραπάνω, ωστόσο, η πιθανοτική φύση του ορισμού της propensity δεν επιτρέπει ακριβείς προβλέψεις των $\mathbf{X}(t)$. Παρόλα αυτά θα μπορούσε να εκτιμηθεί η πιθανότητα

$$P(\mathbf{x}, t | \mathbf{x}_0, t_0) \triangleq \text{Prob}[\mathbf{X}(t) = \mathbf{x}, \text{δεδομένου ότι } \mathbf{X}(t_0) = \mathbf{x}_0]$$

Από την ανωτέρω εξίσωση και από τον ορισμό της τάσης (Propensity) προκύπτει η Κύρια Χημική Εξίσωση (Chemical Master Equation-CME):

$$\frac{\partial P(\mathbf{x}, t | \mathbf{x}_0, t_0)}{\partial t} = \sum_{j=1}^M [a_j(\mathbf{x} - \mathbf{S}_j) P(\mathbf{x} - \mathbf{S}_j, t | \mathbf{x}_0, t_0) - a_j(\mathbf{x})] P(\mathbf{x}, t | \mathbf{x}_0, t_0)$$

Παρότι η CME ορίζει πλήρως την $P(\mathbf{x}, t | \mathbf{x}_0, t_0)$, μόνο σε πολύ λίγες περιπτώσεις μπορεί να δεχθεί αναλυτική λύση. Έτσι η λύση που προτείνεται είναι η παραγωγή αριθμητικών πραγματοποιήσεων της $\mathbf{X}(t)$ μέσω προσομοίωσης. Για να πραγματοποιηθεί κάτι τέτοιο, χρειάζεται να οριστεί μία νέα πιθανότητα:

$p(\tau, j | \mathbf{x}, t) d\tau \triangleq$ η πιθανότητα, δεδομένου ότι $\mathbf{X}(t) = \mathbf{x}$, η επόμενη αντίδραση στο σύστημα να συμβεί στο απειροστό χρονικό διάστημα $[t + \tau, t + \tau + d\tau)$ και να είναι η R_j

η οποία ουσιαστικά είναι η από κοινού πυκνότητα πιθανότητας των τυχαίων μεταβλητών του χρόνου για την πραγματοποίηση της επόμενης αντίδρασης (τ) και τύπου της αντίδρασης αυτής (j).

Όπως σημειώνεται στην [73] και στις εκεί αναφερόμενες εργασίες, για την συγκεκριμένη πιθανότητα ισχύει:

$$p(\tau, j | \mathbf{x}, t) = a_j(\mathbf{x}) \exp(-a_0(\mathbf{x})\tau)$$

όπου

$$\alpha_0(\mathbf{x}) = \sum_{j'=1}^M a_{j'}(\mathbf{x})$$

Η παραπάνω εξίσωση αποτελεί τη βάση για τη Στοχαστική Προσομοίωση των Δικτύων Βιοχημικών Αντιδράσεων, ορίζοντας ότι:

- η τ ακολουθεί εκθετική κατανομή με μέση τιμή και διακύμανση $\frac{1}{\alpha_0(\mathbf{x})}$
- η j είναι μία στατιστικώς ανεξάρτητη ακέραια τυχαία μεταβλητή με σημειακές πιθανότητες $\frac{a_j(\mathbf{x})}{\alpha_0(\mathbf{x})}$

Για την λήψη δειγμάτων από τις κατανομές των τ και j μπορούν να ακολουθηθούν διάφορες ακριβείς Monte Carlo διαδικασίες [75]. Η σημαντικότερη αλλά και απλούστερη είναι ο Αλγόριθμος Στοχαστικής Προσομοίωσης/Ευθεία Μεθόδος (Stochastic Simulation Algorithm - SSA / Direct Method) ο οποίος παρουσιάζεται στην επόμενη ενότητα.

2.4.1.2.2.1 Ο Αλγόριθμος SSA

Ο αλγόριθμος SSA, Direct Method [76], [77] αποτελεί τον πρώτο αλγόριθμο προσομοίωσης Βιοχημικών Αντιδράσεων. Τα βήματα του αλγορίθμου έχουν ως εξής:

1. Υπολογίζονται δύο τυχαίοι αριθμοί r_1 και r_2 στο διάστημα $[0,1]$ από την ομοιόμορφη κατανομή και λαμβάνονται τα δείγματα:

$$j = \text{o μικρότερος ακέραιος που ικανοποιεί την } \sum_{j'=1}^j a_{j'}(\mathbf{x}) > r_2 \alpha_0(\mathbf{x})$$

2. Αρχικοποιείται ο χρόνος σε $t = t_0$ και το Σύστημα στην κατάσταση $\mathbf{x} = \mathbf{x}_0$
3. Για το Σύστημα στην κατάσταση \mathbf{x} τη χρονική στιγμή t , υπολογίζονται οι τιμές των συναρτήσεων propensity $a_j(\mathbf{x})$ και το άθροισμα αυτών $\alpha_0(\mathbf{x})$
4. Λαμβάνονται δύο δείγματα μέσω των

$$\tau = \frac{1}{\alpha_0(\mathbf{x})} \ln\left(\frac{1}{r_1}\right)$$

και

$$j = \text{o μικρότερος ακέραιος που ικανοποιεί την } \sum_{j'=1}^j a_{j'}(\mathbf{x}) > r_2 \alpha_0(\mathbf{x})$$

5. Πραγματοποιείται η επόμενη αντίδραση αντικαθιστώντας:
 - $t \leftarrow t + \tau$
 - $\mathbf{x} \leftarrow \mathbf{x} + \mathbf{S}_j$
6. Επιστροφή στο βήμα 1 ή τερματισμός της Προσομοίωσης.

Μία ισοδύναμη υλοποίηση του SSA είναι αυτή με τη μέθοδο της Πρώτης Αντίδρασης (First Reaction Method) η οποία παρουσιάστηκε στην ίδια εργασία [77] με αυτή που παρουσιάστηκε ο SSA Direct Method. Η μέθοδος Πρώτης Αντίδρασης, αποτελεί έναν εναλλακτικό τρόπο για την δημιουργία των τιμών των τ και j . Στη συγκεκριμένη μέθοδο, αρχικά λαμβάνονται M τυχαίοι αριθμοί r_1, \dots, r_M από την ομοιόμορφη κατανομή στο διάστημα $[0,1]$ και υπολογίζεται η

$$\tau_{j'} = \frac{1}{a_{j'}^l(\mathbf{x})} \ln\left(\frac{1}{r_{j'}}\right) \quad (j' = 1, \dots, M)$$

Στη συνέχεια λαμβάνονται:

$$\left. \begin{aligned} \tau &= \text{ο μικρότερος μεταξύ των } \{\tau_{j'}\} \\ l &= \text{ο αύξων αριθμός του μικρότερου } \{\tau_{j'}\} \end{aligned} \right\}$$

και τέλος:

$$j = \text{ο μικρότερος ακέραιος που ικανοποιεί την } \sum_{j'=1}^j a_{j'}^l(\mathbf{x}) > r_{L+1} a_0^l(\mathbf{x})$$

Θα πρέπει να σημειωθεί ιδιαίτερα ότι ο SSA και η CME έχουν λογική ισοδυναμία μεταξύ τους. Ωστόσο, παρότι ο SSA καταφέρνει να δώσει λύση στο δισεπίλυτο της CME, είναι υπολογιστικά ιδιαίτερα απαιτητικός παρά την ιδιαίτερα απλή μορφή του. Γι' αυτό το λόγο έχουν προταθεί διάφορες παραλλαγές και εναλλακτικές/προσεγγιστικές λύσεις του, οι οποίες παρουσιάζονται στις ακόλουθες ενότητες.

2.4.1.2.2.2 Η μέθοδος της Επόμενης Αντίδρασης

Δεδομένης της υψηλής υπολογιστικής πολυπλοκότητας που παρουσιάζουν και η Direct Method αλλά και η First Reaction Method, υπήρξε η ανάγκη μείωσης αυτής. Κάτι τέτοιο επετεύχθη από του Gibson και Bruck μέσω της ευφυέστερης διαχείρισης των δομών δεδομένων, προτείνοντας τον αλγόριθμο της Επόμενης Αντίδρασης (Next Reaction) [78]. Ο συγκεκριμένος αλγόριθμος παρουσιάζει τα εξής ιδιαίτερα χαρακτηριστικά:

- Δημιουργείται ένας γράφος εξαρτήσεων μεταξύ των αντιδράσεων. Σε αυτό το γράφο υπάρχει ένας κόμβος για κάθε αντίδραση R_j και μία ακμή (j, k) από τον συγκεκριμένο κόμβο σε κάθε άλλο κόμβο R_k στην οποία συμμετέχει αντιδρών το οποίο αποτελεί αντιδρών ή προϊόν της R_j . Με τη χρήση αυτού του γράφου, αποφεύγεται ο άσκοπος επανυπολογισμός αρκετών από τις τάσεις

(propensities) a_j όταν αυτές δεν επηρεάζονται από τις πραγματοποιηθείσες αντιδράσεις.

- Χρησιμοποιούνται απόλυτοι χρόνοι πραγματοποίησης αντίδρασης από την αρχή της προσομοίωσης αντί των σχετικών από την πραγματοποίηση της τελευταίας αντίδρασης. Επίσης, οι τυχαίοι αριθμοί «ανακυκλώνονται» κατά τη διάρκεια πραγματοποίησης την ανανέωσης του χρόνου πραγματοποίησης αντίδρασης.
- Χρησιμοποιείται μία ουρά προτεραιότητας που περιέχει όλες τις αντιδράσεις, ταξινομημένες βάσει του χρόνου πραγματοποίησης τους. Έτσι η επόμενη αντίδραση μπορεί να βρεθεί μέσω της συγκεκριμένης δομής σε σταθερό χρόνο.

Τα βήματα του αλγορίθμου έχουν ως εξής:

1. Αρχικοποίηση:

- $t = t_0$ και $x = x_0$
- Δημιουργία του γράφου εξαρτήσεων G .
- Υπολογισμός των συναρτήσεων propensity $a_j(x)$ στην κατάσταση x
- Για κάθε j , υπολογισμός του χρόνου τ_j μέχρι την επόμενη R_j αντίδραση

$$\tau_j = t + \frac{1}{a_j(x)} \ln \left(\frac{1}{r_j} \right)$$

με r_j τυχαίους αριθμούς από την ομοιόμορφη κατανομή στο διάστημα $[0,1]$

- Αποθήκευση των χρόνων τ_j στην ουρά προτεραιότητας Q

2. Αν R_μ είναι η αντίδραση με τον μικρότερο χρόνο στην Q , τ_μ
3. Τίθεται t ίσο με τ_μ
4. Υπολογισμός της νέας κατάστασης του συστήματος μετά την πραγματοποίηση της αντίδρασης R_μ , $t = \tau$ και $x = x + S_\mu$
5. Για κάθε ακμή (μ, α) στο γράφο εξαρτήσεων G :
 - Τίθεται $a_{\alpha,old} = a_\alpha$ και ανανεώνεται η a_α
 - Αν $\alpha \neq \mu$, τίθεται $\tau_\alpha = \left(\frac{a_{\alpha,old}}{a_\alpha} \right) (\tau_\alpha - t) + t$
 - Αν $\alpha = \mu$, λαμβάνεται ένας τυχαίος αριθμός r_μ και υπολογίζεται:

$$\tau_{\mu} = \tau + \frac{1}{a_{\mu}(x)} \ln\left(\frac{1}{r_{\mu}}\right)$$

6. Αν t είναι μεγαλύτερος από τον χρόνο προσομοίωσης, τερματισμός
7. Καταγραφή της (x, t) και επιστροφή στο βήμα 2.

2.4.1.3 Εναλλακτικές Μέθοδοι Προσομοίωσης

Όπως συμπερασματικά μπορεί να εξαχθεί από τις προηγούμενες ενότητες, η κάθε κατηγορία μαθηματικής μοντελοποίησης και προσομοίωσης των Δικτύων Βιοχημικών Αντιδράσεων έχει σημαντικά πλεονεκτήματα αλλά και μειονεκτήματα. Στα πλεονεκτήματα της ντετερμινιστικής προσομοίωσης μπορούν να συγκαταλεχθούν οι σχετικά μειωμένες απαιτήσεις σε υπολογιστικό χρόνο, ωστόσο σε πολλές περιπτώσεις παρουσιάζει μειωμένη ακρίβεια. Απεναντίας, η Στοχαστική Προσομοίωση, παρότι ιδιαίτερα ακριβής καθότι λαμβάνει υπόψη της την πραγματοποίηση κάθε αντίδρασης σε διακριτά πλαίσια, παρουσιάζει ιδιαίτερα υψηλές απαιτήσεις υπολογιστικής ισχύος. Στην προσπάθεια, λοιπόν, να συνδυαστούν τα πλεονεκτήματα των δύο προσεγγίσεων, στην βιβλιογραφία έχει προταθεί ένας μεγάλος αριθμός μεθόδων προς αυτό το στόχο, οι οποίες συνοψίζονται στις εργασίες [73] και [79]. Οι δύο κύριες κατηγορίες αυτών των μεθόδων αποτελούν οι Προσεγγιστικά Στοχαστικοί (Approximate Stochastic) και οι Υβριδικοί Αλγόριθμοι. Σημαντικός εκπρόσωπος της πρώτης κατηγορίας αποτελεί ο αλγόριθμος Προσομοίωσης με άλματα (Tau-leaping) ενώ για την δεύτερη κατηγορία, παρότι έχει προταθεί ένας μεγάλος αριθμός αλγορίθμων, οι οποίοι συνοψίζονται στην [79], υπάρχουν ακόμα σημαντικά ανοιχτά ζητήματα. Στις ακόλουθες ενότητες παρουσιάζονται ο αλγόριθμος Tau-leaping και οι βασικές αρχές των Υβριδικών αλγορίθμων.

2.4.1.3.1 Προσομοίωση με άλματα

Ο προσεγγιστικά στοχαστικός αλγόριθμος προσομοίωσης με άλματα, Tau-leaping ή τ -Leap, προτάθηκε από τον ίδιο επιστήμονα από τον οποίο προτάθηκε και ο SSA στην εργασία [80]. Στην συγκεκριμένη μέθοδο αποφεύγεται η προσομοίωση κάθε ενός ξεχωριστού γεγονότος αντίδρασης, αλλά πραγματοποιείται ένα άλμα στο χρόνο, μήκους τ , στο οποίο συμπεριλαμβάνονται πολλά γεγονότα αντίδρασης.

Προϋπόθεση για την εκτέλεση του αλγορίθμου είναι το τ να πληροί την συνθήκη να είναι αρκούτσως μικρό ώστε να μην συμβαίνει σημαντική αλλαγή στις propensities a_j των αντιδράσεων κατά το χρονικό διάστημα $[t, t + \tau]$. Αν αυτή η συνθήκη καλύπτεται, τότε, τα κανάλια αντιδράσεων αποδεσμεύονται μεταξύ τους και ο αριθμός των πραγματοποιήσεων

κάθε αντίδρασης K_j , κατά τη διάρκεια του τ , με αρχική κατάσταση την κατάσταση x και αρχικό χρόνο t , μπορεί να προσεγγιστεί από τυχαίες μεταβλητές που ακολουθούν Poisson κατανομή:

$$K_j(\tau; x, t) = P(a_j(x), \tau)$$

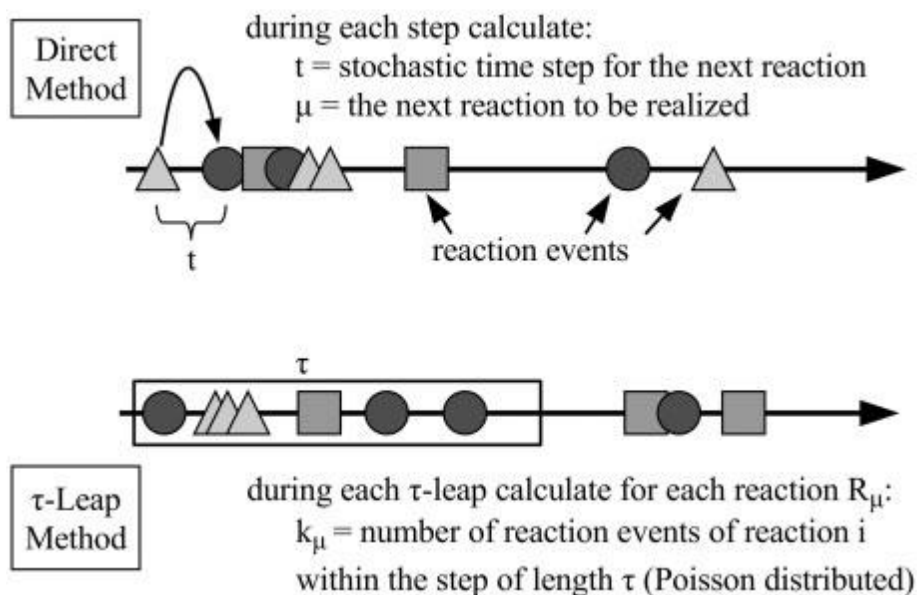
$$\text{με } Prob\{P(a_j, \tau) = k\} = \frac{(a_j \tau)^k}{k!} e^{-a_j \tau}$$

Έτσι, σε κάθε βήμα της προσομοίωσης, λαμβάνεται ένας τυχαίος αριθμός από την κατανομή Poisson k_j για κάθε αντίδραση R_j και η κατάσταση του συστήματος ανανεώνεται βάσει της:

$$x(t + \tau) = x(t) + \sum_{j=1}^M k_j \nu_j$$

Θα πρέπει να σημειωθεί, ότι, λόγω της λήψης δειγμάτων από Poisson κατανομή, κάθε βήμα του Tau-Leaping, είναι υπολογιστικά πιο απαιτητικό σε σχέση με ένα βήμα του SSA (είτε Direct είτε First Reaction Method). Ωστόσο, επειδή πολλά βήματα της προσομοίωσης μπορούν να «προσπεραστούν» όταν το τ είναι αρκούντως μεγάλο, η προσομοίωση γενικά μπορεί να εκτελεστεί πολύ πιο γρήγορα.

Στο ακόλουθο σχήμα, συμπυκνώνεται η φιλοσοφία του αλγορίθμου.



Εικόνα 2.12 Η φιλοσοφία του αλγορίθμου Tau-leaping

2.4.1.3.2 Προσομοίωση μέσω Στοχαστικών Διαφορικών Εξισώσεων

Στην προσπάθεια ακόμα μεγαλύτερης μείωσης του υπολογιστικού κόστους της Στοχαστικής Προσομοίωσης, αν η τιμή του τ μπορεί να επιλεγεί αρκούντως μεγάλη έτσι ώστε κάθε κανάλι αντίδρασης, κατά μέσο όρο, να πραγματοποιεί ένα πολύ μεγάλο αριθμό

πραγματοποιήσεων ($P(a_j(x), \tau) = a_j(x)\tau \gg 1$) ενώ ταυτόχρονα ικανοποιείται η συνθήκη του Tau-Leaping, τότε οι τυχαίες μεταβλητές Poisson μπορούν να προσεγγιστούν από τυχαίες μεταβλητές N_j που ακολουθούν την κανονική κατανομή [81] και έτσι:

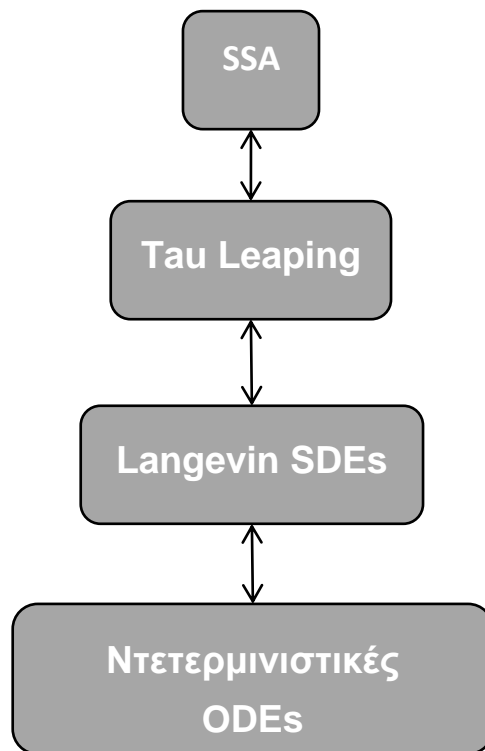
$$\begin{aligned} K_j(\tau; x, t) &= P_j(a_j(x), \tau) \\ &\approx N_j(a_j(x)\tau, a_j(x)\tau) \\ &= a_j(x)\tau + \sqrt{a_j(x)\tau} \cdot N_j(0,1) \end{aligned}$$

Αν ισχύουν τα παραπάνω, τότε οι μεταβάσεις των καταστάσεων του Συστήματος δίνονται από την (χημική) Langevin εξίσωση, η οποία είναι μία Στοχαστική Διαφορική Εξίσωση (Stochastic Differential Equation/SDE) [82], δηλαδή:

$$x(t + \tau) = x(t) + \sum_{j=1}^M v_j a_j(x)\tau + \sum_{j=1}^M v_j \times \sqrt{a_j(x)\tau} \cdot n_j$$

όπου n_j κανονικές τυχαίες μεταβλητές στο διάστημα $[0,1]$.

Όπως είναι εύκολο να παρατηρηθεί, αν $a_j(x)\tau \rightarrow \infty$, τότε ο τελευταίος όρος της παραπάνω εξίσωσης γίνεται ιδιαίτερα μικρότερος σε σχέση με τον δεύτερο. Σε αυτή την περίπτωση, οι επικρατούντες όροι είναι ίδιοι με αυτούς της μεθόδου Euler για την επίλυση ODE. Έτσι όπως σημειώνεται στην [79], υπάρχει μία σύνδεση μεταξύ του SSA και της ντετερμινιστικής προσομοίωσης μέσω ODEs η οποία έχει τα εξής στάδια:

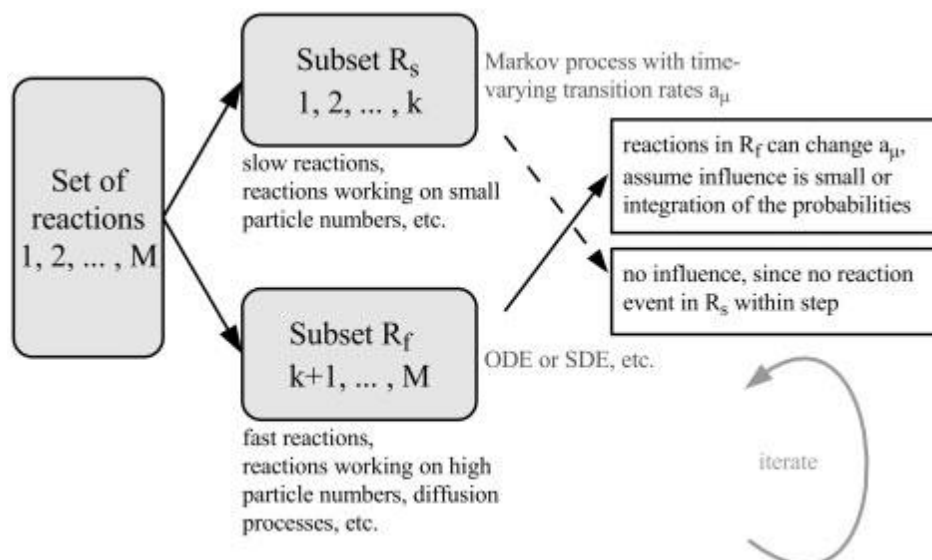


Εικόνα 2.13 Η σχέση μεταξύ των διαφόρων μεθόδων προσομοίωσης [79]

2.4.1.3.3 Υβριδική Προσομοίωση

Όπως προαναφέρθηκε, οι εναλλακτικές μέθοδοι προσομοίωσης, έχουν ως στόχο τον συνδυασμό των πλεονεκτημάτων της ντετερμινιστικής και της στοχαστικής Προσομοίωσης. Σε αντίθεση με τον Tau-Leaping αλλά και την μέθοδο Langevin, οι οποίες τροποποιούν σε σημαντικό βαθμό την μέθοδο με την οποία προσομοιώνεται η πραγματοποίηση των αντιδράσεων, έχουν προταθεί στη βιβλιογραφία (αλλά και εφαρμόζονται στην πράξη ακόμα και από εργαλεία όπως το COPASI) οι Υβριδικοί Αλγόριθμοι, κατά τους οποίους χρησιμοποιούνται ταυτόχρονα αυτοί οι αλγόριθμοι ντετερμινιστικής και στοχαστικής Προσομοίωσης (ωστόσο μπορούν να γίνουν και άλλου τύπου συνδυασμοί όπως συνδυασμός προσομοίωσης με SDE με την προσομοίωση μέσω του SSA).

Για να εφαρμοστεί η φιλοσοφία αυτή, θα πρέπει το σύστημα που επιθυμούμε να προσομοιώσουμε να διαχωριστεί σε δύο τμήματα, έτσι ώστε κάθε προσέγγιση προσομοίωσης να αναλάβει την προσομοίωση ενός από αυτά. Ο πλέον συνηθισμένος διαχωρισμός αφορά στο διαχωρισμό γρήγορων και αργών αντιδράσεων. Οι γρήγορες αντιδράσεις συνήθως αφορούν σε αντιδράσεις που εμπλέκονται οντότητες με υψηλά επίπεδα (όπως π.χ. στον μεταβολισμό) ενώ οι αργές αντιδράσεις σε αντιδράσεις με μικρά επίπεδα οντοτήτων (όπως στην σηματοδότηση ή την έκφραση γονιδίων). Στο ακόλουθο σχήμα δίνεται η γενικότερη φιλοσοφία αυτής της κατηγορίας αλγορίθμων:



Εικόνα 2.14 Η φιλοσοφία των υβριδικών αλγορίθμων προσομοίωσης [79]

Μία λεπτομερής περιγραφή των διαφόρων προσεγγίσεων για τους υβριδικούς αλγορίθμους, ξεφεύγει από τα πλαίσια της παρούσας εργασίας, ενώ δεδομένων των αρκετών ανοιχτών ζητημάτων που παρουσιάζονται ακόμα για αυτούς, δεν υπάρχουν κάποιες συγκεκριμένες μέθοδοι που θα μπορούσαν να περιγραφούν, ως πληρέστερες, αντί άλλων. Τα ανοιχτά ζητήματα αλλά και οι μέχρι σήμερα προταθείσες μέθοδοι παρουσιάζονται στην [79].

2.4.1.4 Ανάλυση και Εκπαίδευση Δικτύων Βιοχημικών Αντιδράσεων

Έχοντας πλέον ορίσει σε μαθηματικό πλαίσιο τα Δίκτυα Βιοχημικών Αντιδράσεων, μπορούμε να εκμεταλλευτούμε ένα μεγάλο φάσμα μαθηματικών εργαλείων έτσι ώστε να πραγματοποιήσουμε διαδικασίες ανάλυσης της δομής και της δυναμικής συμπεριφοράς αυτών. Επίσης, ιδιαίτερα σημαντικό κέρδος δίνεται από την δυνατότητα να εκπαιδευτούν τα Δίκτυα Βιοχημικών Αντιδράσεων έτσι ώστε να παρουσιάζουν συμπεριφορά που μπορεί να ταυτιστεί με την διαθέσιμη ποιοτική και ποσοτική γνώση που προέρχεται από το πραγματικό σύστημα. Στις ακόλουθες ενότητες, περιγράφονται οι σημαντικότερες διαδικασίες ανάλυσης των Δικτύων Βιοχημικών Αντιδράσεων καθώς και η διαδικασία Εκπαίδευσης αυτών η οποία αποτελεί κεντρική διαδικασία για την δημιουργία ενός έγκυρου μοντέλου.

2.4.1.4.1 Μέθοδοι Ανάλυσης Βιοχημικών Δικτύων

2.4.1.4.1.1 Ανάλυση Ευαισθησίας

Όπως παρουσιάστηκε στις προηγούμενες ενότητες, ένα μοντέλο Συστημικής Βιολογίας, όπως τα Δίκτυα Βιοχημικών Αντιδράσεων, απαρτίζονται από έναν αριθμό

παραμέτρων (σταθερές αντιδράσεων, αρχικά επίπεδα οντοτήτων κ.λπ). Σε πολλές περιπτώσεις οι ακριβείς τιμές αυτών των παραμέτρων δεν είναι γνωστές ενώ επίσης δεν είναι γνωστή και η επιρροή που έχει κάθε μία από αυτές στην συμπεριφορά του συστήματος. Η μέθοδος Ανάλυσης Ευαισθησίας, έχει ως στόχο ακριβώς αυτού του βαθμού επιρροής κάθε παραμέτρου στην συμπεριφορά του συστήματος. Κάτι τέτοιο μπορεί να φανεί χρήσιμο για διάφορους λόγους μερικοί εκ των οποίων είναι:

- Σε πολλές περιπτώσεις, οι τιμές αρκετών παραμέτρων είναι άγνωστες, ενώ επίσης η εργαστηριακή τους μέτρηση μπορεί να είναι άμεση αλλά μέχρι και αρκετά δύσκολη. Χρησιμοποιώντας την ανάλυση ευαισθησίας, μπορούμε να πληροφορηθούμε για το αν είναι σημαντικό να γνωρίζουμε/βρούμε την ακριβή τιμή μιας παραμέτρου ή αν μία πιο προσεγγιστική τιμή της αρκεί για τις ανάγκες της μοντελοποίησης. Έτσι μπορούν να αποφασιστούν νέα εργαστηριακά πειράματα αλλά και να ιεραρχηθούν ως προς την κρισιμότητά τους.
- Δεδομένου του στόχου των Μεθόδων Συστημικής Βιολογίας να συνδράμουν στην βιολογική έρευνα, αν στόχος αποτελεί η αλλαγή της συμπεριφοράς ενός συστήματος (παραδείγματος χάριν μέσω ενός φαρμάκου), τότε η ανάλυση ευαισθησίας μπορεί να υποδείξει τους πλέον σημαντικούς παράγοντες για να γίνει η σχετική επέμβαση.
- Η ευρωστία (robustness) των βιολογικών συστημάτων σχετικά με εξωτερικού παράγοντες και επιρροές είναι ιδιαίτερο χαρακτηριστικό τους, δεδομένου ότι οι ζώντες οργανισμοί καλούνται να αντιμετωπίσουν ένα μεγάλο φάσμα περιβαλλοντικών συνθηκών. Κάτι τέτοιο σημαίνει ότι οι βιολογικές διαδικασίες θα πρέπει να παρουσιάζουν σχετικά μειωμένη ευαισθησία στις αλλαγές παραμέτρων. Από την άλλη όμως, οι οργανισμοί θα πρέπει να αντιδρούν και στο περιβάλλον, έτσι κάποιες άλλες διαδικασίες θα πρέπει να είναι ευαίσθητες στις εξωτερικές επιρροές. Οπότε, η ευρωστία, ή η απουσία της, είναι βασικό χαρακτηριστικό των βιολογικών συστημάτων και η ανάλυση ευαισθησίας μπορεί να συμβάλει και στην εξακρίβωση αυτών αλλά και στην κατάλληλη ρύθμιση ενός υπολογιστικού μοντέλου ώστε να αποκτήσει αυτά τα χαρακτηριστικά.

Οι μέθοδοι ανάλυσης ευαισθησίας μπορούν να διαχωριστούν σε δύο βασικές κατηγορίες. Τις Μεθόδους Τοπικής Ανάλυσης Ευαισθησίας (Local Sensitivity Analysis) και τις Μεθόδους Ολικής Ανάλυσης Ευαισθησίας (Global Sensitivity Analysis). Ωστόσο, οι

Μέθοδοι Τοπικής Ανάλυσης Ευαισθησίας είναι αυτές που μέχρι σήμερα έχουν χρησιμοποιηθεί ευρέως για την σχετική ανάλυση των Συστημάτων ενώ υποστηρίζονται από ένα μεγάλο φάσμα εργαλείων συμπεριλαμβανόμενων των COPASI και SimBiology Toolbox. Έτσι, ακολούθως παρουσιάζονται οι βασικές αρχές των Μεθόδων Τοπικής Ανάλυσης Ευαισθησίας (οι οποίες χρησιμοποιήθηκαν στα πλαίσια της παρούσας εργασίας), ενώ ο ενδιαφερόμενος αναγνώστης παραπέμπεται στις [83], [84] για τη μελέτη του ζητήματος της Ολικής Ανάλυσης Ευαισθησίας.

Τα δύο βασικότερα πλαίσια πραγματοποίησης της διαδικασίας της Ανάλυσης Ευαισθησίας, είναι Η Ανάλυση Μεταβολικού Ελέγχου (Metabolic Control Analysis/MCA) και οι Γενικές ή Γενικευμένες Ευαισθησίες (Generic/Generalizes Sensitivities). Ωστόσο, οι δύο αυτές κατηγορίες ορίζονται κάτω από το ίδιο μαθηματικό πλαίσιο. Θα πρέπει να σημειωθεί ωστόσο ότι η χρήση της MCA γενικεύτηκε πριν την αντίστοιχη χρήση των γενικευμένων ευαισθησιών (generic sensitivities), λόγω του πεδίου εφαρμογής της, και έτσι συνηθίζεται ακόμα και για αυτές ο φορμαλισμός να βασίζεται στον αντίστοιχο της MCA ο οποίος δίνεται ακολούθως.

Η ανάλυση Μεταβολικού Ελέγχου είναι μία μέθοδος ποσοτικοποίησης του ελέγχου ή της επιρροής των μεταβολικών μεταβλητών, όπως η συγκέντρωση ενός μεταβολίτη ή ροή μίας σταθερής κατάστασης ενός μεταβολικού συστήματος, από τις παραμέτρους αυτού. Ο έλεγχος αυτό μετριέται μέσω των Συντελεστών Αντίδρασης (Response Coefficients) οι οποίοι ορίζονται ως ο λόγος μεταξύ της σχετικής αλλαγής σε μία μεταβλητή του συστήματος (η αντίδραση αυτού) και της σχετικής αλλαγής στην τροποποιούμενη παράμετρο. Ο ακριβής φορμαλισμός της MCA δίνεται [85] όταν οι response coefficients εκφράζονται ως μερικές παράγωγοι:

$$R_p^Y = \frac{\frac{\partial Y}{Y}}{\frac{\partial P}{P}} = \frac{\partial \ln Y}{\partial \ln P}$$

όπου Y μία μεταβλητή του συστήματος και P η τροποποιούμενη παράμετρος.

Ιδιαίτερα συνηθισμένη εφαρμογή της MCA, είναι η μελέτη του ελέγχου των σταθερών καταστάσεων (steady-state) ροών (fluxes) και των συγκεντρώσεων μεταβολιτών από τις δράσεις των βιοχημικών αντιδράσεων του συστήματος. Σε αυτή την περίπτωση ορίζεται ένα υποσύνολο των response coefficients, οι συντελεστές ελέγχου (control coefficients) οι οποίοι ορίζονται ως εξής:

$$C_v^{[X]_{ss}} = \frac{\partial \ln[X]_{ss}}{\partial \ln v}$$

$$C_v^{J_{ss}} = \frac{\partial \ln J_{ss}}{\partial \ln v}$$

όπου $[X]$ η συγκέντρωση των μεταβολιτών, J η ροή και v ο ρυθμός της αντίδρασης. Θα πρέπει να σημειωθεί ότι οι control coefficients αποτελούν χαρακτηριστικά του συστήματος και εξαρτώνται από όλα τα μέρη αυτού.

Όσον αφορά τις ιδιότητες ευαισθησίας που αφορούν ένα ένζυμο (βασική μορφή βιομορίων στο μεταβολισμό), οι οποίες καλούνται και «τοπικές»(local) ευαισθησίες οι οποίες είναι κρίσιμες για τον έλεγχο, αυτές είναι οι συντελεστές ελαστικότητας (elasticity coefficients). Μέσω των elasticities, μετρείται η σχετική αλλαγή στον ρυθμό μιας ενζυματικής διαδικασίας λόγω μίας σχετικής αλλαγής στα επίπεδα/συγκέντρωση μίας πρωτεΐνης τελεστή (effector protein) [86]. Οι elasticities δίνονται από τον παρακάτω τύπο:

$$\varepsilon_x^v = \frac{\frac{\partial v}{v}}{\frac{\partial x}{x}} = \frac{\partial \ln v}{\partial \ln x}$$

Για την περίπτωση τώρα των γενικευμένων ευαισθησιών (generic sensitivities), μέσω των οποίων επιθυμούμε να υπολογίσουμε την επιρροή οποιασδήποτε παραμέτρου σε οποιαδήποτε μεταβλητή, χρησιμοποιούνται οι response coefficients. Αφού καθοριστούν οι λίστες των στοιχείων του μοντέλου που θα διαφοριστούν (συναρτήσεις) και των μεταβλητών, πραγματοποιείται αριθμητική διαφορίση (numerical differentiation) μέσω πεπερασμένων διαφορών (finite difference). Η τιμή των διαφορών αυτών επιλέγεται από τον χρήστη της εκάστοτε εφαρμογής.

Θα πρέπει να επισημανθεί ξανά ότι και οι response coefficients αλλά και οι control coefficients και οι elasticities παρέχουν πληροφόρηση μόνο για μικρές αλλαγές του μοντέλου (π.χ. της τάξεως του 5%).

Τέλος, σημειώνεται ότι όλα τα παραπάνω μεγέθη (MCA και γενικευμένες ευαισθησίες) μπορούν να αποδοθούν είτε κλιμακωτά (scaled) είτε όχι (unscaled). Οι κλιμακούμενες τιμές περιγράφουν ποσοστιαίες αλλαγές, για παράδειγμα η scaled sensitivity τιμής 0.5 σημαίνει ότι αν αυξηθεί η παράμετρος κατά 10% τότε η τιμή στόχος θα αυξηθεί κατά 5% (0.5 φορές το 10%). Στην unscaled περίπτωση, περιγράφονται οι αλλαγές σε απόλυτες τιμές βάσει των μονάδων που περιγράφονται οι διάφορες οντότητες του συστήματος.

Όσον αφορά το λογισμικό COPASI προσφέρει τον υπολογισμό των ευαισθησιών (sensitivities) και στο πλαίσιο του MCA και στα πλαίσια των γενικευμένων ευαισθησιών (generic sensitivities). Ειδικότερα, το συγκεκριμένο λογισμικό επιτρέπει τον υπολογισμό των γενικευμένων ευαισθησιών και για τις σταθερές καταστάσεις του συστήματος (steady-state) αλλά και για τις χρονοσειρές (time-series) που προκύπτουν από την προσομοίωση αυτού, κάτι που όπως αναφέρεται στην σχετική τεκμηρίωση του εργαλείου, αφορά στην ανάλυση ευαισθησίας για το τελευταίο βήμα της προσομοίωσης (δηλαδή για την κατάσταση που καταλήγει το σύστημα).

2.4.1.4.1.2 Άλλες μέθοδοι Ανάλυσης

Εκτός τις ανάλυσης ευαισθησίας, ένας μεγάλος αριθμός μεθόδων ανάλυσης Δικτύων Βιοχημικών Αντιδράσεων έχει προταθεί στη βιβλιογραφία. Η λεπτομερής παρουσίαση των μεθόδων αυτών ξεφεύγει από τα όρια της παρούσας εργασίας, δεδομένου και ότι δεν χρησιμοποιήθηκε κάποια από αυτές για την δημιουργία του προτεινόμενου μοντέλου, τουλάχιστον στην συγκεκριμένη φάση ανάπτυξής του. Ωστόσο στην ακόλουθη παράγραφο αναφέρονται οι κυριότερες από αυτές, με γνώμονα επιλογής την ευκολία πρόσβασης μέσω των διαφόρων εργαλείων λογισμικού και κυρίως του COPASI μέσω της προσέγγισης του οποίου για αυτές πραγματοποιείται η ακόλουθη παρουσίαση.

Μία ιδιαίτερα χρήσιμη μέθοδος ανάλυσης είναι η Αναζήτηση Παραμέτρων. Μέσω της Αναζήτησης Παραμέτρων, πραγματοποιούνται πολλαπλές προσομοιώσεις ενός μοντέλου με στόχο τη παρατήρηση των διαφορών μεταξύ αυτών. Οι προσομοιώσεις αυτές μπορεί να είναι απλές επαναλήψεις της διαδικασίας προσομοίωσης, με τις ίδιες τιμές παραμέτρων (κάτι ιδιαίτερα χρήσιμο για την περίπτωση της Στοχαστικής Προσομοίωσης) αλλά επίσης μπορούν να αποτελούν και εκτελέσεις με διαφορετικές τιμές παραμέτρων. Οι τιμές των παραμέτρων μπορούν να λαμβάνονται τυχαία από μία κατανομή πιθανότητας για αυτές.

Μία παραπλήσιων στόχων μέθοδος, είναι αυτή του υπολογισμού των εκθετών Lyapunov (Lyapunov exponents) [87]. Οι Lyapunov exponents ενός δυναμικού συστήματος, είναι ποσότητες που χαρακτηρίζουν το ρυθμό διαφοροποίησης μεταξύ δύο απειροστά παρόμοιων ακολουθιών αποτελεσμάτων προσομοίωσης. Λεπτομέρειες για τον τρόπο εφαρμογής στα Δίκτυα Βιοχημικών Αντιδράσεων δίνονται στην [88].

Η μέθοδος Διαχωρισμού μέσω Κλίμακας Χρόνου (Time Scale Separation) αποτελεί μία μέθοδο που συμβάλει στην απλοποίηση των μοντέλων Συστημικής Βιολογίας με στόχο την επιτάχυνση των προσομοιώσεων αλλά και την διαίρεση των Δικτύων Βιοχημικών Αντιδράσεων σε υπο-μοντέλα, δίνοντας έτσι την δυνατότητα απομονωμένης και

αποδοτικότερης ανάλυσης τους. Λεπτομέρειες για την συγκεκριμένη μέθοδο ανάλυσης δίνονται στις [89], [90].

Τέλος, μία ιδιαίτερα χρήσιμη μέθοδος ανάλυσης, που όμως δεν έχει νόημα εφαρμογής σε όλα τα μοντέλα, είναι αυτή του υπολογισμού Σταθερής Κατάστασης (Steady State). Μία σταθερή κατάσταση, αποτελεί η κατάσταση εκεί όπου οι μεταβλητές κατάστασης του μοντέλου για παράδειγμα οι συγκεντρώσεις των οντοτήτων του μοντέλου, δεν αλλάζουν πλέον παρά την εξέλιξη του χρόνου. Λεπτομέρειες για την συγκεκριμένη μέθοδο δίνονται στην [91].

2.4.1.4.2 Εκπαίδευση Βιοχημικών Δικτύων-Εκτίμηση Παραμέτρων

Όπως παρουσιάστηκε στις προηγούμενες ενότητες, βασική προϋπόθεση για την δημιουργία ενός πλήρους μοντέλου Συστημικής Βιολογίας, είναι η γνώση των τιμών των παραμέτρων των χημικών αντιδράσεων. Αν και από μόνα τους τα μοντέλα μπορούν να οδηγήσουν σε σημαντικά συμπεράσματα, κυρίως ποιοτικά, η όσο το δυνατόν ακριβέστερη γνώση των συγκεκριμένων παραμέτρων, μπορεί να οδηγήσει στην δημιουργία ενός μοντέλου που θα προσομοιώνει με σημαντική ακρίβεια και ικανοποιητικής αριθμητική προσέγγισης, τα παρατηρούμενα βιολογικά φαινόμενα. Κεντρική οδός για την απόκτηση αυτών των τιμών των παραμέτρων, πέραν της αναζήτησης στη σχετική εργαστηριακή βιβλιογραφία, αποτελεί η διαδικασία εκτίμησης παραμέτρων.

Η διαδικασία εκτίμησης παραμέτρων αναφέρεται σε εκείνη τη διαδικασία όπου δεδομένου ενός Βιοχημικού Δικτύου Συστημικής Βιολογίας και ενός συνόλου εργαστηριακών ποσοτικοποιημένων δεδομένων για κάποιες ή όλες τις οντότητες του μοντέλου, επιχειρείται μέσω υπολογιστικών μεθόδων να υπολογιστούν οι παράμετροι των αντιδράσεων όπως αυτές παρουσιάστηκαν στις προηγούμενες ενότητες, με στόχο την δημιουργία ενός μοντέλου που οι προσομοιώσεις τους θα επιστρέφουν αποτελέσματα όσο το δυνατόν ομοιότερα με τα διαθέσιμα εργαστηριακά δεδομένα.

Θα πρέπει να σημειωθεί ότι όπως και οι αλγόριθμοι Προσομοίωσης, χωρίζονται σε δύο κύριες κατηγορίες, του Ντετερμινιστικούς και του Στοχαστικούς, έτσι και οι διάφορες προσεγγίσεις της διαδικασίας Εκτίμησης Παραμέτρων χωρίζονται σε:

- μεθόδους που είναι κατάλληλες για μοντελοποιήσεις ντετερμινιστικής προσέγγισης και έχουν ως κυριότερο εργαλείο ορισμού του προβλήματος τις Συνήθειες Διαφορικές Εξισώσεις (ODEs)
- μεθόδους που είναι κατάλληλες για μοντελοποιήσεις Στοχαστικής προσέγγισης με κυριότερο εργαλείο ορισμού του προβλήματος την Chemical Master Equation.

Σε κάθε περίπτωση, ωστόσο, όπως συμβαίνει και με τους αλγορίθμους Προσομοίωσης, η καταλληλότητα της κάθε προσέγγισης εξαρτάται από τα χαρακτηριστικά του κατά περίπτωση μοντέλου, χωρίς να αποκλείεται κατ' ανάγκην η εφαρμογή κάποιων μεθόδων.

Τέλος, θα πρέπει να επισημανθεί ότι ενώ για τον Ντετερμινιστικό χώρο η διαδικασία της εκτίμησης παραμέτρων θεωρείται σε ένα βαθμό ώριμη, για τον Στοχαστικό χώρο, επί του παρόντος, είναι λίγες συγκριτικά οι εργασίες που προτείνουν λύσεις και έτσι δεν θεωρείται διεργασία ρουτίνας [92].

Στις ακόλουθες ενότητες παρουσιάζονται οι κυριότερες προσεγγίσεις που έχουν προταθεί στην πρόσφατη βιβλιογραφία για την διαδικασία της Εκτίμησης Παραμέτρων.

2.4.1.4.2.1 Εκτίμηση Παραμέτρων μέσω μεθόδων Βελτιστοποίησης

Η Εκτίμηση Παραμέτρων Βιοχημικών Δικτύων μέσω μεθόδων Βελτιστοποίησης αποτελεί την πλέον διαδεδομένη πρακτική για την συγκεκριμένη διαδικασία. Ουσιαστικά αποτέλεσε την πρώτη μορφή Εκτίμησης Παραμέτρων ενώ λόγω της γενικότητάς της, όπως θα παρουσιαστεί ακολούθως, μπορεί συνεχώς να εξελίσσεται όσο εξελίσσεται το πεδίο των Μεθόδων Βελτιστοποίησης. Στις παρακάτω ενότητες δίνεται το θεωρητικό πλαίσιο της εφαρμογής των μεθόδων αυτών στην Εκτίμηση Παραμέτρων ενώ επίσης παρουσιάζονται οι κυριότερες μέθοδοι που έχουν εφαρμοστεί με επιτυχία αλλά και σημαντικά συμπεράσματα από την βιβλιογραφία.

2.4.1.4.2.1.1 Θεωρητικό Πλαίσιο Εκτίμησης Παραμέτρων μέσω Μεθόδων Βελτιστοποίησης

Όπως αναφέρθηκε και προηγουμένως, τα συστήματα (Βιοχημικά Δίκτυα) των οποίων τις παραμέτρους θέλουμε να εκτιμήσουμε, χρησιμοποιώντας διαθέσιμα εργαστηριακά δεδομένα, είναι μη γραμμικά. Στα πλαίσια των Μεθόδων Βελτιστοποίησης το πρόβλημα ορίζεται ως Πρόβλημα Μη-Γραμμικού Προγραμματισμού (NLP) με διαφορικούς-αλγεβρικούς περιορισμούς (DAEs). Οι περιορισμοί εξάγονται από τις Διαφορικές Εξισώσεις που περιγράφουν το σύστημα, όπως παρουσιάστηκαν στις προηγούμενες ενότητες. Επίσης, ανάλογα με τις ανάγκες του προβλήματος εκτίμησης παραμέτρων, μπορούν να οριστούν και άλλοι αλγεβρικοί περιορισμοί, πρακτική που ακολουθείται και στα πλαίσια της παρούσας εργασίας για την ανάπτυξη του προταθέντος μοντέλου.

Στον πλήρη ορισμό του προβλήματος της Εκτίμησης Παραμέτρων μέσω Μεθόδων Βελτιστοποίησης, αναζητούμε τις τιμές ενός συνόλου παραμέτρων θ , ώστε να ελαχιστοποιείται η παρακάτω συνάρτηση κόστους:

$$J = \int_0^{t_f} \left(y_{exp}(t) - y_{sim}(\boldsymbol{\theta}, t) \right)^T W(t) (y_{exp}(t) - y_{sim}(\boldsymbol{\theta}, t))$$

Η παραπάνω συνάρτηση κόστους υπόκειται στους παρακάτω περιορισμούς:

- $\mathbf{f} \left(\frac{dx}{dt}, \mathbf{x}, \mathbf{y}, \boldsymbol{\theta}, \mathbf{v}, t \right) = 0$
- $\mathbf{x}(t_0) = \mathbf{x}_0$
- $\mathbf{h}(\mathbf{x}, \mathbf{y}, \boldsymbol{\theta}, \mathbf{v}) = 0$
- $\mathbf{g}(\mathbf{x}, \mathbf{y}, \boldsymbol{\theta}, \mathbf{v}) \leq 0$
- $\boldsymbol{\theta}_{l_{limit}} \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_{u_{limit}}$

Όπου:

- $\boldsymbol{\theta}$: Διάνυσμα Παραμέτρων – μεταβλητές προβλήματος βελτιστοποίησης
- $\mathbf{W}(t)$: Πίνακας βαρών
- \mathbf{x} : Οι μεταβλητές κατάστασης του Συστήματος (συγκεντρώσεις-αριθμός μορίων)
- \mathbf{v} : Διάνυσμα με παραμέτρους που δεν εκτιμούνται και συνήθως δεν διαφοροποιούνται στο χρόνο
- \mathbf{f} : Το σύνολο των περιορισμών που περιγράφουν τη δυναμική του Συστήματος
- \mathbf{h}, \mathbf{g} : Πιθανοί περιορισμοί που περιγράφουν πρόσθετα χαρακτηριστικά του Συστήματος

Στην πράξη (π.χ. στο εργαλείο COPASI) η συνάρτηση κόστους λαμβάνει την παρακάτω μορφή αθροίσματος τετραγώνων με βάρη:

$$J = \sum_{i,j} w_j \left(y_{exp_{i,j}} - y_{sim_{i,j}}(\boldsymbol{\theta}) \right)^2$$

Όπου με i συμβολίζονται οι γραμμές του συνόλου δεδομένων (συνήθως χρονικές στιγμές) και με j οι στήλες (διαφορετικές οντότητες του Συστήματος).

Με w_j συμβολίζονται τα βάρη που χρησιμοποιούνται για κάθε στήλη (οντότητα) του συνόλου δεδομένων. Οι πιο συνήθεις τρόποι υπολογισμού αυτών των βαρών (που υιοθετούνται και στο COPASI) είναι:

- Mean: $w_j = \frac{1}{|\langle x_j \rangle|}$
- Mean Square: $w_j = \frac{1}{\sqrt{\langle x_j^2 \rangle}}$
- Standard Deviation $w_j = \frac{1}{\langle x_j^2 \rangle - \langle x_j \rangle^2}$

Θα πρέπει να σημειωθεί ότι για κάθε ένα διαφορετικό εργαστηριακό πείραμα (αν υπάρχουν διαθέσιμα δεδομένα πολλαπλών πειραμάτων για το ίδιο σύστημα), υπολογίζεται η τιμή του παραπάνω αθροίσματος ξεχωριστά για κάθε πείραμα και η τελική τιμή της συνάρτησης κόστους υπολογίζεται από το άθροισμα των επιμέρους. Επίσης, τα βάρη κλιμακώνονται έτσι ώστε για κάθε πείραμα το μέγιστο βάρος να είναι 1.

2.4.1.4.2.1.2 Καταλληλότητα Εφαρμογής μεθόδων Βελτιστοποίησης και Κυριότερες Μέθοδοι

Όπως προαναφέρθηκε, τα μοντέλα Συστημικής Βιολογίας είναι στην γενική περίπτωση μη-γραμμικά και περιέχουν περιορισμούς. Έτσι, το πρόβλημα Εκτίμησης Παραμέτρων είναι συχνά πολυτροπικό (multimodal) ή μη-κυρτό (nonconvex) πρόβλημα Βελτιστοποίησης (δηλαδή η συνάρτηση κόστους μπορεί να παρουσιάζει πολλαπλά τοπικά ελάχιστα). Έτσι, οι κλασσικές μέθοδοι τοπικής αναζήτησης (ακόμα και αυτές που ακολουθούν τυχαίες επανεκκινήσεις), δεν μπορούν να εγγυηθούν την εύρεση του ολικού ελαχίστου αλλά μπορούν να αποδειχθούν και μη αποδοτικές λόγω επανάληψης των ίδιων σεναρίων. Έτσι, στα συγκεκριμένα προβλήματα, ακολουθούνται μέθοδοι Ολικής Βελτιστοποίησης (Global Optimization-GO) [93].

Οι Μέθοδοι Ολικής Βελτιστοποίησης μπορούν να ομαδοποιηθούν σε δύο κύριες κατηγορίες:

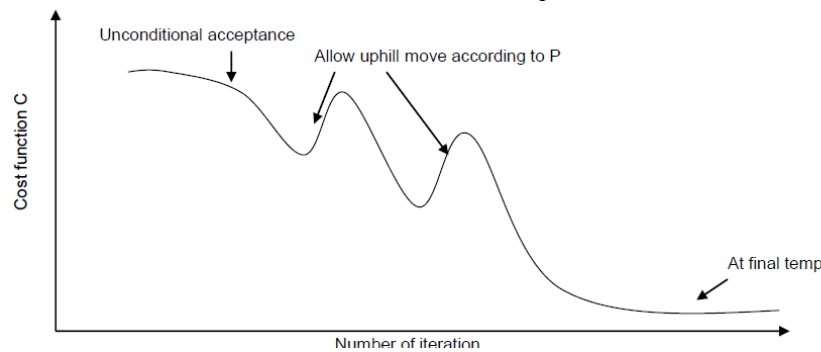
- **Ντετερμινιστικές Μέθοδοι:** Μπορούν να εγγυηθούν την εύρεση ολικού ελαχίστου και για προβλήματα Μη-Γραμμικών Δυναμικών Συστημάτων, ωστόσο παρουσιάζουν ιδιαίτερα υψηλές απαιτήσεις σε χρόνο.
- **Στοχαστικές Μέθοδοι:** Επειδή βασίζονται σε πιθανοτικές προσεγγίσεις δεν εγγυώνται την εύρεση του Ολικού Ελαχίστου ωστόσο μπορούν να εντοπίσουν την περιοχή όπου αυτό βρίσκεται ιδιαίτερα αποδοτικά. Γι' αυτό το λόγο είναι οι κύριες μέθοδοι Βελτιστοποίησης που χρησιμοποιούνται στην Εκτίμηση Παραμέτρων Βιοχημικών Δικτύων [94].

Με βάση τα συμπεράσματα που δίνονται στις [94], [95] (εργασίες στις οποίες παραπέμπεται ο αναγνώστης για περαιτέρω λεπτομέρειες και αναφορές), ακολουθώς παρουσιάζονται, εν συντομία, οι κυριότερες και πλέον διαδεδομένες Στοχαστικές Μέθοδοι Ολικής Βελτιστοποίησης οι οποίες είναι:

- **Προσαρμοζόμενες Στοχαστικές Μέθοδοι (ή Προσαρμοζόμενη Τυχαία Αναζήτηση) (Adaptive Stochastic Methods (ή adaptive random search)):** Βασίζεται στην απλούστερη μορφή Στοχαστικής Βελτιστοποίησης, την Τυχαία

Αναζήτηση (Random Search), ωστόσο εστιάζει την αναζήτηση στην περιοχή όπου έχει βρεθεί η μέχρι τώρα καλύτερη λύση.

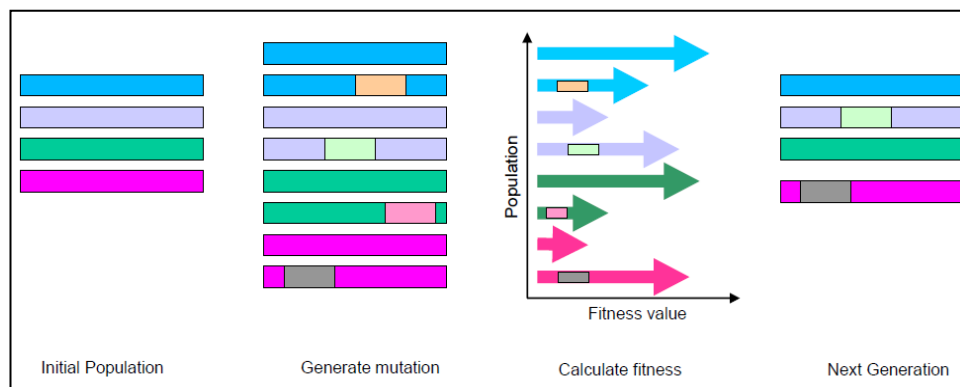
- **Μέθοδοι Ομαδοποίησης (Clustering Methods)** : Μέθοδοι που βασίζονται στην ιδέα των τυχαίων επανεκκινήσεων από διαφορετικές αρχικές καταστάσεις. Ωστόσο προσπαθούν να εντοπίσουν τις περιοχές κοντά στο τοπικό βέλτιστο και έτσι να αποφύγουν την πολλαπλή εύρεση των ίδιων τοπικών βέλτιστων.
- **Προσομοιωμένη Ανόπτηση (Simulated Annealing) (SA)**: Μέθοδος εμπνευσμένη από την διαδικασία της ανόπτησης στην μεταλλουργεία. Επιτρέπει να πραγματοποιούνται κινήσεις που δεν βελτιώνουν τη λύση πάντα ώστε να αποφεύγεται ο εγκλωβισμός σε τοπικά βέλτιστα. Η πιθανότητα αποδοχής τέτοιου είδους κίνησης μειώνεται συνεχώς (ανάλογο της θερμοκρασίας). Στο ακόλουθο σχήμα δίνεται ένα παράδειγμα εκτέλεσης του συγκεκριμένου αλγορίθμου.



Εικόνα 2.15 Παράδειγμα εκτέλεσης του αλγορίθμου της προσομοιωμένης ανόπτησης (simulated annealing)

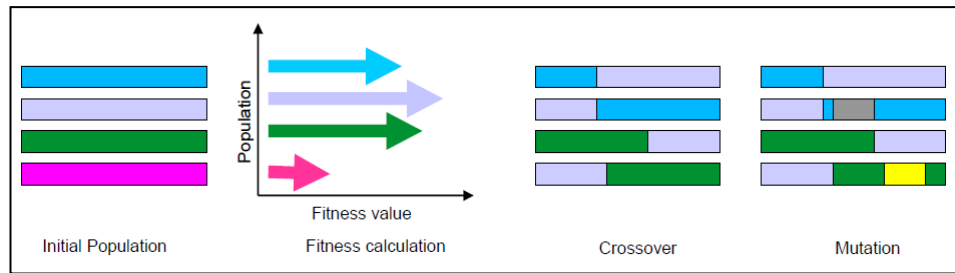
- **Εξελικτικός Υπολογισμός (Evolutionary Computation) (EC)**: Μέθοδοι που βασίζονται στην ιδέα της βιολογικής εξέλιξης μέσω των μηχανισμών αναπαραγωγής και μετάλλαξης και του δόγματος της επιβίωσης του ισχυρότερου. Οι συγκεκριμένες μέθοδοι βρίσκουν συνεχώς καλύτερες λύσεις, δημιουργώντας νέες γενιές. Ακολούθως παρουσιάζονται οι βασικότεροι αλγόριθμοι αυτής της κατηγορίας έχοντας ως βάση την παρουσίαση που δίνεται στα συνοδευτικά κείμενα του λογισμικού COPASI [96].
 - **Εξελικτικός Προγραμματισμός (Evolutionary Programming) (EP)**: υπολογιστική τεχνική που μιμείται την εξέλιξη και βασίζεται στην αναπαραγωγή και την επιλογή. Μία εκτέλεση του αλγορίθμου αποτελείται από οντότητες που αναπαράγονται και ανταγωνίζονται. Κάθε μία αποτελεί μια πιθανή λύση για το πρόβλημα της βελτιστοποίησης και εκπροσωπείται από ένα «γονιδίωμα», όπου κάθε γονίδιο αντιστοιχεί σε μία από τις ρυθμιζόμενες παραμέτρους. Σε κάθε γενιά της εκτέλεσης του EP, κάθε

οντότητα αναπαράγεται, δηλαδή διαιρείται σε δύο νέες οντότητες. Η μία από αυτές περιέχει ακριβώς το ίδιο «γονιδίωμα» με την οντότητα γονέα, ενώ η άλλη επιδέχεται κάποιες μεταλλάξεις (μικρές αλλαγές στις τιμές παραμέτρων). Στο τέλος κάθε γενιάς, ο αλγόριθμος έχει διπλασιάσει τον αριθμό των οντοτήτων. Στη συνέχεια, κάθε μία από τις οντότητες συγκρίνεται με μία σειρά άλλων για τον υπολογισμό των νικών της. Ο αριθμός των νικών είναι ο αριθμός των ανταγωνιστών αυτών που αντιπροσωπεύουν χειρότερες λύσεις από ότι η ίδια. Όλες οι οντότητες βαθμολογούνται ανάλογα με τον αριθμό των νικών τους, και ο πληθυσμός μειώνεται και πάλι στον αρχικό αριθμό οντοτήτων, με την εξάλειψη εκείνων που αντιπροσωπεύουν τις χειρότερες λύσεις. Ακολουθώς δίνεται ένα γραφικό παράδειγμα της φιλοσοφίας του αλγορίθμου:



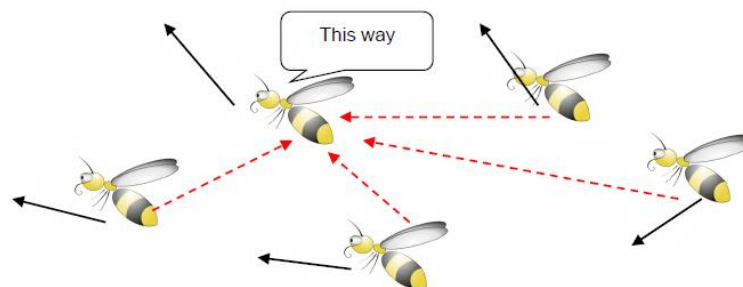
Εικόνα 2.16 Η φιλοσοφία του αλγορίθμου του εξελικτικού προγραμματισμού (evolutionary programming) [94]

- **Εξελικτική Στρατηγική (Evolutionary Strategy) (SRES):** Αλγόριθμος παρόμοιο με τον EP. Ωστόσο, κάθε γονέας δημιουργεί πολλαπλούς απογόνους σε κάθε γενιά και κάθε απόγονος περιέχει ένα συνδυασμό γονιδίων που προέρχονται και από άλλο γονέα αλλά και από μεταλλάξεις.
- **Γενετικός Αλγόριθμος (Genetic Algorithm) (GA):** Αλγόριθμος ο οποίος είναι σε μεγάλο βαθμό παρόμοιος με τον EP, ωστόσο διαφέρουν στο ότι ο GA χρησιμοποιεί εκτός από τις μεταλλάξεις και τον συνδυασμό (Crossover) του «γονιδιώματος» δύο γονέων για την παραγωγή των νέων οντοτήτων. Στο ακόλουθο σχήμα δίνεται μία συνοπτική περιγραφή των βημάτων του αλγορίθμου.



Εικόνα 2.17 Η φιλοσοφία του αλγορίθμου Genetic Algorithm [94]

- **Γενετικός Αλγόριθμος με Στοχαστική Κατάταξη (Genetic Algorithm with Stochastic Ranking) (GASR):** Αλγόριθμος σχεδόν ταυτόσημος με τον Genetic Algorithm, διαφέρει ωστόσο σε κάποια δομικά σημεία, όπως ο τρόπος που δημιουργούνται οι μεταλλάξεις αλλά και η μέθοδος που γίνεται η επιλογή, όπου στην συγκεκριμένη περίπτωση γίνεται μέσω μιας bubble sort διαδικασίας που χρησιμοποιεί τυχαίους παράγοντες.
- **Σμήνος Σωματιδίων (Particle Swarm) (PS):** Εμπνευσμένος από τον τρόπο που ένα σμήνος πουλιών ή ένα κοπάδι ψαριών αναζητούν τροφή. Κάθε μέλος της ομάδας αποτελεί ένα σωματίδιο (Particle). Κάθε σωματίδιο χαρακτηρίζεται από μία θέση X_i και από μία ταχύτητα U_i στον χώρο των παραμέτρων. Επιπλέον, κάθε σωματίδιο, αποθηκεύει την καλύτερη τιμή της χαρακτηριστικής συνάρτησης που έχει επιτύχει και την θέση στην οποία την επέτυχε. Με βάση τώρα την γνώση κάθε σωματίδιο αλλά και την θέση των γειτόνων του (ένα τυχαίο υποσύνολο αυτών), υπολογίζεται κάθε φορά η νέα ταχύτητα του σωματιδίου και κατ' αυτό τον τρόπο κινείται προς την νέα θέση στο χώρο παραμέτρων. Ακολουθώντας δίνεται μία γραφική αναπαράσταση την ιδέα που διέπει τον Particle Swarm.



Εικόνα 2.18 Αναπαράσταση της κεντρικής ιδέας του αλγορίθμου Particle Swarm [94]

2.4.1.4.2.1.3 Συμπεράσματα από τη Βιβλιογραφία, Πλεονεκτήματα και Μειονεκτήματα των μεθόδων Βελτιστοποίησης

Βάσει των όσων παρουσιάστηκαν προηγουμένως, όσον αφορά την εφαρμογή των Μεθόδων Βελτιστοποίησης στην Εκτίμηση Παραμέτρων Δικτύων Βιοχημικών Αντιδράσεων, μπορούν να εξαχθούν κάποια βασικά Πλεονεκτήματα και Μειονεκτήματα τα οποία παρατίθενται ακολούθως:

- Πλεονεκτήματα
 - Αποτελούν ιδιαίτερα μελετημένες μεθόδους με εφαρμογές σε διάφορους τομείς και πλούσια βιβλιογραφία.
 - Υλοποιούνται σε μεγάλο αριθμό γενικών αλλά και εξειδικευμένων εργαλείων (π.χ. COPASI ή COmplex PAthway SIMulator)
 - Δεν επηρεάζονται από τις λεπτομέρειες του Βιολογικού-Βιοχημικού Συστήματος
 - Δεν επεμβαίνουν στην επιλογή αλγορίθμου προσομοίωσης του Συστήματος, αν και στην πραγματικότητα απαιτούν τη χρήση Ντετερμινιστικών αλγορίθμων ώστε να είναι ισοδύναμες οι συγκρίσεις μεταξύ των επαναλήψεων (λόγω της αβεβαιότητας στα αποτελέσματα των Στοχαστικών Αλγορίθμων).
- Μειονεκτήματα
 - Ανάλογα με την επιλογή αλγορίθμου Βελτιστοποίησης, αναζητούν λίγο έως πολύ εξαντλητικά τον χώρο των παραμέτρων (πιθανούς συνδυασμούς τιμών για το διάνυσμα θ)
 - Σε πολλές περιπτώσεις έχουν υψηλές απαιτήσεις υπολογιστικών πόρων [97]
 - Δεν παρουσιάζουν καλά αποτελέσματα όταν ο θόρυβος μέτρησης (κατά τη δημιουργία του συνόλου εργαστηριακών δεδομένων) είναι σημαντικός [97]
 - Στον ορισμό της Συνάρτησης Κόστους δεν ενσωματώνεται ο πιθανός Θόρυβος των Μετρήσεων αλλά και ο Εσωτερικός Θόρυβος του Συστήματος

Παρόλα τα μειονεκτήματά τους, ωστόσο, αποτελούν την αμεσότερη μέθοδο Εκτίμησης Παραμέτρων και λόγω της διαθεσιμότητας του σε διάφορα εργαλεία αποτελούν μία άμεση λύση, τουλάχιστον για τις πρώτες προσπάθειες εκτίμησης παραμέτρων σε ένα υπό δημιουργία μοντέλο. Ωστόσο, στην βιβλιογραφία έχει προταθεί ένα μεγάλος αριθμός εναλλακτικών μεθόδων για την εκτίμηση παραμέτρων, αντιπροσωπευτικά και κυριότερα δείγματα των οποίων παρουσιάζονται, εν συντομία, στην ακόλουθη ενότητα.

Όσον αφορά τη σύγκριση των διαφόρων αλγορίθμων Βελτιστοποίησης, αν και σίγουρα τα αποτελέσματα τους εξαρτώνται και από το πρόβλημα στο οποίο εφαρμόζονται, μία εικόνα για τα χαρακτηριστικά απόδοσης των αλγορίθμων EP,GA,SA και PSO μπορεί να σχηματιστεί από τα αποτελέσματα που δίνονται στην [94]. Στην συγκεκριμένη εργασία εκτιμούνται 4 παράμετροι από μοντέλο 15 παραμέτρων και τα συμπεράσματα είναι ότι την καλύτερη ακρίβεια στην εκτίμηση παραμέτρων πετυχαίνει ο PS αλλά ο χρόνος που απαιτεί είναι σχεδόν πενταπλάσιος αυτού που απαιτείται από τους EP και GA. Τέλος, αν και ο SA καταφέρνει να βρεθεί σε μία περιοχή παραμέτρων ιδιαίτερα κοντά στις πραγματικές σε πάρα πολύ γρήγορο χρόνο, χρειάζεται τον περισσότερο χρόνο από όλους τους δοκιμαζόμενους αλγορίθμους ώστε να ολοκληρώσει την διαδικασία εκτίμησης παραμέτρων. Με βάση αυτά τα συμπεράσματα, όπως και αντίστοιχα που δίνονται σε άλλες εργασίες όπως η [95], έγιναν οι επιλογές αλγορίθμων εκτίμησης παραμέτρων για την εκπαίδευση των μοντέλων που δημιουργήθηκε στα πλαίσια της εργασίας και περιγράφονται με λεπτομέρεια στα Κεφάλαια 3,4 και 5.

2.4.1.4.2.2 Εναλλακτικές μέθοδοι Εκτίμησης Παραμέτρων

Όπως αναφέρθηκε προηγουμένως, η εκτίμηση παραμέτρων μέσω Μεθόδων Βελτιστοποίησης, παρουσιάζει κάποια σημαντικά μειονεκτήματα. Λύση στα μειονεκτήματα αυτά έρχονται να δώσουν πιο εξεζητημένες μέθοδοι εκτίμησης παραμέτρων οι οποίες συνεχώς αυξάνονται σε αριθμό στη σχετική βιβλιογραφία. Οι συγκεκριμένες μέθοδοι βασίζονται σε αλγορίθμους και μεθόδους που αναπτύχθηκαν στο χώρο της Μηχανικής Μάθησης και του Αυτομάτου Ελέγχου, οι οποίοι, βέβαια, αποτελούν συγγενικούς χώρους. Ειδικότερα, όσον αφορά τις μεθόδους Μηχανικής Μάθησης, οι κυριότερες μέθοδοι που έχουν προταθεί στην πρόσφατη βιβλιογραφία μπορούν να ομαδοποιηθούν σε δύο κύριες κατηγορίες, αυτές που για την εκτίμηση των παραμέτρων βασίζονται στην Μεγιστοποίηση της Πιθανοφάνειας (Likelihood) [98] και αυτές που χρησιμοποιούν τις αρχές της Bayesian Στατιστικής [99] και της σχετικής μεθοδολογίας Συμπερασμού (Inference) [100]. Από τις μεθόδους Αυτομάτου Ελέγχου, οι σχετικές εφαρμογές αναφέρονται στην χρήση των Kalman Filter [101] και των Particle Filter [102], οι οποίες ωστόσο παρουσιάζουν αρκετά κοινά στη φιλοσοφία του σε σχέση με τις Bayesian μεθόδους.

Στα πλαίσια της εκπόνησης της παρούσας εργασίας, πραγματοποιήθηκε μία ευρεία ανασκόπηση της βιβλιογραφίας που αφορά την εφαρμογή των παραπάνω κατηγοριών μεθόδων στην Εκτίμηση Παραμέτρων Δικτύων Βιοχημικών Αντιδράσεων. Επειδή, όμως, οι συγκεκριμένες μέθοδοι, επί του παρόντος, δεν έχουν δοκιμαστεί ευρέως και δεν έχουν διαδοθεί σημαντικά στην κοινότητα της Συστημικής Βιολογίας, δεν επιλέχθηκε η εφαρμογή

τους, στα πρώτα στάδια ανάπτυξης του προτεινόμενου μοντέλου. Έτσι η παρουσίαση της συγκεκριμένης ανασκόπησης (η οποία βάσει της διαθέσιμης γνώσης δεν έχει παρουσιαστεί ξανά στην ελληνική γλώσσα), δίνεται στο Παράρτημα I, ενώ στην ακόλουθη υπό-ενότητα, δίνονται τα γενικότερα συμπεράσματα αυτής. Σε κάθε περίπτωση, ωστόσο, η αξιοποίηση των συγκεκριμένων μεθόδων αποτελεί μελλοντικό στόχο της εργασίας, όπως περιγράφεται και στο σχετικό Κεφάλαιο, με στόχο την αναζήτηση όλων των περιθωρίων βελτίωσης του μοντέλου.

2.4.1.4.2.3 Συμπεράσματα ανασκόπησης μεθόδων Εκτίμησης Παραμέτρων

Στις παραπάνω ενότητες, παρουσιάστηκαν οι πλέον αντιπροσωπευτικές και οι πιο διαδεδομένες προσεγγίσεις του προβλήματος της Εκτίμησης Παραμέτρων Δικτύων Βιοχημικών Αντιδράσεων. Για την εξαγωγή συμπερασμάτων, θα πρέπει να συνεκτιμήσουμε την υφιστάμενη κατάσταση γύρω από την Μοντελοποίηση και την Προσομοίωση Βιοχημικών Δικτύων, την διαθεσιμότητα και την μορφή των διαθέσιμων δεδομένων αλλά και την καταλληλότητα εφαρμογής των διαφόρων μεθόδων.

Στις μέρες μας η εξέλιξη στις μεθόδους Βιοτεχνολογίας έχει δώσει τη δυνατότητα σε πολλές περιπτώσεις οι διαθέσιμες μετρήσεις να αναφέρονται στο επίπεδο του ενός κυττάρου. Επίσης ο ρυθμός παραγωγής δεδομένων έχει αυξηθεί ραγδαία ενώ το αντίστοιχος κόστος έχει μειωθεί αντιστοίχως. Ωστόσο, σε πολλές περιπτώσεις τα λαμβανόμενα δεδομένα είναι ελλιπή. Συνεχίζοντας, όσον αφορά τις ανάγκες χρήσης Στοχαστικής προσέγγισης για την μοντελοποίηση και προσομοίωση των βιολογικών φαινομένων, αυτή καθίσταται όλο και περισσότερο πιο αναγκαία λόγω και της διαθεσιμότητας λεπτομερών δεδομένων όπως αναφέρθηκε προηγουμένως. Με βάση τα παραπάνω εξάγονται τα ακόλουθα συμπεράσματα για κάθε κατηγορία.

Μέθοδοι Βελτιστοποίησης: παρότι χρησιμοποιήθηκαν με επιτυχία επί δεκαετίες και σε πολλά διαφορετικά συστήματα, όσον αφορά το μέγεθος και την φύση, παρουσιάζουν τα εξής αρνητικά σημεία:

- Παρουσιάζονται σε πολλές περιπτώσεις ιδιαίτερα απαιτητικές σε πόρους μιας και λίγο έως πολύ εξερευνούν όλο το διαθέσιμο χώρο παραμέτρων. Ωστόσο αν γνωρίζουμε την περιοχή όπου κινούνται οι τιμές των παραμέτρων, ο χώρος αναζητήσεων μπορεί να μειωθεί δραστικά.
- Δεν αποδίδουν καλά σε περιπτώσεις όπου ο θόρυβος μετρήσεων είναι υψηλός [97]
- Δεν μπορούν να αξιοποιήσουν δεδομένα από συστήματα που εμφανίζουν εγγενή Θόρυβο και έτσι χρίζουν Στοχαστικής Προσομοίωσης [103]

Μέθοδοι ML: δημιουργείται το κατάλληλο πλαίσιο ώστε να χρησιμοποιηθεί η Στοχαστική Μοντελοποίηση στην εκτίμηση παραμέτρων είτε προσεγγίζοντας την ML είτε εκτιμώντας τη μέσω προσομοιώσεων και στη συνέχεια χρησιμοποιώντας Μεθόδους Βελτιστοποίησης για την εύρεση των παραμέτρων. Επίσης οι ML μέθοδοι μπορούν να ανταπεξέρθουν σε δεδομένα με Θόρυβο Μετρήσεων που δεν είναι κατ' ανάγκη Λευκός Gaussian [74]. Ωστόσο και εδώ υπάρχουν κάποια αρνητικά σημεία:

- Δεν υπάρχει κάποια προφανής συνάρτηση απόστασης (όπως υπάρχει στις Μεθόδους Βελτιστοποίησης) και έτσι θα πρέπει να δοκιμάζονται διάφορες προτάσεις.
- Οι προτεινόμενες μέθοδοι απαιτούν δεδομένα για όλες τις συμμετέχουσες στο σύστημα οντότητες.

Μέθοδοι Αυτόματου Ελέγχου: Επιτρέπουν την χρήση ενός ιδιαίτερα διαδεδομένου αλγορίθμου, του Kalman Filter, το οποίο πέραν της απόδοσης, προσφέρει και ευκολία στον ορισμό των προβλημάτων. Η εφαρμογή αυτή θα μπορούσαμε να πούμε ότι παρουσιάζουν τα εξής πλεονεκτήματα:

- Λαμβάνει υπόψη της τον εγγενή Θόρυβο του Συστήματος αλλά και τον Θόρυβο Μετρήσεων
- Μπορεί να αξιοποιήσει ελλιπή δεδομένα ορίζοντας συνάρτηση που περιγράφει την σχέση των μετρήσεων με τις οντότητες του Συστήματος

Ως μειονεκτήματα των μέχρι τώρα προτεινόμενων εφαρμογών μπορούν να προσμετρηθούν:

- Η απουσία κάποιας μεθόδου που κάνει χρήση της ακριβούς Στοχαστικής Προσομοίωσης
- Σε όλες τις εργασίες γίνεται παραδοχή Λευκού θορύβου κάτι που στην πραγματικότητα μπορεί να μην ισχύει [104]
- Επίσης, ο Θόρυβος Συστήματος και ο Θόρυβος Μετρήσεων θεωρούνται ανεξάρτητοι κάτι που πάλι μπορεί να μην ισχύει [104]
- Τέλος, η τεχνική επέκτασης καταστάσεων αλλά και γενικότερα η φύση των Kalman Filter όταν εφαρμόζονται σε μη-γραμμικά προβλήματα μπορεί να οδηγήσει σε μη αξιόπιστα αποτελέσματα λόγω των αναγκαστικών προσεγγιστικών μεθόδων, όπως παρουσιάστηκαν παραπάνω

Bayesian Μέθοδοι: έρχονται να επιλύσουν πολλά από τα παραπάνω προβλήματα, όντας ικανές να λειτουργήσουν σε Στοχαστικά Πλαίσια (ακόμα και ακριβή), να ανταπεξέλθουν των ελλιπών δεδομένων αλλά και του Θορύβου Μετρήσεων και να

συνυπολογίσουν τον εγγενή Θόρυβο του Συστήματος χωρίς συγκεκριμένες παραδοχές για αυτόν. Επίσης, σημαντικό πλεονέκτημα αυτών των μεθόδων είναι ότι εξάγουν την πλήρη κατανομή πιθανοτήτων για τις παραμέτρους και όχι απλά μία σημειακή εκτίμηση. Ωστόσο για τις συγκεκριμένες μεθόδους θα πρέπει να σημειωθεί ότι:

- Καθίστανται πρακτικά ανέφικτες στη χρήση ακριβούς Στοχαστικής Προσομοίωσης λόγω του ιδιαίτερα αυξημένου υπολογιστικού κόστους για προβλήματα ρεαλιστικού μεγέθους (κάτι που αντιμετωπίζεται με την υιοθέτηση προσεγγιστικών μεθόδων όπως η diffusion approximation)
- Απαιτούν την ανάπτυξη και χρήση σύνθετων και απαιτητικών σε υπολογιστικούς πόρους MCMC μεθόδων Προσομοίωσης και Δειγματοληψίας

2.5 In Silico Ογκολογία και Ογκοπροσομοιωτής

Ο χώρος της In Silico Ογκολογίας (ως ειδική περιοχή της In Silico Ιατρικής) αποτελεί τον επιστημονικό τομέα εκείνο που στοχεύει στην εφαρμογή μεθόδων προερχόμενων από τις επιστήμες των υπολογιστών, της φυσικής και των μαθηματικών προκειμένου να υποστηριχθεί η κατανόηση της ασθένειας και του σύνθετου φυσικού φαινομένου του καρκίνου σε ποσοτική βάση μέσω της ανάπτυξης πολυκλιμακωτών και σύνθετων υπολογιστικών μοντέλων. Ο βασικότερος στόχος του τομέα αυτού είναι η υποστήριξη της βελτιστοποίησης της θεραπείας του καρκίνου σε εξατομικευμένη βάση.

Σημαντικός αριθμός ερευνητικών ομάδων εργάζονται προς αυτή την κατεύθυνση παγκοσμίως. Βασικός διαχωρισμός της φιλοσοφίας των προσεγγίσεων αποτελεί η μεθοδολογία μοντελοποίησης του φαινομένου με κύριες κατηγορίες αυτές της χρήσης συνεχών μαθηματικών, της χρήσης διακριτών μαθηματικών αλλά και τις υβριδικές μεθόδους.

Οι προσεγγίσεις συνεχών μαθηματικών βασίζονται κυρίως στη χρήση εξισώσεων διάχυσης οι οποίες εφαρμόζονται προκειμένου να περιγραφεί η διάχυση μορίων όπως αυτά του οξυγόνου και της γλυκόζης ή/και της διήθησης των καρκινικών κυττάρων στον περιβάλλον ιστό (στην περίπτωση διηθητικών καρκίνων). Παραδείγματα προσπαθειών αυτής της προσέγγισης δίνονται στις [105]–[111].

Οι προσεγγίσεις διακριτών μαθηματικών βασίζονται κυρίως στην μελέτη διαφόρων διακριτών καταστάσεων στις οποίες μπορεί να βρεθούν τα κύτταρα αλλά και των μεταβάσεων μεταξύ αυτών. Για παράδειγμα οι καταστάσεις αυτές μπορεί να είναι η φάση του κυτταρικού κύκλου που βρίσκεται το κύτταρο ή ο βαθμός διαφοροποίησής του. Οι μεταβάσεις μεταξύ των καταστάσεων υλοποιούνται με τη χρήση «μεθόδων απόφασης»

όπως τα κυτταροκινητικά διαγράμματα (cytokinetic diagrams), οι τεχνικές πράκτορα (agent-based) κ.α. Παραδείγμα χρήσης της διακριτής προσέγγισης δίνονται στις [112]–[121].

Οι υβριδικές προσεγγίσεις επιχειρούν να αξιοποιήσουν τις δυνατότητες των δύο παραπάνω προσεγγίσεων προκειμένου να μελετηθούν σε μεγαλύτερο βάθος διάφοροι σχετικοί με τον καρκίνο δυναμικοί μηχανισμοί. Ένα παράδειγμα αυτού του συνδυασμού αποτελεί η συνδυαστική χρήση ενός φαρμακοκινητικού μοντέλου, που συνήθως περιγράφεται με συνήθεις διαφορικές εξισώσεις με ένα μοντέλο διακριτών καταστάσεων και γεγονότων (discrete state, discrete event) προκειμένου να προσομοιωθεί η απόκριση μεγάλων απεικονίσιμων όγκων σε χημειοθεραπευτικά σχήματα [116]. Επιπλέον παράδειγμα αποτελεί η μοντελοποίηση της διήθησης του καρκίνου σε ιστό [122].

Μεταξύ των μοντέλων που αναφέρθηκαν στις προηγούμενες παραγράφους συγκαταλέγεται και η οικογένεια μοντέλων που αναπτύσσεται από την Ομάδα για την In Silico Ογκολογία και την In Silico Ιατρική του ΕΠΙΣΕΥ-ΕΜΠ (<http://in-silico-oncology.iccs.ntua.gr/>) με την γενικότερη ονομασία «Ογκοπροσομοιωτής» (Oncosimulator). Πολλαπλές μορφές Ογκοπροσομοιωτών που εστιάζουν σε διάφορες μορφές καρκίνου (Μαστού, Πνεύμονα, Νεφροβλάστωμα, Γλοιοβλάστωμα κ.α) έχουν αναπτυχθεί τα τελευταία χρόνια και στα πλαίσια χρηματοδοτούμενων από την Ευρωπαϊκή Επιτροπή ερευνητικών προγραμμάτων όπως τα ACGT (FP6-2005-IST-026996), Contra Cancrum (FP7-ICT-2007-2- 223979), TUMOR (FP7-ICT-2009.5.4-247754), p-medicine (FP7-ICT-2009.5.3 -.270089) και CHIC (FP7-ICT-2011-9, Grant agreement no: 600841). Ενδεικτικές εργασίες όπου παρουσιάζεται η φιλοσοφία, οι αρχές λειτουργίας και τα αποτελέσματα εκτέλεσης και ανάλυσης των Ογκοπροσομοιωτών είναι οι [117], [123]–[129]. Οι μέχρι σήμερα αναπτυσσόμενοι ογκοπροσομοιωτές χωρίζονται σε δύο βασικές κατηγορίες, αυτές της χωρικής προσέγγισης που αφορά την τετραδιάστατη προσομοίωση συμπαγών όγκων και της μη χωρικής προσέγγισης που εφαρμόζεται είτε σε περιπτώσεις μη συμπαγών όγκων είτε σε περιπτώσεις συμπαγών όγκων στους οποίους όμως δεν δίνεται ιδιαίτερη βαρύτητα στη χωρική τους εξέλιξη. Στα πλαίσια της παρούσας διατριβής ιδιαίτερη εστίαση δόθηκε στον Ογκοπροσομοιωτή για την περίπτωση της Οξείας Λεμφοβλαστική Λευχαιμία (Acute Lymphoblastic Leukemia) ο οποίος συγκαταλέγεται στην κατηγορία της μη χωρικής προσέγγισης και στο εξής αναφέρεται ως Ογκοπροσομοιωτής Λευχαιμίας (ALL Oncosimulator). Ο Ογκοπροσομοιωτής Λευχαιμίας αναπτύσσεται επίσης από την Ομάδα για την In Silico Ογκολογία και την In Silico Ιατρική του ΕΠΙΣΕΥ-ΕΜΠ και παρουσιάστηκε για πρώτη φορά στην [130].

Στις ακόλουθες ενότητες παρουσιάζονται οι βασικές αρχές και οι λεπτομέρειες υλοποίησης των δύο κατηγοριών ογκοπροσομοιωτή.

2.5.1 Ογκοπροσομοιωτής χωρικής προσέγγισης

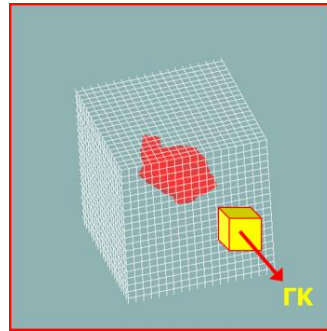
Η ανάπτυξη του Ογκοπροσομοιωτή βασίζεται στην αρκούντως τεκμηριωμένη παραδοχή ότι η ανάπτυξη και συντήρηση των καρκινικών όγκων είναι αποτέλεσμα της ύπαρξης καρκινικών βλαστικών κυττάρων (cancer stem cells), δηλαδή κυττάρων με θεωρητικώς απεριόριστη δυνατότητα πολλαπλασιασμού. Γενικότερα, οι περισσότεροι τύποι καρκίνου συν-αποτελούνται από ετερογενείς πληθυσμούς κυττάρων με διαφορές στην δυνατότητα πολλαπλασιασμού. Τα καρκινικά βλαστικά κύτταρα συνήθως αποτελούν ένα μικρό ποσοστό του συνολικού αριθμού κυττάρων που απαρτίζουν ένα όγκο.

Βασικό χαρακτηριστικό των βλαστικών κυττάρων είναι η δυνατότητα τους να αυτό-ανανεώνονται, δηλαδή να δημιουργούν μέσω διαίρεσης νέα βλαστικά κύτταρα. Η διαίρεση αυτή μπορεί να είναι είτε συμμετρική είτε α-συμμετρική [131]. Στην πρώτη περίπτωση, η διαίρεση έχει ως αποτέλεσμα δύο βλαστικά κύτταρα. Στην δεύτερη περίπτωση, το ένα εκ των δύο κυττάρων δεν είναι βλαστικό αλλά ξεκινάει την διαδρομή προς την διαφοροποίηση ενώ ταυτόχρονα χάνει την δυνατότητα της απεριόριστης δυνατότητας πολλαπλασιασμού (limited mitotic potential cell – LIMP cell). Συγκεκριμένα, οι ακόλουθες κλάσεις/κατηγορίες καρκινικών κυττάρων λαμβάνονται υπόψη από τον Ογκοπροσομοιωτή:

1. Βλαστικά ή στελεχειαία κύτταρα (stem cells)
2. Κύτταρα πεπερασμένης δυνατότητας πολλαπλασιασμούς (Limited Mitotic Potential cells - LIMP). Κύτταρα τα οποία δύνανται να πραγματοποιήσουν ένα περιορισμένο αριθμό διαιρέσεων προτού φτάσουν στην τελική διαφοροποίηση.
3. Διαφοροποιημένα κύτταρα (differentiated cells)
4. Αποπτωτικά κύτταρα (Apoptotic cells). Κύτταρα τα οποία έχουν οδηγηθεί στον κυτταρικό θάνατο μέσω απόπτωσης.
5. Νεκρωτικά κύτταρα (Necrotic cells). Κύτταρα τα οποία έχουν οδηγηθεί στον κυτταρικό θάνατο μέσω νέκρωσης.

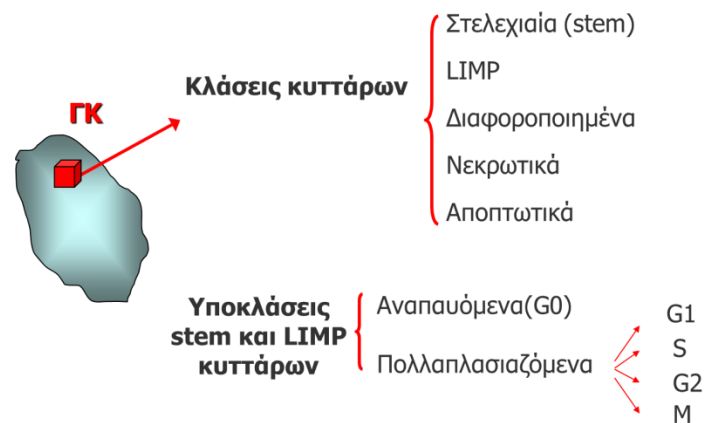
Στην περίπτωση της χωρικής προσέγγισης η περιοχή του όγκου, η οποία μπορεί να έχει ανακτηθεί από απεικονιστικά δεδομένα (π.χ. μαγνητική τομογραφία), διακριτοποιείται με χρήση ενός τρισδιάστατου κυβικού πλέγματος. Το ελάχιστο (κυβικό) στοιχείο όγκου του πλέγματος ονομάζεται γεωμετρική κυψελίδα (ΓΚ, geometrical cell - GC) όπως παρουσιάζεται στην Εικόνα 2.19. Οι διαστάσεις της γεωμετρικής κυψελίδας επιλέγονται με

βάση το μέγεθος του καρκινικού όγκου, τους διαθέσιμους υπολογιστικούς πόρους και την εκάστοτε διακριτική ικανότητα των απεικονιστικών δεδομένων όταν αυτά είναι διαθέσιμα.



Εικόνα 2.19 Τρισδιάστατο πλέγμα διακριτοποίησης που καλύπτει την περιοχή του όγκου (ΓΚ: Γεωμετρική Κυψελίδα).

Σε κάθε γεωμετρική κυψελίδα που ανήκει στην περιοχή του όγκου περιέχονται καρκινικά κύτταρα, καθένα από τα οποία μπορεί να ανήκει σε μία πέντε κλάσεις/κατηγορίες κυττάρων όπως δόθηκαν προηγουμένως. Οι φάσεις του κυτταρικού κύκλου (G1, S, G2, M) καθώς και η φάση ανάπαυσης (G0) αποτελούν υποκλάσεις στις οποίες μπορεί να βρίσκονται τα στελεχιαία ή τα LIMP κύτταρα (Εικόνα 2.20).



Εικόνα 2.20 Κάθε Γεωμετρική Κυψελίδα (ΓΚ) μπορεί να περιέχει κύτταρα από πέντε κλάσεις κυττάρων. Οι φάσεις του κυτταρικού κύκλου (G1, S, G2, M) και η φάση ανάπαυσης (G0) αποτελούν υποκλάσεις των στελεχιαίων και LIMP κυττάρων.

Κάθε γεωμετρική κυψελίδα περιέχει κατά την έναρξη μιας προσομοίωσης έναν αριθμό βιολογικών κυττάρων (ΑΒΚ) που καθορίζεται με βάση τυπικές τιμές της πυκνότητας των κυττάρων στους όγκους, ή με βάση ειδική χωρική/απεικονιστική πληροφορία για έναν δεδομένο όγκο, αν είναι διαθέσιμη. Τα αρχικά αυτά κύτταρα μπορούν

να κατανεμηθούν στις πέντε κλάσεις κυττάρων (στελεχιαία, LIMP, διαφοροποιημένα, νεκρωτικά, αποπτωτικά) με διαφορετικούς τρόπους, ανάλογα με το προφίλ του όγκου που θέλει να εξετάσει ο χρήστης (να οριστούν δηλαδή διαφορετικά τα ποσοστά των πέντε βασικών κυτταρικών πληθυσμών). Για παράδειγμα, όγκοι στους οποίους κυριαρχούν ποσοστιαία τα πολλαπλασιαζόμενα κύτταρα, αναμένεται να έχουν πιο επιθετική συμπεριφορά από τους όγκους στους οποίους υπερισχύουν τα διαφοροποιημένα κύτταρα. Στην περίπτωση που υπάρχουν διαθέσιμα απεικονιστικά δεδομένα, τότε αυτά μπορούν να χρησιμοποιηθούν για την αρχικοποίηση του όγκου.

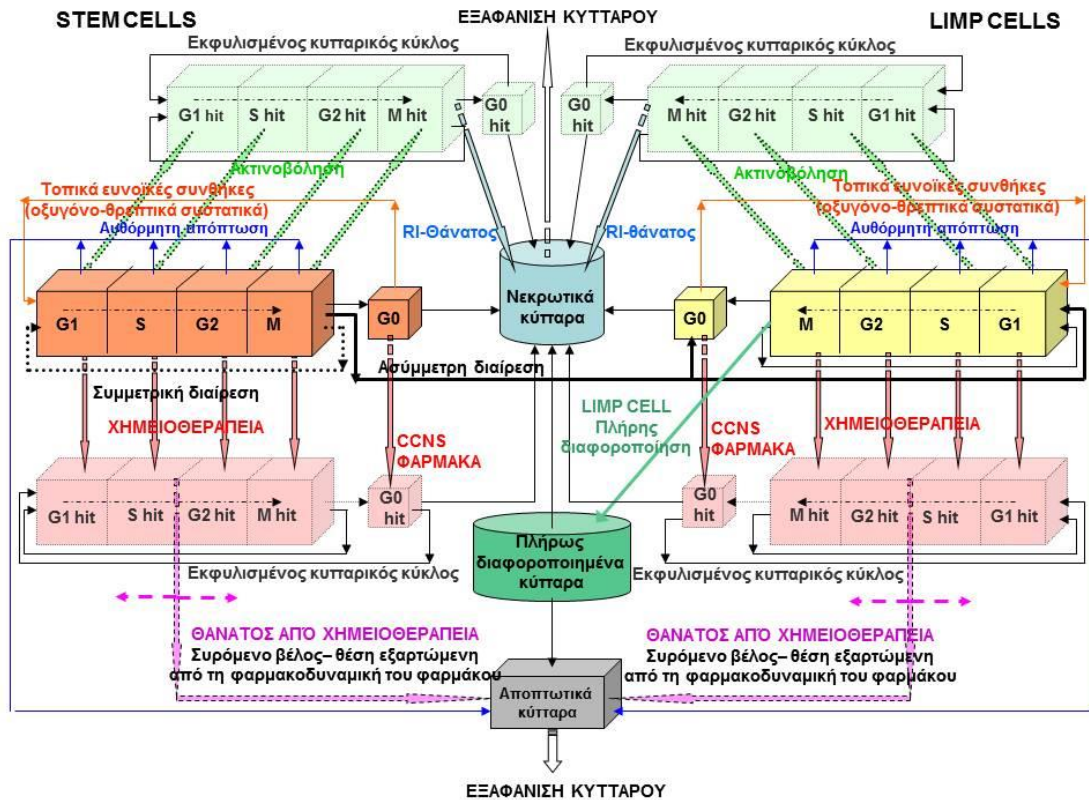
Σε κάθε χρονικό βήμα της προσομοίωσης (μία ώρα) πραγματοποιούνται δύο σάρωσεις του πλέγματος στην περιοχή του όγκου. Με κατάλληλους κανόνες, που περιγράφονται στη συνέχεια, στην πρώτη σάρωση προσομοιώνεται η βιολογική συμπεριφορά των καρκινικών κυττάρων, ενώ στη δεύτερη μοντελοποιείται η χωρική αυξομείωση του όγκου.

Η έξοδος του μοντέλου σε δεδομένη χρονική είναι η κατανομή των διάφορων κυτταρικών πληθυσμών στην περιοχή του όγκου. Διάφορες ποσότητες ενδιαφέροντος (μέγεθος του όγκου, κλάσμα ανάπτυξης, υποξικό κλάσμα κλπ.) μπορούν επίσης να υπολογιστούν.

Ακολούθως δίνονται οι δύο βασικές κατηγορίες βημάτων της ροής εκτέλεσης του Ογκοπροσομοιωτή χωρικής προσέγγισης.

1^η σάρωση του γεωμετρικού πλέγματος.

Η βιολογική συμπεριφορά των καρκινικών κυττάρων στο χρόνο καθορίζεται από το θεωρούμενο κυτταροκινητικό μοντέλο. Το κυτταροκινητικό μοντέλο που άπτεται της γενικής περίπτωσης του Ογκοπροσομοιωτή χωρικής προσέγγισης δίνεται στην ακόλουθη εικόνα.



Εικόνα 2.21 Γενικευμένο κυτταροκινητικό μοντέλο που μπορεί να προσαρμοστεί ανάλογα με το εκάστοτε θεραπευτικό πρωτόκολλο. Τα βέλη αναπαριστούν πιθανές μεταβάσεις μεταξύ κυτταρικών καταστάσεων και αντιστοιχούν σε παραμέτρους του μοντέλου. STEM CELLS: στελεχιαία καρκινικά κύτταρα. LIMP CELLS: κύτταρα περιορισμένου μιτωτικού δυναμικού. G1: Gap 1 φάση, S: Φάση σύνθεσης DNA, G2: Gap 2 φάση, M: μίτωση, G0: φάση ανάπαυσης, hit: κύτταρα που έχουν χτυπηθεί θανάσιμα από την ακτινοβολία ή τη χημειοθεραπεία. CCNS φάρμακα: Cell Cycle Non Specific φάρμακα.

Κατά τη διάρκεια μιας προσομοίωσης, σε κάθε χρονικό βήμα, σαρώνεται το πλέγμα του όγκου και εφαρμόζονται οι κανόνες των μεταβάσεων μεταξύ των διάφορων κλάσεων και υποκλάσεων κυττάρων. Ανάλογα με τον προσομοιωτικό στόχο της κάθε υπολογιστικής υπομονάδας στο κυτταροκινητικό μοντέλο μπορεί να περιλαμβάνονται κάποιοι ή και όλοι οι παρακάτω βιολογικοί μηχανισμοί:

1. Διέλευση πολλαπλασιαζόμενου κυττάρου από τις διαδοχικές φάσεις (G1, S, G2, M) του κυτταρικού κύκλου.
2. Συμμετρική και ασύμμετρη διαίρεση βλαστικού κυττάρου.
3. Σταδιακή διαφοροποίηση κυττάρου ύστερα από έναν αριθμό μιτωτικών διαιρέσεων.
4. Μετάβαση πολλαπλασιαζόμενου κυττάρου στη φάση ανάπαυσης (G0) λόγω μη επαρκούς τροφοδοσίας με οξυγόνο και θρεπτικά συστατικά.
5. Επιστροφή αναπαυόμενου κυττάρου στον ενεργό κυτταρικό κύκλο λόγω αποκατάστασης της τροφοδοσίας με οξυγόνο και θρεπτικά συστατικά.
6. Κυτταρικός θάνατος λόγω αυθόρμητης απόπτωσης

7. Κυτταρική νέκρωση λόγω ακτινοβολίας
8. Κυτταρική απόπτωση λόγω χημειοθεραπείας

Η κάθε υπολογιστική μονάδα χρησιμοποιεί τους απαραίτητους βιολογικούς μηχανισμούς για τη μοντελοποίηση του θεραπευτικού πρωτοκόλλου που εξετάζεται. Για παράδειγμα σε μία υπολογιστική υπό-μονάδα που προσομοιώνει την ελεύθερη ανάπτυξη ενός όγκου, χωρίς θεραπευτική παρέμβαση, θα συμπεριλαμβάνονται οι μηχανισμοί 1,2,3,4,5 και 6, ενώ σε μια υπολογιστική υπό-μονάδα που προσομοιώνει μόνο χημειοθεραπεία θα συμπεριλαμβάνονται οι μηχανισμοί 1,2,3,4,5, 6 και 8.

Τις παραμέτρους του μοντέλου που άπτονται των βιολογικών μηχανισμών που δίνονται στην Εικόνα 2.21 παρουσιάζει ο Πίνακας 2.1.

Πίνακας 2.1 Παράμετροι του Ογκοπροσομοιωτή που σχετίζονται με τους υπό μοντελοποίηση βιολογικούς μηχανισμούς

Παράμετρος (μονάδα)	Περιγραφή
T_c (h)	Διάρκεια κυτταρικού κύκλου
T_{G0} (h)	Διάρκεια φάσης ανάπαυσης G0
T_N (h)	Διάρκεια νέκρωσης
T_A (h)	Διάρκεια απόπτωσης
N_{LIMP} (no.)	Αριθμός μιώσεων για την πλήρη διαφοροποίηση των κυττάρων περιορισμένου μιτωτικού δυναμικού (LIMP, Limited Mitotic Potential)
α (Gy^{-1})	Παράμετρος α του μοντέλου LQ
β (Gy^{-2})	Παράμετρος β του μοντέλου LQ
CKR	Πιθανότητα απόπτωσης λόγω χημειοθεραπείας
R_A (h^{-1})	Ρυθμός απόπτωσης
R_{NDiff} (h^{-1})	Ρυθμός νέκρωσης των διαφοροποιημένων κυττάρων
R_{ADiff} (h^{-1})	Ρυθμός απόπτωσης των διαφοροποιημένων κυττάρων
P_{G0toG1} (h^{-1})	Κλάσμα αναπαυόμενων κυττάρων που επανέρχονται στον ενεργό κυτταρικό κύκλο στη μονάδα του χρόνου
P_{sleep}	Κλάσμα κυττάρων που περνούν στη φάση ανάπαυσης μετά τη μίτωση
P_{sym}	Κλάσμα στελεχιαίων (stem) κυττάρων που πραγματοποιούν συμμετρική διαίρεση

Θα πρέπει να τονιστεί ιδιαίτερα το γεγονός ότι οι παράμετροι αυτές δεν αποτελούν αυθαίρετες μαθηματικές ποσότητες, αλλά συνδέονται άμεσα με τους συγκεκριμένους βιολογικούς μηχανισμούς. Σημειώνεται ακόμη ότι οι παραπάνω μηχανισμοί αναφέρονται

στο κυτταρικό επίπεδο , αλλά ουσιαστικά καθορίζονται από γονιδιακούς παράγοντες που μπορεί να διαφέρουν από μια κλινική περίπτωση σε μία άλλη και έτσι να διαφοροποιείται και η απόκριση του όγκου στη θεραπεία. Έτσι το μοντέλο έχει τη μελλοντικής ενσωμάτωσης πληροφορίας από το υπό-κυτταρικό επίπεδο (παράδειγμα αποτελούν τα Κεφάλαια 3 και 0 της παρούσας διατριβής).

Επίσης θα πρέπει να τονιστεί ότι προκειμένου να αποκλειστούν συνδυασμοί τιμών παραμέτρων που δύνανται να οδηγήσουν σε μη αποδεκτή συμπεριφορά προσομοίωσης (π.χ. αυτό-συρρικνούμενους όγκους) ελέγχεται η ακόλουθη συνθήκη για τις επιλεγμένες παραμέτρους:

$$(1 + P_{sym}) \left(1 - P_{sleep} + P_{sleep} \frac{\frac{P_{G0toG1}}{T_{G0}}}{R_A + \frac{1}{T_{G0}}} \right) e^{-R_A T_C} \geq 1$$

η οποία εφόσον λαμβάνει τιμές μεγαλύτερες του 1 θεωρείται ένδειξης ορθότητας της παραμετροποίησης του μοντέλου.

Η βασική αρχή για τη μοντελοποίηση της ακτινοθεραπείας είναι το Γραμμικό-Τετραγωνικό Μοντέλο ή Μοντέλο LQ (Linear Quadratic model, LQ model), που χρησιμοποιείται ευρέως στην πειραματική και κλινική ραδιοβιολογία [6], [132]. Για την περίπτωση της εξωτερικής ακτινοβολίας χρησιμοποιείται η σχέση:

$$SF(d) = \exp[-(\alpha d + \beta d^2)]$$

όπου $SF(d)$ είναι το κλάσμα επιβίωσης ύστερα από μία (ομοιόμορφη) δόση ακτινοβολίας d (Gy) σε ένα πληθυσμό κυττάρων. Οι παράμετροι α (Gy^{-1}) and β (Gy^{-2}) σχετίζονται με την ακτινοευαισθησία του κυττάρου. Είναι γνωστό ότι η ακτινοευαισθησία των κυττάρων διαφέρει ανάλογα με τη φάση του κυτταρικού κύκλου στην οποία βρίσκονται, παρατήρηση που μπορεί ενσωματώνεται εύκολα στη μοντελοποίηση μέσω του Ογκοπροσομοιωτή λόγω του διακριτού χαρακτήρα του μοντέλου και της θεώρησης των διάφορων υποκλάσεων που αντιστοιχούν στις φάσεις του κυτταρικού κύκλου. Τα κύτταρα που έχουν πληγεί από την ακτινοβολία και υφίστανται νεκρωτικό θάνατο θεωρείται ότι ολοκληρώνουν δύο κύκλους μίτωσης πριν πεθάνουν.

Η παραπάνω μορφή του μοντέλου LQ μπορεί να χρησιμοποιηθεί για τις περιπτώσεις ακτινοθεραπευτικών σχημάτων στα οποία τα διαδοχικά κλάσματα απέχουν μεταξύ τους τουλάχιστον 6 ώρες, ώστε να πραγματοποιείται επιδιόρθωση των υπο-θανατηφόρων βλαβών στο χρονικό διάστημα μεταξύ των κλασμάτων [132] . Σε διαφορετική περίπτωση

(πχ. βραχυθεραπεία), χρησιμοποιούνται τροποποιημένες εκδοχές του απλού LQ μοντέλου, που λαμβάνουν υπόψη τους τα χρονικά χαρακτηριστικά της ακτινοβολήσης και τον χρόνο επιδιόρθωσης των κυτταρικών βλαβών στον θεωρούμενο ιστό [6], [132].

Όσον αφορά στη μοντελοποίηση της χημειοθεραπείας, στόχο αποτελεί ο υπολογισμός του ποσοστού των στελεχιαίων και LIMP κυττάρων που πεθαίνουν λόγω του θεωρούμενου φαρμάκου, και οδηγούνται σε απόπτωση. Μέσω της θεώρησης των διαφορετικών κλάσεων και υποκλάσεων κυττάρων, υπάρχει η δυνατότητα προσομοίωσης διαφόρων κατηγοριών φαρμάκων: φαρμάκων που δρουν μόνο σε πολλαπλασιαζόμενα κύτταρα (cell cycle - specific), φαρμάκων που δρουν και στα πολλαπλασιαζόμενα κύτταρα αλλά και στη φάση ανάπαυσης (cell cycle – non specific) και φαρμάκων που δρουν ή όχι σε συγκεκριμένη φάση του κυτταρικού κύκλου (cell cycle phase - specific ή cell cycle phase – non specific). Ο χρόνος απομάκρυνσης των χτυπημένων από τη χημειοθεραπεία κυττάρων εξαρτάται από τις φαρμακοκινητικές και φαρμακοδυναμικές ιδιότητες του θεωρούμενου σκευάσματος στον συγκεκριμένο ιστό, και μελετάται ξεχωριστά για κάθε φάρμακο.

Σημειώνεται ότι η προσέγγιση αυτή, που ξεκινάει από τα μακροσκοπικά απεικονιστικά δεδομένα του ασθενούς (στο επίπεδο του ιστού), και προχωράει προς τα χαμηλότερα επίπεδα βιοπολυπλοκότητας (στο επίπεδο του κυττάρου) μέσω της θεώρησης εμπλεκόμενων βιολογικών μηχανισμών και του ορισμού αντίστοιχων παραμέτρων ονομάζεται top-down.

2^η σάρωση του πλέγματος του γεωμετρικού πλέγματος.

Η δεύτερη σάρωση του πλέγματος σχετίζεται με τη **χωρική συμπεριφορά των κυττάρων**, και βασίζεται στον καθορισμό ενός μεγίστου ($ABK_{πάνω}$) και ενός ελαχίστου ($ABK_{κάτω}$) ορίου στον επιτρεπτό αριθμό κυττάρων μέσα σε κάθε γεωμετρική κυψελίδα:

$$ABK_{πάνω} = ABK + fr(ABK), \quad ABK_{κάτω} = ABK - fr(ABK)$$

όπου ο παράγοντας $fr(ABK)$ συμβολίζει κάποιο ποσοστό (fraction) του αριθμού ABK , και μπορεί να ρυθμιστεί ώστε να επιτυγχάνεται κατά προσέγγιση ομοιόμορφη κατανομή κυττάρων στο πλέγμα. Τα κατώφλια $ABK_{πάνω}$ και $ABK_{κάτω}$ αποτελούν τη βάση του μηχανισμού με τον οποίο μπορεί να αναπαρασταθεί η **χωρική αυξομείωση των όγκων**. Συγκεκριμένα, σε κάθε χρονικό βήμα της προσομοίωσης, κατά τη δεύτερη σάρωση του πλέγματος υπολογίζεται ο συνολικός αριθμός των κυττάρων σε κάθε γεωμετρική κυψελίδα. Αν ο αριθμός αυτός βρεθεί υψηλότερος του $ABK_{πάνω}$ ή χαμηλότερος του $ABK_{κάτω}$, χρησιμοποιούνται ειδικά σχεδιασμένοι αλγόριθμοι που δημιουργούν νέες κυψελίδες (για να

τοποθετηθεί το πλεονάζον περιεχόμενο των αρχικών υπερφορτωμένων κυψελίδων) ή διαγράφουν πολύ άδειες κυψελίδες, αντίστοιχα, και πραγματοποιούν σειριακή μετακίνηση κυττάρων από κυψελίδα σε κυψελίδα. Αυτή η διαδικασία δημιουργίας νέων ή διαγραφής γεωμετρικών κυψελίδων στην περιοχή του όγκου, οδηγεί αντίστοιχα σε αύξηση ή μείωση του όγκου.

2.5.2 Ογκοπροσομοιωτής μη χωρικής προσέγγισης

Ο Ογκοπροσομοιωτής μη χωρικής προσέγγισης [130], [133] αποτελεί παραλλαγή αυτού της χωρικής προσέγγισης και μπορεί να εφαρμοστεί στην περίπτωση τόσο συμπαγών όγκων (μεταβολικά ομοιογενών ή μη ομοιογενών), όταν δεν δίνεται έμφαση στη χωρική εξέλιξή τους, όσο και μη συμπαγών όγκων, όπως η λευχαιμία (βλ. Ενότητα 5.4). Οι σημαντικές διαφοροποιήσεις της μη χωρικής από τη χωρική προσέγγιση αφορούν την απαλοιφή των αλγορίθμων προσομοίωσης της χωρικής εξέλιξης του όγκου και τον ορισμό μεγαλύτερου αριθμού διαμερισμάτων για τα βλαστικά και προγονικά κύτταρα.

Όπως παρουσιάστηκε, στην χωρική προσέγγιση, η περιοχή του όγκου αντιμετωπίζεται σαν ένα πλέγμα γεωμετρικών κυψελίδων. Οι μορφολογικοί κανόνες που διέπουν την απαλοιφή ή τη δημιουργία μιας νέας κυψελίδας οδηγούν σε συρρίκνωση ή επέκταση του όγκου σύμμορφη με το αρχικό του σχήμα. Στην μη χωρική προσέγγιση, τα πολλαπλασιαζόμενα και τα αναπαυόμενα καρκινικά κύτταρα κατανέμονται όχι σε τέσσερα και ένα διαμερίσματα αλλά σε τόσα διαμερίσματα όσο και η διάρκεια της αντίστοιχης φάσης του κυτταρικού κύκλου σε ώρες. Κάθε διαμέρισμα αντιστοιχεί στο χρονικό διάστημα μιας ώρας. Ο χειρισμός μπορεί να εξαλείψει σφάλματα κβαντοποίησης τα οποία εμφανίζονται στο χωρικό μοντέλο και τα οποία σχετίζονται με τη θεώρηση περιορισμένου αριθμού διαμερισμάτων (G1, S, G2, M, G0), κάτι που καθορίζεται και από τις υπολογιστικές απαιτήσεις αυτού. Επιπλέον, η απόκλιση της χωρικής εξέλιξης του όγκου δεν επηρεάζει τη χρονική εξέλιξη των επιμέρους και του συνολικού κυτταρικών πληθυσμών. Το μοντέλο είναι διακριτό ως προς τον χρόνο με χρονικό βήμα ίσο με μία ώρα, δηλαδή τη διάρκεια της βραχύτερης φάσης του κυτταρικού κύκλου (μίτωση), όπως και στη χωρική υλοποίηση.

Η διαφοροποίηση του κυτταροκινητικού μοντέλου της μη χωρικής προσέγγισης εστιάζεται κυρίως στο πλήθος των (υπο)-διαμερισμάτων που κατανέμονται τα πολλαπλασιαζόμενα και αναπαυόμενα κύτταρα. Στη μη χωρική προσέγγιση, τα κύτταρα εντός του ενεργού κυτταρικού κύκλου (βλαστικά και προγονικά) χωρίζονται, αντί των τεσσάρων διαμερισμάτων που αντιστοιχούν στις τέσσερις φάσεις (G1, S, G2, M) του κυτταρικού κύκλου, σε ένα πλήθος υποδιαμερισμάτων που ισούται με τη διάρκεια του

κυτταρικού κύκλου. Κάθε υποδιαμέρισμα αντιστοιχεί σε κάθε ώρα του κυτταρικού κύκλου και επομένως περιέχει τα κύτταρα που διανύουν την αντίστοιχη ώρα. Όμοια τα αναπαυόμενα κύτταρα (βλαστικά και προγονικά), αντί για ένα, κατανέμονται σε ένα πλήθος υποδιαμερισμάτων που αντιστοιχούν σε κάθε ώρα της διάρκειας της φάσης ηρεμίας.

Επιπλέον συζήτηση για τον Ογκοπροσομοιωτή μη χωρικής προσέγγισης, με εστίαση στην εφαρμογή του στην περίπτωση της Οξείας Λεμφοβλαστικής Λευχαιμίας γίνεται στο Κεφάλαιο 0.

2.6 Φιλοσοφία εφαρμογής μεθόδων

Παρότι οι υπό μελέτη τομείς παρουσιάζουν σημαντική αυτοτέλεια και έχουν οδηγήσει από μόνοι τους σε σημαντικά επιστημονικά ευρήματα, στα πλαίσια της συγκεκριμένης διατριβής έχει τεθεί ο στόχος της συν-αξιοποίησης τους ώστε να παραχθούν αλγόριθμοι και μοντέλα που θα επιτρέψουν την εξαγωγή νέων αποτελεσμάτων/ευρημάτων συγκεκριμένα για την περιοχή του καρκίνου (και ειδικότερα της ALL) αλλά και την περαιτέρω επέκταση του Ογκοπροσομοιωτή (κυρίως του ALL Ογκοπροσομοιωτή του οποίου ο κώδικας, στη γλώσσα C++, και η αντίστοιχη εκτελέσιμη μορφή έχουν γίνει διαθέσιμα για τις ανάγκες της παρούσας διατριβής από την ομάδα για την In Silico Ογκολογία). Η επέκταση αυτή κυρίως αφορά στην δυνατότητα προσομοίωσης φαινομένων αλλά και χρήσης δεδομένων από το μοριακό επίπεδο. Η προσέγγιση που ακολουθείται στηρίζεται στην φιλοσοφία «σύνοψης και μεταπήδησης» (“summarize and jump”) όπως προτείνεται στην [134]. Η συγκεκριμένη προσέγγιση, συνοπτικά, ορίζει πως τα αποτελέσματα της εκτέλεσης προβλεπτικών μοντέλων και αλγορίθμων που αφορούν σε φαινόμενα που συμβαίνουν σε χαμηλότερα επίπεδα βιο-πολυπλοκότητας θα πρέπει να εξάγουν τα συμπεράσματά τους σε μορφή που μπορεί άμεσα να αξιοποιηθεί από αντίστοιχες διαδικασίες που αναφέρονται σε υψηλότερα επίπεδα. Επί παραδείγματι, ένα μοντέλο προσομοίωσης βιοχημικών δικτύων (όπως παρουσιάστηκαν στην Ενότητα 2.4.1), που άπτεται του μοριακού επιπέδου, θα πρέπει μετά το πέρας της εκτέλεσης να εξάγει κάποια πρόβλεψη για την τιμή κάποιας ή κάποιων από τις παραμέτρους του Ογκοπροσομοιωτή, του οποίου μεγάλος αριθμός παραμέτρων αφορούν το κυτταρικό επίπεδο.

Στα ακόλουθα κεφάλαια, παρουσιάζεται η ανάπτυξη μοντέλων και αλγορίθμων κατά την οποία ακολουθήθηκε η παραπάνω προσέγγιση.

3. Μοντελοποίηση της βιοχημικής ρύθμισης του κυτταρικού κύκλου στην Οξεία Λεμφοβλαστική Λευχαιμία

3.1 Εισαγωγή

Στα πλαίσια της εφαρμογής μεθόδων συστημικής βιολογίας, μελετάται ο κυτταρικός κύκλος των λευχαιμικών κυττάρων ασθενών παιδικής ηλικίας που πάσχουν από Οξεία Λεμφοβλαστική Λευχαιμία (Acute Lymphoblastic Leukemia-ALL). Ειδικότερα υπήρξε εστίαση (όσο αυτό ήταν εφικτό) στην υποκατηγορία Πρόδρομος-B Οξεία Λεμφοβλαστική Λευχαιμία (Precursor-B ALL, BCP-ALL). Η επιλογή αυτή έγινε όχι μόνο λόγω του υψηλού ρυθμού εμφάνισης της συγκεκριμένης υποκατηγορίας [135], [136], αλλά επίσης λόγω του σημαντικού όγκου γνώσης που είναι διαθέσιμος στην βιβλιογραφία σε σχέση με άλλες υποκατηγορίες.

Κύριο υπολογιστικό μέσω μελέτης του συγκεκριμένου βιολογικού συστήματος, στα πλαίσια της διατριβής, είναι τα μοντέλα προσομοίωσης δικτύων βιοχημικών αντιδράσεων (biochemical reaction network models) όπως παρουσιάστηκαν στο Κεφάλαιο 2, μέσω της ανάπτυξης ενός νέο προτεινόμενου μοντέλου.

Πέραν του αυτόνομου στόχου για καλύτερη κατανόηση του βιολογικού αυτού συστήματος και της απορρύθμισή του στην περίπτωση της ALL, η απόκτηση ενός μοντέλου που θα είναι ικανό να προσομοιώσει τον κυτταρικό κύκλο των BCP-ALL κυττάρων και πιθανόν να προβλέψει την διάρκεια αυτού, θεωρείται ιδιαίτερα σημαντική δεδομένης της ύπαρξης της αντίστοιχης παραμέτρου (διάρκεια κυτταρικού κύκλου, T_c) του Ογκοπροσομοιωτή. Η συγκεκριμένη παράμετρος αποτελεί δυνητικό σημείο επικοινωνίας μεταξύ του Ογκοπροσομοιωτή και του μοντέλου που αναπτύχθηκε και παρουσιάζεται στο παρόν Κεφάλαιο, στα πλαίσια της στρατηγικής “summarize and jump”, όπως αυτή παρουσιάστηκε στην Ενότητα 2.6.

Στις ακόλουθες ενότητες, αρχικά παρουσιάζονται βασικές πληροφορίες για το κυτταρικό κύκλο και την βιοχημική ρύθμιση αυτού ενώ ακολούθως παρουσιάζονται τα ειδικά χαρακτηριστικά αυτού στην Οξεία Λεμφοβλαστική Λευχαιμία και η μοντελοποίηση αυτών.

3.2 Η βιοχημική ρύθμιση του κυτταρικού κύκλου

Στην παρούσα υπό-ενότητα παρουσιάζονται οι λεπτομέρειες της βιοχημικής ρύθμισης του κυτταρικού κύκλου, δηλαδή τα φαινόμενα που παρατηρούνται στον υπό-κυτταρικό χώρο και έχουν ως αποτέλεσμα της παρατηρούμενη σε κυτταρικό επίπεδο συμπεριφορά που περιγράφηκε στην Ενότητα 1.3.1. Αρχικά δίνεται μία συνοπτική περιγραφή για τα φαινόμενα που παρατηρούνται σε όλη την έκταση του κυτταρικού κύκλου ενώ ακολούθως δίνεται ιδιαίτερη έμφαση στις σχετικές με την φάση G1 και το Σημείο Περιορισμού λεπτομέρειες. Όπως θα παρουσιαστεί λεπτομερώς στη συνέχεια, οι κυριότερες σχετικές με την Οξεία Λεμφοβλαστική Λευχαιμία παρατηρούνται σε αυτή τη φάση του κυτταρικού κύκλου.

Η βιοχημική ρύθμιση του κυτταρικού κύκλου περιλαμβάνει κρίσιμες διαδικασίες για την επιβίωση του κυττάρου (π.χ. ανίχνευση και την διόρθωση της γενετικής βλάβης, πρόληψη ανεξέλεγκτης κυτταρικής διαίρεσης). Στη φυσιολογική περίπτωση τα βιοχημικά/μοριακά γεγονότα που ελέγχουν τον κυτταρικό κύκλο είναι καθορισμένα και κατευθυνόμενα με τέτοιο τρόπο ώστε οι διάφορες διαδικασίες να πραγματοποιούνται με συγκεκριμένη σειρά και να είναι μη αναστρέψιμες [137], [138].

Τρεις ομάδες μοριακών οντοτήτων ελέγχουν την ρύθμιση του κυτταρικού κύκλου:

- Οι Κυκλίνες (Cyclins)
- Οι Εξαρτώμενες από Κυκλίνες Κινάσες (Cyclin Dependent Kinases, CDKs ή Cdk)
- Οι αναστολείς των Εξαρτώμενων από Κυκλίνες Κινάσεων (Cyclin Dependent Kinase Inhibitors, CDKIs)

Οι κινάσες αποτελούν είδη ενζύμων που δρουν ως καταλύτης για τη φωσφορυλίωση μορίων ή πρωτεϊνών, τροποποιώντας έτσι τη λειτουργία τους.

3.2.1 Ο ρόλος των κυκλινών και των εξαρτώμενων από κυκλίνες κινασών

Οι Κυκλίνες και οι Εξαρτώμενες από Κυκλίνες Κινάσες αποτελούν τις δύο βασικές κατηγορίες μορίων που καθορίζουν σε μεγάλο βαθμό την πρόοδο του κυτταρικού κύκλου. Πολλά από τα γονίδια τα οποία κωδικοποιούν τις κυκλίνες και τις CDKs είναι υψηλά συντηρημένα μεταξύ όλων των ευκαρυωτικών κυττάρων.

Η ομάδα των κυκλινών που εμπλέκονται στον κυτταρικό κύκλο περιλαμβάνει τέσσερις διαφορετικές πρωτεΐνες οι οποίες είναι οι:

- Κυκλίνη A (Cyclin A)

- Κυκλίνη B (Cyclin B)
- Κυκλίνη D (Cyclin D)
- Κυκλίνη E (Cyclin E)

Και οι τέσσερις περιέχουν αλληλουχία 100 αμινοξέων στην οποία γίνεται η πρόσδεση με τις αντίστοιχες κινάσες (CDKs). Εν γένει, τα επίπεδα συγκέντρωσης ή ενεργοποίησης των κυκλινών διαφέρουν σημαντικά στις διάφορες φάσεις του κύκλου λόγω της αποδόμησης και επανασύνθεσης ή ενεργοποίησης και απενεργοποίησης τους τους σε διαφορετικά χρονικά σημεία. Ακριβώς σε αυτή την περιοδική αλλαγή συγκέντρωσης ή ενεργοποίησης οφείλουν την ονομασία τους. Συγκεκριμένα, αύξηση συγκέντρωσης (ή ενεργοποίησης) της Κυκλίνης E παρατηρείται κατά τη φάση G1 και πριν την είσοδο στην S φάση, της Κυκλίνης A πριν την είσοδο στην S φάση και κατά τη διάρκεια αυτής και της Κυκλίνης B πριν τη φάση M. Εξαιρεση αποτελεί η Κυκλίνη D της οποίας τα επίπεδα παραμένουν γενικά σταθερά καθ' όλη τη διάρκεια του κύκλου. Όπως θα συζητηθεί λεπτομερώς σε επόμενη υπό-ενότητα, η περιοδική ή όχι ενεργοποίηση της Κυκλίνης D αποτελεί ακόμα ανοιχτό ζήτημα έρευνας.

Οι εξαρτώμενες από κυκλίνες κινάσες (CDKs) είναι κινάσες σερίνης-θρεονίνης με καταλυτική δράση και αποτελούνται από τις κινάσες CDK2, CDK4, CDK6 και την cdc2 (ή CDK1). Στο μόριό τους περικλείουν μία κοινή αλληλουχία αμινοξέων που αντιστοιχεί στο 40% περίπου της δομής τους. Στο τμήμα αυτό συμπεριλαμβάνονται η αγκύλη ενεργοποίησης (activation loop) καθώς και η έλικα PSTAIRE που αποτελεί σημείο σύνδεσης με την αντίστοιχη κυκλίνη. Τα επίπεδα τους παραμένουν εν γένει σταθερά σε όλη τη διάρκεια του κυτταρικού κύκλου και επομένως η δραστηριότητα των συμπλόκων εξαρτάται από την συγκέντρωση των κυκλινών στις διάφορες φάσεις του κύκλου.

Ο βασικός ρόλος των CDKs όταν ενεργοποιούνται από την συνδεδεμένη κυκλίνη είναι να εκτελούν την βιοχημική αντίδραση της φωσφορυλίωσης επί συγκεκριμένων πρωτεϊνών στόχων (με αποτέλεσμα την ενεργοποίηση ή την απενεργοποίηση αυτών). Οι πρωτεΐνες-στόχος αυτές ρυθμίζουν την έξοδο του κυττάρου από την προηγούμενη φάση και την είσοδο του στην επόμενη φάση του κυτταρικού κύκλου [138]. Όπως προαναφέρθηκε, απαιτείται η σύνδεση των CDKs με κυκλίνες, προκειμένου να εκτελεστεί η προαναφερθείσα λειτουργία, σε συγκεκριμένα ετεροδιμερή τα οποία είναι:

- Κυκλίνη A:Cdk2 (Cyclin A:Cdk2)
- Κυκλίνη-B:Cdc2 (Cyclin B:Cdc2)

- Κυκλίνη-D:CDK4 και Κυκλίνη-D:CDK6 (Cyclin D:Cdk4, Cyclin D:Cdk6, σημειώνονται και ως Κυκλίνη-D:CDK4,6, Cyclin D:Cdk4,6)
- Κυκλίνη-E:CDK2 (Cyclin E:Cdk2)

Τα σύμπλοκα αυτά παρουσιάζουν την βασική τους δραστηριότητα σε συγκεκριμένες φάσεις του κυτταρικού κύκλου. Συγκεκριμένα, στις φάσεις G0 και G1 δρα το σύμπλοκο Cyclin D:Cdk4,6, στη φάση G1 και S δρα το σύμπλοκο Cyclin E:Cdk2, στη φάση S το σύμπλοκο Cyclin A:Cdk2, στη φάση G2 το σύμπλοκο Cyclin A:Cdc2 ενώ στη φάση M το σύμπλοκο Cyclin B:Cdc2.

3.2.2 Αναστολείς των συμπλεγμάτων κυκλινών και εξαρτώμενων από κυκλίνες κινασών

Τρεις οικογένειες γονιδίων δρουν ανασταλτικά στο ρόλο των CDKs οι οποίες ονομάζονται Αναστολείς των εξαρτώμενων από κυκλίνες κινασών (Cyclin Dependent Kinase Inhibitors, CDKIs).

- η οικογένεια WAF1/CIP1/(γονίδιο p21)
- η οικογένεια KIP (γονίδια p27,p57)
- η οικογένεια INK4a (γονίδια p14,p15,p16 και p18).

Λόγω του καθοριστικού ρόλου που έχουν τα ανωτέρω αναφερόμενα γονίδια στην πρόληψη σχηματισμού όγκων, συχνά αναφέρονται και ως ογκοκατασταλτικά γονίδια.

Οι οικογένειες CIP και KIP αναστέλλουν κυρίως τη δράση των κινασών CDK2 και cdc2 αδρανοποιώντας τα αντίστοιχα συμπλέγματα κυκλινών-κινασών στην φάση G1 του κυτταρικού κύκλου ενώ η οικογένεια INK4a δρα ανασταλτικά κυρίως στη δράση των κυκλινών CDK4/6. Ωστόσο σημαντικό ρόλο στην καταπίεση της CDK4 παίζει και το p27.

Με την επίδραση ανασταλτικών ερεθισμάτων (π.χ. σύνδεση του μετατρεπτικού παράγοντα των όγκων TGFβ στους υποδοχείς επιθηλιακών κυττάρων) αυξάνονται τα επίπεδα του παράγοντα p15, ο οποίος αναστέλλει την σύνθεση και την δράση των ετεροδιμερών κυκλίνης D-CDK4/6, προκαλώντας παύση του κυτταρικού κύκλου στην φάση G1 και επομένως μη μετάβαση στην φάση S. Με την επίδραση των ιδίων ερεθισμάτων αυξάνουν και τα επίπεδα των παραγόντων p21 και p27, προκαλώντας αναστολή δράσης των αντίστοιχων ετεροδιμερών κυκλίνης E-CDK2, κυκλίνης A-CDK2, κυκλίνης A-CDC2 και κυκλίνης B-CDC2.

Αντίθετα μιτογόνα ερεθίσματα μειώνουν την ενδοπυρηνική συγκέντρωση των παραγόντων p21 και p27 με αποτέλεσμα να μειώνεται η ανασταλτική τους δράση επί των ετεροδιμερών E-CDK2, A-CDK2, A-CDC2 και B-CDC2 και κατά συνέπεια να επιτρέπεται η πρόοδος του κυτταρικού κύκλου [2,8].

3.2.3 Σημεία Ελέγχου του Κυτταρικού Κύκλου

Όπως συζητήθηκε και σε προηγούμενες υπο-ενότητες, τα σημεία ελέγχου εξασφαλίζουν την παρακολούθηση από το κύτταρο της προόδου του κυτταρικού κύκλου και επιτρέπουν την ρύθμιση αυτής. Τα σημεία ελέγχου εμποδίζουν την πρόοδο του κυτταρικού κύκλου σε καθορισμένα σημεία επιτρέποντας την επαλήθευση των απαραίτητων διαδικασιών και την επιδιόρθωση των βλαβών του DNA ενώ το κύτταρο δεν μπορεί να προχωρήσει στην επόμενη φάση μέχρι την ολοκλήρωση των διαδικασιών αυτών. Στη βιβλιογραφία αναφέρονται τα εξής σημεία ελέγχου μετάβασης φάσεων :

- **Σημείο Ελέγχου G1-S** («Σημείο Περιορισμού», Restriction Point, RP). Το σημείο αυτό είναι από τα κυριότερα και εντοπίζεται προς το τέλος της φάσης G1. Βασική του αποστολή είναι ο έλεγχος της ακεραιότητας του DNA που πρόκειται να αναδιπλασιαστεί ώστε να αποφευχθεί η αντιγραφή παθολογικού DNA και η μεταφορά του στα θυγατρικά κύτταρα. Στο σημείο αυτό, εφόσον διαπιστωθεί παθολογία, σταματά ο κύκλος μέχρι να ολοκληρωθεί η επιδιόρθωση της βλάβης του DNA. Η πρωτεΐνη η οποία είναι κυρίως υπεύθυνη για την φύλαξη της πύλης μετάβασης G1/S είναι η πρωτεΐνη του γονιδίου του ρετινοβλαστώματος (retinoblastoma protein, pRb) και ελέγχει την πρόοδο του κύκλου ανάλογα με τα επίπεδα φωσφορυλίωσής της. Κατά την διάρκεια της φάσης G1 η pRb είναι συνδεδεμένη με την ομάδα μεταγραφικών παραγόντων E2F, μέλη της οποίας συνδέονται με το DNA, αλλά είναι βρίσκονται σε ανενεργή κατάσταση λόγω της ταυτόχρονης δέσμευσής τους με την pRb. Στην υπέρ-φωσφορυλιωμένη μορφή η pRb αποδεσμεύεται από το σύμπλεγμα E2F με αποτέλεσμα το τελευταίο να μπορεί να προάγει τη μεταγραφή γονιδίων. Η ενεργοποίηση των E2F έχει σαν αποτέλεσμα την έκφραση διαφόρων γονιδίων όπως Κυκλίνη E, Κυκλίνη A, DNA πολυμεράση κ.α., όπως και των ίδιων. Η έκφραση των γονιδίων αυτών είναι απολύτως απαραίτητη για την περαιτέρω πρόοδο του κύκλου.
- **Σημείο Ελέγχου G2**. Ενεργοποιείται όταν η σύνθεση του DNA είναι αλλοιωμένη ή παρεμποδίζεται. Στη φάση αυτή δημιουργείται το σύμπλεγμα

Κυκλίνης B:Cdc2, το οποίο φωσφορυλιώνει διάφορες πρωτεΐνες αναγκαίες για την μίτωση. Ανασταλτικό ρόλο έχει η πρωτεΐνη p21 που ανήκει στην οικογένεια αναστολέων WAF1/CIP1

- **Σημείο Ελέγχου M.** Σε αυτό το σημείο ελέγχεται η ορθή καθήλωση των χρωματοσωμάτων στην μιτωτική άτρακτο και η ίση κατανομή τους στα δύο θυγατρικά κύτταρα

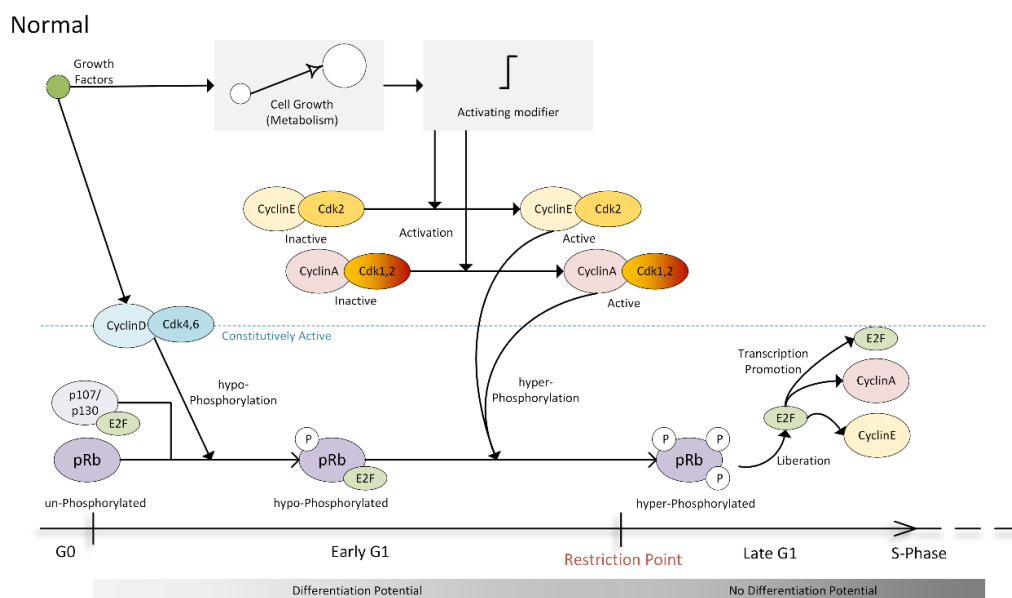
Εστιάζοντας περαιτέρω στο Σημείο Περιορισμού (RP), όπως παρουσιάζεται στην [139], στην βιβλιογραφία εμφανίζονται δύο βασικές θεωρίες σχετικά με την βιοχημική ρύθμιση του RP. Η πρώτη η οποία αποτελεί το «τρέχον παράδειγμα» (“current paradigm”), η οποία κυρίως προσπαθεί να εξηγήσει πειράματα με βασικό άξονα την αποστέρηση των κυττάρων από ορό (serum) και κατόπιν την επαναδιέγερση τους με αυτόν [140]. Η δεύτερη νέο-εισαχθείσα θεωρία (αναφερόμενη ως «νέα θεωρία για το RP») [139], ωστόσο, βασίζεται σε πειραματικά δεδομένα από κύτταρα στα οποία ήταν συνεχώς διαθέσιμοι αυξητικοί παράγοντες (growth factors).

Ειδικότερα, με βάση το «τρέχον παράδειγμα», η διέγερση των αναπαυόμενων (resting) κυττάρων με αυξητικούς παράγοντες οδηγεί σε σταδιακή εμφάνιση ενεργών CyclinD:Cdk4,6 συμπλόκων. Ωστόσο η «νέα θεωρία για το RP» υποστηρίζει πως τα σύμπλοκα αυτά εκφράζονται συνεχώς και παραμένουν ενεργά καθ' όλη τη διάρκεια του κυτταρικού κύκλου. Όσον αφορά τις επιδράσεις των οντοτήτων αυτών στην εκτέλεση του RP, το «τρέχον παράδειγμα» υποστηρίζει πως τα φαινόμενα αυτά οδηγούν σε μερική απενεργοποίηση της pRb μέσω της υπό-φωσφορυλίωσης (hypo-phosphorylation) της. Η υποστηριζόμενη αυτή απενεργοποίηση ευνοεί την έκφραση της Κυκλίνης E (Cyclin E) και τον σχηματισμό ενεργών CyclinE:Cdk2 συμπλόκων, ως αποτέλεσμα της σταδιακής απελευθέρωσης των παραγόντων μεταγραφής E2F (οι οποίοι είναι κρίσιμοι για την έναρξη της διαδικασίας διπλασιασμού του DNA) και της απομόνωσης (sequestration) της p27 αντίστοιχα, η οποία p27 δρα ως Αναστολέας Εξαρτώμενης από Κυκλίνες Κινάσης (Cyclin Dependent Kinase Inhibitor, CDKI). Η εμφάνιση των ενεργών συμπλόκων CyclinE:Cdk2, εν τέλει, οδηγεί σε πλήρη απενεργοποίηση της pRb μέσω υπέρ-φωσφορυλίωσης.

Ωστόσο, για την υπό-φωσφορυλιωμένη (hypo-phosphorylated) μορφή της pRb, αναφέρεται όλο και συχνότερα στη βιβλιογραφία ότι παρουσιάζει δυνατότητες καταπίεσης της αύξησης του κυττάρου (growth suppression capabilities), κυρίως μέσω της καταπίεσης των E2F [141]–[143]. Έτσι, στη «νέα θεωρία για το RP», υιοθετείτε το συγκεκριμένο εύρημα και απορρίπτονται τα όσα υποστηρίζει το «τρέχον παράδειγμα» σχετικά με τη δράση των διαφόρων καταστάσεων φωσφορυλίωσης της pRb.

Σχετικά με την ενεργοποίηση των συμπλόκων CyclinE:Cdk2, η νέα θεωρία εισάγει ένα ακόμα μη πλήρως καλά καθορισμένο τροποποιητή ενεργοποίησης (activating modifier) ο οποίος ενεργοποιεί τις Cdk2 με κριτήριο την μεταβολική κατάσταση του κυττάρου. Ο μηχανισμός αυτός θεωρείται πως λειτουργεί με παρόμοιο τρόπο με τον ενεργοποιητή Bck2 που συναντάται στην G1 φάση του κυτταρικού κύκλου του ζυμομύκητα (yeast). Επιπλέον, με βάση αυτή τη θεωρία, τα σύμπλοκα CyclinE:Cdk2 εκφράζονται συνεχώς ωστόσο εμφανίζονται ανενεργά κατά την πρώιμη (early) G1 φάση και ενεργά κατά την ύστερη (late) G1 φάση. Το πρότυπο αυτό της ενεργοποίησης, παρουσιάζεται να είναι σε συσχέτιση με το αντίστοιχο πρότυπο ταλάντωσης της ενεργούς (υπό-φωσφορυλιωμένης) και της ανενεργούς (υπέρ-φωσφορυλιωμένης) μορφής της pRb.

Προκειμένου να επιβεβαιώσουν περαιτέρω την «νέα θεωρία για το RP», οι συγγραφείς της [139], ανέπτυξαν επίσης και ένα μαθηματικό μοντέλο στο οποίο η αλληλεπίδραση του μεταβολισμού του κυττάρου με την ενεργοποίηση της Cdk2 υλοποιήθηκε με την εισαγωγή ενός χρόνο-εξαρτώμενου (time-dependent) μηχανισμού διακόπτη (τροποποιητής ενεργοποίησης-activating modifier). Ο διακόπτης αυτός τροποποιεί τους ρυθμούς των σχετικών με την ενεργοποίηση της Cdk2 βιοχημικών αντιδράσεων (τους δίνει μη μηδενική τιμή) ύστερα από ένα προ-καθορισμένο χρονικό διάστημα το οποίο σχετίζεται με το ρυθμό αύξησης του όγκου του κυττάρου (cell growth rate). Οι βασικές αρχές της «νέας θεωρίας για το RP» παρουσιάζονται στην ακόλουθη εικόνα (από την οποία έχουν απαλειφθεί λεπτομέρειες όπως η καταπίεση των Cyclin:Cdk συμπλόκων από τις CDKIs για λόγους απλοποίησης).



Εικόνα 3.1 Κύρια μέρη του βιολογικού μονοπατιού της βιοχημικής ρύθμισης του G1/S «Σημείου Περιορισμού» (Restriction Point) στον φυσιολογικό (μη λευχαιμικό, Normal) κυτταρικό κύκλο.

Βάσει της πλέον πρόσφατης θεωρίας, όπως παρουσιάζεται στην [139], για την λειτουργικότητα του RP, η παρουσία αυξητικών παραγόντων (growth factors) οδηγεί σε συνεχή ενεργοποιημένη κατάσταση τα σύμπλοκα μεταξύ Κυκλίνης D (Cyclin D) και των Εξαρτώμενων από Κυκλίνες Κινάσεων 4 και 6 (Cyclin Dependent Kinases 4,6 – Cdk4,6). Η ενεργοποίηση των συμπλόκων αυτών, οδηγεί στην υπό-φωσφορυλίωση (hypo-phosphorylation) της pRb στην πρόωμη G1 φάση (Early G1). Η hypo-phosphorylated pRb διατηρεί την δυνατότητα καταπίεσης των παραγόντων μεταγραφής (transcription factors) E2F, οι οποίοι έχουν κομβικό ρόλο στην εξέλιξη του κυτταρικού κύκλου, ρυθμίζοντας την έκφραση των Κυκλινών A και E (Cyclin A και Cyclin E αντίστοιχα). Η παρουσία των Growth Factors επίσης οδηγεί στην ενεργοποίηση του μεταβολικού μηχανισμού του κυττάρου και στην αύξηση της μάζας/όγκου αυτού. Όταν η τελευταία φθάσει σε είναι κρίσιμο κατώφλι, τα σύμπλοκα Cyclin E:Cdk2 και Cyclin A:Cdk1,2 ενεργοποιούνται οδηγώντας στην υπέρ-φωσφορυλίωση (hyper-phosphorylation) της pRb στην τελική φάση της G1 (late G1), από την οποία και μετά, βάσει της θεωρίας για το RP, το κύτταρο χάνει την δυνατότητα διαφοροποίησης στα πλαίσια του παρόντος κύκλου. Τέλος, η αύξηση των επιπέδων της hyper-phosphorylated pRb, οδηγεί στην σταδιακή αύξηση των επιπέδων των E2F οι οποίοι με τη σειρά τους οδηγούν στην αύξηση των επιπέδων της Cyclin A. Τα επίπεδα αυτής της Κυκλίνης είναι ενδεικτικά για την μετάβαση των κυττάρων στην S-Phase [144].

3.3 Ειδικά χαρακτηριστικά της βιοχημικής ρύθμισης του κυτταρικού κύκλου στην Οξεία Λεμφοβλαστική Λευχαιμία

Ο κυτταρικός κύκλος αλλά και γενικότερα ο πολλαπλασιασμός των λευχαιμικών κυττάρων, παρουσιάζουν κάποια ιδιαίτερα αξιοσημείωτα χαρακτηριστικά. Ξεκινώντας από το επίπεδο του κυτταρικού πληθυσμού, η υπόθεση για την ύπαρξη βλαστικών καρκινικών κυττάρων, όπως έχει υιοθετηθεί σε διάφορες μορφές καρκίνου [145]–[147], πρόσφατα αμφισβητήθηκε για την περίπτωση της ALL [148], [149]. Ωστόσο, κυτταρικοί υποπληθυσμοί με σημαντικά διαφορετικά χαρακτηριστικά κυτταρικού κύκλου βρίσκονται σε δείγματα μυελού των οστών από ασθενείς με ALL [150]. Ειδικότερα, CD19+ κύτταρα (κύτταρα που εκφράζουν το B-λεμφοκυτταρικό αντιγόνο CD19) είναι ο επικρατέστερος και ταχύτερα πολλαπλασιαζόμενος πληθυσμός, αποτελώντας το 90% του συνολικού πληθυσμού των κυττάρων. Ωστόσο εμφανίζεται και ένα μικρό ποσοστό CD19- κυττάρων,

με μικρό ρυθμό πολλαπλασιασμού και αραιή παρουσία στην S-φάση που τα καθιστά υποψήφια για το ρόλο των προγονικών/βλαστικών κυττάρων [150].

Εστιάζοντας στο κυτταρικό επίπεδο, τα λευχαιμικά κύτταρα των ασθενών με BCP-ALL εμφανίζουν μέση τιμή (\pm τυπική απόκλιση) διάρκειας κυτταρικού κύκλου T_c της τάξεων των $112.5 (\pm 46.8)$ h σε σύγκριση με τις $65.5 (\pm 3.5)$ h των φυσιολογικών ομολόγων τους (μη νεοπλασματικών πρόδρομων B λεμφοκυττάρων) [151]. Όσον αφορά την κατανομή των BCP-ALL κυττάρων στις διάφορες φάσεις του κυτταρικού κύκλου, έχει δειχθεί πως η πλειοψηφία των κυττάρων αυτών βρίσκονται στην G1 φάση (περισσότερα του 80%) και μόνο ένα μικρό μέρος (~7-10%) βρίσκονται στις S, G2 και M φάσεις [152]–[154]. Επιπροσθέτως, το ποσοστό των κυττάρων που βρίσκονται στη φάση ανάπαυσης (G0) είναι σημαντικά μικρό (περίπου 2%) [142], [155]. Τέλος, ένα ακόμα χαρακτηριστικό που εξηγεί την σχεδόν πλήρη επικράτηση των λευχαιμικών κυττάρων στον μυελό των οστών των ασθενών, παρά τον αργό ρυθμό πολλαπλασιασμού τους, είναι η σημαντικά μειωμένος ρυθμός αποπτωτικού θανάτου, όπου μέσω αυτής της οδού κυτταρικού θανάτου μειώνεται ο πληθυσμός των λευχαιμικών κυττάρων κατά 4%, ενώ η δημιουργία νέων κυττάρων οδηγεί κατά μέσο όρο σε καθημερινή επέκταση της ασθένειας κατά 10-11%.

Εμβαθύνοντας σε διαδικασίες που άπτονται του μοριακού επιπέδου, αρχικά αξίζει να σημειωθεί πως η συχνά παρατηρούμενη στου καρκίνους μετάβαση στην αερόβια γλυκόλυση (γνωστή και ως φαινόμενο του Warburg), έχει επίσης παρατηρηθεί στην περίπτωση της ALL [156]. Εστιάζοντας τώρα σε μηχανισμούς που επηρεάζουν απευθείας τον κυτταρικό κύκλο, ένα εύρημα σημαντικής αξίας είναι η σχεδόν πλήρης υπέρ-φωσφορυλιωμένη (hyper-phosphorylated) κατάσταση της πρωτεΐνης Retinoblastoma (pRb) [142], [144], [152]. Η αξία του ευρήματος πηγάζει από ευρέως αποδεκτό και κεντρικό ρόλο που η διαδοχική φωσφορυλίωση της pRb (ή τον συγγενικών πρωτεϊνών p107 και p130) κατέχει στην έναρξη της μετάβασης από την φάση G1 στην φάση S [140]–[142], [144], [152], [153], [157]–[161] και η οποία συνολικά ελέγχεται από το βιολογικό μονοπάτι «Σημείο περιορισμού» (Restriction Point-RP) [140].

Βάσει των όσων παρουσιάστηκαν στην ενότητα 3.2, η συμπεριφορά των βιοχημικών οντοτήτων που ρυθμίζουν τον κυτταρικό κύκλο αντιβαίνει στη θεωρία του RP. Παρόλο που η υπέρ-φωσφορυλιωμένη κατάσταση της pRb ορίζει μία γρήγορη και χωρίς επιστροφή μετάβαση στην S φάση, ο ρυθμός μετάβασης στην BCP-ALL είναι απρόσμενα χαμηλός και μόνο ένα μικρό ποσοστό κυττάρων βρίσκονται στην S φάση, σε αντίθεση με τα ομόλογα φυσιολογικά κύτταρα [144], [151]–[153], [162]. Για παράδειγμα, σε κύτταρα που εκφράζουν το αντιγόνο αιμοποιητικών προγονικών κυττάρων CD34 (κύτταρα CD34+), η

διέγερση τους με κυτοκίνες (μικρά πρωτεϊνικά μόρια, τα οποία παράγονται από διαφορετικούς τύπους κυττάρων και εξασφαλίζουν την επικοινωνία μεταξύ των κυττάρων του ανοσοποιητικού συστήματος) οδηγεί σε επικράτηση της υπέρ-φωσφορυλιωμένης μορφής της pRb και σε σημαντική αύξηση του ποσοστού των κυττάρων που βρίσκονται στις φάσεις S και G2/M του κυτταρικού κύκλου. Αντίθετα, μη διεγερμένα κύτταρα του ίδιου τύπου, στα οποία μόνο η μη και η υπό-φωσφορυλιωμένη μορφή της pRb ανιχνεύεται, κυρίως βρίσκονται στις φάσεις G0 και πρόιμη-G1 φάση [142]. Η παρατηρούμενη υπέρ-φωσφορυλιωμένη κατάσταση της pRb πιθανόν να υποδεικνύει ότι τα BCP-ALL κύτταρα βρίσκονται συγκεκριμένα στην ύστερη-G1 φάση, ακριβώς στο RP ή και πέρα από αυτό. Η κατάσταση αυτή θα μπορούσε να εξηγήσει, με βάση και τα όσα παρουσιάστηκαν και στο Κεφάλαιο 1 για την σχέση του κυτταρικού κύκλου με την διαφοροποίηση, τον φαινότυπο αντίστασης σε διαφοροποίηση που εμφανίζουν τα λευχαιμικά κύτταρα [153]. Οπότε, είναι εξαιρετικά πιθανό κάποια αλλοίωση στους μηχανισμούς της G1 φάσης και της G1/S μετάβασης να έχει συμβεί στα BCP-ALL κύτταρα λόγω της καρκινικής τους μετάλλαξης.

Αναζητώντας τα μοριακά εκείνα στοιχεία τα οποία μπορεί να έχουν επηρεαστεί, αρχικά θα πρέπει να σημειωθεί ότι η G1/S μετάβαση των BCP-ALL κυττάρων, έχειδειχθεί να είναι ανεπηρέαστη όσον αφορά την παραγωγή και την ενεργοποίηση της Cyclin A [144]. Αυτό είναι εν γένει μη αναμενόμενο, λαμβάνοντας υπόψη ότι το γονίδιο που κωδικοποιεί την Cyclin A (CCNA1) αποτελεί έναν από τους γνωστούς στόχους των παραγόντων μεταγραφής E2F και ότι τα επίπεδα της παραγόμενης πρωτεΐνης έχουν συσχετιστεί με την μετάβαση των κυττάρων, συμπεριλαμβανομένων των BCP-ALL κυττάρων, στην S φάση [139], [144]. Οπότε μια απορρύθμιση που απευθείας και μοναδικά αναφέρεται στην Cyclin A, πρέπει να αποκλειστεί.

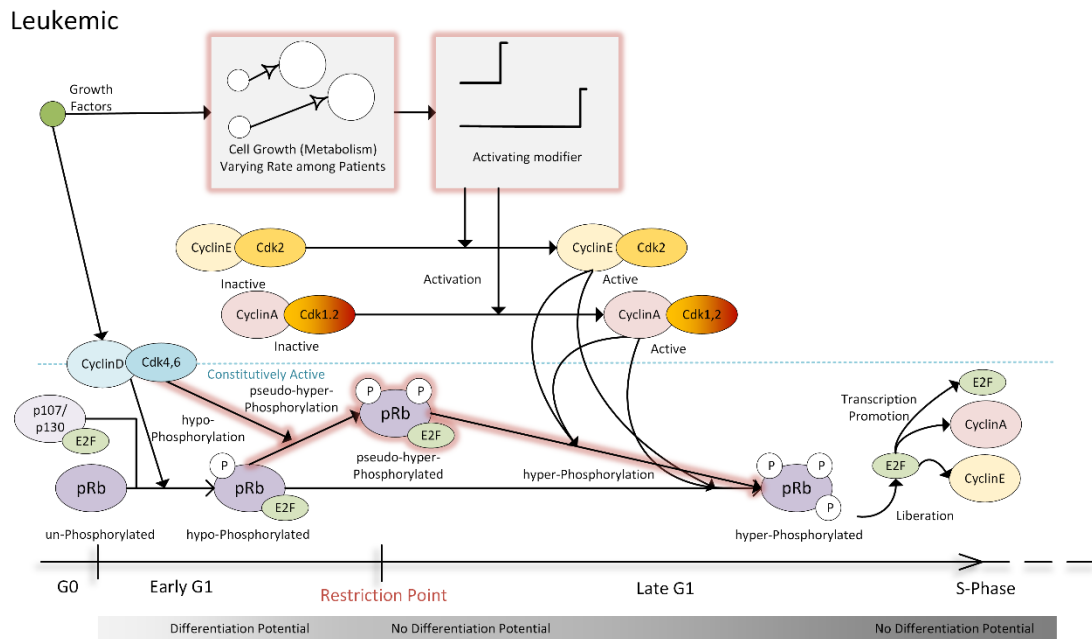
Ψάχνοντας σε προηγούμενα στάδια από αυτό της έκφρασης και ενεργοποίησης της Cyclin A στο RP, έχειδειχθεί στη βιβλιογραφία πως η διαδοχική φωσφορυλίωση της pRb είναι απορυθμισμένη στην BCP-ALL και πως παρατηρείται ένα περιορισμένος πλεονασμός μεταξύ Cdk2 και Cdk4 σε αυτό το φαινόμενο [153]. Πιο συγκεκριμένα, έχει παρατηρηθεί στην κυτταρική σειρά NALM-6 αλλά και σε κύτταρα ασθενών που πάσχουν από BCP-ALL ότι οι εξειδικεύσεις υποστρώματος (substrate specificities) των Cdk4,6 είναι απορυθμισμένες στην BCP-ALL κάτι που υποστηρίζεται από το εύρημα ότι αυτές οι κινάσες, έχουν βρεθεί να μπορούν παραδόξως να φωσφορυλιώνουν το σημείο φωσφορυλίωσης Serine 612 (ser612) της pRb για το οποίο θεωρείται γενικότερα ότι εξειδικεύεται η Cdk2 [153], [163]–[166]. Επιπλέον, έχει αναφερθεί πως ακόμα και σε hyper-phosphorylated κατάσταση, η pRb διατηρεί την πρόσδεση της στον πυρήνα του

κυττάρου και συνεχίζει μερικώς να καταπιέζει του E2F (τουλάχιστον τον E2F-1) [144], [153], [166]. Ιδιαίτερη σημασία επίσης έχει το γεγονός ότι τα παραπάνω φαινόμενα έχουν παρατηρηθεί να συμβαίνουν σε διαφορετική έκταση μεταξύ των διαφόρων παρατηρούμενων ασθενών, θέτοντας τα έτσι ως υποψήφια να συνεισφέρουν στην σημαντική διακύμανση που παρουσιάζει ο ρυθμός πολλαπλασιασμού των λευχαιμικών κυττάρων μεταξύ αυτών.

Μια υπόθεση που έχει κατατεθεί στην βιβλιογραφία [153], σχετικά με τις πιθανές συνέπειες που μπορεί να έχει η παραπάνω αναφερόμενη απορρύθμιση, αφορά στην πιθανότητα η επέμβαση των Cdk4,6 στα προτιμώμενα από την Cdk2 σημεία φωσφορυλίωσης, να οδηγεί στην δημιουργία μεγάλων φωσφο-ομάδων στην pRb πριν την ενεργοποίηση και την συμμετοχή της Cdk2 (η οποία είναι ενδεικτική του περάσματος στην late-G1 κατά την φυσιολογική εκτέλεση του κυτταρικού κύκλου) στη ρύθμιση του κυτταρικού κύκλου. Επιπλέον, έχει αναφερθεί [142] πως σε αναλύσεις με ανοσοαποτυπώματα τύπου Western (Western Blots) δειγμάτων κυττάρων BCP-ALL ασθενών, πολλαπλές μορφές της pRb, μεταξύ δύο και πέντε, μπορούν να ανιχνευθούν. Τρεις από αυτές τις μορφές, που παρουσιάζουν διαφορετικές κινητικότητες σε πειράματα ηλεκτροφόρησης πρωτεϊνών (SDS-PAGE, Sodium Dodecyl Sulfate – PolyAcrylamide Gel Electrophoresis) ανάλογα με το επίπεδο φωσφορυλίωσης, πιθανότατα αντιπροσωπεύουν τις μη-, υπό και υπέρ- φωσφορυλιωμένες (un-, hypo- και hyper-phosphorylated) μορφές της pRb. Ωστόσο, στα πλαίσια της παρούσας διατριβής, γίνεται η υπόθεση πως τουλάχιστον μία από τις δύο εναπομένουσες μορφές μπορεί να αποτελεί το αποτέλεσμα αυτής της ιδιόμορφης φωσφορυλίωσης από την Cdk4. Η υπόθεση αυτή είναι σε συμφωνία με σχετική υπόθεση που γίνεται στην βιβλιογραφία [144], όπου υποστηρίζεται η πιθανή μερική απενεργοποίηση της λειτουργικότητας της pRb στην BCP-ALL, οδηγώντας στην απώλεια της δυνατότητας του κυττάρου να διαφοροποιηθεί, χωρίς ωστόσο να είναι εξασφαλισμένη η διέλευση από το RP και η ολοκλήρωση του κυτταρικού κύκλου. Η κατάσταση αυτή της pRb που παρουσιάζει υψηλά επίπεδα φωσφορυλίωσης χωρίς ωστόσο να έχει τις ιδιότητες της υπέρ-φωσφορυλιωμένης μορφής την ονομάζεται (για πρώτη φορά από την παρούσα διατριβή) ψευδό-υπέρ-φωσφορυλιωμένη κατάσταση (pseudo-hyper-phosphorylated pRb). Η εξακρίβωση των πιθανών συνεπειών της απορρύθμισης της ακολουθιακής (sequential) φωσφορυλίωσης στην λειτουργικότητα της pRb μπορεί να χαρακτηριστεί ως σύνθετη διαδικασία. Παρά τα ιδιαίτερα σημαντικά ευρήματα που αποκαλύπτουν το ρόλο του κάθε σημείου φωσφορυλίωσης στην ρύθμιση της λειτουργίας της pRb [143], [167]–[169], η πλήρης εικόνα δεν έχει αποκαλυφθεί ακόμα. Σε υποστήριξη της υπόθεσής μας, ωστόσο,

έχει αναφερθεί σε ετερογενή πειράματα ότι η φωσφορυλίωση της θέσης φωσφορυλίωσης ser612 τουλάχιστον δεν μειώνει την δυνατότητα της pRb να καταπιέζει τους E2F [170], [171].

Λαμβάνοντας υπόψη τα παραπάνω, οι βασικές απορρυθμίσεις που πιστεύεται ότι είναι σημαντικές ώστε να εξηγηθεί η εκτέλεση του RP στα λευχαιμικά κύτταρα δίνονται στην Εικόνα 3.2.



Εικόνα 3.2 Τροποποιήσεις και απορρυθμίσεις του G1/S «Σημείου Περιορισμού» (Restriction Point) στην BCP-ALL

Οι βασικές διαφορές (επισημασμένες με κόκκινο χρώμα) μεταξύ των όσων παρουσιάστηκαν στην Εικόνα 3.1 (φυσιολογική συμπεριφορά του «Σημείου περιορισμού») και των όσων παρουσιάζονται εδώ αφορούν στην δυνατότητα των συμπλόκων CyclinD:Cdk4,6, εκτός από το να υπό-φωσφορυλιώνουν την pRb, να οδηγούν αυτόνομα την πρωτεΐνη σε ένα ενδιάμεσο στάδιο φωσφορυλίωσης (ψευδο-υπέρ-φωσφορυλίωση, pseudo-hyper-phosphorylation). Στο στάδιο αυτό η πρωτεΐνη διατηρεί την δυνατότητα να καταπιέζει τους E2F παράγοντες μεταγραφής παρά το αυξημένο περιεχόμενό της σε φώσφορο. Επίσης η συγκεκριμένη μορφή της pRb πιστεύεται ότι έχει απωλέσει τις σχετικές με την διαφοροποίηση του κυττάρου ιδιότητες της [144] και έτσι η συγκέντρωση της μπορεί να θεωρηθεί ότι αποτελεί ένδειξη παρουσίας του κυττάρου ακριβώς επί ή πέραν του «Σημείου Περιορισμού». Επιπλέον, δεδομένου ότι με βάση την «νέα θεωρία για το RP», τα σύμπλοκα CyclinD:Cdk4,6 είναι συνεχώς ενεργά κατά τη διάρκεια της G1 φάσης, σε αντίθεση με τα σύμπλοκα CyclinE:Cdk2, μπορούν να συμβάλουν στη δημιουργία ανώτερων μορφών φωσφορυλίωσης της pRb από τις πρώτες ώρες της εκτέλεσης του

κυτταρικού κύκλου, κάτι που δύναται να εξηγήσει την σχεδόν πλήρη επικράτηση των μορφών αυτών στα κύτταρα των ασθενών με BCP-ALL. Ωστόσο, η πλήρης απελευθέρωση των E2F δεν είναι εφικτή μέχρι η pRb να υπέρ-φωσφορυλιωθεί πλήρως (ως αποτέλεσμα της ενεργοποίησης των συμπλόκων CyclinE:Cdk2 και CyclinA:Cdk1,2) κάτι που μπορεί να γίνει είτε μετά την ψευδο-υπέρ-φωσφορυλίωση είτε απευθείας μέσω της φυσιολογικής οδού η οποία είναι επίσης ενεργή στα BCP-ALL cells (η Cdk2 συνεχίζει να μπορεί να φωσφορυλιώσει το σημείο ser612) [152], [153].

Τέλος, λαμβάνοντας υπόψη τις σημαντικές διαφορές στον ρυθμό μεταβολισμού της γλυκόζης μεταξύ των ασθενών που παρουσιάζουν καλή ή κακή απόκριση στη θεραπεία με Γλυκοκορτικοειδή (Prednisone Sensitive και Prednisone Resistant ασθενείς) [154], ο μεταβλητός ρυθμός μεταβολισμού, που μπορεί να οδηγήσει σε διαφοροποιημένη ρύθμιση της ενεργοποίησης των CyclinE:Cdk2 συμπλόκων, καθυστερημένη ή επιταχυμένη και διαμέσου της λειτουργίας του τροποποιητή ενεργοποίησης (activating modifier), θεωρείται κρίσιμος ώστε η παρατηρούμενη διακύμανση στον ρυθμό πολλαπλασιασμού των λευχαιμικών κυττάρων να εξηγηθεί.

Δεδομένου ότι η δοθείσα περιγραφή για την βιοχημική ρύθμιση του κυτταρικού κύκλου στην BCP-ALL αντιβαίνει ακόμα και τις ευρέως αποδεκτές αρχές για το RP, η απευθείας προσομοίωση της μέσω ήδη δημοσιευμένων μοντέλων για τον κυτταρικό κύκλο [139], [158], [158], [172]–[177], τροποποιώντας απλά τις τιμές κάποιων παραμέτρων τους θεωρείται αδύνατη. Προκύπτει λοιπόν η ανάγκη δημιουργίας ενός νέου μοντέλου, που θα παρουσιάζει δομικές αλλαγές σε σχέση με τα υπάρχοντα. Ωστόσο, στα πλαίσια της διατριβής αναζητήθηκε η δυνατότητα το μοντέλο αυτό να αναπτυχθεί έχοντας ως βάση κάποιο από αυτά (μοντέλο αναφοράς – reference model), με βάση κριτήρια τα οποία παρουσιάζονται σε επόμενη ενότητα.

Σε αυτά τα πλαίσια, κεντρικό στόχο της παρούσας προσπάθειας αποτέλεσε η αναζήτηση του κατά πόσο η εισαγωγή των παραπάνω ευρημάτων σε ένα ήδη καλώς ορισμένο μοντέλο για τον κυτταρικό κύκλο μέσω της τροποποίησης της δομής του και μέσω επαναπροσαρμογής των παραμέτρων του θα είναι ικανή έτσι ώστε να τροποποιήσει την συμπεριφορά του με τέτοιο τρόπο ώστε να προσομοιώνονται τα παρατηρούμενα στην BCP-ALL φαινόμενα (τα σχετικά με τα υπό-κυτταρικά φαινόμενα και την κινητική του κυτταρικού κύκλου). Επίσης, μία επιτυχής προσαρμογή του μοντέλου θα μπορούσε να το καταστήσει υποψήφιο σημαντικό κομμάτι ενός μηχανισμού προσωποποιημένης πρόβλεψης της διάρκειας του κυτταρικού κύκλου (T_c).

3.4 Βιβλιογραφική ανασκόπηση της μοντελοποίησης της βιοχημικής ρύθμισης του κυτταρικού κύκλου και επιλογή μοντέλου αναφοράς

Η επιλογή το μοντέλου αναφοράς βασίστηκε σε συγκεκριμένα κριτήρια με βάση τα οποία αξιολογήθηκαν ένας σημαντικός αριθμός μοντέλων που παρουσιάζονται στη διεθνή βιβλιογραφία για την βιοχημική ρύθμιση του κυτταρικού κύκλου. Ο Πίνακας 3.1 παρουσιάζει τα συγκεκριμένα κριτήρια και τα αποτελέσματα αξιολόγησης των μοντέλων. Στις ακόλουθες παραγράφους γίνεται λεπτομερής παρουσίαση των κριτηρίων.

Πίνακας 3.1 Κριτήρια επιλογής του μοντέλου αναφοράς και αξιολόγηση υποψήφιων μοντέλων

Μοντέλο	Κριτήρια						
	Λεπτομερής μοντελοποίηση των βημάτων φωσφορυλίωσης της pRb	Μοντελοποίηση των επιπέδων της Κυκλίνης A (Cyclin A)	Μοντελοποίηση της επιρροής του μεταβολισμού στην εκτέλεση του κυτταρικού κύκλου	Πλήρης μοντελοποίηση του μηχανισμού του κυτταρικού κύκλου	Αναφορά σε πειράματα με συνεχή έκθεση σε αυξητικούς παράγοντες	Μοντελοποίηση των επιπέδων των p16 και p27	Διαθεσιμότητα σε αναγνώσιμη από υπολογιστή μορφή
Yao et al. [178]	✗	✗	✗	✗	✗	✗/✗	✓
Tyson and Novak [172] and Conradie et al. [179]	✗	✓	✓	✓	✗	✗/✓	✓
Haberichter et al. [139]	✓	✓	✓	✗	✓	✗/✓	✓
Swat et al. [176]	✓	✗	✗	✗	✗	✗/✗	✓
Iwamoto et al. (2008) [173]	✓	✓	✗	✗	-	✓/✓	✗
Iwamoto et al. (2011) [174]	✓	✓	✗	✓	-	✓/✓	✗
Pfeuty [175]	✓	✗	✗	✗	✗	✗/✓	✗

Aguda and Tang [180]	×	×	×	×	-	✓/✓	✓
- : Δεν ορίζεται/Δεν είναι γνωστό							

Αρχικά, η λεπτομερής μοντελοποίηση του μηχανισμού του RP, συμπεριλαμβανομένων των διακριτών και ξεχωριστών βημάτων της φωσφορυλίωσης της pRb (σε κάποια μοντέλα τα βήματα υπό- και υπέρ-φωσφορυλίωσης είναι ενοποιημένα, Πίνακας 3.1) τέθηκε ως βασική προϋπόθεση. Όπως παρουσιάστηκε σε προηγούμενες υπό-ενότητες, βασικές απορρυθμίσεις του κυτταρικού κύκλου είναι γνωστό ότι επηρεάζουν τον μηχανισμό αυτό. Εν συνεχεία, γνωρίζοντας ότι τα επίπεδα της Κυκλίνης A (Cyclin A) συσχετίζονται με το ποσοστό των λευχαιμικών κυττάρων που εισέρχονται στην S φάση [144], προτιμήθηκαν μοντέλα που προσομοιώνουν την έκφραση της Κυκλίνης A και την αλληλεπίδραση της με άλλες βιοχημικές οντότητες που αφορούν τον κυτταρικό κύκλο.

Όσον αφορά την συμμετοχή του μεταβολισμού στην ρύθμιση του κυτταρικού κύκλου, όπως συζητήθηκε σε προηγούμενες υπο-ενότητες, ο ρυθμός γλυκόλυσης (glycolysis rate) έχει βρεθεί να διαφέρει μεταξύ δύο βασικών υπό-κατηγοριών ασθενών (Ασθενείς με Καλή ή Κακή απόκριση στην Πρεδνιζόνη, Good vs Poor Prednisone responders) και να συσχετίζεται με την κυτταρική αύξηση και τον ρυθμό πολλαπλασιασμού των κυττάρων. Έτσι ένα μοντέλο το οποίο λαμβάνει υπόψη το συγκεκριμένο ρυθμιστικό μηχανισμό θα επέτρεπε την μελέτη της επιρροής του μεταβολισμού στη δυναμική κυτταρικού κύκλου (cell cycle dynamics) των λευχαιμικών κυττάρων.

Ακολούθως, η διαθεσιμότητα του προς επιλογή μοντέλου σε μία αναγνώσιμη από τον υπολογιστή γλώσσα (π.χ. σε SBML όπως αυτή παρουσιάστηκε στο Κεφάλαιο 2), θεωρείται σημαντική ώστε να εξασφαλιστεί η ορθή εκ νέου υλοποίηση του μοντέλου και η αναπαραγωγή των αρχικών του αποτελεσμάτων.

Επιπροσθέτως, η δημιουργία ενός μοντέλου με βάση πειράματα στα οποία τα υπό-μελέτη κύτταρα εκτίθεντο σε συνεχή τροφοδοσία με αυξητικούς παράγοντες, θεωρείται εγγύτερη στην in vivo κατάσταση (in vivo setting) και γι' αυτό είναι προτιμητέα.

Η G1 φάση του κυτταρικού κύκλου θεωρείται ως η πιο ευέλικτη φάση, όσον αφορά την διάρκειά της [181]. Όπως παρουσιάζεται λεπτομερώς και σε ακόλουθες ενότητες, υπάρχουν σημαντικές ενδείξεις ότι αυτό ισχύει και στην περίπτωση της BCP-ALL. Επιπροσθέτως, τα κύρια ευρήματα σχετικά με την απορρύθμιση του κυτταρικού κύκλου στη ALL αφορούν αυτή τη φάση. Οπότε, παρόλο που ένα μεγάλος αριθμός μοντέλων που περιγράφουν/μοντελοποιούν όλη την έκταση του μηχανισμού του κυτταρικού κύκλου (και

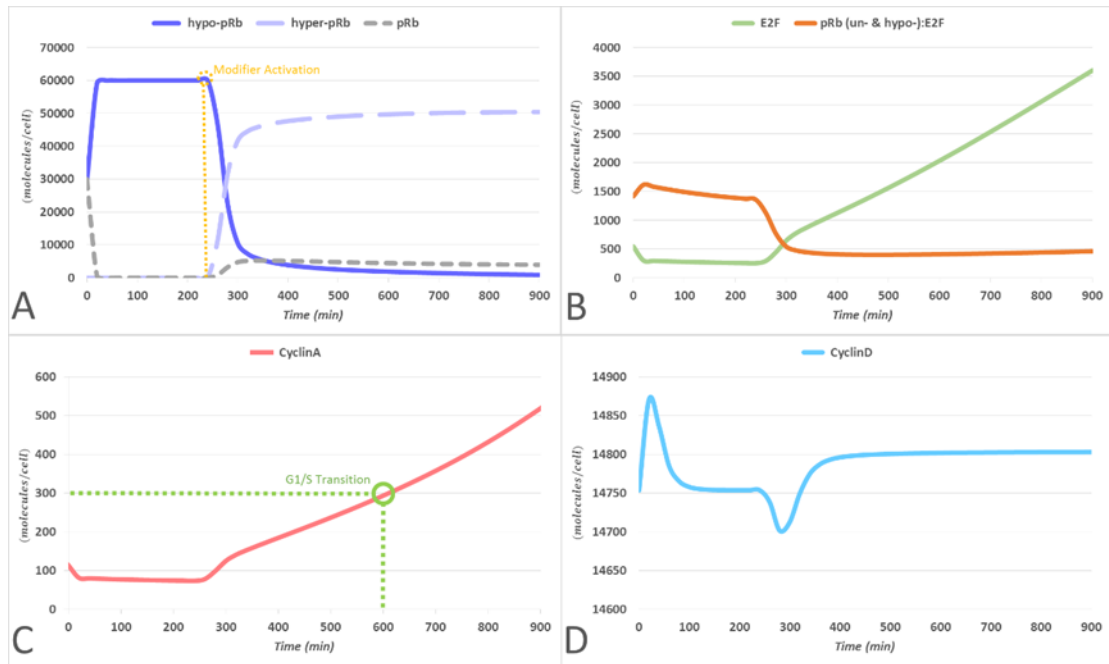
όχι μόνο συγκεκριμένες φάσεις) είναι διαθέσιμος, αυτό το χαρακτηριστικό δεν θεωρείται απολύτως απαραίτητο.

Τέλος, η απενεργοποίηση του γονιδίου p16 είναι ένα αρκετά συχνό φαινόμενο στην ALL, είτε μέσω διαγραφής του γονιδίου (gene deletion) είτε μέσω μεθυλίωσης (methylation) [182]. Έτσι, όσον αφορά την μοντελοποίηση των σχετικών με τις CDKI p16 και p27 καταπιεστικών φαινομένων επί των συμπλόκων CyclinD:Cdk4,6 και CyclinE:Cdk2, η πιθανή απουσία αυτού του μέρους του βιολογικού μονοπατιού (p16) δεν θεωρείται ικανή να επηρεάσει την ανάλυση που πραγματοποιείται στα πλαίσια της διατριβής, σε αντίθεση με την p27.

Με βάση τα ανωτέρω κριτήρια, το μοντέλο που επιλέχθηκε να αποτελέσει βάση για το νέο προτεινόμενο από τη διατριβή μοντέλο ήταν αυτό που παρουσιάζεται στην [139] (Habericther et al.), δεδομένου ότι πληροί όλες τις προϋποθέσεις που αξιολογήθηκαν ως σημαντικές. Το συγκεκριμένο μοντέλο για τη συνέχεια του κειμένου θα αναφέρεται ως «μοντέλο αναφοράς» (reference model).

Μεταξύ άλλων, ιδιαίτερα σημαντικό χαρακτηριστικό του μοντέλου αυτού αποτελεί η μοντελοποίηση της παρεμβολής του μεταβολισμού στην ενεργοποίηση των CyclinE:Cdk2 μέσω ενώ μηχανισμού όπου μετά την παρέλευση συγκεκριμένου χρονικού διαστήματος αυξάνει τους ρυθμούς των αντιδράσεων που σχετίζονται με την ενεργοποίηση αυτή (τροποποιητής ενεργοποίησης, activating modifier). Τα αποτελέσματα προσομοίωσης του μοντέλου αναφοράς δίνονται στην Εικόνα 3.3 ώστε να είναι άμεσα συγκρίσιμα με τα αποτελέσματα του προτεινόμενου μοντέλου όπως θα παρουσιαστούν στις ακόλουθες υπό-ενότητες της διατριβής. Στην συγκεκριμένη προσομοίωση, χρησιμοποιούνται οι παράμετροι του μοντέλου όπως υπολογίστηκαν από τους συγγραφείς της [139] και όπως δίνονται στην σχετική για το συγκεκριμένο μοντέλο καταχώρηση στη βάση μοντέλων συστημικής βιολογίας BIOMODELS [183] (<https://www.ebi.ac.uk/biomodels-main/BIOMD0000000109>). Οι προσομοιώσεις έγιναν με τη χρήση του συμβατού με την SBML εργαλείου μοντελοποίησης και προσομοίωσης COPASI [184], με τη χρήση του Ντετερμινιστικού Αλγορίθμου LSODA με βήμα επίλυσης (interval) 10 min.

Θα πρέπει τέλος να σημειωθεί πως η προσαρμογή του συγκεκριμένου μοντέλου έγινε με βάση εργαστηριακά δεδομένα που προήλθαν από κυτταρική σειρά ανθρώπινων κυττάρων HCT166 καρκίνου του μεγάλου εντέρου (colon carcinoma cells) τα οποία καλλιεργήθηκαν σε συνθήκες συνεχούς παροχής αυξητικών παραγόντων (growth factors).



Εικόνα 3.3 Αποτελέσματα προσομοίωσης του μοντέλου αναφοράς [139] για 900 min (15 h)

Όπως παρουσιάζεται στην Εικόνα 3.3Α, η υπό-φωσφορυλιωμένη μορφή της pRb (hypo-pRb, μωβ) ξεκάθαρα επικρατεί για τα πρώτα 300 λεπτά (min) της προσομοίωσης λόγω της φωσφορυλίωσης των μη φωσφορυλιωμένων αντιγράφων της pRb (γκρι) με την μεσολάβηση των CyclinD:Cdk4,6 και παραμένουν σταθερά σε υψηλά επίπεδα μέχρι την ενεργοποίηση του τροποποιητή ενεργοποίησης στα 240 min (κάτι που έχει ως αποτέλεσμα την ενεργοποίηση των συμπλόκων CyclinE:Cdk2 και CyclinA:Cdk1,2).

Για την ίδια χρονική διάρκεια, τα επίπεδα των ελεύθερων E2F (Εικόνα 3.3Β, πράσινο) και της Cyclin A (Εικόνα 3.3C, κόκκινο) δεν παρουσιάζουν τάσεις αύξησης λόγω της καταπίεσης των πρώτων από την pRb (Εικόνα 3.3Β, πορτοκαλί). Ωστόσο, όταν τα επίπεδα της υπέρ-φωσφορυλιωμένης pRb ξεκινούν να ανεβαίνουν (Εικόνα 3.3Α, ανοιχτό μωβ) ως αποτέλεσμα της ενεργοποίησης των CyclinE:Cdk2, τα επίπεδα τους αυξάνονται ραγδαία. Αυτό έχει ως αποτέλεσμα την αύξηση των επιπέδων της CyclinA (Εικόνα 3.3C, κόκκινο), τα οποία φτάνουν στο όριο των 300 (molecules/cell) περίπου στα 600 min (σημειώνεται με πράσινο). Όπως συζητείται λεπτομερώς και σε ακόλουθες ενότητες, το συγκεκριμένο όριο προκύπτει με βάση την εργαστηριακή παρατήρηση του χρόνου κατά τον οποίο γίνεται η μετάβαση στην S φάση στα πειράματα τα οποία διενεργήθηκαν για την ανάπτυξη του μοντέλου αναφοράς.

Τέλος τα επίπεδα της CyclinD (γαλάζιο) δεν παρουσιάζουν σημαντικές διακυμάνσεις κατά την εκτέλεση του μοντέλου.

Η συγκεκριμένη συμπεριφορά είναι τυπική για ένα μοντέλο που προσομοιώνει το RP. Δεδομένου όμως ότι η παραπάνω εικόνα δεν συναντάται στα BCP-ALL κύτταρα, γεννάται η ανάγκη για την δημιουργία του προτεινόμενου νέου μοντέλου. Τα βήματα που ακολουθήθηκαν για τον ορισμό και την προσαρμογή του δίνονται στις ακόλουθες ενότητες.

3.5 Δημιουργία και υλοποίηση προτεινόμενου μοντέλου

Έχοντας ως βάση τις απορρυθμίσεις του κυτταρικού κύκλου στην BCP-ALL, όπως παρουσιάστηκαν στις προηγούμενες ενότητες, συγκεκριμένες προσθήκες και τροποποιήσεις πραγματοποιήθηκαν στο μοντέλο αναφοράς που αφορούν την δομή του (βιοχημικές αντιδράσεις) αλλά και τις παραμέτρους τους (ρυθμοί αντιδράσεων, reaction rates). Ο Πίνακας 3.2 παρουσιάζει το σύνολο των βιοχημικών αντιδράσεων ενώ ο Πίνακας 3.3 το σύνολο των μοριακών οντοτήτων του μοντέλου αναφοράς όπως και τα νέο-εισαχθέντα αντίστοιχα. Επίσης στις ακόλουθες παραγράφους δίνεται λεπτομερής περιγραφή της κάθε μίας καινούργιας βιοχημικής αντίδρασης ή οντότητας ενώ στην Εικόνα 3.4, το μέρος της δομής του μοντέλου το οποίο τροποποιήθηκε παρουσιάζεται σε Systems Biology Graphical Notation (SBGN) [185] (θα πρέπει να σημειωθεί ότι η πλήρης απεικόνιση όλου του μοντέλου αναφοράς θα ήταν αδύνατη λόγω του σημαντικού αριθμού βιοχημικών οντοτήτων και αντιδράσεων).

Πίνακας 3.2 Βιοχημικές αντιδράσεις που μοντελοποιούνται στο προτεινόμενο μοντέλο και οι σχετικές κινητικές παράμετροι

A/A Αντίδρασης	Αντίδραση	Κινητικός Νόμος	Όνομα παραμέτρου και τιμή ή ορισμός
1	-> CyclinD	Constant Flux*	$k_{sCyclinD} = 1761.08 \text{ ((molecules/cell)/min)}$
2	CyclinD ->	Mass Action [#]	$k_{dCyclinD} = 0.05 \text{ (min)}^{-1}$
3	Cyclin D:Cdk 4 -> Cdk 4	Mass Action	$k_{dCyclinD} = 0.05 \text{ (min)}^{-1}$
4	p27:Cyclin D:Cdk 4 -> p27:Cdk 4	Mass Action	$k_{dCyclinD} = 0.05 \text{ (min)}^{-1}$
5	-> p27	Constant Flux	$k_{sp27} = 195.472 \text{ ((molecules/cell)/min)}$
6	p27 ->	Mass Action	$k_{dp27} = 0.001575 \text{ (min)}^{-1}$
7	p27:Cdk 4 -> Cdk 4	Mass Action	$k_{dp27} = 0.001575 \text{ (min)}^{-1}$
8	p27:Cyclin D:Cdk 4 -> Cyclin D:Cdk 4	Mass Action	$k_{dp27} = 0.001575 \text{ (min)}^{-1}$
9	p27:Cdk2 -> Cdk 2	Mass Action	$k_{dp27} = 0.001575 \text{ (min)}^{-1}$
10	p27:Cyclin E:Cdk2 -> Cyclin E:Cdk2	Mass Action	$k_{dp27} = 0.001575 \text{ (min)}^{-1}$
11	p27:Cyclin A:Cdk2 -> Cyclin A:Cdk2	Mass Action	$k_{dp27} = 0.001575 \text{ (min)}^{-1}$
12	p27:Cdk2(M) -> Cdk2(M)	Mass Action	$k_{dp27} = 0.001575 \text{ (min)}^{-1}$
13	p27:Cyclin E:Cdk2(M) -> Cyclin E:Cdk2(M)	Mass Action	$k_{d1p27} = 0.071149 \text{ (min)}^{-1}$
14	p27:Cyclin A:Cdk2(M) -> Cyclin A:Cdk2(M)	Mass Action	$k_{dp27} = 0.001575 \text{ (min)}^{-1}$
15	-> CyclinE	Constant Flux	$k_{sCyclinE} = k_{s0CyclinE} + \frac{k_{s1CyclinE} * [E2F]}{k_{sMCyclinE} + [E2F]}$

			$k_{s0CyclinE} = 254.0742$ $k_{s1CyclinE} = 980.611$ $k_{sMCyclinE} = 9992.647$
16	CyclinE ->	Mass Action	$k_{dCyclinE} = 0.05 \text{ (min)}^{-1}$
17	Cyclin E:Cdk2 -> Cdk 2	Mass Action	$k_{dCyclinE} = 0.05 \text{ (min)}^{-1}$
18	p27:Cyclin E:Cdk2 -> p27:Cdk2	Mass Action	$k_{dCyclinE} = 0.05 \text{ (min)}^{-1}$
19	Cyclin E:Cdk2(M) -> Cdk2(M)	Mass Action	$k_{dCyclinE} = 0.05 \text{ (min)}^{-1}$
20	p27:Cyclin E:Cdk2(M) -> p27:Cdk2(M)	Mass Action	$k_{dCyclinE} = 0.05 \text{ (min)}^{-1}$
21	-> CyclinA	Constant Flux	$k_{sCyclinA} = k_{s0CyclinA} + \frac{k_{s1CyclinA} * [E2F]}{k_{sMCyclinA} + [E2F]}$ $k_{s0CyclinA} = 499.9437$ $k_{s1CyclinA} = 7999.996$ $k_{sMCyclinA} = 4064.384$
22	CyclinA ->	Mass Action	$k_{dCyclinA} = 0.05 \text{ (min)}^{-1}$
23	Cyclin A:Cdk1 -> Cdk1	Mass Action	$k_{dCyclinA} = 0.05 \text{ (min)}^{-1}$
24	Cyclin A:Cdk1(M) -> Cdk1(M)	Mass Action	$k_{dCyclinA} = 0.05 \text{ (min)}^{-1}$
25	Cyclin A:Cdk2 -> Cdk 2	Mass Action	$k_{dCyclinA} = 0.05 \text{ (min)}^{-1}$
26	p27:Cyclin A:Cdk2 -> p27:Cdk2	Mass Action	$k_{dCyclinA} = 0.05 \text{ (min)}^{-1}$
27	Cyclin A:Cdk2(M) -> Cdk2(M)	Mass Action	$k_{dCyclinA} = 0.05 \text{ (min)}^{-1}$
28	p27:Cyclin A:Cdk2(M) -> p27:Cdk2(M)	Mass Action	$k_{dCyclinA} = 0.05 \text{ (min)}^{-1}$
29	-> E2F	Constant Flux	$k_{sE2F} = k_{s0E2F} + \frac{k_{s1E2F} * [E2F]}{k_{sME2F} + [E2F]}$ $k_{s0E2F} = 6.927086$ $k_{s1E2F} = 65.44282$ $k_{sME2F} = 9818.78$
30	E2F ->	Mass Action	$k_{dE2F} = 0.002229 \text{ (min)}^{-1}$
31	E2F:pRb -> pRb	Mass Action	$k_{dE2F} = 0.006465 \text{ (min)}^{-1}$
32	E2F:hypo-pRb -> hypo-pRb	Mass Action	$k_{dE2F} = 0.006465 \text{ (min)}^{-1}$
33	E2F:hyper-pRb -> hyper-pRb	Mass Action	$k_{dE2F} = 0.006465 \text{ (min)}^{-1}$
33_leuk	E2F:pseudo-hyper-pRb -> pseudo-hyper-pRb	Mass Action	$k_{dE2Fleuk} = pr_{dE2F} * k_{dE2F}$ $pr_{dE2F} = 1$
34	-> Emi1	Constant Flux	$k_{sEmi1} = k_{s0Emi1} + \frac{k_{s1Emi1} * [E2F]}{k_{sMEmi1} + [E2F]}$ $k_{s0Emi1} = 2.004744$ $k_{s1Emi1} = 1788.517$ $k_{sMEmi1} = 9608.162$
35	Emi1 ->	Mass Action	$k_{dEmi1} = 0.018158 \text{ (min)}^{-1}$
36	APCC:Emi1 -> APCC	Mass Action	$k_{dEmi1} = 0.018158 \text{ (min)}^{-1}$
37	Cdk 4 + CyclinD -> Cyclin D:Cdk 4	Mass Action	$k_{bCyclinDCdk4} = 1.43 \cdot e^{-6} \left(\frac{\text{molecules}}{\text{cell}} * \text{min} \right)^{-1}$
38	Cyclin D:Cdk 4 -> Cdk 4 + CyclinD	Mass Action	$k_{uCyclinDCdk4} = 0.1 \text{ (min)}^{-1}$
39	p27:Cdk 4 + CyclinD -> p27:Cyclin D:Cdk 4	Mass Action	$k_{bCyclinDCdk4} = 1.43 \cdot e^{-6} \left(\frac{\text{molecules}}{\text{cell}} * \text{min} \right)^{-1}$
40	p27:Cyclin D:Cdk 4 -> p27:Cdk 4 + CyclinD	Mass Action	$k_{uCyclinDCdk4} = 0.1 \text{ (min)}^{-1}$
41	Cdk 4 + p27 -> p27:Cdk 4	Mass Action	$k_{bp27Cdk4} = 6.34 \cdot e^{-6} \left(\frac{\text{molecules}}{\text{cell}} * \text{min} \right)^{-1}$
42	p27:Cdk 4 -> Cdk 4 + p27	Mass Action	$k_{up27Cdk4} = 0.1 \text{ (min)}^{-1}$
43	Cyclin D:Cdk 4 + p27 -> p27:Cyclin D:Cdk 4	Mass Action	$k_{bp27Cdk4} = 6.34 \cdot e^{-6} \left(\frac{\text{molecules}}{\text{cell}} * \text{min} \right)^{-1}$
44	p27:Cyclin D:Cdk 4 -> Cyclin D:Cdk 4 + p27	Mass Action	$k_{bp27Cdk4} = 6.34 \cdot e^{-6} \left(\frac{\text{molecules}}{\text{cell}} * \text{min} \right)^{-1}$
45	Cdk 2 + p27 -> p27:Cdk2	Mass Action	$k_{bp27Cdk2} = 1.23 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} * \text{min} \right)^{-1}$
46	p27:Cdk2 -> Cdk 2 + p27	Mass Action	$k_{up27Cdk2} = 0.1 \text{ (min)}^{-1}$
47	Cyclin E:Cdk2 + p27 -> p27:Cyclin E:Cdk2	Mass Action	$k_{bp27Cdk2} = 1.23 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} * \text{min} \right)^{-1}$
48	p27:Cyclin E:Cdk2 -> Cyclin E:Cdk2 + p27	Mass Action	$k_{up27Cdk2} = 0.1 \text{ (min)}^{-1}$
49	Cyclin A:Cdk2 + p27 -> p27:Cyclin A:Cdk2	Mass Action	$k_{bp27Cdk2} = 1.23 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} * \text{min} \right)^{-1}$
50	p27:Cyclin A:Cdk2 -> Cyclin A:Cdk2 + p27	Mass Action	$k_{up27Cdk2} = 0.1 \text{ (min)}^{-1}$

51	Cdk2(M) + p27 -> p27:Cdk2(M)	Mass Action	$k_{bp27cdk2} = 1.23 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} \right) * \text{min}^{-1}$
52	p27:Cdk2(M) -> Cdk2(M) + p27	Mass Action	$k_{up27cdk2} = 0.1 \text{ (min)}^{-1}$
53	Cyclin E:Cdk2(M) + p27 -> p27:Cyclin E:Cdk2(M)	Mass Action	$k_{bp27cdk2} = 1.23 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} \right) * \text{min}^{-1}$
54	p27:Cyclin E:Cdk2(M) -> Cyclin E:Cdk2(M) + p27	Mass Action	$k_{up27cdk2} = 0.1 \text{ (min)}^{-1}$
55	Cyclin A:Cdk2(M) + p27 -> p27:Cyclin A:Cdk2(M)	Mass Action	$k_{bp27cdk2} = 1.23 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} \right) * \text{min}^{-1}$
56	p27:Cyclin A:Cdk2(M) -> Cyclin A:Cdk2(M) + p27	Mass Action	$k_{up27cdk2} = 0.1 \text{ (min)}^{-1}$
57	Cdk 2 -> Cdk2(M)	Mass Action	$k_{act} = \begin{cases} \text{Time} < \text{ModifierTime}, 0 \\ \text{Time} \geq \text{ModifierTime}, 0.0175 \end{cases} \text{ (min)}^{-1}$
58	p27:Cdk2 -> p27:Cdk2(M)	Mass Action	$k_{act} = \begin{cases} \text{Time} < \text{ModifierTime}, 0 \\ \text{Time} \geq \text{ModifierTime}, 0.0175 \end{cases} \text{ (min)}^{-1}$
59	Cyclin E:Cdk2 -> Cyclin E:Cdk2(M)	Mass Action	$k_{act} = \begin{cases} \text{Time} < \text{ModifierTime}, 0 \\ \text{Time} \geq \text{ModifierTime}, 0.0175 \end{cases} \text{ (min)}^{-1}$
60	p27:Cyclin E:Cdk2 -> p27:Cyclin E:Cdk2(M)	Mass Action	$k_{act} = \begin{cases} \text{Time} < \text{ModifierTime}, 0 \\ \text{Time} \geq \text{ModifierTime}, 0.0175 \end{cases} \text{ (min)}^{-1}$
61	Cyclin A:Cdk2 -> Cyclin A:Cdk2(M)	Mass Action	$k_{act} = \begin{cases} \text{Time} < \text{ModifierTime}, 0 \\ \text{Time} \geq \text{ModifierTime}, 0.0175 \end{cases} \text{ (min)}^{-1}$
62	p27:Cyclin A:Cdk2 -> p27:Cyclin A:Cdk2(M)	Mass Action	$k_{act} = \begin{cases} \text{Time} < \text{ModifierTime}, 0 \\ \text{Time} \geq \text{ModifierTime}, 0.0175 \end{cases} \text{ (min)}^{-1}$
63	Cdk 2 + CyclinE -> Cyclin E:Cdk2	Mass Action	$k_{bCyclinECdk2} = 5.01 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} \right) * \text{min}^{-1}$
64	Cyclin E:Cdk2 -> Cdk 2 + CyclinE	Mass Action	$k_{uCyclinECdk2} = 0.1 \text{ (min)}^{-1}$
65	p27:Cdk2 + CyclinE -> p27:Cyclin E:Cdk2	Mass Action	$k_{bCyclinECdk2} = 5.01 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} \right) * \text{min}^{-1}$
66	p27:Cyclin E:Cdk2 -> p27:Cdk2 + CyclinE	Mass Action	$k_{uCyclinECdk2} = 0.1 \text{ (min)}^{-1}$
67	Cdk2(M) + CyclinE -> Cyclin E:Cdk2(M)	Mass Action	$k_{bCyclinECdk2} = 5.01 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} \right) * \text{min}^{-1}$
68	Cyclin E:Cdk2(M) -> Cdk2(M) + CyclinE	Mass Action	$k_{uCyclinECdk2} = 0.1 \text{ (min)}^{-1}$
69	p27:Cdk2(M) + CyclinE -> p27:Cyclin E:Cdk2(M)	Mass Action	$k_{bCyclinECdk2} = 5.01 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} \right) * \text{min}^{-1}$
70	p27:Cyclin E:Cdk2(M) -> p27:Cdk2(M) + CyclinE	Mass Action	$k_{uCyclinECdk2} = 0.1 \text{ (min)}^{-1}$
71	Cdk 2 + CyclinA -> Cyclin A:Cdk2	Mass Action	$k_{bCyclinACdk2} = 9.52 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} \right) * \text{min}^{-1}$
72	Cyclin A:Cdk2 -> Cdk 2 + CyclinA	Mass Action	$k_{uCyclinACdk2} = 0.1 \text{ (min)}^{-1}$
73	p27:Cdk2 + CyclinA -> p27:Cyclin A:Cdk2	Mass Action	$k_{bCyclinACdk2} = 9.52 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} \right) * \text{min}^{-1}$
74	p27:Cyclin A:Cdk2 -> p27:Cdk2 + CyclinA	Mass Action	$k_{uCyclinACdk2} = 0.1 \text{ (min)}^{-1}$
75	Cdk2(M) + CyclinA -> Cyclin A:Cdk2(M)	Mass Action	$k_{bCyclinACdk2} = 9.52 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} \right) * \text{min}^{-1}$
76	Cyclin A:Cdk2(M) -> Cdk2(M) + CyclinA	Mass Action	$k_{uCyclinACdk2} = 0.1 \text{ (min)}^{-1}$
77	p27:Cdk2(M) + CyclinA -> p27:Cyclin A:Cdk2(M)	Mass Action	$k_{bCyclinACdk2} = 9.52 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} \right) * \text{min}^{-1}$
78	p27:Cyclin A:Cdk2(M) -> p27:Cdk2(M) + CyclinA	Mass Action	$k_{uCyclinACdk2} = 0.1 \text{ (min)}^{-1}$
79	Cdk1 -> Cdk1(M)	Mass Action	$k_{act} = \begin{cases} \text{Time} < \text{ModifierTime}, 0 \\ \text{Time} \geq \text{ModifierTime}, 0.0175 \end{cases} \text{ (min)}^{-1}$
80	Cyclin A:Cdk1 -> Cyclin A:Cdk1(M)	Mass Action	$k_{act} = \begin{cases} \text{Time} < \text{ModifierTime}, 0 \\ \text{Time} \geq \text{ModifierTime}, 0.0175 \end{cases} \text{ (min)}^{-1}$
81	Cdk1 + CyclinA -> Cyclin A:Cdk1	Mass Action	$k_{bCyclinACdk1} = 6.48 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} \right) * \text{min}^{-1}$
82	Cyclin A:Cdk1 -> Cdk1 + CyclinA	Mass Action	$k_{uCyclinACdk1} = 0.1 \text{ (min)}^{-1}$
83	Cdk1(M) + CyclinA -> Cyclin A:Cdk1(M)	Mass Action	$k_{bCyclinACdk1} = 6.48 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} \right) * \text{min}^{-1}$
84	Cyclin A:Cdk1(M) -> Cdk1(M) + CyclinA	Mass Action	$k_{uCyclinACdk1} = 0.1 \text{ (min)}^{-1}$
85	pRb + Cyclin D:Cdk 4 -> Cyclin D:Cdk 4_pRb_hypo-pRb_Int	Mass Action	$k_{bD4pRb} = 3.1 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} \right) * \text{min}^{-1}$
85_leuk_1	E2F:hypo-pRb + Cyclin D:Cdk 4 -> Cyclin D:Cdk 4_E2F:hypo-pRb_E2F:pseudo-hyper-pRb_Int	Mass Action	$k_{bD4pRb_{leuk}} = pr_{bD4} * k_{bD4pRb}$ $pr_{bD4} = 0.8416992397$
85_leuk_2	hypo-pRb + Cyclin D:Cdk 4 -> Cyclin D:Cdk 4_hypo-pRb_pseudo-hyper-pRb_Int	Mass Action	$k_{bD4pRb_{leuk}} = pr_{bD4} * k_{bD4pRb}$ $pr_{bD4} = 0.8416992397$
86	Cyclin D:Cdk 4_pRb_hypo-pRb_Int -> pRb + Cyclin D:Cdk 4	Mass Action	$k_{uD4pRb} = 0.1 \text{ (min)}^{-1}$
86_leuk_1	Cyclin D:Cdk 4_E2F:hypo-pRb_E2F:pseudo-hyper-pRb_Int -> E2F:hypo-pRb + Cyclin D:Cdk 4	Mass Action	$k_{uD4pRb_{leuk}} = pr_{uD4} * k_{uD4pRb}$ $pr_{uD4} = 1$
86_leuk_2	Cyclin D:Cdk 4_hypo-pRb_pseudo-hyper-pRb_Int -> hypo-pRb + Cyclin D:Cdk 4	Mass Action	$k_{uD4pRb_{leuk}} = pr_{uD4} * k_{uD4pRb}$ $pr_{uD4} = 1$

87	Cyclin D:Cdk 4_pRb_hypo-pRb_Int -> hypo-pRb + Cyclin D:Cdk 4	Mass Action	$k_{uD4pRb} = 1.69466 \text{ (min)}^{-1}$
87_leuk_1	Cyclin D:Cdk 4_E2F:hypo-pRb_E2F:pseudo-hyper-pRb_Int -> E2F:pseudo-hyper-pRb + Cyclin D:Cdk 4	Mass Action	$k_{uD4pRbLeuk} = pr_{uD4} * k_{uD4pRb}$ $pr_{uD4} = 0.484117838$
87_leuk_2	Cyclin D:Cdk 4_hypo-pRb_pseudo-hyper-pRb_Int -> pseudo-hyper-pRb + Cyclin D:Cdk 4	Mass Action	$k_{uD4pRbLeuk} = pr_{uD4} * k_{uD4pRb}$ $pr_{uD4} = 1$
88	E2F:pRb + Cyclin D:Cdk 4 -> Cyclin D:Cdk 4_E2F:pRb_E2F:hypo-pRb_Int	Mass Action	$k_{bD4pRb} = 3.1 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} \right) * \text{min}^{-1}$
89	Cyclin D:Cdk 4_E2F:pRb_E2F:hypo-pRb_Int -> E2F:pRb + Cyclin D:Cdk 4	Mass Action	$k_{uD4pRb} = 0.1 \text{ (min)}^{-1}$
90	Cyclin D:Cdk 4_E2F:pRb_E2F:hypo-pRb_Int -> E2F:hypo-pRb + Cyclin D:Cdk 4	Mass Action	$k_{uD4pRb} = 1.69466 \text{ (min)}^{-1}$
91	hypo-pRb + Cyclin E:Cdk2(M) -> CyclinE:Cdk2(M)_hypo-pRb_hyper-pRb_Int	Mass Action	$k_{bE2pRb} = 5.74 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} \right) * \text{min}^{-1}$
91_leuk	pseudo-hyper-pRb + Cyclin E:Cdk2(M) -> CyclinE:Cdk2(M)_pseudo-hyper-pRb_hyper-pRb_Int	Mass Action	$k_{bE2pRbLeuk} = pr_{bE2A2A1} * k_{bE2pRb}$ $pr_{bE2A2A1} = 0.1301449788$
92	Cyclin E:Cdk2(M)_hypo-pRb_hyper-pRb_Int -> hypo-pRb + Cyclin E:Cdk2(M)	Mass Action	$k_{uE2pRb} = 0.1 \text{ (min)}^{-1}$
92_leuk	Cyclin E:Cdk2(M)_pseudo-hyper-pRb_hyper-pRb_Int -> pseudo-hyper-pRb + Cyclin E:Cdk2(M)	Mass Action	$k_{uE2pRbLeuk} = pr_{uE2A2A1} * k_{uE2pRb}$ $pr_{uE2A2A1} = 1$
93	Cyclin E:Cdk2(M)_hypo-pRb_hyper-pRb_Int -> hyper-pRb + Cyclin E:Cdk2(M)	Mass Action	$k_{upE2pRb} = 4.78271 \text{ (min)}^{-1}$
93_leuk	Cyclin E:Cdk2(M)_pseudo-hyper-pRb_hyper-pRb_Int -> hyper-pRb + Cyclin E:Cdk2(M)	Mass Action	$k_{upE2pRbLeuk} = pr_{upE2A2A1} * k_{upE2pRb}$ $pr_{upE2A2A1} = 0.428204395$
94	E2F:hypo-pRb + Cyclin E:Cdk2(M) -> Cyclin E:Cdk2(M)_E2F:hypo-pRb_E2F:hyper-pRb_Int	Mass Action	$k_{bE2pRb} = 5.74 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} \right) * \text{min}^{-1}$
94_leuk	E2F:pseudo-hyper-pRb + Cyclin E:Cdk2(M) -> Cyclin E:Cdk2(M)_E2F:pseudo-hyper-pRb_E2F:hyper-pRb_Int	Mass Action	$k_{bE2pRbLeuk} = pr_{bE2A2A1} * k_{bE2pRb}$ $pr_{bE2A2A1} = 0.1301449788$
95	Cyclin E:Cdk2(M)_E2F:hypo-pRb_E2F:hyper-pRb_Int -> E2F:hypo-pRb + Cyclin E:Cdk2(M)	Mass Action	$k_{uE2pRb} = 0.1 \text{ (min)}^{-1}$
95_leuk	Cyclin E:Cdk2(M)_E2F:pseudo-hyper-pRb_E2F:hyper-pRb_Int -> E2F:pseudo-hyper-pRb + Cyclin E:Cdk2(M)	Mass Action	$k_{uE2pRbLeuk} = pr_{uE2A2A1} * k_{uE2pRb}$ $pr_{uE2A2A1} = 1$
96	Cyclin E:Cdk2(M)_E2F:hypo-pRb_E2F:hyper-pRb_Int -> E2F:hyper-pRb + Cyclin E:Cdk2(M)	Mass Action	$k_{upE2pRb} = 4.78271 \text{ (min)}^{-1}$
96_leuk	Cyclin E:Cdk2(M)_E2F:pseudo-hyper-pRb_E2F:hyper-pRb_Int -> E2F:hyper-pRb + Cyclin E:Cdk2(M)	Mass Action	$k_{upE2pRbLeuk} = pr_{upE2A2A1} * k_{upE2pRb}$ $pr_{upE2A2A1} = 0.428204395$
97	hypo-pRb + Cyclin A:Cdk2(M) -> Cyclin A:Cdk2(M)_hypo-pRb_hyper-pRb_Int	Mass Action	$k_{bA2pRb} = 6.25 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} \right) * \text{min}^{-1}$
97_leuk	pseudo-hyper-pRb + Cyclin A:Cdk2(M) -> Cyclin A:Cdk2(M)_pseudo-hyper-pRb_hyper-pRb_Int	Mass Action	$k_{bA2pRbLeuk} = pr_{bE2A2A1} * k_{bA2pRb}$ $pr_{bE2A2A1} = 0.1301449788$
98	Cyclin A:Cdk2(M)_hypo-pRb_hyper-pRb_Int -> hypo-pRb + Cyclin A:Cdk2(M)	Mass Action	$k_{uA2pRb} = 0.1 \text{ (min)}^{-1}$
98_leuk	Cyclin A:Cdk2(M)_pseudo-hyper-pRb_hyper-pRb_Int -> pseudo-hyper-pRb + Cyclin A:Cdk2(M)	Mass Action	$k_{uA2pRbLeuk} = pr_{uE2A2A1} * k_{uA2pRb}$ $pr_{uE2A2A1} = 1$
99	Cyclin A:Cdk2(M)_hypo-pRb_hyper-pRb_Int -> hyper-pRb + Cyclin A:Cdk2(M)	Mass Action	$k_{upA2pRb} = 0.200091 \text{ (min)}^{-1}$
99_leuk	Cyclin A:Cdk2(M)_pseudo-hyper-pRb_hyper-pRb_Int -> hyper-pRb + Cyclin A:Cdk2(M)	Mass Action	$k_{upA2pRbLeuk} = pr_{upE2A2A1} * k_{upA2pRb}$ $pr_{upE2A2A1} = 0.428204395$
100	E2F:hypo-pRb + Cyclin A:Cdk2(M) -> Cyclin A:Cdk2(M)_E2F:hypo-pRb_E2F:hyper-pRb_Int	Mass Action	$k_{bA2pRb} = 6.25 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} \right) * \text{min}^{-1}$
100_leuk	E2F:pseudo-hyper-pRb + Cyclin A:Cdk2(M) -> Cyclin A:Cdk2(M)_E2F:pseudo-hyper-pRb_E2F:hyper-pRb_Int	Mass Action	$k_{bA2pRbLeuk} = pr_{bE2A2A1} * k_{bA2pRb}$ $pr_{bE2A2A1} = 0.1301449788$
101	Cyclin A:Cdk2(M)_E2F:hypo-pRb_E2F:hyper-pRb_Int -> E2F:hypo-pRb + Cyclin A:Cdk2(M)	Mass Action	$k_{uA2pRb} = 0.1 \text{ (min)}^{-1}$
101_leuk	Cyclin A:Cdk2(M)_E2F:pseudo-hyper-pRb_E2F:hyper-pRb_Int -> E2F:pseudo-hyper-pRb + Cyclin A:Cdk2(M)	Mass Action	$k_{uA2pRbLeuk} = pr_{uE2A2A1} * k_{uA2pRb}$ $pr_{uE2A2A1} = 1$
102	Cyclin A:Cdk2(M)_E2F:hypo-pRb_E2F:hyper-pRb_Int -> E2F:hyper-pRb + Cyclin A:Cdk2(M)	Mass Action	$k_{upA2pRb} = 0.200091 \text{ (min)}^{-1}$
102_leuk	Cyclin A:Cdk2(M)_E2F:pseudo-hyper-pRb_E2F:hyper-pRb_Int -> E2F:hyper-pRb + Cyclin A:Cdk2(M)	Mass Action	$k_{upA2pRbLeuk} = pr_{upE2A2A1} * k_{upA2pRb}$ $pr_{upE2A2A1} = 0.428204395$
103	hypo-pRb + Cyclin A:Cdk1(M) -> Cyclin A:Cdk1(M)_hypo-pRb_hyper-pRb_Int	Mass Action	$k_{bA1pRb} = 6.73 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} \right) * \text{min}^{-1}$
103_leuk	pseudo-hyper-pRb + Cyclin A:Cdk1(M) -> Cyclin A:Cdk1(M)_pseudo-hyper-pRb_hyper-pRb_Int	Mass Action	$k_{bA1pRbLeuk} = pr_{bE2A2A1} * k_{bA1pRb}$ $pr_{bE2A2A1} = 0.1301449788$
104	Cyclin A:Cdk1(M)_hypo-pRb_hyper-pRb_Int -> hypo-pRb + Cyclin A:Cdk1(M)	Mass Action	$k_{uA1pRb} = 0.1 \text{ (min)}^{-1}$
104_leuk	Cyclin A:Cdk1(M)_pseudo-hyper-pRb_hyper-pRb_Int -> pseudo-hyper-pRb + Cyclin A:Cdk1(M)	Mass Action	$k_{uA1pRbLeuk} = pr_{uE2A2A1} * k_{uA1pRb}$ $pr_{uE2A2A1} = 1$
105	Cyclin A:Cdk1(M)_hypo-pRb_hyper-pRb_Int -> hyper-pRb + Cyclin A:Cdk1(M)	Mass Action	$k_{upA1pRb} = 0.202132 \text{ (min)}^{-1}$
105_leuk	Cyclin A:Cdk1(M)_pseudo-hyper-pRb_hyper-pRb_Int -> hyper-pRb + Cyclin A:Cdk1(M)	Mass Action	$k_{upA1pRbLeuk} = pr_{upE2A2A1} * k_{upA1pRb}$ $pr_{upE2A2A1} = 0.428204395$
106	E2F:hypo-pRb + Cyclin A:Cdk1(M) -> Cyclin A:Cdk1(M)_E2F:hypo-pRb_E2F:hyper-pRb_Int	Mass Action	$k_{bA1pRb} = 6.73 \cdot e^{-5} \left(\frac{\text{molecules}}{\text{cell}} \right) * \text{min}^{-1}$
106_leuk	E2F:pseudo-hyper-pRb + Cyclin A:Cdk1(M) -> Cyclin A:Cdk1(M)_E2F:pseudo-hyper-pRb_E2F:hyper-pRb_Int	Mass Action	$k_{bA1pRbLeuk} = pr_{bE2A2A1} * k_{bA1pRb}$ $pr_{bE2A2A1} = 0.1301449788$
107	Cyclin A:Cdk1(M)_E2F:hypo-pRb_E2F:hyper-pRb_Int -> E2F:hypo-pRb + Cyclin A:Cdk1(M)	Mass Action	$k_{uA1pRb} = 0.1 \text{ (min)}^{-1}$
107_leuk	Cyclin A:Cdk1(M)_E2F:pseudo-hyper-pRb_E2F:hyper-pRb_Int -> E2F:pseudo-hyper-pRb + Cyclin A:Cdk1(M)	Mass Action	$k_{uA1pRbLeuk} = pr_{uE2A2A1} * k_{uA1pRb}$ $pr_{uE2A2A1} = 1$
108	Cyclin A:Cdk1(M)_E2F:hypo-pRb_E2F:hyper-pRb_Int -> E2F:hyper-pRb + Cyclin A:Cdk1(M)	Mass Action	$k_{upA1pRb} = 0.202132 \text{ (min)}^{-1}$
108_leuk	Cyclin A:Cdk1(M)_E2F:pseudo-hyper-pRb_E2F:hyper-pRb_Int -> E2F:hyper-pRb + Cyclin A:Cdk1(M)	Mass Action	$k_{upA1pRbLeuk} = pr_{upE2A2A1} * k_{upA1pRb}$ $pr_{upE2A2A1} = 0.428204395$

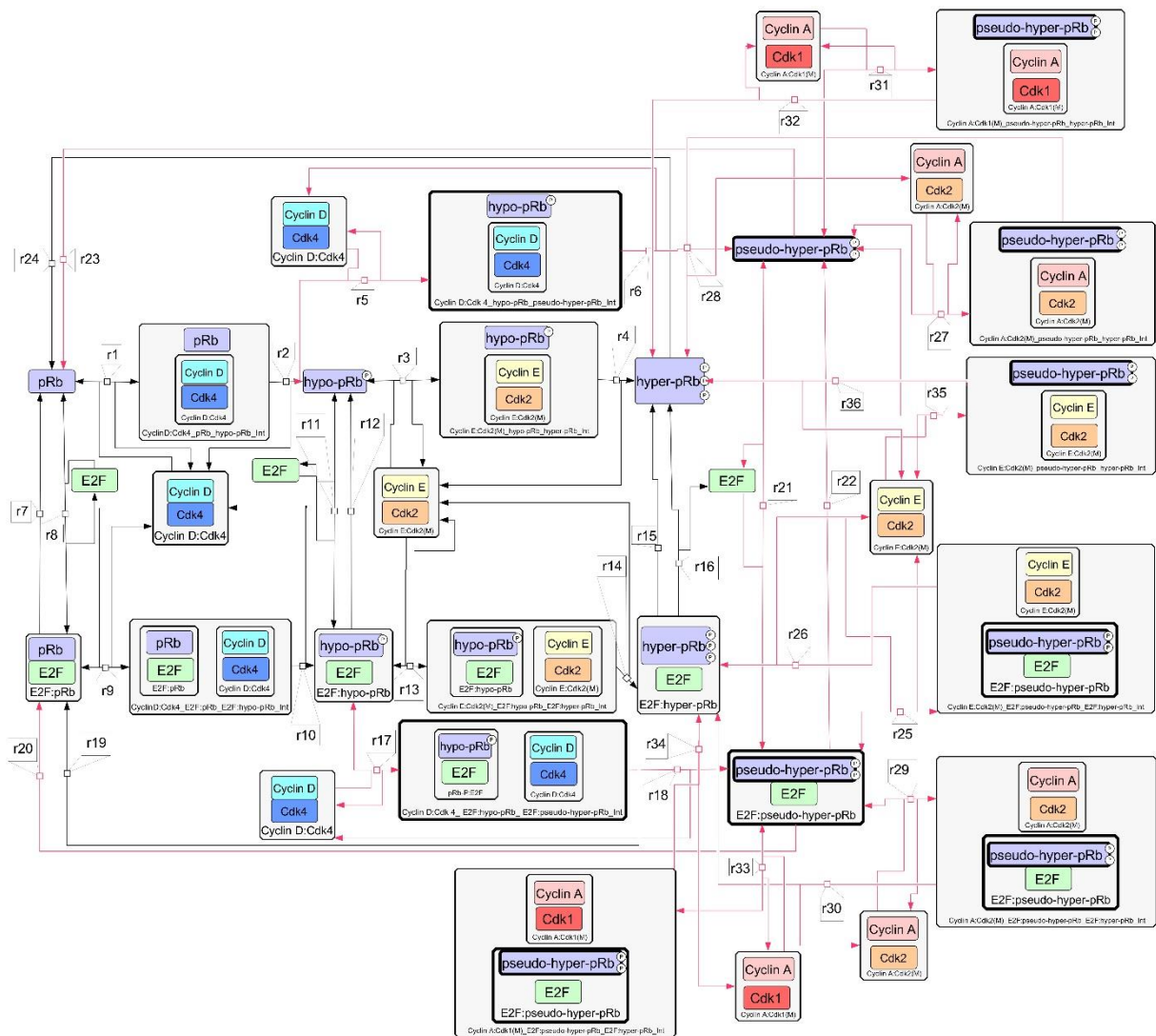
109	hyper-pRb -> pRb	Mass Action	$k_{tpRbDphos} = 0.023194 \text{ (min)}^{-1}$
109_leuk	pseudo-hyper-pRb -> pRb	Mass Action	$k_{tpRbDphosleuk} = pr_{tpRbDphos} * k_{tpRbDphos}$ $pr_{tpRbDphos} = 1$
110	E2F:hyper-pRb -> E2F:pRb	Mass Action	$k_{tpRbDphos} = 0.023194 \text{ (min)}^{-1}$
110_leuk	E2F:pseudo-hyper-pRb -> E2F:pRb	Mass Action	$k_{tpRbDphosleuk} = pr_{tpRbDphos} * k_{tpRbDphos}$ $pr_{tpRbDphos} = 1$
111	pRb + E2F -> E2F:pRb	Mass Action	$k_{bE2FpRb} = 9.66 \cdot e^{-6} \left(\frac{\text{molecules}}{\text{cell}} * \text{min} \right)^{-1}$
112	E2F:pRb -> pRb + E2F	Mass Action	$k_{uE2FpRb} = 0.1 \text{ (min)}^{-1}$
113	hypo-pRb + E2F -> E2F:hypo-pRb	Mass Action	$k_{bE2FpRb} = 9.66 \cdot e^{-6} \left(\frac{\text{molecules}}{\text{cell}} * \text{min} \right)^{-1}$
114	E2F:hypo-pRb -> hypo-pRb + E2F	Mass Action	$k_{uE2FpRb} = 0.1 \text{ (min)}^{-1}$
115	E2F:hyper-pRb -> hyper-pRb + E2F	Mass Action	$k_{uE2FpRb} = 0.1 \text{ (min)}^{-1}$
115_leuk	E2F:pseudo-hyper-pRb -> pseudo-hyper-pRb + E2F	Mass Action	$k_{uE2FpRb,leuk} = pr_{uE2F} * k_{uE2FpRb}$ $pr_{uE2F} = 0.6127679966$
115_leuk_re verse	pseudo-hyper-pRb + E2F -> E2F:pseudo-hyper-pRb	Mass Action	$k_{bE2FpRb,leuk} = pr_{bE2F} * k_{bE2FpRb}$ $pr_{bE2F} = 1.88406857$
116	APCC + Emi1 -> APCC_Emi1	Mass Action	$k_{bEmi1APCC} = 0.0001 \text{ (min)}^{-1}$
117	APCC_Emi1 -> APCC + Emi1	Mass Action	$k_{uEmi1APCC} = 0.1 \text{ (min)}^{-1}$
118	CyclinA + APCC -> APCC_CyclinA_Int	Mass Action	$k_{bAPCCcyclinA} = 1.61 \cdot e^{-6} \left(\frac{\text{molecules}}{\text{cell}} * \text{min} \right)^{-1}$
119	APCC_CyclinA_Int -> CyclinA + APCC	Mass Action	$k_{uAPCCcyclinA} = 0.1 \text{ (min)}^{-1}$
120	APCC_CyclinA_Int -> APCC	Mass Action	$k_{udAPCCcyclinA} = 4.99955 \text{ (min)}^{-1}$
121	Cyclin A:Cdk2 + APCC -> APCC_Cdk2_Cyclin A:Cdk2_Int	Mass Action	$k_{bAPCCcyclinA} = 1.61 \cdot e^{-6} \left(\frac{\text{molecules}}{\text{cell}} * \text{min} \right)^{-1}$
122	APCC_Cdk2_Cyclin A:Cdk2_Int -> Cyclin A:Cdk2 + APCC	Mass Action	$k_{uAPCCcyclinA} = 0.1 \text{ (min)}^{-1}$
123	APCC_Cdk2_Cyclin A:Cdk2_Int -> Cdk2 + APCC	Mass Action	$k_{udAPCCcyclinA} = 4.99955 \text{ (min)}^{-1}$
124	p27:Cyclin A:Cdk2 + APCC -> APCC_p27:Cdk2_p27:Cyclin A:Cdk2_Int	Mass Action	$k_{bAPCCcyclinA} = 1.61 \cdot e^{-6} \left(\frac{\text{molecules}}{\text{cell}} * \text{min} \right)^{-1}$
125	APCC_p27:Cdk2_p27:Cyclin A:Cdk2_Int -> p27:Cyclin A:Cdk2 + APCC	Mass Action	$k_{uAPCCcyclinA} = 0.1 \text{ (min)}^{-1}$
126	APCC_p27:Cdk2_p27:Cyclin A:Cdk2_Int -> p27:Cdk2 + APCC	Mass Action	$k_{udAPCCcyclinA} = 4.99955 \text{ (min)}^{-1}$
127	Cyclin A:Cdk2(M) + APCC -> APCC_Cdk2(M)_Cyclin A:Cdk2(M)_Int; p27:Cyclin A:Cdk2	Mass Action	$k_{bAPCCcyclinA} = 1.61 \cdot e^{-6} \left(\frac{\text{molecules}}{\text{cell}} * \text{min} \right)^{-1}$
128	APCC_Cdk2(M)_Cyclin A:Cdk2(M)_Int -> Cyclin A:Cdk2(M) + APCC	Mass Action	$k_{uAPCCcyclinA} = 0.1 \text{ (min)}^{-1}$
129	APCC_Cdk2(M)_Cyclin A:Cdk2(M)_Int -> Cdk2(M) + APCC	Mass Action	$k_{udAPCCcyclinA} = 4.99955 \text{ (min)}^{-1}$
130	p27:Cyclin A:Cdk2(M) + APCC -> APCC_p27:Cdk2(M)_p27:Cyclin A:Cdk2(M)_Int	Mass Action	$k_{bAPCCcyclinA} = 1.61 \cdot e^{-6} \left(\frac{\text{molecules}}{\text{cell}} * \text{min} \right)^{-1}$
131	APCC_p27:Cdk2(M)_p27:Cyclin A:Cdk2(M)_Int -> p27:Cyclin A:Cdk2(M) + APCC	Mass Action	$k_{uAPCCcyclinA} = 0.1 \text{ (min)}^{-1}$
132	APCC_p27:Cdk2(M)_p27:Cyclin A:Cdk2(M)_Int -> p27:Cdk2(M) + APCC	Mass Action	$k_{udAPCCcyclinA} = 4.99955 \text{ (min)}^{-1}$
133	Cyclin A:Cdk1 + APCC -> APCC_Cdk1_Cyclin A:Cdk1_Int	Mass Action	$k_{bAPCCcyclinA} = 1.61 \cdot e^{-6} \left(\frac{\text{molecules}}{\text{cell}} * \text{min} \right)^{-1}$
134	APCC_Cdk1_Cyclin A:Cdk1_Int + APCC -> Cyclin A:Cdk1	Mass Action	$k_{uAPCCcyclinA} = 0.1 \text{ (min)}^{-1}$
135	APCC_Cdk1_Cyclin A:Cdk1_Int -> Cdk1 + APCC	Mass Action	$k_{udAPCCcyclinA} = 4.99955 \text{ (min)}^{-1}$
136	Cyclin A:Cdk1(M) + APCC -> APCC_Cdk1(M)_Cyclin A:Cdk1(M)_Int	Mass Action	$k_{bAPCCcyclinA} = 1.61 \cdot e^{-6} \left(\frac{\text{molecules}}{\text{cell}} * \text{min} \right)^{-1}$
137	APCC_Cdk1(M)_Cyclin A:Cdk1(M)_Int + APCC -> Cyclin A:Cdk1(M)	Mass Action	$k_{uAPCCcyclinA} = 0.1 \text{ (min)}^{-1}$
138	APCC_Cdk1(M)_Cyclin A:Cdk1(M)_Int -> Cdk1(M) + APCC	Mass Action	$k_{udAPCCcyclinA} = 4.99955 \text{ (min)}^{-1}$
Επιπλέον αντιδράσεις για την δοκιμή υποθέσεων			
139	-> Drug	Mass Action	r_{drug}
140	Drug + Cyclin D -> Drug:Cyclin D	Mass Action	$r_{drugBinding}$
			ModifierTime = 675.9932642 min

*Constant-flux: Σταθερή ροή, #Mass action: Νόμος δράσης των μαζών, (M): Ενεργοποιημένες οντότητες λόγω της ενεργοποίησης του τροποποιητή ενεργοποίησης

Πίνακας 3.3 Αρχικά επίπεδα βιοχημικών οντοτήτων που απαρτίζουν το προτεινόμενο μοντέλο

Όνομα Οντότητας	Αρχικά επίπεδα (<i>molecules/cell</i>)
APCC	24582.9
APCC_Cdk1_Cyclin A:Cdk1_Int	104.388
APCC_Cdk1(M)_Cyclin A:Cdk1(M)_Int	0
APCC_Cdk 2_Cyclin A:Cdk2_Int	52.8191
APCC_Cdk2(M)_Cyclin A:Cdk2(M)_Int	0
APCC_p27:Cdk2_p27:Cyclin A:Cdk2_Int	90.5091
APCC_p27:Cdk2(M)_p27:Cyclin A:Cdk2(M)_Int	0
APCC_CyclinA_Int	8.79462
APCC_Emi1	5160.61
Cdk1	98550.6
Cyclin A:Cdk1	1345.01
Cdk1(M)	0
Cyclin A:Cdk1(M)	0
Cyclin A:Cdk1(M)_hypo-pRb_hyper-pRb_Int	0
Cyclin A:Cdk1(M)_E2F:hypo-pRb_E2F:hyper-pRb_Int	0
Cdk 2	33942
Cyclin E:Cdk2	2176.46
Cyclin A:Cdk2	680.557
Cdk2(M)	0
Cyclin E:Cdk2(M)	0
Cyclin E:Cdk2(M)_hypo-pRb_hyper-pRb_Int	0
Cyclin E:Cdk2(M)_E2F:hypo-pRb_E2F:hyper-pRb_Int	0
Cyclin A:Cdk2(M)	0
Cyclin A:Cdk2(M)_hypo-pRb_hyper-pRb_Int	0
Cyclin A:Cdk2(M)_E2F:hypo-pRb_E2F:hyper-pRb_Int	0
p27:Cdk2	58162
p27:Cyclin E:Cdk2	3729.51
p27:Cyclin A:Cdk2	1166.18
p27:Cdk2(M)	0
p27:Cyclin E:Cdk2(M)	0
p27:Cyclin A:Cdk2(M)	0
Cdk 4	46551.9
Cyclin D:Cdk 4	6547.64
Cyclin D:Cdk 4_pRb_hypo-pRb_Int	0
Cyclin D:Cdk 4_E2F:pRb_E2F:hypo-pRb_Int	0
p27:Cdk 4	41117.2
p27:Cyclin D:Cdk 4	5783.23
Cyclin A	113.316
Cyclin D	19264
Cyclin E	191.985
E2F	546.211

Emi1	248.046
p27	14150.9
pRb	58583.4
E2F:pRb	1416.59
hypo-pRb	0
E2F:hypo-pRb	0
hyper-pRb	0
E2F:hyper-pRb	0
pseudo-hyper-pRb	0
E2F:pseudo-hyper-pRb	0
Cyclin D:Cdk 4_E2F:hypo-pRb_E2F:pseudo-hyper-pRb_Int	0
Cyclin D:Cdk 4_hypo-pRb_pseudo-hyper-pRb_Int	0
Cyclin E:Cdk2(M)_pseudo-hyper-pRb_hyper-pRb_Int	0
Cyclin E:Cdk2(M)_E2F:pseudo-hyper-pRb_E2F:hyper-pRb_Int	0
Cyclin A:Cdk2(M)_pseudo-hyper-pRb_hyper-pRb_Int	0
Cyclin A:Cdk2(M)_E2F:pseudo-hyper-pRb_E2F:hyper-pRb_In	0
Cyclin A:Cdk1(M)_pseudo-hyper-pRb_hyper-pRb_Int	0
Cyclin A:Cdk1(M)_E2F:pseudo-hyper-pRb_E2F:hyper-pRb_Int	0
Επιπλέον οντότητες για την δοκιμή υποθέσεων	
Drug	0
Drug_CyclinD	0



Εικόνα 3.4 Μέρος της δομής (νέες αντιδράσεις και οντότητες) του προτεινόμενου μοντέλου

Συνολικά, 10 νέες μοριακές οντότητες και 29 νέες βιοχημικές αντιδράσεις προστέθηκαν στις 49 μοριακές οντότητες και στις 138 αντιδράσεις του μοντέλου αναφοράς. Στην Εικόνα 3.4, οι νέο-εισαχθείσες μοριακές οντότητες είναι περικυκλωμένες από εντονότερο πλαίσιο ενώ οι νέες αντιδράσεις παρουσιάζονται με κόκκινες συνδέσεις. Το σύμβολο (M) δίπλα στο όνομα οντοτήτων ή συμπλόκων υποδεικνύει την ενεργοποίησή τους μέσω του τροποποιητή ενεργοποίησης.

Όπως φαίνεται στην εικόνα, η pRb μπορεί να βρεθεί σε τέσσερις διαφορετικές καταστάσεις φωσφορυλίωσης (phosphorylation states). Οι καταστάσεις αυτές είναι:

- Η μη φωσφορυλιωμένης (pRb)
- Η υπό-φωσφορυλιωμένης (hypo-pRb)
- Η υπέρ-φωσφορυλιωμένη (hyper-pRb)
- Η ψευδο-υπέρ-φωσφορυλιωμένη (pseudo-hyper-pRb)

Οι πρώτες τρεις καταστάσεις έχουν υιοθετηθεί από το μοντέλο αναφοράς. Ωστόσο η τελευταία ορίζεται de novo στα πλαίσια της διατριβής.

Προκειμένου να υλοποιηθούν οι μεταβάσεις μεταξύ των παλαιών και της νέο-οριζόμενης κατάστασης, σχηματίζονται επίσης ενδιάμεσα σύμπλοκα (complexes) μεταξύ των pRb, hypo-pRb ή pseudo-hyper-pRb και των συμπλόκων Cyclin D:Cdk4,6, Cyclin E:Cdk2, Cyclin A:Cdk2 ή Cyclin A:Cdk1 (Αντιδράσεις: r1,r3,r5,r23,r27,r31). Ωστόσο, οι μεταβάσεις στην επόμενη κατάσταση (Αντιδράσεις: r2, r4, r6, r36, r28, r32) πραγματοποιούνται μερικώς και ένα ποσοστό από αυτά τα σύμπλοκα διασπώνται αφήνοντας τις pRb, hypo-pRb ή την pseudo-hyper-pRb στην προηγούμενη τους κατάσταση (αντίστροφο σκέλος των αντιδράσεων: : r1, r3, r5, r35, r27, r31). Παρομοίως και παράλληλα, η μετάβαση μεταξύ των διαφόρων καταστάσεων φωσφορυλίωσης μπορεί να πραγματοποιηθεί και όταν οι E2F είναι προσδεμένοι στις διάφορες μορφές της πρωτεΐνης pRb (Αντιδράσεις: r9, r13, r17, r25, r29, r33 και r10, r14, r18, r26, r30, r34)

Στο κομμάτι του μοντέλου το οποίο υιοθετήθηκε από το μοντέλο αναφοράς, οι pRb και hypo-pRb μορφές μπορούν να δεσμεύσουν τους E2F παράγοντες μεταγραφής μέσω μία απλής αντίδρασης (r8, r11), σχηματίζοντας τα pRb:E2F και hypo-pRb:E2F σύμπλοκα αντίστοιχα. Αντιθέτως, η hyper-pRb δεν μπορεί να δεσμεύει de novo τον E2F και έτσι όποιες οντότητες E2F είναι δεσμευμένες σε hyper-pRb (hyper-pRb:E2F σύμπλοκα) ήδη πριν το βήμα υπέρ-φωσφορυλίωσης μπορούν μόνο να απελευθερωθούν (r16). Το συγκεκριμένο βήμα απελευθέρωσης, μπορεί επίσης να συμβεί και για τις pRb και hypo-pRb μορφές (αντίστροφο σκέλος των r8,r11). Ωστόσο, όπως σημειώθηκε προηγουμένως, αυτές οι μορφές τις pRb μπορούν να επαναδεσμεύουν τους E2F.

Για την περίπτωση της pseudo-hyper-pRb, αυτή η μορφή της πρωτεΐνης μπορεί να δεσμεύσει de novo τους ελεύθερους E2F, διαμορφώνοντας τα pseudo-hyper-pRb:E2F σύμπλοκα (αντίδραση r21). Τέλος, οι μεταφραστικοί παράγοντες E2F που είναι συνδεδεμένοι με οποιαδήποτε μορφή της pRb μπορούν να αποδομηθούν απευθείας (Αντιδράσεις: r7,r12,r15,r22).

Σχετικά τώρα με τον μαθηματικό ορισμό των νέων μερών του μοντέλου, επιλέχθηκαν κινητικοί νόμοι που βασίζονται στον νόμο δράσης των μαζών (mass action kinetic laws) ομοίως με την πρωτότυπη εργασία στην οποία ορίστηκε το μοντέλο αναφοράς. Επιπροσθέτως, αυτές οι νέες αντιδράσεις είναι εν γένει τροποποιημένες εκδόσεις υπάρχοντων αντιδράσεων στο μοντέλο αναφοράς. Σε αυτά τα πλαίσια, υποτέθηκε μία γραμμική σχέση μεταξύ των ρυθμών αντιδράσεων τους και αυτών των αντίστοιχων

αντιδράσεων στο μοντέλο αναφοράς. Για την υλοποίηση αυτής της σχέσης, ορίστηκαν συγκεκριμένες σταθερές αναλογίας. Ο Πίνακας 3.4 παρουσιάζει αυτές τις σταθερές.

Πίνακας 3.4 Μαθηματικός ορισμός των παραμέτρων-ρυθμών αντιδράσεων των νέο-εισαχθέντων αντιδράσεων

Αριθμός σχετικής πρωτότυπης αντίδρασης στην Εικόνα 3.4	Αριθμός νέας αντίδρασης στην Εικόνα 3.4	Ονομασία παραμέτρου στην πρωτότυπη αντίδραση*	Ορισμός νέας παραμέτρου
r15	r22	k_{dE2F}	$k_{dE2F_{leuk}} = pr_{dE2F} * k_{dE2F}$
r1	r5	k_{bD4pRb}	$k_{bD4pRb_{leuk}} = pr_{bD4} * k_{bD4pRb}$
r9	r17		
r1 (Αντίστροφη)	r5 (Αντίστροφη)	k_{uD4pRb}	$k_{uD4pRb_{leuk}} = pr_{uD4} * k_{uD4pRb}$
r9 (Αντίστροφη)	r17 (Αντίστροφη)		
r2	r6	$k_{upD4pRb}$	$k_{upD4pRb_{leuk}} = pr_{upD4} * k_{upD4pRb}$
r10	r18		
r24	r23	$k_{tpRbDephos}$	$k_{tpRbDephos_{leuk}} = pr_{tpRbDephos} * k_{tpRbDephos}$
r19	r20		
r16	r21	$k_{uE2FpRb}$	$k_{uE2FpRb_{leuk}} = pr_{uE2F} * k_{uE2FpRb}$
r11	r21 (Αντίστροφη)	$k_{bE2FpRb}$	$k_{bE2FpRb_{leuk}} = pr_{bE2F} * k_{bE2FpRb}$
r3	r35	k_{bE2pRb}	$k_{bE2pRb_{leuk}} = pr_{bE2A2A1} * k_{bE2pRb}$
r13	r25		
n.s.	r27	k_{bA2pRb}	$k_{bA2pRb_{leuk}} = pr_{bE2A2A1} * k_{bA2pRb}$
n.s.	r29		
n.s.	r31	k_{bA1pRb}	$k_{bA1pRb_{leuk}} = pr_{bE2A2A1} * k_{bA1pRb}$
n.s.	r33		
r3 (Αντίστροφη)	r35 (Αντίστροφη)	k_{uE2pRb}	$k_{uE2pRb_{leuk}} = pr_{uE2A2A1} * k_{uE2pRb}$
r13 (Αντίστροφη)	r25 (Αντίστροφη)		
n.s.	r27 (Αντίστροφη)	k_{uA2pRb}	$k_{uA2pRb_{leuk}} = pr_{uE2A2A1} * k_{uA2pRb}$
n.s.	r29 (Αντίστροφη)		
n.s.	r31 (Αντίστροφη)	k_{uA1pRb}	$k_{uA1pRb_{leuk}} = pr_{uE2A2A1} * k_{uA1pRb}$
n.s.	r33 (Αντίστροφη)		

r4	r36	$k_{upE2pRb}$	$k_{upE2pRbleuk} = pr_{upE2A2A1} * k_{upE2pRb}$
r14	r26		
n.s.	r28	$k_{upA2pRb}$	$k_{upA2pRbleuk} = pr_{upE2A2A1} * k_{upA2pRb}$
n.s.	r30		
n.s.	r32	$k_{upA1pRb}$	$k_{upA1pRbleuk} = pr_{upE2A2A1} * k_{upA1pRb}$
n.s.	r34		

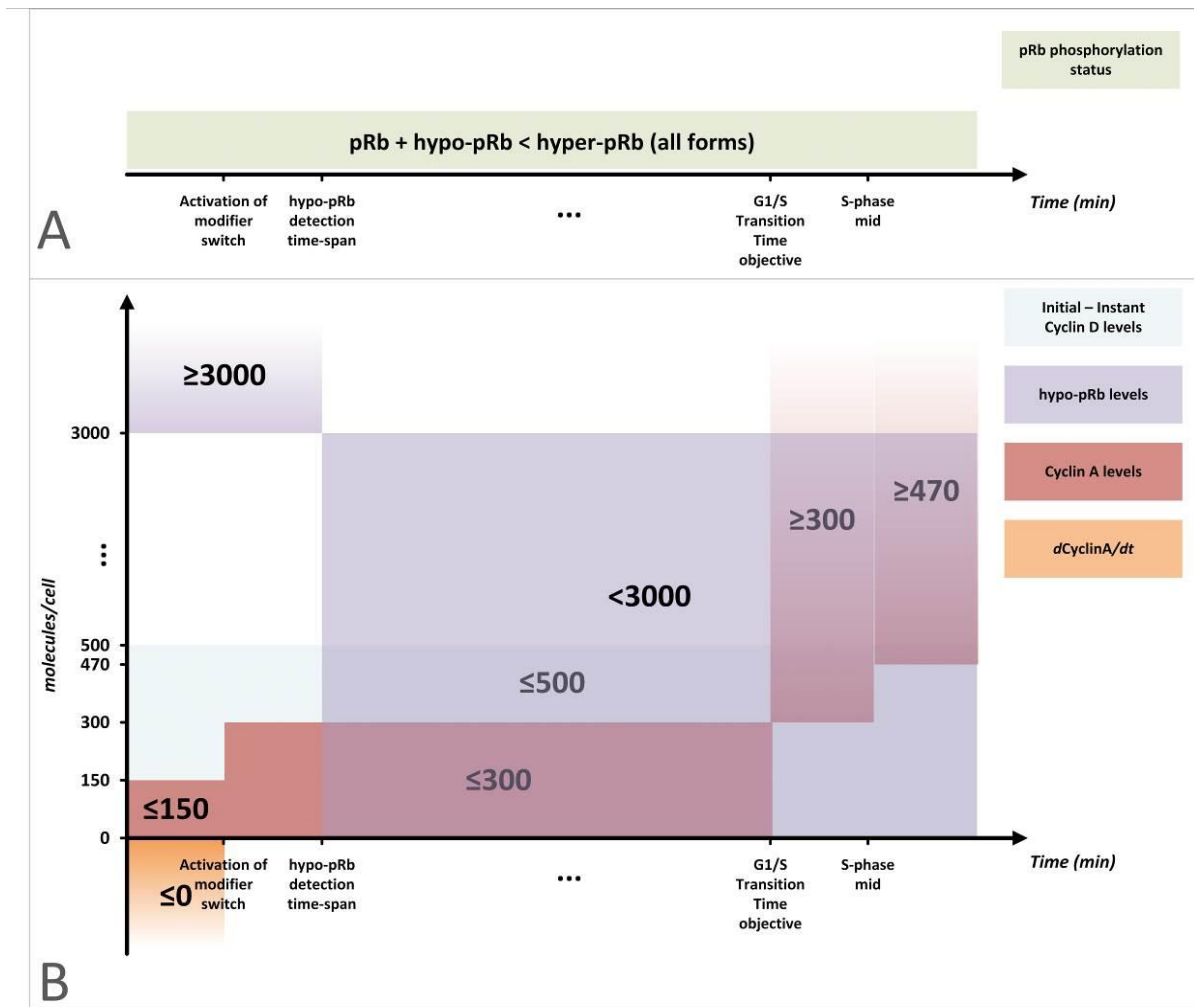
n.s.: δεν εμφανίζεται στην Εικόνα 3.4.

*Τα ονόματα των παραμέτρων για τις ήδη ορισμένες στο μοντέλο αναφοράς παραμέτρους παρέμειναν όπως ορίστηκαν στην [139]. Οι αντίστοιχες νέο-εισαχθείσες παράμετροι ονομάστηκαν παρομοίως αλλά με την προσθήκη του δείκτη *leuk*. Οι σταθερές αναλογίας (proportionality constants) ονομάζονται με τη χρήση του συμβόλου *pr* και ενός δείκτη με ονομασία σχετική με την αντίστοιχη αντίδραση.

Για την εισαγωγή των νέων αντιδράσεων και οντοτήτων στο μοντέλο αναφοράς, χρησιμοποιήθηκε το συμβατό με την SBML [186] εργαλείο COPASI [184] (Εκδόσεις 4.12, 4.13 και 4.14). Το μοντέλο αναφοράς λήφθηκε από την βάση μοντέλων BIOMODELS [183] χρησιμοποιώντας το μοναδικό κλειδί BIOMD0000000109 και επιβεβαιώθηκε χρησιμοποιώντας τις Συνήθειες Διαφορικές Εξισώσεις όπως δίνονται στην [139].

3.6 Στρατηγική προσαρμογής του προτεινόμενου μοντέλου και υλοποίησή της

Πριν την προσαρμογή του μοντέλου, οι στόχοι που τέθηκαν για την συγκεκριμένη διαδικασία σε σχέση με την επιθυμητή δυναμική συμπεριφορά του μοντέλου παρουσιάζονται στην συγκεκριμένη ενότητα. Οι στόχοι αυτοί ορίζονται κυρίως σε ημι-ποσοτικά/ποιοτικά πλαίσια λόγω της απουσίας ποσοτικών δεδομένων χρονοσειρών για τα επίπεδα των μοριακών οντοτήτων που μοντελοποιούνται. Τα κριτήρια αυτά παρουσιάζονται στην Εικόνα 3.5. Επίσης λεπτομερή παράθεση των κριτηρίων αυτών μαζί με την σχετική ποσοτική πληροφορία και της σχετικές πηγές παρουσιάζει ο Πίνακας 3.5. Βάσει των συγκεκριμένων κριτηρίων ορίστηκαν τμήματα συναρτήσεων κόστους (objective functions) και περιορισμοί βελτιστοποίησης όπως θα παρουσιαστεί σε επόμενη ενότητα. Επίσης στις ακόλουθες υπο-ενότητες γίνεται μία ποιοτική-συνοπτική παρουσίαση των κριτηρίων αλλά και λεπτομερής ορισμός τους.



Εικόνα 3.5 (Ημί-)ποσοτικά/ποιοτικά κριτήρια που ορίστηκαν για την κατάσταση και τα επίπεδα κεντρικών οντοτήτων του μοντέλου που χρησιμοποιήθηκαν για την προσαρμογή του προτεινόμενου μοντέλου μέσω βελτιστοποίησης

Πίνακας 3.5 (Ημί-)Ποσοτική πληροφορία εξαχθείσα από την βιβλιογραφία που χρησιμοποιήθηκε για την προσαρμογή του μοντέλου

Ιδιότητα		Τιμή	Πηγή
1. Επίπεδα φωσφορυλίωσης pRb			
1.α	Επικράτηση της hyper-pRb στα BCP-ALL κύτταρα	ποιοτική	<u>Συγκεκριμένη για την ALL:</u> [142]
1.β	hypo-pRb θετικά BCP-ALL κύτταρα	16% (\pm 13.2%)	
1.γ	Όριο ανίχνευσης hypo-pRb	3000 (<i>molecules/cell</i>)	<u>Γενική:</u> εξαχθείσα από την προσομοίωση του μοντέλου αναφοράς ([139] και υπό-ενότητα 3.6.2)
2. Κινητική κυττάρων (Cell Kinetics)			
2.α	Κύτταρα εντός του κυτταρικού κύκλου	94%	<u>Συγκεκριμένη για την ALL:</u> [142]

	(Cycling cells)		
2.β	Κύτταρα σε ηρεμία (Quiescent cells)	2%	<u>Συγκεκριμένη για την ALL:</u> [142], [155]
2.γ	Αποπτωτικά κύτταρα (Apoptotic cells)	4%	<u>Συγκεκριμένη για την ALL:</u> [162]
2.δ	Κύτταρα στις φάσεις S/G2/M	10%	<u>Συγκεκριμένη για την ALL:</u> [152]–[154]
2.ε	Κύτταρα στη φάση G1	84%	<u>Συγκεκριμένη για την ALL:</u> Εξάγεται από τις 2.α-δ (βλ. υπό-ενότητα 3.6.2)
2.ζ	Κύτταρα στην πρώιμη-G1 φάση (early G1-phase)	13.7%	<u>Συγκεκριμένη για την ALL:</u> Εξάγεται από τις 1.β. και 2.α. (βλ. υπό-ενότητα 3.6.2)
3. Χρονισμοί Κυτταρικού Κύκλου			
3.α	Μέση διάρκεια κυτταρικού κύκλου στα BCP-ALL κύτταρα (T_c)	112.5 h (± 46.8 h)*	<u>Συγκεκριμένη για την ALL:</u> [151]
3.β	Διάρκεια της S-φάσης στα BCP-ALL κύτταρα (T_s)	18 h	
3.γ	Διάρκεια G2-φάσης	6 h	<u>Γενική:</u>
3.δ	Διάρκεια M-φάσης	1 h	[181]
3.ε	Μέση διάρκεια της G1-φάσης στα BCP-ALL κύτταρα	87,5 (± 46.8 h)	Εξάγεται από 3.α-δ (βλ. υπό-ενότητα 3.6.2)
	Εκτιμώμενο ρεαλιστικό χρονικό διάστημα για την G1/S μετάβαση στην BCP-ALL	~2400 - ~8000 min	
3.ζ.	Η διάρκεια της πρώιμης-G1 φάσης στα BCP-ALL κύτταρα	~800 min	Εξάγεται από τις 2.ζ και 3.ε (βλ. υπό-ενότητα 3.6.2)
3.η.	Μέσω της S-φάσης	9 h μετά την μετάβαση G1/S	<u>Συγκεκριμένη για την ALL:</u> (βλ. υπό-ενότητα 3.6.2)
4. Επίπεδα Κυκλίνης A (Cyclin A)			
4.α	Επίπεδα για την μετάβαση στην S-φάση	300 (<i>molecules/cell</i>)	<u>Γενική:</u> εξαχθείσα από την προσομοίωση του μοντέλου αναφοράς ([139] και υπό-ενότητα 3.6.2)
4.β	Επίπεδα (ρυθμός) πριν την ενεργοποίηση του τροποποιητή ενεργοποίησης	<150 (<=0)	

4.γ	Επίπεδα (ρυθμός) μετά την ενεργοποίηση του τροποποιητή ενεργοποίησης	<300	
4.δ	Επίπεδα στο μέσο της S-φάσης (9 h)	470 (<i>molecules/cell</i>)	
5. Επίπεδα Κυκλίνης D (Cyclin D)			
	Μέγιστη διαφορά μεταξύ αρχικών και τρεχόντων-στιγμιαίων επιπέδων.	500 (<i>molecules/cell</i>)	Γενική: εξαχθείσα από την προσομοίωση του μοντέλου αναφοράς [139]

*Μέση τιμή (\pm τυπική απόκλιση)

3.6.1 Ποιοτική παρουσίαση κριτηρίων προσαρμογής

Με βάση τα οριζόμενα κριτήρια και όπως εξάγεται από την σχετική βιβλιογραφία που παρουσιάστηκε σε προηγούμενες ενότητες, σε γενικές γραμμές το προτεινόμενο μοντέλο θα πρέπει την ίδια στιγμή να προβλέπει την συμπεριφορά της Κυκλίνης A και των E2F να είναι ίδια με αυτή που παρατηρείται στο μοντέλο αναφορά όμως περαιτέρω εκτεινόμενη στο χρόνο λόγω της καθυστερημένης G1/S μετάβασης. Αντιθέτως, για τις υπό και υπέρ-φωσφορυλιωμένες μορφές της pRb η συμπεριφορά θα πρέπει να είναι ανεστραμμένη στις πρώτες ώρες της G1 φάσης, όπου η πρόιμη-G1 φάση κανονικά εμφανίζεται.

Σε αυτά τα πλαίσια, υπάρχουν 5 κατηγορίες κριτηρίων που πρέπει να καλυφθούν. Ξεκινώντας με την κατάσταση φωσφορυλίωσης της pRb (Εικόνα 3.5A και σύνολο ιδιοτήτων 1-Πίνακας 3.5), οι υπέρ-φωσφορυλιωμένες μορφές της pRb (pseudo-hyper-pRb και hyper-pRb) θα πρέπει εν γένει να επικρατούν των συνολικών επιπέδων αυτής της πρωτεΐνης. Ωστόσο, η υπο-φωσφορυλιωμένη pRb είναι επίσης ανιχνεύσιμη στα BCP-ALL κύτταρα (Εικόνα 3.5B και Πίνακας 3.5 ιδιότητα 1.β). Η παρουσία της hypo-pRb είναι ενδεικτική για την παρουσία των κυττάρων στην πρόιμη G1-φάση. Οπότε, θα μπορούσε να γίνει η υπόθεση ότι τα λευχαιμικά κύτταρα εμφανίζουν επίσης μία ομοιάζουσα με πρόιμη-G1 φάση (early-G1-like) κατά την διάρκεια των πρώτων ωρών του κυτταρικού του κύκλου. Σε αυτά τα πλαίσια, κατά την διαδικασία προσαρμογής του μοντέλου, συγκεκριμένοι περιορισμοί τέθηκαν έτσι ώστε τα επίπεδα της hypo-pRb να είναι τουλάχιστον ανώτερα των επιπέδων ανίχνευσης (Πίνακας 3.5 ιδιότητα 1.γ) για ένα προκαθορισμένο χρονικό διάστημα (Πίνακας 3.5 ιδιότητα 1.γ και hypo-pRb detection threshold Εικόνα 3.5B). Τα ανωτέρω ορίστηκαν με βάση την υπόθεση ότι η hypo-pRb είναι πλέον μη ανιχνεύσιμη όταν τα κύτταρα φεύγουν από την πρόιμη-G1 φάση.

Θα πρέπει να σημειωθεί πως τα παρότι τα επίπεδα ανιχνευσιμότητας που εμφανίζει ο Πίνακας 3.5 έχουν διαμορφωθεί με βάση πειράματα που πραγματοποιήθηκαν σε κύτταρα διαφορετικού τύπου (κυτταρική σειρά καρκινώματος μεγάλου εντέρου) [139], πιστεύεται πως αποτελούν μια ρεαλιστική προσέγγιση των επιπέδων ανίχνευσης των υπο-μελέτη πρωτεϊνών για κάθε παρόμοιο πείραμα.

Η κινητική των BCP-ALL κυττάρων (cell kinetics) (σύνολο ιδιοτήτων 2 Πίνακας 3.5) εξήχθη από την σχετική βιβλιογραφία προκειμένου να υπο-βοηθήσει την εκτίμηση των χρονισμών του κυτταρικού κύκλου (cell cycle timing, σύνολο ιδιοτήτων 3 Πίνακας 3.5) τους οποίους το μοντέλο θα πρέπει να αναπαράγει μετά την προσαρμογή του. Μεταξύ αυτών των χρονισμών, ξεχωρίζουν σε ενδιαφέρον:

- Η μέση διάρκεια της G1-φάσης στα BCP-ALL κύτταρα (Πίνακας 3.5, Ιδιότητα 3.ε)
- Το εκτιμώμενο ρεαλιστικό χρονικό διάστημα για την G1/S μετάβαση στην BCP-ALL
- Η διάρκεια της πρόιμης-G1 φάσης στα BCP-ALL κύτταρα (Πίνακας 3.5, Ιδιότητα 3.ζ)

Η τελευταία ιδιότητα χρησιμοποιείται ως ένας χρονικός περιορισμός για την ύπαρξη ανιχνεύσιμων επιπέδων της hypo-pRb, ενώ οι πρώτες δύο καθορίζουν τη χρονική στιγμή κατά την οποία η μετάβαση στην S φάση πραγματοποιείται. Έτσι, οι ιδιότητες αυτές παίζουν κεντρικό ρόλο στην προσαρμογή του μοντέλου.

Δεδομένου ότι τα επίπεδα της Κυκλίνης A συσχετίζονται με το ποσοστό των BCP-ALL κυττάρων στην S-φάση [144], ως κριτήριο για την G1/S μετάβαση τέθηκε η αύξηση της Κυκλίνης A σε συγκεκριμένα επίπεδα (Πίνακας 3.5, Ιδιότητα 4.α). Οπότε, προκειμένου να γίνει αποδεκτό ότι μία δοθείσα παραμετροποίηση του μοντέλου προβλέπει μία συγκεκριμένη διάρκεια για το κυτταρικό κύκλο, τα επίπεδα της Κυκλίνης A θα πρέπει να φτάσουν αυτό το προκαθορισμένο κατώφλι εντός του αναμενόμενου χρόνου για την G1/S μετάβαση (G1/S Transition Time objective στην Εικόνα 3.5). Επιπροσθέτως, γνωρίζοντας ότι η διαδικασία της G1/S μετάβασης έχει βρεθεί αναλλοίωτη, τουλάχιστον όσον αφορά τα σχετικά με την Κυκλίνη A φαινόμενα [144], τέθηκαν συγκεκριμένα κριτήρια προκειμένου να διασφαλιστεί ότι οι τάσεις της δυναμικής (dynamics) της Κυκλίνης A παραμένουν αναλλοίωτες σε σχέση με τις αντίστοιχες στο μοντέλο αναφοράς (σύνολο ιδιοτήτων 4, Πίνακας 3.5).

Ο τελευταίος στόχος που τέθηκε για την προσαρμογή του μοντέλου, αφορά στα επίπεδα της Κυκλίνης D. Οι χαμηλές τάσεις διακύμανσής της, αποτελούν κεντρικό

χαρακτηριστικό της ανάλυσης που έγινε κατά την ανάπτυξη του μοντέλου αναφοράς [139]. Έτσι, προστέθηκε ένα κριτήριο που αφορά την μέγιστη διαφορά μεταξύ των αρχικών και των στιγμιαίων επιπέδων της Κυκλίνης D, η οποία δεν θα πρέπει να ξεπερνά το όριο των 500 (*molecules/cell*). Η επιλογή της συγκεκριμένης τιμής επηρεάστηκε από την δυναμική συμπεριφορά της συγκεκριμένης πρωτεΐνης στο μοντέλο αναφοράς (Εικόνα 3.3).

3.6.2 Λεπτομερής παρουσίαση κριτηρίων προσαρμογής

Στην παρούσα ενότητα δίνεται μία πιο λεπτομερής παρουσίαση των κριτηρίων προσαρμογής του μοντέλου (Πίνακας 3.5, Εικόνα 3.3) η οποία παρατίθεται για λόγους πληρότητας του κειμένου.

Ιδιότητες 1 α-γ

Παρόλου που το μεγαλύτερο ποσοστό της pRb βρίσκεται σε υπερ-φωσφορυλιωμένη μορφή στα BCP-ALL κύτταρα (ιδιότητα 1.α), η υπο-φωσφορυλιωμένη μορφή είναι ανιχνεύσιμη σε αυτά τα κύτταρα και μάλιστα σε ένα σημαντικό ποσοστό αυτών ($16\% \pm 13.2\%$) [142] (ιδιότητα 1.β). Κάτι τέτοιο μπορεί να είναι ενδεικτικό του ότι για την μέση περίπτωση, αυτή η μορφή της pRb είναι παρούσα, τουλάχιστον σε ανιχνεύσιμα επίπεδα, για ένα μέρος της διάρκειας της G1-φάσης. Επιπλέον, η παρουσία της hypo-pRb είναι ενδεικτική για την παρουσία των κυττάρων στην πρώιμη G1-φάση. Με βάση τα ανωτέρω, γίνεται η υπόθεση ότι και τα λευχαιμικά κύτταρα παρουσιάζουν επίσης μία όμοια με πρώιμη G1 φάση κατά τις πρώτες ώρες του κυτταρικού τους κύκλου. Δυστυχώς, λόγω της γενικής απουσίας από τη βιβλιογραφία ποσοτικών και εκτεινόμενων στο χρόνο δεδομένων για τα ακριβή επίπεδα των διαφόρων μορφών της pRb στα BCP-ALL κύτταρα, η μορφή της πρωτεΐνης που επικρατεί κατά τη διάρκεια αυτής της υποτιθέμενης φάσης δεν μπορεί να διευκρινιστεί. Έτσι, κατά την διάρκεια της προσαρμογής του μοντέλου, συγκεκριμένη περιορισμοί τέθηκαν έτσι ώστε η hypo-pRb να είναι τουλάχιστον υψηλότερη σε συγκέντρωση από ένα κατώφλι ανίχνευσης και να μην εξαφανίζεται πλήρως πριν παρέλθει ένα προ-καθορισμένο χρονικό διάστημα (*hypo-pRb detection time span*, Εικόνα 3.3), του οποίου ο υπολογισμός παρουσιάζεται στην συνέχεια του κειμένου. Το κατώφλι ανίχνευσης για την hypo-pRb υπολογίστηκε μέσω προσομοίωσης του μοντέλου αναφοράς. Συγκεκριμένα, στα πειράματα που παρουσιάζονται στην [139], η hypo-pRb εξαφανίζεται 8 ώρες μετά την επανέναρξη του κυτταρικού κύκλου. Έτσι τα επίπεδα που επιστρέφει το μοντέλο αναφοράς για αυτό το χρονικό σημείο, υιοθετήθηκαν ως τα επίπεδα από τα οποία και πέρα δεν μπορεί να ανιχνευθεί η πρωτεΐνη (κατώφλι ανίχνευσης, περίπου 3000

(*molecules/cell*). Παρόλου που αυτό το κριτήριο διαμορφώθηκε με βάση πειράματα που πραγματοποιήθηκαν σε διαφορετικό κυτταρικό τύπο (κυτταρική σειρά καρκίνου του μεγάλου εντέρου), πιστεύεται ότι αποτελεί μία ρεαλιστική προσέγγιση για τα επίπεδα ανίχνευσης της *hypo-pRb* σε οποιοδήποτε παρόμοιο πείραμα.

Ιδιότητες 2ζ και 3α-η

Η μετάβαση από την *S* φάση θα πρέπει να προβλέπεται να συμβαίνει εντός συγκεκριμένου χρονικού διαστήματος. Το συγκεκριμένο χρονικό διάστημα καθορίζεται από την παρατηρούμενη διακύμανση στη διάρκεια του κυτταρικού κύκλου στα BCP-ALL κύτταρα. Παρόλο που υπάρχει διαθέσιμη ποσοτική πληροφορία για την διάρκεια του πλήρους κυτταρικού κύκλου αλλά και συγκεκριμένη της *S* φάσης σε αυτόν τον τύπο κυττάρου, μία συγκεκριμένη μέτρηση για την διάρκεια της *G1* φάσης δεν είναι διαθέσιμη στην βιβλιογραφία. Επομένως, με δεδομένο ότι το μοντέλο που επιλέχθηκε ως μοντέλο αναφοράς καλύπτει μόνο την εξέλιξη της *G1* φάσης και την μετάβαση στην *S* φάση, θα πρέπει να γίνει μία εκτίμηση της διάρκειας της *G1* φάσης (ή αντίστοιχα της χρονικής στιγμής όπου πραγματοποιείται η μετάβαση *G1/S*) με βάσης συγκεκριμένες παραδοχές για την διάρκεια των *M* και *G2* φάσεων. Κάτι τέτοιο είναι εφικτό με δεδομένου ότι υπάρχουν σαφής ενδείξεις πως, γενικά, η *G1* είναι η πιο ευμετάβλητη χρονικά φάση του κυτταρικού κύκλου και συμβάλει κεντρικά στην μεταβλητότητα της συνολικής διάρκειας του κυτταρικού κύκλου σε αντίθεση με τις *S*, *G2* και *M* φάσεις οι οποίες δεν εμφανίζουν σημαντικές χρονικές διακυμάνσεις [181]. Τα παραπάνω υποστηρίζεται περαιτέρω από το εύρημα ότι η διάρκεια της *S* φάσης στα BCP-ALL κύτταρα είναι ίση με αυτή των αντίστοιχων υγιών κυττάρων [151].

Προκειμένου να υπολογιστούν η μέση τιμή όπως επίσης και ένα ρεαλιστικό διάστημα για την χρονική στιγμή στην οποία η μετάβαση στην *S* φάση μπορεί να πραγματοποιηθεί στα BCP-ALL κύτταρα, λήφθηκαν υπόψη οι συγκεκριμένες τιμές για την μέση διάρκεια του κυτταρικού κύκλου και των σχετικών φάσεων στα BCP-ALL κύτταρα, όπως δίδονται στην σχετική βιβλιογραφία (ιδιότητες 3.α-3.δ). Έτσι, λαμβάνοντας υπόψη την μέση τιμή της T_c , η σχετική μέση διάρκεια της *G1* φάσης (ή αντίστοιχα η χρονική στιγμή στην οποία πραγματοποιείται η μετάβαση στην *S* φάση) μπορεί να εκτιμηθεί αφαιρώντας την διάρκεια των *S*, *G2* και *M* φάσεων από την πλήρη διάρκεια του κυτταρικού κύκλου (ιδιότητα 3.ε). Επιπλέον, προσθέτοντας ή αφαιρώντας την τιμή απόκλιση από την μέση τιμή της T_c , τα άκρα για το ρεαλιστικό διάστημα της διάρκειας της *G1* φάσης (ιδιότητα 3.ζ) μπορούν να εκτιμηθούν λαμβάνοντας υπόψη τις κυτταροκινητικές ιδιότητες

των BCP-ALL κυττάρων. Συγκεκριμένα, γνωρίζοντας ότι περίπου το 94% των BCP-ALL κυττάρων βρίσκονται εντός του κυτταρικού κύκλου, το 2% βρίσκονται σε ηρεμία και το 4% είναι αποπτωτικά [142] (ιδιότητες 2.α-γ) και αφαιρώντας το ποσοστό των κυττάρων στις S/G2/M φάσεις (το οποίο είναι περίπου 10%, ιδιότητα 2.δ), το ποσοστό των κυττάρων στην G1 φάση μπορεί να εκτιμηθεί στο 84% (ιδιότητα 2.ε), κάτι που συμφωνεί και με τις πληροφορίες που είναι διαθέσιμες στην βιβλιογραφία [142], [144], [151], [155]. Επιπροσθέτως, γνωρίζουμε ότι η *hypo-pRb* είναι ανιχνεύσιμη περίπου στο 16.3% των κυττάρων (ιδιότητα 1.β). Επομένως, δεχόμενοι ότι η *hypo-pRb* εμφανίζεται σχεδόν αποκλειστικά στην G1 φάση, το ποσοστό αυτό μπορεί να μετατραπεί, μέσω μιας απλής αναλογίας, στο ότι από το σύνολο των κυττάρων που βρίσκονται στην G1, το 13.7% βρίσκεται στην όμοια με πρόωμη-G1 φάση (ιδιότητα 2.ζ). Δεχόμενοι ότι τα κύτταρα είναι ανάλογα κατανομημένα στη διάρκεια της G1 φάσης, για τη μέση περίπτωση τα κύτταρα θα πρέπει να βρίσκονται στην όμοια με πρόωμη-G1 φάση για το 13.7% της διάρκειας της G1 φάσης, το οποίο έπειτα από την εφαρμογή απλής αναλογίας και στρογγυλοποίησης, υπολογίζεται να είναι 800 λεπτά (ιδιότητα 3.ζ).

Ιδιότητες 4α-4.δ

Το κριτήριο για την μετάβαση στην S φάση τέθηκε να είναι η αύξηση της Κυκλίνης A σε επίπεδα παρόμοια με αυτά που η συγκεκριμένη κυκλίνη λαμβάνει κατά την μετάβαση στην S φάση στο μοντέλο αναφοράς (300 Cyclin A *molecules/cell* στις 10 ώρες, ιδιότητα 4.α). Όπως σημειώθηκε και προηγουμένως, τα επίπεδα αυτά πιστεύεται πως αποτελούν μία ρεαλιστική προσέγγιση για τα επίπεδα της Κυκλίνης A που είναι απαιτητά για την μετάβαση στην S φάση, ασχέτως με τον κυτταρικό τύπο. Επιπλέον, η μετάβαση G1/S έχει βρεθεί αναλλοίωτη όσον αφορά τα σχετικά με την Κυκλίνη A φαινόμενα [144]. Επομένως, συγκεκριμένα κριτήρια έπρεπε να τεθούν προκειμένου να διασφαλιστεί ότι οι τάσεις της δυναμικής συμπεριφοράς της Κυκλίνης A θα παραμείνουν αναλλοίωτες σε σχέση με τα αποτελέσματα προσομοίωσης του μοντέλου αναφοράς. Συγκεκριμένα, όπως φαίνεται στην Εικόνα 3.3, τα επίπεδα της Κυκλίνης A είναι σταθερά ή ακόμα και μειώνονται μέχρι την ενεργοποίηση του τροποποιητή ενεργοποίησης (*activating modifier*). Επομένως για το προτεινόμενο μοντέλο, όπως φαίνεται και στην Εικόνα 3.5B, εισάχθηκε συγκεκριμένο κριτήριο προκειμένου τα επίπεδα της Κυκλίνης A να παραμείνουν χαμηλότερα από ένα συγκεκριμένο όριο (150 *molecules/cell*) για αυτό το χρονικό διάστημα. Επίσης, για να διασφαλιστεί ότι η Κυκλίνη A δεν θα εμφανίζει τάσεις αύξησης στο ίδιο χρονικό διάστημα, ο ρυθμός αύξησης της Κυκλίνης A περιορίστηκε έτσι ώστε να παραμείνει αρνητικός ή

τουλάχιστον ίσος με μηδέν (ιδιότητα 4.β). Μετά την ενεργοποίηση του activating modifier, τα επίπεδα της Κυκλίνης A θα πρέπει να ξεκινήσουν να αυξάνονται, ωστόσο να παραμείνουν χαμηλότερα από 300 *molecules/cell*, με δεδομένο ότι στα επίπεδα αυτά θα πρέπει να φτάσει μόνο όταν ο χρόνος για την μετάβαση G1/S έχει παρέλθει (ιδιότητα 4.γ). Τέλος, αφού τα επίπεδα της Κυκλίνης A περάσουν το κατώφλι της G1/S μετάβασης, θα πρέπει να συνεχίσουν να αυξάνονται, δεδομένου ότι ο συγκεκριμένος τύπος Κυκλίνης εκφράζεται κυρίως κατά τις S και G2 φάσεις του κυτταρικού κύκλου (κάτι που παρατηρείται και κατά την εκτέλεση του μοντέλου αναφοράς). Για τον ορισμό ενός σχετικού ημί-ποσοτικού κριτηρίου, λήφθηκε δείγμα για τα επίπεδα της Κυκλίνης A που επιστρέφονται από το μοντέλο αναφοράς στο μέσο της S φάσης (S-phase mid στην Εικόνα 3.5, όπου για την περίπτωση του μοντέλου αναφοράς ήταν 3 ώρες). Τα επίπεδα αυτά στο μοντέλο αναφοράς είναι 470 *molecules/cell* (ιδιότητα 4.δ) τα οποία και τίθενται ως στόχος για το προτεινόμενο μοντέλο. Ωστόσο στην περίπτωση του προτεινόμενου μοντέλου αυτό θα πρέπει να γίνει στις 9 ώρες μετά την μετάβαση στην S φάση (το μισό της διάρκειας της S φάσης στην BCP-ALL).

3.7 Προσαρμογή του προτεινόμενου μοντέλου

Για την προσαρμογή του μοντέλου χρησιμοποιήθηκε το εργαλείο COPASI και συγκεκριμένα η λειτουργία βελτιστοποίησης (optimization) που προσφέρει. Για τον έλεγχο της κάλυψης των κριτηρίων προσαρμογής, ορίστηκε μία συνάρτηση κόστους (Objective function) ακολουθώντας την μέθοδο της γραμμικής κλιμάκωσης (linear scalarization) για την πολύ-κριτηριακή βελτιστοποίηση (multi-objective optimization).

Για τον ορισμό της συνάρτησης κόστους χρησιμοποιήθηκε η λειτουργία «γεγονότων» (events) του COPASI. Συγκεκριμένα, τέσσερα διαφορετικά γεγονότα ορίστηκαν, τα οποία παρατηρώντας τον χρόνο της προσομοίωσης, αλλάζουν τιμές σε 7 διαφορετικούς σημαφόρους (flags). Οι σημαφόροι αυτοί αφορούν το ξεχωριστά χρονικά διαστήματα όπως δίνονται στην Εικόνα 3.5. Οι τιμές των σημαφόρων για κάθε χρονικό διάστημα δίνονται στον ακόλουθο πίνακα:

Πίνακας 3.6 Ορισμός των σημαφόρων προσαρμογής

Χρονικό Διάστημα	Σημαφόροι						
	<i>Pre_{Modifier}</i>	<i>Post_{Modifier}</i>	<i>Flag₀</i>	<i>Flag₁</i>	<i>Flag₂</i>	<i>Flag₃</i>	<i>Flag₄</i>
(<i>Time</i> ≥ 0) & (<i>Time</i> < <i>ModifierTime</i>)	1	0	1	0	0	0	0
(<i>Time</i> ≥ <i>ModifierTime</i>) & (<i>Time</i> < <i>hypo-pRb detection time span</i>)	0	1	0	1	0	0	0
(<i>Time</i> ≥ <i>hypo-pRb detection time span</i>) & (<i>Time</i> < <i>G1/S Transition Time Objective</i>)	0	1	0	0	1	0	0
(<i>Time</i> ≥ <i>G1/S Transition Time Objective</i>) & (<i>Time</i> < <i>S-phase-mid</i>)	0	1	0	0	0	1	0
<i>Time</i> ≥ <i>S-phase-mid</i>	0	1	0	0	0	0	1

Χρησιμοποιώντας τους ανωτέρω οριζόμενους σημαφόρους, ορίζονται συγκεκριμένες συναρτήσεις κόστους που αφορούν τις βιοχημικές οντότητες που λαμβάνονται υπόψη στα ημι-ποσοτικά/ποιοτικά κριτήρια. Οι συναρτήσεις αυτές δίνονται στους ακόλουθους πίνακες.

Πίνακας 3.7 Ορισμός τους κόστους προσαρμογής για τα επίπεδα της Κυκλίνης A και της Κυκλίνης D

Επίπεδα Κυκλίνης A (Cyclin A Levels)	
$CycA_{penalty_i} = \sum_{Time=0}^{Time=Simulation\ Time} CycA_{instant\ penalty_i}$ $i = 0,1,2,3,4$	$CycA_{instant\ penalty_0} = \begin{cases} CyclinA > 150 \text{ (molecules/cell)}, Flag_0 \\ 0 \end{cases}$
	$CycA_{instant\ penalty_1} = \begin{cases} CyclinA > 300 \text{ (molecules/cell)}, Flag_1 \\ 0 \end{cases}$
	$CycA_{instant\ penalty_2} = \begin{cases} CyclinA > 300 \text{ (molecules/cell)}, Flag_2 \\ 0 \end{cases}$
	$CycA_{instant\ penalty_3} = \begin{cases} CyclinA < 300 \text{ (molecules/cell)}, Flag_3 \\ 0 \end{cases}$
	$CycA_{instant\ penalty_4} = \begin{cases} CyclinA < 470 \text{ (molecules/cell)}, Flag_4 \\ 0 \end{cases}$

$CycA_{penalty_{sum}} = \sum_i CycA_{penalty_i}, i = 0,1,2,3,4$	
Ρυθμός Αύξησης Κυκλίνης A (Cyclin A Rate)	
$CycA Rate_{penalty}$ $= \sum_{Time=0}^{Time=Simulation Time} CycA Rate_{instant penalty_0}$	$CycA Rate_{instant penalty_0}$ $= \begin{cases} \frac{d CycA}{dt} > 0 ((molecules/cell)/min), Flag_0 \\ 0 \end{cases}$
Επίπεδα Κυκλίνης D (Cyclin D Levels)	
$CycD Var_{penalty}$ $= \sum_{Time=0}^{Time=Simulation Time} CycD Var_{instant penalty}$	$CycD Var_{instant penalty}$ $= \begin{cases} CycD_{initial levels} - CycD_{instant levels} > 500(molecules/cell), 1 \\ 0 \end{cases}$

Πίνακας 3.8 Ορισμός τους κόστους προσαρμογής για τα επίπεδα της pRb

Επίπεδα υπό-φωσφορυλιωμένης pRb (hypo-pRb Levels)	
$hypo-pRb_{penalty_i} = \sum_{Time=0}^{Time=Simulation Time} hypo-pRb_{instant penalty_i}$ $i = 0,1,2,3,4$	$hypo-pRb_{instant penalty_0}$ $= \begin{cases} hypo - pRb < hypo-pRb_{detection levels}, Flag_0 \\ 0 \end{cases}$
	$hypo-pRb_{instant penalty_1}$ $= \begin{cases} hypo - pRb \geq hypo-pRb_{detection levels}, Flag_1 \\ 0 \end{cases}$
	$hypo-pRb_{instant penalty_2}$ $= \begin{cases} hypo - pRb \geq hypo-pRb_{detection levels}, Flag_2 \\ 0 \end{cases}$
	$hypo-pRb_{instant penalty_3}$ $= \begin{cases} hypo - pRb \geq hypo-pRb_{detection levels}, Flag_3 \\ 0 \end{cases}$
	$hypo-pRb_{instant penalty_4}$ $= \begin{cases} hypo - pRb \geq hypo-pRb_{detection levels}, Flag_4 \\ 0 \end{cases}$
$hypo-pRb_{penalty_{sum}} = \sum_i hypo-pRb_{penalty_i}, i = 0,1,2,3,4$	
Επίπεδα pRb και υπό-φωσφορυλιωμένης pRb σε σύγκριση με τα επίπεδα της υπέρ-φωσφορυλιωμένης pRb (pRb + hypo-pRb) vs. hyper-pRb (all forms) levels	
$(pRb+hypo-pRb) vs. hyper-pRb_{penalty_i}$ $= \sum_{Time=0}^{Time=Simulation Time} (pRb+hypo-pRb) vs. hyper-pRb_{instant penalty_i}$ $i = 0,1$	$(pRb+hypo-pRb) vs. hyper-pRb_{instant penalty_0}$ $= \begin{cases} hyper-pRb (all forms) > pRb+hypo-pRb, Pre_Modifier, 0 \\ \end{cases}$
	$(pRb+hypo-pRb) vs. hyper-pRb_{instant penalty_1}$ $= \begin{cases} hyper-pRb (all forms) > pRb+hypo-pRb, Post_Modifier, 0 \\ \end{cases}$

Με βάση τις ανωτέρω συναρτήσεις κόστους, ορίζεται η συνολική συνάρτηση κόστους η οποία και χρησιμοποιήθηκε για την εκτίμηση παραμέτρων μέσω της ελαχιστοποίησης της τιμής της:

Objective Function (\mathbf{x}_j)

$$= CycA_{penalty_{sum}}(\mathbf{x}_j) + CycA\ Rate_{penalty}(\mathbf{x}_j) \\ + hypo-pRb_{penalty_{sum}}(\mathbf{x}_j) + CycD\ Var_{penalty}(\mathbf{x}_j)$$

Η οποία υπόκειται στους ακόλουθους περιορισμούς:

$$0 \leq (pRb+hypo-pRb)\ vs.\ hyper-pRb_{penalty_0} \leq 300$$

$$0 \leq (pRb+hypo-pRb)\ vs.\ hyper-pRb_{penalty_1} \leq 100$$

όπου \mathbf{x}_j μία παραμετροποίηση του μοντέλου.

Θα πρέπει να σημειωθεί ότι μία ακριβής μέτρηση για την μορφή της πρωτεΐνης pRb που επικρατεί στις διάφορες υπό-φάσεις της G1 δεν είναι διαθέσιμη στην γνωστή βιβλιογραφία. Επομένως τα άνω όριο στους παραπάνω περιορισμούς αποτελούν εμπειρικές εκτιμήσεις του μέγιστου χρόνου για τον οποίο η υπέρ-φωσφορυλιωμένη μορφή της pRb εμφανίζει χαμηλότερα επίπεδα σε σχέση με αυτά των μη- και υπό-φωσφορυλιωμένων μορφών.

Η ανωτέρω οριζόμενη συνάρτηση κόστους, εισήχθη στο εργαλείο COPASI και συγκεκριμένα στην λειτουργία βελτιστοποίησης (optimization task) και η τιμή της ελαχιστοποιήθηκε με τη χρήση του αλγορίθμου ολικής βελτιστοποίησης Particle Swarm [187], όπως αυτός παρουσιάστηκε και στο Κεφάλαιο 2, δεχόμενοι τις προκαθορισμένες τιμές για τον αλγόριθμο αυτόν.

Κατά την εκτίμηση παραμέτρων, για κάθε χρονική στιγμή όπου ένα από τα κριτήρια δεν καλύπτεται, μία μοναδιαία ποσότητα κόστους (penalty) προστίθεται στο συνολικό άθροισμα κόστους για το σύνολο του χρονικού πλαισίου προσομοίωσης και για μία δοθείσα παραμετροποίηση του μοντέλου (όπως αυτή έχει προκύψει από την αναζήτηση του αλγορίθμου βελτιστοποίησης στον χώρο παραμέτρων).

Οι παράμετροι οι οποίες επιλέχθηκαν να εκτιμηθούν (και άρα οι τιμές τους να είναι μεταβλητές κατά την διαδικασία βελτιστοποίησης-εκτίμησης παραμέτρων) δίνονται στον ακόλουθο πίνακα.

Πίνακας 3.9 Παράμετροι του μοντέλου των οποίων οι τιμές εκτιμήθηκαν και διαστήματα τιμών αυτών

Όνομα Παραμέτρου	Διάστημα τιμών στο χώρο παραμέτρων
pr_{bD4}	[0 – 1]
pr_{upD4}	[0 – 1]
$pr_{tpRbDephos}$	[0 – 1]
pr_{uE2F}	[0 – 1]
pr_{bE2F}	[0 – 2]
$pr_{bE2A2A1}$	[0 – 1]
$pr_{upE2A2A1}$	[0 – 1]
<i>ModifierTime</i>	120-780 min
$K_{sCyclinD}$	100-2000 ((molecules/cell)/min)
Αρχικά επίπεδα Cyclin D	10000-20000 (molecules/cell)

Όπως παρουσιάζει ο Πίνακας 3.9, η πλειοψηφία των νέο-οριζόμενων σταθερών αναλογίας (Πίνακας 3.4) επιλέχθηκε προς εκτίμηση. Μία εξαίρεση έγινε για την pr_{uE2F} δεδομένου ότι δεν υπάρχουν ενδείξεις στην γνωστή βιβλιογραφία για τροποποιημένη αποδόμηση του E2F στην BCP-ALL. Επομένως η τιμή της τέθηκε σταθερή και μοναδιαία. Παρομοίως οι pr_{uD4} και $pr_{uE2A2A1}$ τέθηκαν επίσης σταθερές και με μοναδιαία τιμή σε αντιστοιχία με τις k_{uD4pRb} και k_{uE2pRb} οι οποίες είχαν τεθεί σταθερές στο μοντέλο αναφορά. Στο συγκεκριμένο Πίνακα δίνονται επίσης τα διαστήματα τιμών για τις υπό εκτίμηση παραμέτρους. Επειδή εν γένει η προσομοίωση του μοντέλου θα πρέπει να οδηγεί σε μερική εμφάνιση των μη κανονικών/παθογενών φαινομένων (δεδομένης της συνύπαρξής τους με τα φυσιολογικά φαινόμενα του κυτταρικού κύκλου, όπως παρουσιάστηκε σε προηγούμενες υπό-ενότητες) οι σταθερές αναλογίας επιτράπηκε να λαμβάνουν τιμές στο διάστημα [0,1]. Επιπλέον, για την περίπτωση των pr_{uE2F} και $pr_{tpRbDephos}$, μόνο τιμές μικρότερες της μονάδας (ή τουλάχιστον ίσες με αυτή) μπορούν να εξηγήσουν την συνεχιζόμενη καταπίεση του E2F από την pRb (ακόμα και στην ψευδό-φωσφορυλιωμένη της μορφή) και να λάβουν υπόψη την υπόθεση για πιθανή αλλοίωση της διαδικασίας από-φωσφορυλίωσης της pRb στην BCP-ALL [152]. Ωστόσο, δεν μπορεί να αποκλειστεί ότι η σημαντική καταπίεση της έκφρασης της Κυκλίνης A (όπως καταδεικνύει η καθυστέρηση στην μετάβαση G1/S) είναι αποτέλεσμα όχι μόνο της ίσης με άλλες μορφές αλλά πιθανόν αυξημένης ικανότητας της ψευδό-υπέρ-φωσφορυλιωμένης μορφής της pRb να δεσμεύει de novo τον E2F. Έτσι έγινε μία εξαίρεση για την παράμετρο pr_{bE2F} , της οποίας η τιμή αναζητήθηκε στο διάστημα [0,2].

Σχετικά με την παράμετρο *ModifierTime*, στο μοντέλο αναφοράς η ενεργοποίηση του σχετικού μηχανισμού (τροποποιητής ενεργοποίησης), πραγματοποιείται στο τέλος της πρώιμης-G1 φάσης. Η χρονική στιγμή στην οποία συμβαίνει αυτό εκτιμήθηκε βάσει εργαστηριακών δεδομένων. Ωστόσο στην περίπτωση της BCP-ALL, όπως παρουσιάστηκε σε προηγούμενες ενότητες, η φυσιολογική πρώιμη-G1 φάση παραλείπεται ή τουλάχιστον μειώνεται σε χρονική διάρκεια, η διάρκεια της G1 φάσης παρατείνεται και ο μεταβολισμός της γλυκόζης έχει τροποποιηθεί και διαφέρει μεταξύ ασθενών. Με βάση αυτά, μπορεί να υποστηριχθεί μία υπόθεση ότι η στιγμή της ενεργοποίησης του μηχανισμού στην BCP-ALL μπορεί να είναι διαφορετική σε σχέση με αυτή στο μοντέλο αναφοράς. Επομένως, αυτή η παράμετρος τέθηκε προς εκτίμηση κατά την προσαρμογή του μοντέλου. Επιπλέον, μέσω της παρατήρησης της συμπεριφοράς και του μοντέλου αναφοράς αλλά και των πρώτων εκδόσεων του προτεινόμενου μοντέλου, η παράμετρος *ModifierTime* βρίσκεται να επηρεάζει κυρίως την χρονική στιγμή από την οποία και μετά η *hypo-pRb* γίνεται μη ανιχνεύσιμη. Έτσι, συμπεραίνεται ότι πιθανότατα η τιμή της παραμέτρου θα πρέπει να είναι μικρότερη από την εκτιμώμενη διάρκεια της όμοιας με πρώιμη-G1 φάσης. Σε αυτά τα πλαίσια, το διάστημα τιμών για την παράμετρο αυτή τέθηκε να είναι το $[120, \textit{hypo-pRb detection time span}) \textit{min}$ το οποίο για την μέση περίπτωση μπορεί να μεταφραστεί σε $[120, 800) \textit{min}$, όπως συζητήθηκε προηγουμένως. Για τις ανάγκες της εκτίμησης παραμέτρων το άνω άκρο διάστημα τροποποιήθηκε ελαφρώς (Πίνακας 3.9), προκειμένου η τιμή της παραμέτρου να παραμείνει μικρότερη από $800 \textit{min}$. Επίσης, η παράλειψη ή σμίκρυνση της πρώιμης G1 φάσης στην ALL θα μπορούσε να αποτελεί ένδειξη για ενεργοποίηση του τροποποιητή ενεργοποίησης ακόμα πιο νωρίς σε σχέση με το μοντέλο αναφοράς. Αυτό εξηγεί την επιλογή για σχετική μικρό κάτω άκρο στο ανωτέρω αναφερόμενο διάστημα έτσι ώστε να αναζητηθεί ένα αρκούντως μεγάλος χώρος παραμέτρων.

Προκειμένου τώρα να εξασφαλιστεί ότι το προτεινόμενο μοντέλο θα μπορέσει να καλύψει τα κριτήρια σχετικά με την Κυκλίνη D, ο ρυθμός της παραγωγής της ($k_{SCyclinD}$) όπως επίσης και τα αρχικά επίπεδα της επαν-εκτιμήθηκαν. Στα ανωτέρω συνηγορεί και η παρατήρηση ότι μεταξύ ασθενών με BCP-ALL, η Κυκλίνη D εμφανίζει διακυμάνσεις στην έκφραση [152].

Τέλος, σε σχέση με τα αρχικά επίπεδα των διάφορων μορφών της pRb, αποφασίστηκε (σε αντίθεση με το μοντέλο αναφοράς όπου τα κύτταρα από τα οποία παρήχθησαν τα δεδομένα είχαν συγχρονιστεί στην πρώιμη-G1 φάση) να εκκινεί το

προτεινόμενο μοντέλο με παρουσία μόνο μη-φωσφορυλιωμένων μορφών της pRb. Λαμβάνοντας υπόψη ότι η pRb από-φωσφορυλιώνεται μετά την μίτωση [188], [189], μία τέτοια επιλογή θεωρείται πως επιτρέπει στο προτεινόμενο μοντέλο να προσομοιώνει όλη τη χρονική έκταση της G1 φάσης. Επομένως, το διαθέσιμο σύνολο pRb οντοτήτων και pRb:E2F συμπλόκων διανεμήθηκαν αποκλειστικά σε μη-φωσφορυλιωμένες μορφές. Το αρχικά επίπεδα των υπόλοιπων οντοτήτων παρέμειναν ανεπηρέαστα σε σχέση με το μοντέλο αναφοράς (Πίνακας 3.3).

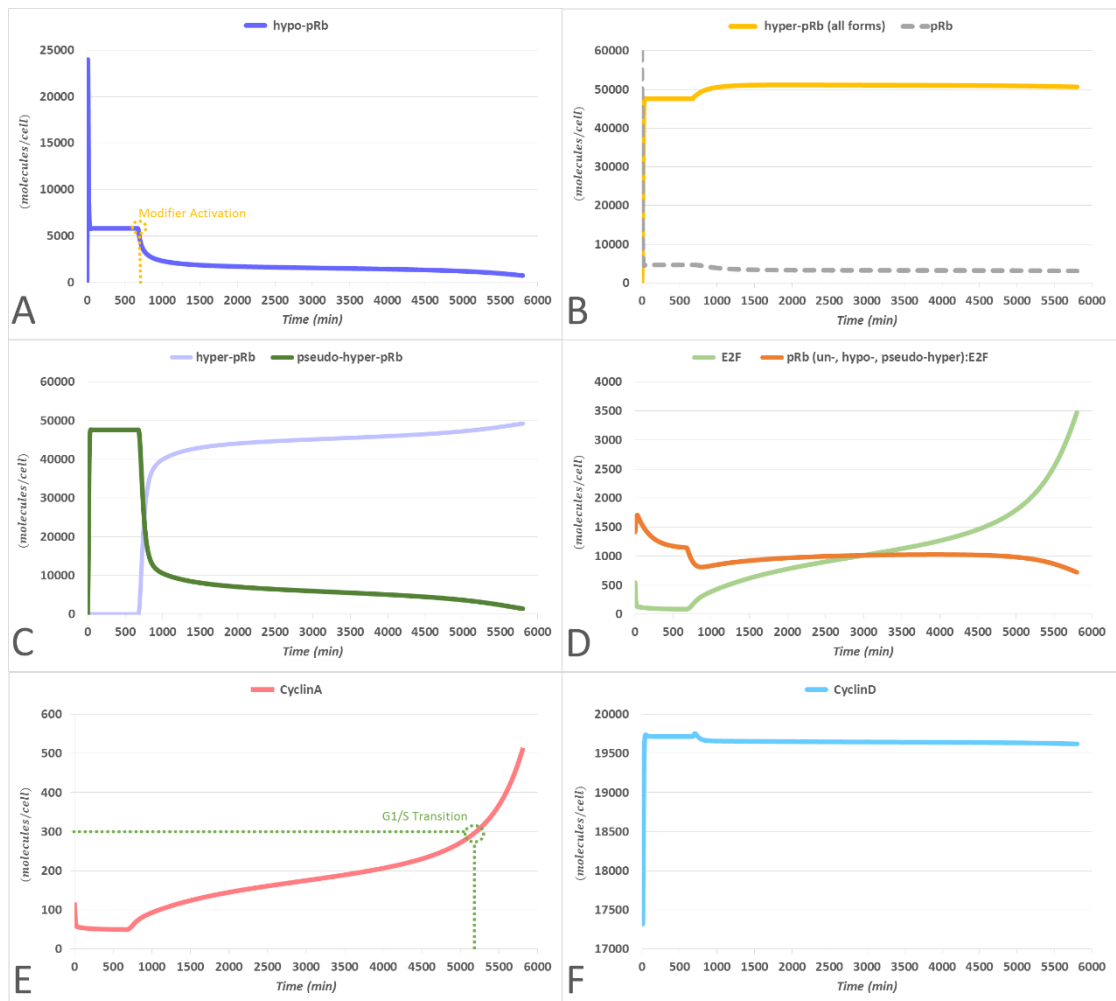
3.8 Αποτελέσματα προσαρμογής, δοκιμές εκτέλεσης και αξιολόγησης του μοντέλου

Εκτελώντας την διαδικασία ολικής στοχαστικής βελτιστοποίησης (ελαχιστοποίηση της συνάρτησης κόστους, δοθέντων των περιορισμών) υπολογίστηκε ένα σύνολο εκτιμήσεων για τις υπο-εκτίμηση παραμέτρους. Θα πρέπει να σημειωθεί ότι κατά την εκκίνηση του αλγορίθμου, οι παράμετροι λάμβαναν τυχαίες τιμές εντός του χώρου παραμέτρων που τις αφορούσε. Η τιμή 28.07651751 επετεύχθη για την τιμή της συνάρτησης κόστους δοθέντος του συνόλου αυτού. Οι τιμές του βέλτιστου (τουλάχιστον με βάση την διαδικασία ελαχιστοποίησης) συνόλου δίνονται στον ακόλουθο Πίνακα.

Πίνακας 3.10 Αποτελέσματα εκτίμησης παραμέτρων

Όνομα παραμέτρου	Εκτιμηθείσα τιμή
pr_{bD4}	0.841699
pr_{upD4}	0.484118
$pr_{tpRbDephos}$	1
pr_{uE2F}	0.612768
pr_{bE2F}	1.88407
$pr_{bE2A2A1}$	0.130145
$pr_{upE2A2A1}$	0.428204
<i>ModifierTime</i>	675.993 <i>min</i>
$K_{sCyclinD}$	1761.08 ((<i>molecules/cell</i>)/ <i>min</i>)
Αρχικά επίπεδα Cyclin D	19264 (<i>molecules/cell</i>)

Τα αποτελέσματα προσομοίωσης του προτεινόμενου μοντέλου με το συγκεκριμένο σύνολο παραμέτρων για 6000 *min* (αρκετές ώρες μετά την στοχευόμενη χρονική στιγμή για την μετάβαση G1/S για την μέση περίπτωση) δίνονται στην Εικόνα 3.6



Εικόνα 3.6 Αποτελέσματα εκτέλεσης του προτεινόμενου μοντέλου μετά την εκτίμηση των παραμέτρων του και για την μέση περίπτωση

Όπως παρουσιάζεται στην Εικόνα 3.6Α και στην Εικόνα 3.6Β, σε αντίθεση με τα αντίστοιχα αποτελέσματα του μοντέλου αναφοράς (Εικόνα 3.3), οι υπέρ-φωσφορυλιωμένες μορφές της pRb (Εικόνα 3.6Β, κίτρινο) ξεκάθαρα επικρατούν των συνολικών επιπέδων της pRb και εμφανίζουν σημαντικά υψηλότερα επίπεδα σε σχέση με την μη φωσφορυλιωμένη μορφή (pRb, Εικόνα 3.6Β, γκρι) της οποίας τα αρχικά αντίγραφα καταναλώνονται άμεσα. Επιπροσθέτως, τα επίπεδα της hypo-pRb (Εικόνα 3.6Α, μωβ) παρόλο που στην έναρξη της G1 εμφανίζουν θεαματική αλλά πρόσκαιρη αύξηση, ακολούθως εμφανίζονται σημαντικά μόνο για ένα περιορισμένο χρονικό διάστημα, σε συμφωνία με το κριτήριο που τέθηκε για την διάρκεια της όμοιας με πρώιμη-G1 φάσης.

Σχετικά τώρα, με τα επίπεδα της ψευδο-υπέρ-φωσφορυλιωμένης pRb (pseudo-hyper-pRb, Εικόνα 3.6C σκούρο πράσινο) αυτά αυξάνονται σημαντικά από τις πρώτες ώρες τις G1 φάσης και αποκλειστικά εκπροσωπούν τις υπέρ-φωσφορυλιωμένες μορφές της pRb μέχρι την ενεργοποίηση του τροποποιητής ενεργοποίησης από όπου και μετά

επικρατούν τα επίπεδα της υπέρ-φωσφορυλιωμένης pRb (Εικόνα 3.6C, hyper-pRb, ανοιχτό μωβ).

Όσον αφορά τα επίπεδα των ελεύθερων E2F παραγόντων μεταγραφής (Εικόνα 3.6D πράσινο), αυτά συναντώνται σε σημαντικά επίπεδα μόνο μετά την ενεργοποίηση του τροποποιητή ενεργοποίησης. Ωστόσο, αρκετά σημαντικά επίπεδα της πρωτεΐνης βρίσκονται δεσμευμένα σε μορφές της pRb που διατηρούν αυτή την ιδιότητα για σημαντικό χρονικό διάστημα (Εικόνα 3.6D, πορτοκαλί).

Η συμπεριφορά των επιπέδων της CyclinA (Εικόνα 3.6E, κόκκινο), είναι σε συμφωνία με τα κριτήρια που τέθηκαν, παρουσιάζοντας φθίνουσα ή σταθερή τάση για τις πρώτες ώρες της προσομοίωσης και αύξουσα μέχρι το τέλος αυτής. Επίσης, φτάνει το απαιτούμενο όριο για την G1/S μετάβαση στα 5200 min.

Τέλος η CyclinD (Εικόνα 3.6F,γαλάζιο) εμφανίζει ασήμαντη διακύμανση στα επίπεδα της για όλη τη διάρκεια της προσομοίωσης.

Εμβαθύνοντας στα αποτελέσματα, ένα ιδιαίτερα ενδιαφέρον συμπέρασμα εξάγεται από τα επίπεδα των ψευδό-υπέρ-φωσφορυλιωμένης και υπέρ-φωσφορυλιωμένης pRb (pseudo-hyper-pRb και hyper-pRb αντίστοιχα) όπως φαίνεται στην Εικόνα 3.6C. Όταν τα επίπεδα των μορφών αυτών συγκριθούν με τα αντίστοιχα των hypo-pRb (υπό-φωσφορυλιωμένη) και hyper-pRb στην Εικόνα 3.3, παρατηρείται μία ανάλογη εικόνα. Στο προτεινόμενο μοντέλο, όμως, η hypo-pRb έχει αντικατασταθεί από την pseudo-hyper-pRb. Αντιθέτως όμως, λόγω της διατήρησης της ικανότητας της pseudo-hyper-pRb να καταπιέζει τους E2F (Εικόνα 3.6D) και παρόλο που η hyper-pRb εκφράζεται άμεσα και άφθονα μετά την ενεργοποίηση του τροποποιητή ενεργοποίησης, η έκφραση της Cyclin A (Εικόνα 3.6E) παρουσιάζει σημαντικά χαμηλότερο ρυθμό, τουλάχιστον για τα πρώτα 4000 min (~65 h). Μόνο όταν απελευθερωθούν σημαντικά επίπεδα του E2F, τα επίπεδα της κυκλίνης αυτής παρουσιάζουν αυξητικές τάσεις οι οποίες αποτελούν ένδειξη και ευνοούν την μετάβαση στην S φάση.

Επομένως, τα αποτελέσματα προσομοίωσης θεωρούνται ως σύμφωνα με την παρατηρούμενη δυναμική συμπεριφορά του κυτταρικού κύκλου (cell cycle dynamics) στην BCP-ALL και με τη μέση χρονική δυναμική (temporal dynamics) για την μετάβαση στη S-φάση. Με βάση τα αποτελέσματα αυτά, το συμπέρασμα που θα μπορούσε να εξαχθεί είναι ότι ο μηχανισμός του RP έχει τροποποιηθεί έτσι στην περίπτωση της BCP-ALL, ώστε η καταπίεση των E2F παραγόντων μεταγραφής να πραγματοποιείται από μία πιο φωσφορυλιωμένη μορφή της pRb. Ωστόσο, η καταπίεση αυτή λαμβάνει χώρα για ένα σημαντικό χρονικό διάστημα, παρόμοια με το αντίστοιχο στην φυσιολογική συμπεριφορά

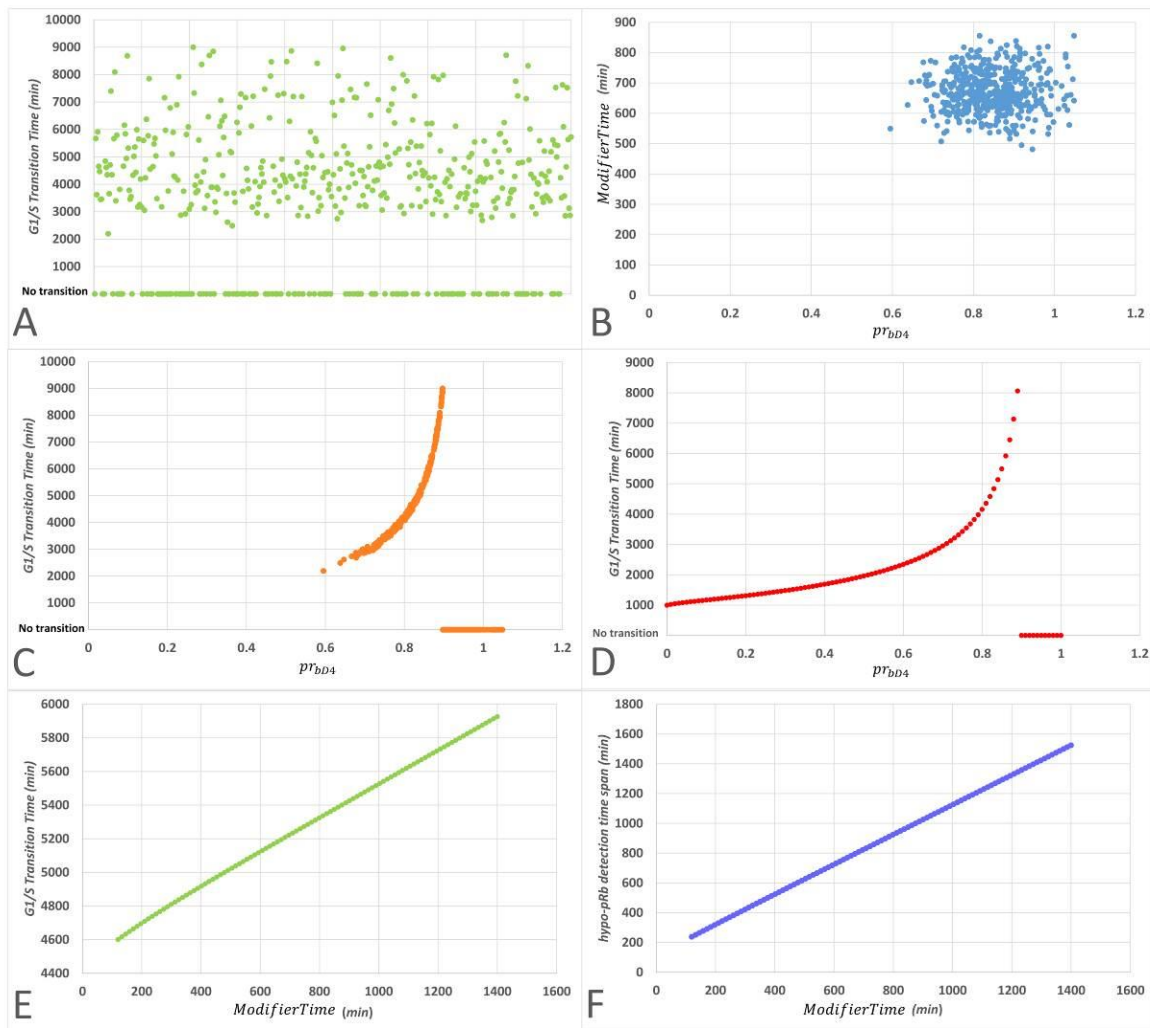
του RP. Όμως, ο ρόλος των μη- και υπό-φωσφορυλιωμένων μορφών της pRb, οι οποίες ευνοούν την διαδικασία διαφοροποίησης-ωρίμανσης του κυττάρου, έχει σχεδόν εξαλειφθεί. Αυτό το εύρημα, συμφωνεί με την αντίσταση στη διαφοροποίηση που εμφανίζουν τα λευχαιμικά κύτταρα.

3.9 Δοκιμή υποθέσεων

Τελευταίο βήμα της ανάλυσης του μοντέλου αποτέλεσε η δοκιμή του για έναν αριθμό επιπλέον σεναρίων και υποθέσεων με βασικό στόχο να μελετηθεί η συμφωνία του με επιπλέον ευρήματα και να παρουσιαστεί η πιθανή εφαρμογή του ως προβλεπτικό εργαλείο.

Ως πρώτο βήμα ακολουθήθηκε η διαδικασία σάρωσης παραμέτρων (parameter scan) χρησιμοποιώντας την αντίστοιχη λειτουργία του εργαλείου COPASI. Οι στόχοι για την συγκεκριμένη διαδικασία ήταν να διαπιστωθεί κατά πόσο το μοντέλο μπορεί να προβλέψει η G1/S μετάβαση να συμβαίνει εντός ενός ρεαλιστικού πλαισίου για την BCP-ALL (Πίνακας 3.5) όταν διαταραχθούν οι τιμές κάποιων παραμέτρων του και κατά πόσο είναι ικανό να εξηγήσει την παρατηρούμενη σημαντική αύξηση στην τυπική απόκλιση της T_c στην BCP-ALL. Οι παράμετροι που επιλέχθηκε να διαταραχθούν ήταν η pr_{bD4} σταθερά αναλογίας και η *ModifierTime*. Όπως συζητήθηκε και σε προηγούμενες ενότητες, οι παράμετροι αυτοί σχετίζονται με δύο φαινόμενα τα οποία εμφανίζουν διακυμάνσεις μεταξύ ασθενών με ALL. Το πρώτο, αφορά στο βαθμό με το οποίο η Cdk4 συμμετέχει στην ψευδο-υπερ-φωσφορυλίωση της pRb ενώ το δεύτερο στον μεταβολικό ρυθμό του κυττάρου.

Για την παρούσα διαδικασία σάρωσης παραμέτρων, πραγματοποιήθηκαν 500 τυχαία βήματα δειγματοληψίας τιμών των παραμέτρων από δύο κανονικές κατανομές (μία για κάθε παράμετρο). Ως μέσες τιμές των κατανομών αυτών τέθηκαν οι τιμές που εκτιμήθηκαν κατά την προσαρμογή του μοντέλου (Πίνακας 3.10), ενώ ως τυπική απόκλιση τέθηκε το 10% των αντίστοιχων μέσων τιμών. Για κάθε βήμα δειγματοληψίας, πραγματοποιήθηκε προσομοίωση 9000 *min* (150 *h*), δηλαδή για αρκετές ώρες μετά την μέγιστη διάρκεια για την G1 φάση όπως αυτή υπολογίζεται σε προηγούμενη υπό-ενότητα (Πίνακας 3.5). Τα αποτελέσματα της διαδικασίας παρουσιάζονται στην ακόλουθη εικόνα.



Εικόνα 3.7 Αποτελέσματα σάρωσης παραμέτρων

Όπως φαίνεται στην Εικόνα 3.7Α, δειγματοληπτώντας τις τιμές παραμέτρων, η μετάβαση στην S-φάση προβλέπεται να λαμβάνει χώρα σε ένα ευρύ φάσμα χρονικών στιγμών το οποίο συμφωνεί σημαντικά με το διάστημα τιμών που παρουσιάζει ο Πίνακας 3.5. Τα δείγματα που λήφθηκαν από τον δι-διάστατο χώρο παραμέτρων δίνονται στην Εικόνα 3.7Β. Για έναν αριθμό προσομοιώσεων, ωστόσο, προβλέπεται αδυναμία μετάβασης στην S-φάση, τουλάχιστον πριν το όριο των 9000 *min*. Οι περιπτώσεις αυτές παρουσιάζονται στην Εικόνα 3.7Α με μηδενικό χρόνο για την G1/S μετάβαση. Αυτό είναι αποτέλεσμα υψηλών τιμών για την παράμετρο pr_{bD4} , όπως είναι εμφανές βλέποντας την Εικόνα 3.7C (τα αποτελέσματα της οποίας παρήχθησαν από το ίδιο *in silico* πείραμα).

Προκειμένου να μελετηθεί η ανεξάρτητη επιρροή των δύο παραμέτρων στο σύστημα, δύο επιπλέον διαδικασίες σάρωσης παραμέτρων πραγματοποιήθηκαν, μία για κάθε μία από τις υπό-διερεύνηση παραμέτρους, κρατώντας την τιμή της άλλης παραμέτρου σταθερή. Στην Εικόνα 3.7D, ένα σύνολο 100 προσομοιώσεων παρουσιάζονται για

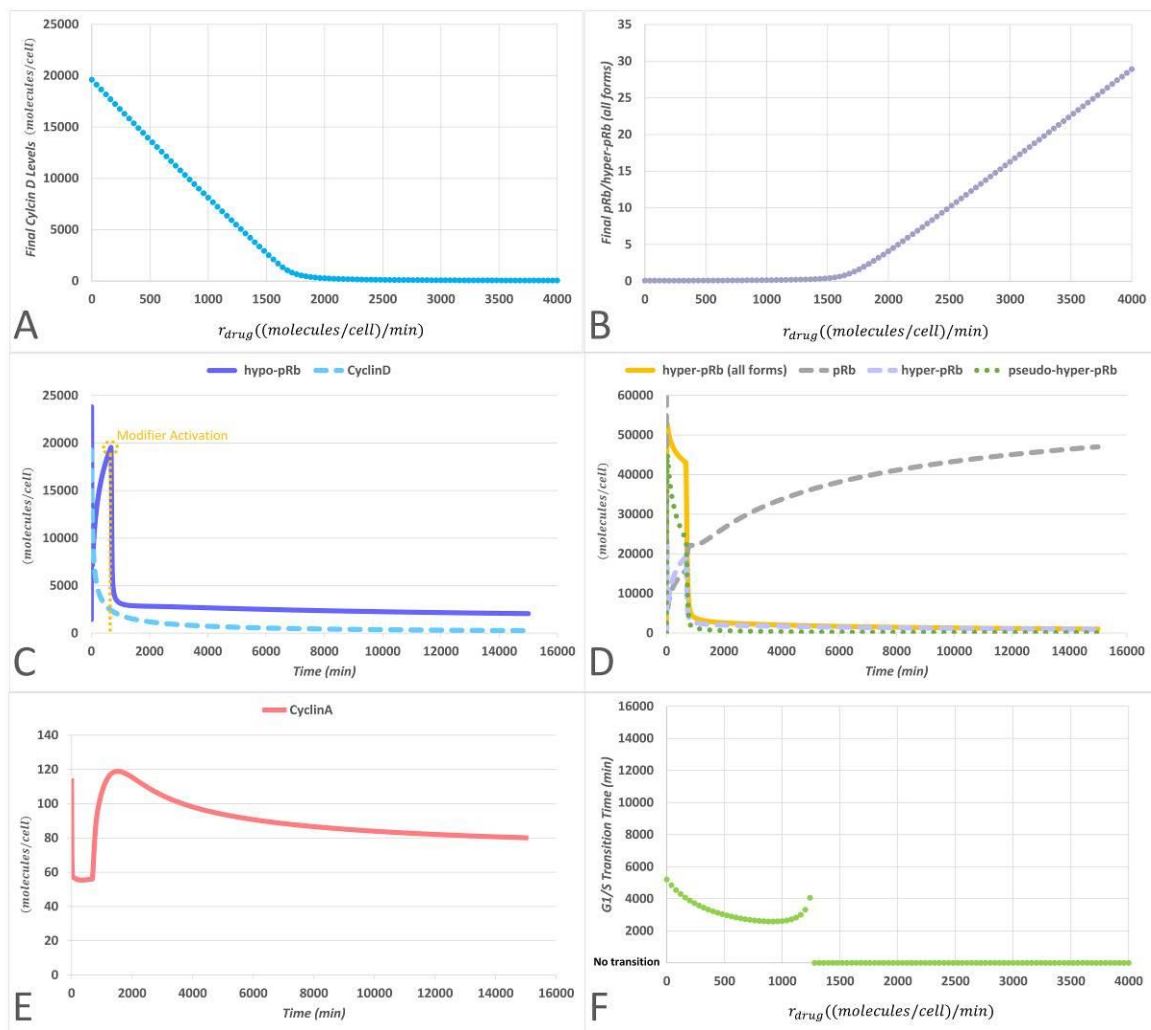
διαφορετικές τιμές της pr_{bD4} , εντός του διαστήματος που ορίζει ο Πίνακας 3.9 για την συγκεκριμένη παράμετρο. Όπως φαίνεται και στην σχετικό εικόνα, η σχέση μεταξύ της τιμής της παραμέτρου και του χρόνου μετάβασης G1/S παρουσιάζει αύξουσα δυναμική η οποία ωστόσο προοδευτικά κορένεται.

Όσον αφορά την σχέση μεταξύ *ModifierTime* και του χρόνου για την G1/S μετάβαση ή του χρονικού διαστήματος όπου η hypo-pRb παραμένει ανιχνεύσιμη (hypo-pRb detection time span), ξεκάθαρα γραμμικές σχέσεις παρατηρούνται στις σχετικές εικόνες (Εικόνα 3.7E,F), όπως αναμενόταν. Οι συγκεκριμένες εικόνες παρήχθησαν μέσω μια διαδικασίας σάρωσης παραμέτρων (100 προσομοιώσεις) για την *ModifierTime*, στο διάστημα [120, 1400] min. Το συγκεκριμένο διάστημα αποφασίστηκε έχοντας κατά νου το αντίστοιχο διάστημα για τα hypo-pRb θετικά κύτταρα ($16\% \pm 13.2\%$) σε πληθυσμούς ασθενών με ALL (Πίνακας 3.5). Το διάστημα αυτό ξεκινάει από σχεδόν πλήρη απουσία της hypo-pRb και καταλήγει σε διπλασιασμό του αριθμού των κυττάρων σε σχέση με την μέση περίπτωση, για την οποία η τιμή της *ModifierTime* εκτιμήθηκε να είναι ~ 675 min κατά την προσαρμογή του μοντέλου. Υποθέτοντας εκ νέου ότι τα κύτταρα είναι ομοιόμορφα κατανομημένα στις διάφορες φάσεις ή υπο-φάσεις του κυτταρικού κύκλου (με βάση πάντα την διάρκεια αυτών), διπλασιάζοντας την τιμή της παραμέτρου για τη μέση περίπτωση προκύπτει η τιμή για το άνω άκρο.

Με βάση τα αποτελέσματα, θα πρέπει να σημειωθεί πως παρόλο που η *ModifierTime* μπορεί να αλλάξει σημαντικά τη χρονική στιγμή στην οποία πραγματοποιείται η G1/S μετάβαση, αυτή η παράμετρος από μόνης της δεν είναι ικανή να αναπαράγει την παρατηρούμενη διακύμανση της T_c στην BCP-ALL, τουλάχιστον εντός του διαστήματος που δοκιμάστηκε. Επιπλέον, θα πρέπει να δοθεί προσοχή στο ότι αυτό το διάστημα δεν μπορεί να αυξηθεί σημαντικά, έχοντας κατά νου το ποσοστό των hypo-pRb θετικών κυττάρων, όπως εξηγήθηκε στην προηγούμενη παράγραφο. Αντιθέτως, για την διαταραχή της pr_{bD4} , υφίσταται μία μη γραμμική σχέση μεταξύ της τιμής της και του χρόνου της G1/S μετάβασης. Αυτό οδηγεί σε σημαντική διαφοροποίηση των αποτελεσμάτων προσομοίωσης που καλύπτουν όλο το εκτιμώμενο φάσμα για την διάρκεια της G1 φάσης. Οπότε, παρόλο που η συνδυασμένη διαταραχή των δύο παραμέτρων μπορεί να εξηγήσει την σημαντική μεταβλητότητα της T_c , αυτό κυρίως οφείλεται στην επιρροή της pr_{bD4} και στα σχετικά ακόλουθα φαινόμενα του σχηματισμού ψευδό-υπερφωσφορυλιωμένης pRb. Όπως έχει ήδη αναφερθεί, το φαινόμενο της τροποποίησης των προτιμήσεων υποστρώματος (substrate specificities) των Cdk4,6 (στο οποίο βασίζεται η

ψευδό-υπέρ-φωσφορυλίωση της pRb), συμβαίνει σε διαφορετικό βαθμό μεταξύ των ασθενών με ALL. Επομένως, μπορεί να εξαχθεί το συμπέρασμα ότι η υιοθέτηση των προτεινόμενων τροποποιήσεων του συστήματος (RP) όχι μόνο επιτρέπει την προσομοίωση της παραλλαγμένης δυναμικής συμπεριφοράς του βιολογικού μονοπατιού του RP, αλλά επίσης δείχνει ότι αυτές θα μπορούσαν να αποτελέσουν εξήγηση για την αυξημένη διακύμανση της T_c στην BCP-ALL.

Το δεύτερο σενάριο το οποίο υλοποιήθηκε, αφορούσε στις πιθανές συνέπειες της χορήγησης ενός φαρμάκου με κυτταροστατική/αντί-πολλαπλασιαστική (anti-proliferative) δράση. Ένα ευρέως χρησιμοποιούμε φάρμακο αυτής της κατηγορίας στην ALL είναι τα Γλυκοκορτικοειδή (Glucocorticoids), κυρίως τα Πρεδνιζόνη (Prednisone) και Δεξαμεθασόνη (Dexamethasone), των οποίων η χορήγηση οδηγεί σε παύση του κυτταρικού κύκλου στην φάση G1 (cell cycle arrest) και αποπτωτικό θάνατο των λευχαιμικών κυττάρων [32]. Σημειώνεται πως η πιθανή συσχέτιση μεταξύ των δύο φαινομένων δεν έχει ακόμα πλήρως αποκαλυφθεί. Ωστόσο, σημαντικά ευρήματα μπορούν να εξαχθούν από την βιβλιογραφία σχετικά με τον μηχανισμό δράσης αυτών των φαρμάκων. Ειδικότερα για το φαινόμενο της παύσης του κυτταρικού κύκλου, έχει δειχθεί ότι οι κυτταροστατικές (cytostatic) ιδιότητες των γλυκοκορτικοειδών κυρίως εμφανίζονται μέσω της καταπίεσης της Κυκλίνης D και διαμέσου της αποφωσφορυλίωσης της pRb [190]–[194]. Προκειμένου να καταστεί το μοντέλο ικανό να προσομοιώσει αυτού του είδους την παρέμβαση με τον απλούστερο δυνατό τρόπο, δύο νέες αντιδράσεις και δύο νέες οντότητες προστέθηκαν σε αυτό όπως παρουσιάζει ο Πίνακας 3.2 (Αντιδράσεις 139 και 140). Η πρώτη αντίδραση εισάγει το μόριο του φαρμάκου στο σύστημα με ρυθμό r_{drug} ενώ η δεύτερη αφορά στην καταπίεση της ελεύθερης Κυκλίνης D (Cyclin D) μέσω μιας μη ανεστρέψιμης (irreversible) δέσμησης στο φάρμακο με ρυθμό $r_{drugBinding}$. Δεδομένης της απουσίας πειραματικών δεδομένων από τη μία και της ανάγκης να τεθεί ρεαλιστική τιμή στην $r_{drugBinding}$, η τιμή της επιλέχθηκε να είναι παρόμοια με το ρυθμό με τον οποίο η Cdk4 δεσμεύει την CyclinD ($k_{bcyclinDcdk4}$, Πίνακας 3.2). Ακολούθως, έγινε μια σάρωση παραμέτρου για την παράμετρο r_{drug} στο διάστημα $[0,4000]$ ((*molecules/cell*)/*min*), με τις τιμές των υπόλοιπων παραμέτρων να παραμένουν αμετάβλητες. Για κάθε βήμα της διαδικασίας σάρωσης παραμέτρου, το μοντέλο προσομοιώθηκε για 15000 min (250 h), προκειμένου να είναι εφικτή η ανίχνευση ακόμα και ιδιαίτερα καθυστερημένης μετάβασης στην S φάση. Τα σχετικά αποτελέσματα δίνονται στην ακόλουθη εικόνα.



Εικόνα 3.8 Σάρωση παραμέτρων και αποτελέσματα προσομοίωσης για το σενάριο χορήγησης φαρμάκου.

Στην Εικόνα 3.8A, παρουσιάζεται η σχέση μεταξύ των επιπέδων της CyclinD (γαλάζιο) στο τέλος της προσομοίωσης και της τιμής της παραμέτρου r_{drug} . Όπως φαίνεται, όταν η τελευταία πλησιάζει σε τιμές μεγαλύτερες του μέσου του διαστήματος το οποίο δοκιμάστηκε, η ελεύθερη CyclinD οδηγείται σε πολύ χαμηλά ή σχεδόν μηδενικά επίπεδα. Μία παρόμοια, ωστόσο ανεστραμμένη, σχέση με την εν λόγω παράμετρο παρουσιάζει ο λόγος μεταξύ μη-φωσφορυλιωμένης pRb (pRb) και υπέρ-φωσφορυλιωμένων μορφών (hyper-pRb all forms) της pRb όπως παρουσιάζεται στην Εικόνα 3.8B. Από την εικόνα αυτή μπορεί να εξαχθεί ότι όταν η ελεύθερη Cyclin D τείνει να λάβει τιμές που την καθιστούν σχεδόν πλήρως καταπιεσμένη, η μη-φωσφορυλιωμένη μορφή της pRb κυριαρχεί των συνολικών επιπέδων της πρωτεΐνης. Αυτό είναι σε συμφωνία με τα όσα σημειώνονται στην βιβλιογραφία για τα αποτελέσματα της χορήγησης αυτών των φαρμάκων [193], [194].

Ένα ενδεικτικό αποτέλεσμα προσομοίωσης του μοντέλου θέτοντας την τιμή της r_{drug} σε 2000 ((molecules/cell)/min), δίδεται στις Εικόνα 3.8C, Εικόνα 3.8D και

Εικόνα 3.8E. Για το συγκεκριμένο *in silico* πείραμα θα πρέπει να σημειωθεί ότι η έναρξη της χορήγησης του φαρμάκου θεωρείται πως είναι συγχρονισμένη με την έναρξη του κυτταρικού κύκλου. Όπως παρατηρείται στις εικόνες, τα επίπεδα της Cyclin D (γαλάζιο) μειώνονται σημαντικά άμεσα ενώ η *hyp*-pRb (μωβ) προβλέπεται να εμφανίζει σημαντικά επίπεδα για ιδιαίτερα μικρό χρονικό διάστημα. Οι υπέρ-φωσφορυλιωμένες μορφές της pRb (σκούρο κίτρινο) αποφωσφορυλιώνονται ταχέως κάτι που οδηγεί σε σημαντική αύξηση της μη φωσφορυλιωμένης μορφής της pRb (Εικόνα 3.8D, γκρι). Τέλος, όπως φαίνεται στην Εικόνα 3.8E, η Cyclin A αποτυγχάνει να φτάσει τα απαραίτητα για την μετάβαση στην S φάση επίπεδα.

Τέλος ένα αποτέλεσμα ιδιαίτερης αξίας δίνεται στην Εικόνα 3.8F. Σε αυτή την εικόνα, παρουσιάζεται η σχέση μεταξύ χρόνου μετάβασης G1/S (G1/S Transition Time) και τιμής της παραμέτρου r_{drug} , για την ίδια διαδικασία σάρωσης παραμέτρου με προηγούμενως. Η σχέση αυτή εμφανίζεται να έχει τρεις διαφορετικές φάσεις δυναμικών τάσεων συμπεριφοράς (dynamic trend).

Συγκεκριμένα, για τιμές της r_{drug} στο διάστημα $(0,1000]$ ($(molecules/cell)/min$), η μετάβαση G1/S προβλέπεται να συμβαίνει νωρίτερα σε σχέση με την περίπτωση που δεν χορηγείται φάρμακο. Ακολούθως, αυξάνοντας την τιμή της παραμέτρου, ο χρόνος για την G1/S τείνει να αποκατασταθεί στα προηγούμενα επίπεδα μέχρι μία ξαφνική παύση του κυτταρικού κύκλου η οποία συμβαίνει όταν η παράμετρο λάβει περίπου την τιμή 1300 ($(molecules/cell)/min$). Λαμβάνοντας υπόψη τις μηχανιστικές λεπτομέρειες του μοντέλου, αυτή η συμπεριφορά μπορεί να εξηγηθεί από μία προοδευτική μείωση των διαθέσιμων οντοτήτων Cyclin D προκειμένου τα σύμπλοκα Cyclin D:Cdk4,6 να μπορούν να σχηματιστούν, τα οποία με τη σειρά του ευνοούν τον σχηματισμό υπό-φωσφορυλιωμένης και ψευδό-υπέρ-φωσφορυλιωμένης pRb. Όπως παρουσιάστηκε, αυτές οι δύο μορφές της pRb είναι ικανές να καθυστερήσουν την εξέλιξη της G1 φάσης μέσω της καταπίεσης των E2F παραγόντων μεταγραφής. Ωστόσο, μία σχεδόν πλήρης εξάλειψη αυτών των οντοτήτων οδηγεί σε αδυναμία σχηματισμού υπέρ-φωσφορυλιωμένων μορφών της pRb και έτσι στην επακόλουθη παύση του κυτταρικού κύκλου. Ιδιαίτερα ενδιαφέρον αποτελεί το γεγονός ότι τα ευρήματα του συγκεκριμένου *in silico* πειράματος είναι σε συμφωνία, τουλάχιστον σε ποιοτικό επίπεδο, με *in vitro* εργαστηριακά ευρήματα όπου το φάρμακο Πρεδνιζολόνη (Prednisolone, ενεργός μεταβολής της Πρεδνιζόνης) βρέθηκε να έχει μιτογόνο (mitogenic) δράση όταν αυτό χορηγείται σε χαμηλές δόσεις [195]. Παρόλο που το άνωθεν συμπέρασμα δεν βασίζεται σε μία ώριμη ανάλυση με χρήση εργαστηριακών

δεδομένων και δεν ενσωματώνει μία λεπτομερή μοντελοποίηση της δράσης του φαρμάκου, πιστεύεται ότι αποτελεί ξεκάθαρα μία καλή ένδειξη για την ορθότητα του προτεινόμενου μοντέλου.

3.10 Σύνοψη, Συμπεράσματα και Μελλοντικοί στόχοι

Στα πλαίσια του παρόντος κεφαλαίου μελετήθηκαν λεπτομερώς οι μηχανιστικές ιδιότητες του βιολογικού μονοπατιού «Σημείο Περιορισμού» (Restriction Point), το οποίο πρωτίστως ρυθμίζει την εξέλιξη της G1 φάσης του κυτταρικού κύκλου και οι πιθανές απορρυθμίσεις αυτού στην BCP-ALL. Μέσω μίας προσεκτικής ανασκόπησης της σχετικής βιβλιογραφίας αποκαλύφθηκε η κεντρικότητα της σχεδόν αποκλειστικής επικράτησης των υπέρ-φωσφορυλιωμένων μορφών της πρωτεΐνης ρετινοβλαστώματος (retinoblastoma protein, pRb) σε αντίθεση με τη μη-φωσφορυλιωμένη και την υπό-φωσφορυλιωμένη μορφή της. Το συγκεκριμένο εύρημα βρίσκεται σε αντίθεση με τον χαμηλό ρυθμό πολλαπλασιασμού των λευχαιμικών κυττάρων ο οποίος είναι μικρότερος των αντίστοιχων φυσιολογικών κυττάρων και με το μικρό ποσοστό των λεμφοβλαστών που βρίσκονται στην S-φάση. Μία εύλογη υπόθεση η οποία τεκμηριώνεται στην βιβλιογραφία προτείνει πως αυτή η σχετική με τον κυτταρικό κύκλο συμπεριφορά πιθανόν αποτελεί αποτελέσματα μίας μερικής απενεργοποίησης των λειτουργιών της pRb, ειδικότερα αυτών που σχετίζονται με την διαδικασία διαφοροποίησης του κυττάρου. Με βάση την παραπάνω υπόθεση, οι υπέρ-φωσφορυλιωμένες μορφές αυτής βιοχημικής οντότητας πιθανόν να διατηρούν την δυνατότητα να αποτρέπουν την μετάβαση στην S φάση, συνεχίζοντας να καταπιέζουν σχετικούς μεταγραφικούς παράγοντες. Κάτι τέτοιο θα είχε ως αποτέλεσμα, οι λευχαιμικοί κυτταρικοί πληθυσμοί να αντιστέκονται στην διαφοροποίηση ενώ ταυτόχρονα να αποφεύγουν ένα εξαιρετικά μεγάλο ρυθμό πολλαπλασιασμού. Ο παρατηρούμενος στην ALL περιορισμένος πλεονασμός (limited redundancy) των κύριων Cyclin:Cdk συμπλόκων στην φωσφορυλίωση της pRb, θεωρείται ότι έχει κεντρικό ρόλο στον επαναπρογραμματισμό του RP.

Προκειμένου τα παραπάνω φαινόμενα και υποθέσεις να μελετηθούν *in silico*, πραγματοποιήθηκαν συγκεκριμένες παρεμβάσεις-τροποποιήσεις σε ένα ήδη δημοσιευμένο μοντέλο που προσομοιώνει την βιοχημική ρύθμιση της G1 φάσης του κυτταρικού κύκλου [139]. Επίσης τέθηκε ο στόχος να ρυθμιστεί το μοντέλο με τρόπο τέτοιο ώστε τα αποτελέσματά του να συμφωνούν με τον παρατηρούμενο στην ALL χρόνο που απαιτεί η μετάβαση στην S φάση. Έπειτα από την προσθήκη επιπλέον αντιδράσεων και μεταβλητών στο μοντέλο, ακολουθήθηκε μία διαδικασία προσαρμογής του μοντέλου προκειμένου να

καταστεί ικανό να προσομοιώσει τα φαινόμενα που σχετίζονται με την απορρύθμιση του RP στην ALL και να έρθει σε αριθμητική συμφωνία με την καταγεγραμμένη στη βιβλιογραφία μέση κυτταροκινητική συμπεριφορά στην BCP-ALL. Για την επίτευξη του στόχου, σχεδιάστηκε μία πρωτότυπη από «πάνω προς τα κάτω» (top-down) και ημι-ποσοτική/ποιοτική διαδικασία προσαρμογής η οποία ενσωματώνει και στοχαστικές μεθόδους ολικής βελτιστοποίησης. Μετά την προσαρμογή των παραμέτρων του, το μοντέλο δείχθηκε να είναι ικανό στο να προβλέψει ότι η μετάβαση στην S φάση λαμβάνει χώρα εντός ενός ρεαλιστικού χρονικού διαστήματος το οποίο είναι σε συμφωνία με την σχετική με την BCP-ALL βιβλιογραφία. Τέλος, δοκιμάζοντας *in silico* την παρέμβαση ενός συγκεκριμένου τύπου αντί-λευχαιμικού φαρμάκου (Γλυκοκορτικοειδή) στον κυτταρικό κύκλο, η συμπεριφορά του συστήματος δείχθηκε να είναι σε συμφωνία με την καταγεγραμμένη στη βιβλιογραφία αντί-πολλαπλασιαστική δράση του φαρμάκου.

Τα παραπάνω αποτελέσματα αποτελούν μία αρκετά καλή ένδειξη για την ορθότητα του προτεινόμενου μοντέλου. Πιθανό επόμενο βήμα στην ανάλυση αυτού θα μπορούσε να αποτελέσει η πιο λεπτομερής προσαρμογή του με τη χρήση εργαστηριακών δεδομένων από το πρωτόωμα λευχαιμικών κυτταρικών πληθυσμών. Τέλος, ένα σημαντικό μελλοντικό βήμα θα περιλάμβανε την συσχέτιση των παραμέτρων του μοντέλου με ευρέως διαθέσιμες στο κλινικό περιβάλλον μετρήσεις-μεταβλητές προκειμένου να υποστηριχθούν οι δυνατότητες του μοντέλου στο να προβλέπει προσωποποιημένες τιμές της διάρκειας του κυτταρικού κύκλου. Μία τέτοια πρόβλεψη θα μπορούσε να αποτελέσει σημαντική επιπλέον γνώση για την προσωποποιημένη παραμετροποίηση μηχανιστικών και πολύ-επίπεδων μοντέλων για τον καρκίνο όπως ο Ογκοπροσομοιωτής.

4. Μοντελοποίηση Φαρμακοκινητικής της από στόματος χρήσης του φαρμάκου Πρεδνιζόνη σε παιδιά με Οξεία Λεμφοβλαστική Λευχαιμία

4.1 Εισαγωγή

Ο τομέας της φαρμακοκινητικής αποτελεί κλάδο της φαρμακολογίας ο οποίος ασχολείται με την μαθηματική μελέτη της πορείας μιας δοθείσας ουσίας στο σώμα ενός ζωντανού οργανισμού. Η ουσία αυτή συνήθως είναι κάποιο φαρμακευτικό σκεύασμα. Παρότι η Φαρμακοκινητική, όντας αρκετά ώριμος και ανεξάρτητα αναπτυσσόμενος χώρος, δεν εντάσσεται άμεσα στους τομείς με τους οποίους ασχολείται η διατριβή, αποτελεί σημαντικό κλάδο της υπολογιστικής μοντελοποίησης βιολογικών συστημάτων ενώ σε πολλές περιπτώσεις εμφανίζει σημαντική συνάφεια σε μεθόδους αλλά και σε στόχους με τους τομείς αυτούς. Επιπροσθέτως, δεδομένου του κεντρικού στόχου του τομέα της In Silico Ογκολογίας γενικότερα και του Ογκοπροσομοιωτή ειδικότερα για την προσομοίωση της απόκρισης του ασθενούς στη θεραπεία, η όσο το δυνατόν ακριβέστερη πρόβλεψη της συγκέντρωσης του δοθέντος ή δοθέντων φαρμάκων στα σημεία ενδιαφέροντος (όπως ο καρκινικός όγκος για την περίπτωση στερεών όγκων και το αίμα για την περίπτωση της ALL) μπορεί να οδηγήσει σε περαιτέρω εξατομίκευση αυτής.

Στα πλαίσια αυτά, επιχειρείται η δημιουργία ενός μοντέλου ικανού να προσομοιώσει την Φαρμακοκινητική του φαρμάκου Πρεδνιζόνη (Prednisone, PRED) για την περίπτωση της από στόματος χορήγησης.

Παρά τον κεντρικό ρόλο που του συγκεκριμένο φάρμακο (PRED) έχει στην θεραπεία της παιδικής ALL, ο αριθμός των εργασιών που αναφέρονται στις φαρμακοκινητικές ιδιότητες του Prednisone ή του ενεργού μεταβολίτη του Prednisolone (PREDNL) είναι σχετικά μικρός και ακόμα λιγότερες εργασίες έχουν ως αντικείμενο την μοντελοποίηση της φαρμακοκινητικής τους. Ειδικότερα, μία εργασία που συζητά την μοντελοποίηση της φαρμακοκινητικής, σε επίπεδο πληθυσμού, του από στόματος χορηγούμενου PRED (Oral PRED) σε παιδιά που πάσχουν από ALL, φαίνεται να μην είναι διαθέσιμη στην διεθνή βιβλιογραφία. Ένα τέτοιο μοντέλο, θα επέτρεπε την πρόβλεψη, σε εξατομικευμένο επίπεδο, της φαρμακοκινητικής των PRED και PREDLN στους ασθενείς των οποίων οι περιπτώσεις προσομοιώνονται μέσω του ALL Ογκοπροσομοιωτή και έτσι να μελετηθεί με ακριβέστερο τρόπο η φαρμακοδυναμική του φαρμάκου.

Ωστόσο, ένα πληθυσμιακό φαρμακοκινητικό μοντέλο για την χορήγηση απευθείας του ενεργού μεταβολίτη PREDNL σε παιδιά με ALL προτείνεται στην [196] ενώ μοντέλα για την από στόματος χορήγηση του PRED, εκτός πλαισίου ALL, προτείνονται στις [197]–[199].

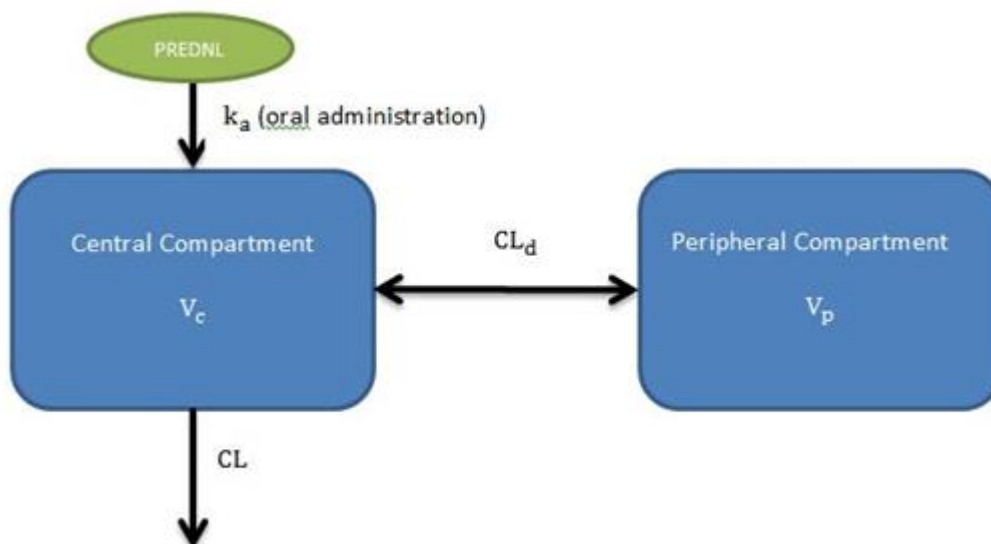
Έτσι, στο συγκεκριμένο κεφάλαιο της διατριβής, τίθεται ο στόχος να δημιουργηθεί ένα μοντέλο για την από στόματος χορήγηση του PRED σε παιδιά με ALL βασιζόμενοι στα προαναφερθέντα μοντέλα μιας και η απουσία φαρμακοκινητικών δεδομένων για αυτόν τον πληθυσμό, δεν επιτρέπει την δημιουργία ενός πλήρως καινούργιου μοντέλου.

4.2 Βιβλιογραφική ανασκόπηση της μοντελοποίησης και της προσομοίωσης της φαρμακοκινητικής του φαρμάκου Πρεδνιζόνη

Όπως συζητήθηκε και στην εισαγωγική ενότητα, ενός αριθμός προ-υπαρχόντων μοντέλων χρησιμοποιήθηκαν ως βάση για την ανάπτυξη του προτεινόμενου μοντέλου. Τα βασικά χαρακτηριστικά των μοντέλων αυτών παρουσιάζονται στις ακόλουθες υπο-ενότητες.

4.2.1 Φαρμακοκινητικό μοντέλο για την χορήγηση Πρεδνιζολόνης σε παιδιά που πάσχουν από Οξεία Λεμφοβλαστική Λευχαιμία

Ένα πληθυσμιακό μοντέλο για την από στόματος (orally) ή την ενδοφλέβια (intravenously) χορήγηση του ενεργού μεταβολίτη Prednisolone προτάθηκε στην [196], ακολουθώντας την προσέγγιση δύο διαμερισμάτων (two-compartment PK model). Για την συνέχεια του κειμένου το συγκεκριμένο μοντέλο θα ονομάζεται ως Μοντέλο 1 (Model 1) ενώ η δομή του παρουσιάζεται στην ακόλουθη εικόνα.



Εικόνα 4.1 Δομή του μοντέλου που προτείνεται στην [196] (Μοντέλο 1) για την προσομοίωση της φαρμακοκινητικής του Prednisolone (Central Compartment: Κεντρικό διαμέρισμα, Peripheral Compartment: Περιφερικό διαμέρισμα, Oral administration: Από στόματος χορήγηση)

Όπως παρουσιάζεται στην Εικόνα 4.1, το από στόματος χορηγούμενο PREDNL προσλαμβάνεται μέσω μία διαδικασίας πρώτης τάξης (first order process) με ρυθμό k_a από το κεντρικό διαμέρισμα (Central Compartment) το οποίο έχει όγκο V_c . Μετά την πρόσληψη (absorption) το φάρμακο μπορεί να μεταφερθεί με ρυθμό CL_d (Inter-compartmental clearance rate) και με αντιστρέψιμο τρόπο από το περιφερικό διαμέρισμα (Peripheral Compartment) το οποίο έχει όγκο V_p είτε να αποβληθεί (elimination) από το Κεντρικό Διαμέρισμα με ρυθμό CL .

Στην συγκεκριμένη εργασία εκτός του παραπάνω μοντέλου, αναπτύχθηκε και ένα μοντέλο παλινδρόμησης το οποίο συσχετίζει τις παραμέτρους του με συγκεκριμένα χαρακτηριστικά του ασθενούς τα οποία είναι η σωματική του επιφάνεια (Body Surface Area/BSA) σε m^2 και το βάρος (WT) του σε κιλά (kg).

Για την εκτίμηση των τιμών παραμέτρων για το φαρμακοκινητικό μοντέλο και των συντελεστών του μοντέλου παλινδρόμησης χρησιμοποιήθηκαν δεδομένα συγκέντρωσης του μη δεσμευμένου PREDNL στο αίμα 23 παιδιών ασθενών με ALL τα οποία έλαβαν $60 mg/m^2$ PREDNL ανά ημέρα σε 3 δόσεις των $20.8 mg/m^2$ ανά 8 ώρες.

Θα πρέπει να σημειωθεί ότι στην παρούσα εργασία ο αντιστρέψιμος μεταβολισμός του PREDNL σε PRED δεν λήφθηκε υπόψη.

Το εν λόγω μοντέλο παλινδρόμησης δίνεται στον ακόλουθο πίνακα

Πίνακας 4.1 Μοντέλο παλινδρόμησης για τις παραμέτρους του Μοντέλου 1 όπως δίδεται στην [196]

Μοντέλο Παλινδρόμησης	Εκτιμηθείσες τιμές συντελεστών	Συντελεστές διακύμανσης
$CL(lh^{-1}) = \theta_1 \times BSA$		
θ_1	30.7	6.5
$V_c(l) = \theta_2 + \theta_3 \times WT$		
θ_2	14.0	34
θ_3	0.73	21
$V_p(l) = \theta_4 + \theta_5 \times WT$		
θ_4	17.3	54
θ_5	1.35	35

Το συγκεκριμένο μοντέλο παλινδρόμησης θεωρείται ιδιαίτερα σημαντικό μιας και μέσω της χρήσης του μπορούν να προκύψουν κάποιες, τουλάχιστον αρχικές, εκτιμήσεις για τις βασικές φαρμακοκινητικές παραμέτρους του ενεργού μεταβολίτη του PRED, PREDNL σε εξατομικευμένο επίπεδο.

Όσον αφορά τις άλλες δύο βασικές παραμέτρους k_a και CL_d οι συγγραφείς αναφέρουν πως η διακύμανση μεταξύ των ασθενών δεν ήταν εφικτό να εκτιμηθεί με ακρίβεια και έτσι παρέχουν μόνο την εκτίμησή τους σε επίπεδο πληθυσμού.

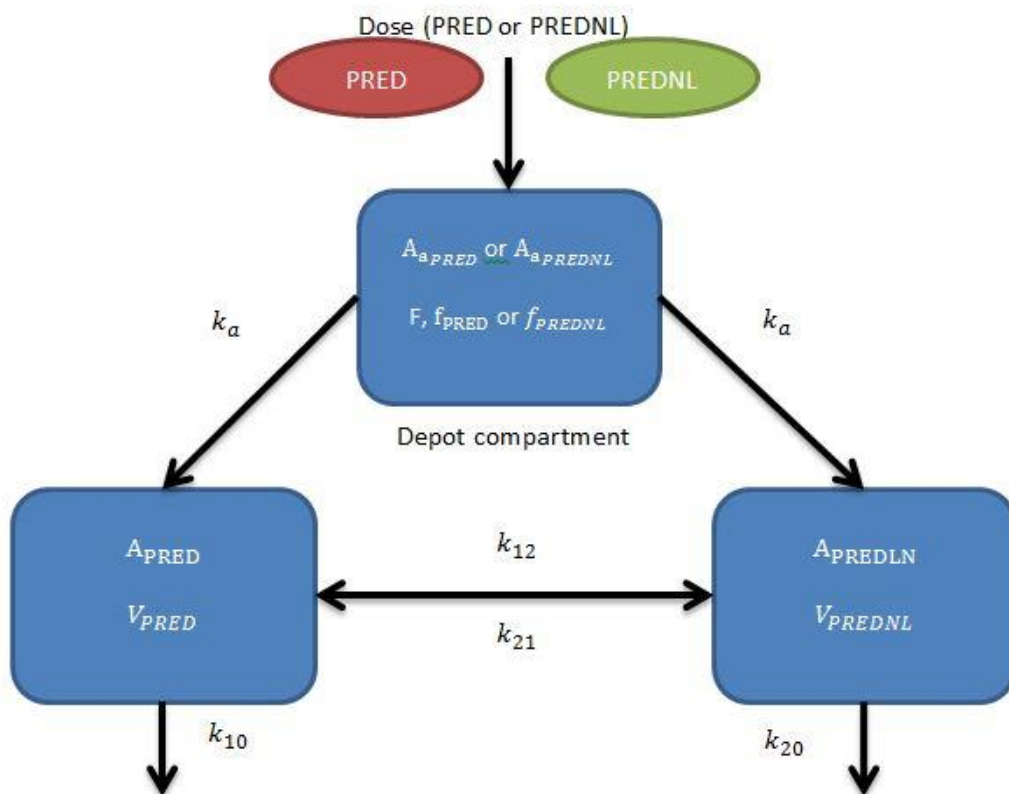
Τέλος οι διάμεσες τιμές και η διακύμανση των φαρμακοκινητικών παραμέτρων όπως εκτιμήθηκαν από την εργασία δίνονται στον ακόλουθο πίνακα.

Πίνακας 4.2 Εκτιμήσεις τιμών παραμέτρων για το φαρμακοκινητικό μοντέλο που προτείνεται στην [196]

Παράμετρος	$k_a (h^{-1})$	$CL ((l/h/m^2))$	$CL_d ((l/h/m^2))$	$V_c (l/m^2)$	V_p
Διάμεσος	1.35 (πληθυσμιακή εκτίμηση)	31.8	15.5	44.3	43.0
Διάστημα	-	24.3-36.0	6.7-20.9	10.5-78.4	33.1-87.2

4.2.2 Φαρμακοκινητικό μοντέλο για την από στόματος χορήγηση του φαρμάκου Πρεδνιζόνη

Στην εργασία [199] προτείνεται ένα φαρμακοκινητικό μοντέλο (στην συνέχεια του κειμένου Μοντέλο 2 ή Model 2) ικανό να προσομοιώσει τα επίπεδα του PRED ή/και του PREDNL στο πλάσμα ανθρώπων μετά την από στόματος ή την ενδοφλέβια χορήγηση και των δύο μορφών του φαρμάκου. Η δομή του μοντέλου για την περίπτωση της από-στόματος χορήγησης δίνεται στην ακόλουθη εικόνα



Εικόνα 4.2 Δομή του μοντέλου που προτείνεται στην [199] για την προσομοίωση της φαρμακοκινητικής του Prednisolone (Depot compartment: αποθηκευτικό διαμέρισμα, Dose: Δόση)

Στο συγκεκριμένο μοντέλο ακολουθήθηκε μια ειδική μέθοδος φαρμακοκινητικής μοντελοποίησης προκειμένου ο αντιστρέψιμος μεταβολισμός του PRED σε PREDNL και η από στόματος χορήγηση των φαρμάκων αυτών να προσομοιωθούν. Όπως φαίνεται στην Εικόνα 4.2, το μοντέλο δεν προσομοιώνει διακριτά τα επίπεδα των φαρμάκων σε ένα Κεντρικό και σε ένα Περιφερικό διαμέρισμα (όπως στο Μοντέλο 1) αλλά χρησιμοποιεί δύο διαμερίσματα (A_{PRED} και A_{PREDNL} με όγκους V_{PRED} και V_{PREDNL} αντίστοιχα) για να μοντελοποιηθούν ο όγκος κατανομής (volume of distribution) και οι συγκεντρώσεις (concentrations) των δύο σταδίων του φαρμάκου και η αντιστρέψιμη μετατροπή τους. Επιπλέον χρησιμοποιείται ένα αποθηκευτικό διαμέρισμα (depot compartment) έτσι ώστε να ληφθεί υπόψη το ποσοστό βιοδιαθεσιμότητας (F) της κάθε μορφής φαρμάκου. Ειδικότερα, τα βήματα που ακολουθούνται από το μοντέλο για την προσομοίωση είναι τα ακόλουθα:

1. Η από στόματος δόση περνάει στο depot compartment
2. Με βάση την βιοδιαθεσιμότητα (F) του κάθε φαρμάκου, μόνο ένα ποσοστό της αρχικής δόσης καταλήγει στο σύστημα. Έτσι όταν χορηγείται PRED, έχουμε ότι $A_{aPRED}(0) = Dose_{PRED} * F_{PRED}$ όπου $A_{aPRED}(0)$ η αρχική συγκέντρωση του

PRED στο depot compartment, $Dose_{PRED}$ η από στόματος χορηγούμενο δόση PRED και F_{PRED} η βιοδιαθεσιμότητα του PRED

3. Επειδή το PRED επιδέχεται ενός άμεσου ηπατικού μεταβολισμού πρώτης διόδου (first-pass conversion in liver), ένα ποσοστό του βιοδιαθέσιμου φαρμάκου καταλήγει στη συστηματική κυκλοφορία ως PRED (f_{PRED}) και ένα ποσοστό ως PREDNL (f_{PREDNL})
4. Με την μετατροπή πρώτης διόδου (first-pass conversion) κάθε τύπος του φαρμάκου προσλαμβάνεται από τα δύο διαμερίσματα μέσω μιας διαδικασίας πρώτης τάξης με ρυθμό k_a
5. Η αποβολή των μορφών του φαρμάκου με ρυθμούς k_{10} και k_{20} και ο αντιστρέψιμος μεταβολισμός του με ρυθμούς k_{21} και k_{12} προσομοιώνεται μέσω δύο διαμερισμάτων A_{PRED} και A_{PREDNL}

Προκειμένου οι προαναφερθείσες παράμετροι να εκτιμηθούν, οι συγγραφείς της εργασίας χρησιμοποίησαν δεδομένα από τη βιβλιογραφία για τα επίπεδα των PRED και PREDNL στο πλάσμα. Οι εκτιμήσεις των παραμέτρων αυτών δίνονται στον ακόλουθο πίνακα:

Πίνακας 4.3 Τιμές παραμέτρων όπως εκτιμήθηκαν στην [199] για το Μοντέλο 2

Παράμετροι (μονάδες)	Περιγραφή	Εκτιμήσεις (%CV)
k_{10} ($hours^{-1}$)	Σταθερά ρυθμού αποβολής πρώτης τάξης για το PRED	0.25 (13.33)
k_{20} ($hours^{-1}$)	Σταθερά ρυθμού αποβολής πρώτης τάξης για το PREDNL	0.33 (13.97)
k_{12} ($hours^{-1}$)	Σταθερά πρώτης τάξης μετατροπής από PRED σε PREDNL	0.23 (17.59)
k_{21} ($hours^{-1}$)	Σταθερά πρώτης τάξης μετατροπής από PREDNL σε PRED	397.3 (12.56)
V_{PRED} (l)	Όγκος του PRED	110.4 (7.38)
K_{aPRED} ($hours^{-1}$)	Σταθερά ρυθμού απορρόφησης πρώτης τάξης για το (Oral PRED)	1.08 (8.24)
F	Βιοδιαθεσιμότητα	Oral PRED: 0.75(3.76) Oral PREDNL: 0.92(11.61)
f_{PRED}	Ποσοστό του PRED που εισέρχεται στο σύστημα	Oral PRED: 0.14
f_{PREDNL}	Ποσοστό του PREDNL που εισέρχεται στο σύστημα	Oral PREDNL: 0.86

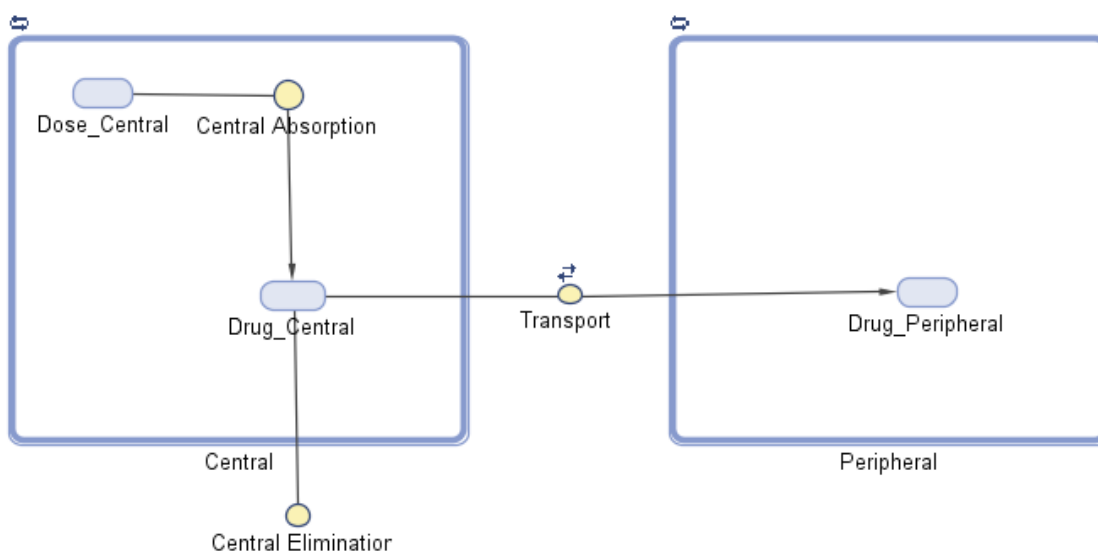
4.3 Ανάπτυξη εκπαίδευση και δοκιμή προτεινόμενου μοντέλου

4.3.1 Ορισμός προβλήματος

Όπως συζητήθηκε και στις προηγούμενες ενότητες, παρότι το Μοντέλο 1 προσομοιώνει τα επίπεδα του PREDNL με εξατομικευμένο και εξειδικευμένο στην περίπτωση των παιδιών που πάσχουν από ALL τρόπο, προκειμένου να προσομοιωθεί επίσης και η πρόσληψη του PRED στην περίπτωση της από στόματος χορήγησης, θα πρέπει να υποστεί συγκεκριμένες τροποποιήσεις. Οι τροποποιήσεις αυτές έγιναν έχοντας ως βάση τις αρχές μοντελοποίησης που υιοθετούνται στο Μοντέλο 2 με τρόπο που παρουσιάζεται στις ακόλουθες ενότητες.

4.3.2 Υλοποίηση προ-υπαρχόντων μοντέλων και έλεγχος αυτής

Προκειμένου να γίνουν οι σχετικές τροποποιήσεις και δεδομένου ότι οι συνήθεις διαφορικές εξισώσεις που περιγράφουν το Μοντέλο 1 δεν παρέχονται από τους συγγραφείς της εργασίας [196], το Μοντέλο 1 υλοποιήθηκε εκ νέου στο εργαλείο SimBiology του MATLAB με βάση την περιγραφή που δίνεται στην εργασία. Η γραφική αναπαράσταση του μοντέλου στο εργαλείο SimBiology δίνεται στην ακόλουθη εικόνα.



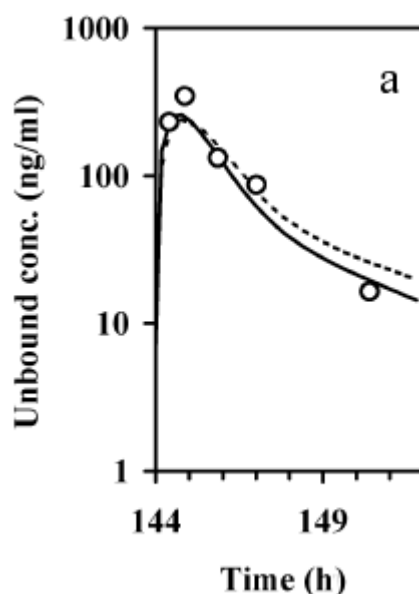
Εικόνα 4.3 Γραφική αναπαράσταση του μοντέλου για την από στόματος χορήγηση PREDNL όπως υλοποιήθηκε στο SimBiology [194] (Central: Κεντρικό διαμέρισμα, Peripheral: Περιφερικό διαμέρισμα, Dose: Δόση, Absorption: Απορρόφηση, Transport: Μεταφορά, Elimination: Αποβολή)

Προκειμένου να εξασφαλιστεί ότι η υλοποίηση έγινε επιτυχώς, επιλέχθηκε ένα παράδειγμα ασθενούς από την [196] προκειμένου να παρουσιαστεί η ικανότητα του μοντέλου να προσομοιώσει τα επίπεδα του μη δεσμευμένου Prednisolone στο πλάσμα το αίματος. Ο ασθενής που επιλέχθηκε είχε τα ακόλουθα χαρακτηριστικά.

Πίνακας 4.4 Χαρακτηριστικά ασθενούς που επιλέχθηκε από την εργασία [196] για την επιβεβαίωση της ορθής υλοποίησης του μοντέλου

Βάρος	Ύψος	CL	CL _d	V _c	V _p
42.0	152	30.0	9.3	31.8	49.3

Τα δείγματα όπως και τα πρωτότυπα αποτελέσματα προσομοίωσης δίνονται στην ακόλουθη εικόνα η οποία έχει ληφθεί από την [196]. Ο συγκεκριμένος ασθενής έλαβε από το στόμα δόση PREDNL στην 144^η ώρα (7^η ημέρα της θεραπείας) και ακολουθούσε ήδη θεραπεία με 20.8mg/m² κάθε 8 ώρες.



Εικόνα 4.4 Τα αποτελέσματα προσομοίωσης του φαρμακοκινητικού μοντέλου όπως δίνονται στην [196] για τον ασθενή που επιλέχθηκε προκειμένου να επιβεβαιωθεί η ορθή υλοποίηση (συμπαγής γραμμή: προσομοιωμένη πρόβλεψη, διακεκομμένη γραμμή: πληθυσμιακή πρόβλεψη, κύκλοι: δείγματα συγκέντρωσης μη δεσμευμένου PREDNL στο πλάσμα του ασθενούς).

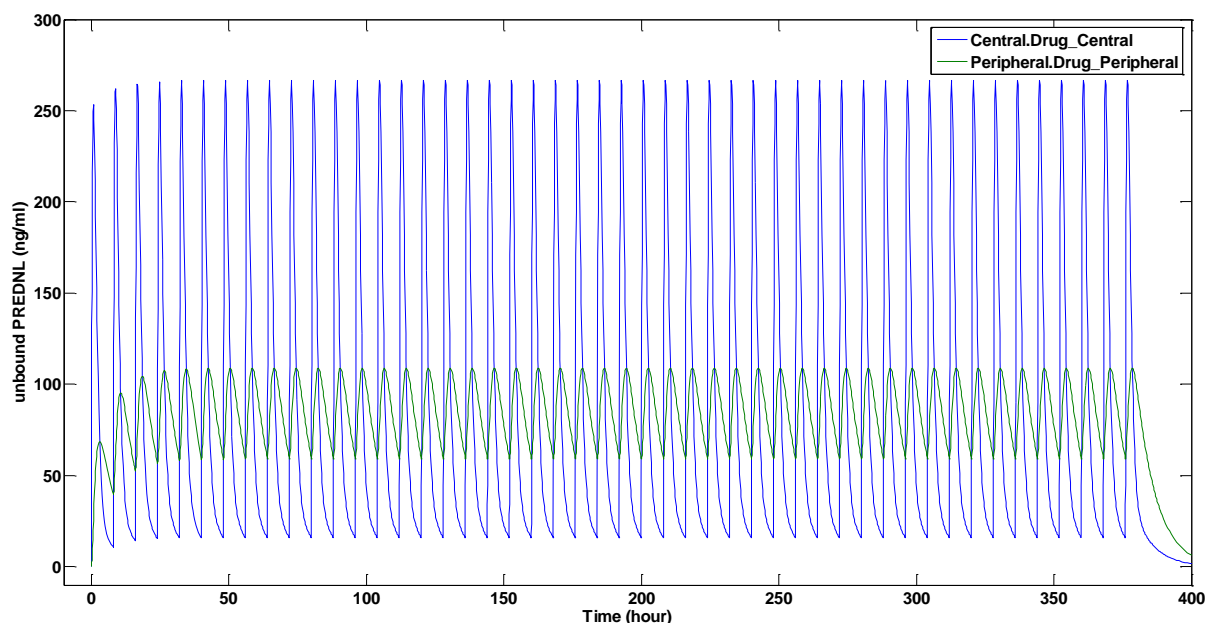
Προκειμένου να είναι εφικτή η σύγκριση της νέας υλοποίησης με τα αποτελέσματα της αρχική υλοποίησης, τα αποτελέσματα που δίνονται στην Εικόνα 4.4 ψηφιοποιήθηκαν με τη χρήση του εργαλείου *GetData Graph Digitizer* (<http://www.getdata-graph-digitizer.com/>) μέσω του οποίου λήφθηκαν 11 δείγματα από την συνεχή γραμμή (αποτέλεσμα προσομοίωσης).

Όσον αφορά τις τιμές παραμέτρων του μοντέλου, χρησιμοποιήθηκε το μοντέλο παλινδρόμησης που δίνει ο Πίνακας 4.1 προκειμένου να υπολογιστούν οι κατάλληλες τιμές για τις CL , CL_d , V_c , V_p . Ωστόσο, επειδή η τιμή για την BSA δεν δίνεται στην [196] συγκεκριμένα για κάθε ασθενή, αυτή υπολογίστηκε με βάση το ύψος και το βάρος του ασθενούς με τη χρήση της φόρμουλας Gehan and George [200], το οποίο υπολογίστηκε σε $1,34 m^2$. Με βάση τα ανωτέρω οι τιμές των παραμέτρων που τέθηκαν δίνονται στον ακόλουθο πίνακα.

Πίνακας 4.5 Οι φαρμακοκινητικές παράμετροι που χρησιμοποιήθηκαν για να προσομοιωθεί η περίπτωση του επιλεγμένου από την [196] ασθενούς.

V_c	42.612 (l)
V_p	66.062 (l)
CL	40.2 (l)
CL_d	12.462 ($l h^{-1}$)
K_a	1.35 (h^{-1})
Δόση(Dose)	27.872 (mg)

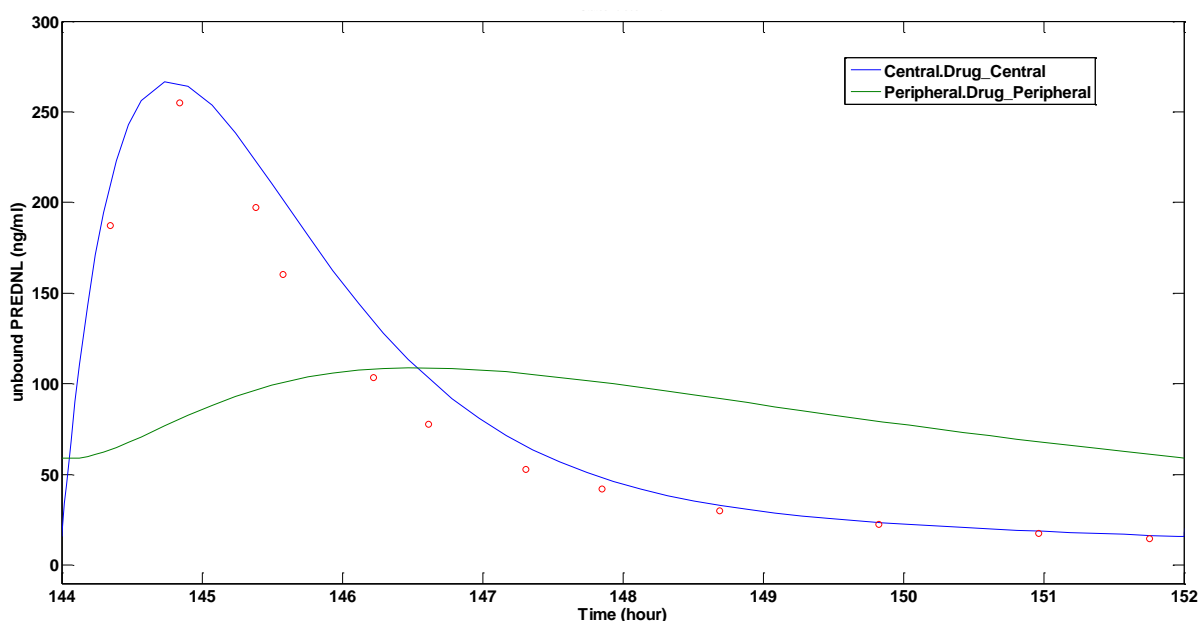
Τα αποτελέσματα προσομοίωσης (με την νέα υλοποίηση) για το πλήρες σχήμα θεραπείας (47 δόσεις ανά δωρο) δίνονται στην ακόλουθη εικόνα.



Εικόνα 4.5 Αποτελέσματα προσομοίωσης της φαρμακοκινητικής του PREDNL για το πλήρες σχήμα θεραπείας που ακολουθείται στην [196] μέσω της νέας υλοποίησης του μοντέλου.

Προκειμένου να ελεγχθεί η συμφωνία των αποτελεσμάτων προσομοίωσης με τα δεδομένα που παρουσιάζονται στην Εικόνα 4.4, εκτελέστηκε προσομοίωση για τις ώρες

144 έως 152 της οποίας τα αποτελέσματα δίνονται στην Εικόνα 4.6 μαζί με τα δείγματα που λήφθηκαν από την Εικόνα 4.4.



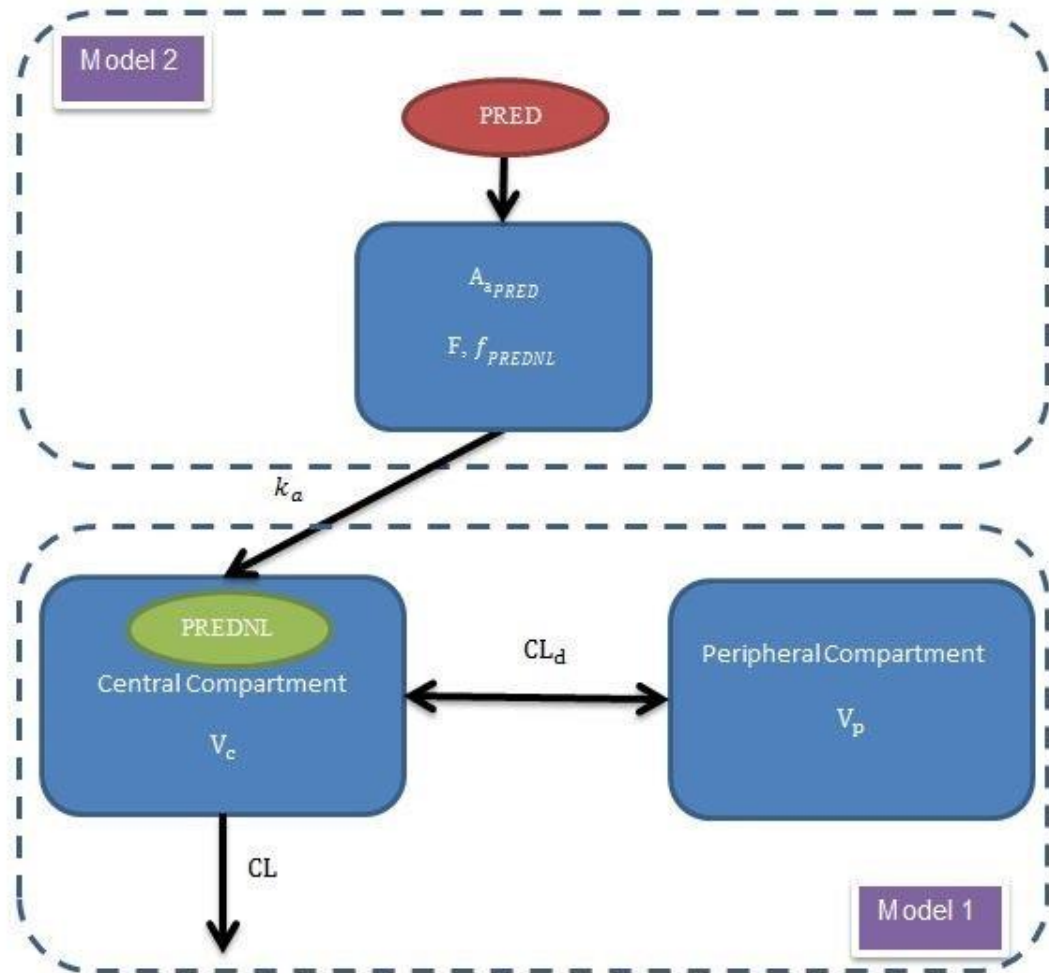
Εικόνα 4.6 Αποτελέσματα προσομοίωσης με την νέα υλοποίηση του μοντέλου για τις ώρες 144 έως 152 σε σύγκριση με τα δεδομένα που λήφθηκαν από την Εικόνα 4.4

Όπως είναι εμφανές, η συγκέντρωση του φαρμάκου προσομοιώνεται επιτυχώς από το νέο μοντέλο. Παρότι μία σχετική απόκλιση σε συγκεκριμένες περιοχές είναι εμφανής, πιστεύεται ότι είναι αποτέλεσμα της πιθανούς διαφορετικής μεθόδου για τον υπολογισμό της BSA, της ακρίβειας στην ψηφιοποίηση για την Εικόνα 4.4 αλλά και της μη προσωποποιημένης τιμής για την παράμετροι k_a .

4.3.3 Δημιουργία της δομής του προτεινόμενου μοντέλου

Δεδομένου ότι βασικός στόχος της παρούσας διαδικασίας μοντελοποίησης ήταν να διατηρηθεί το Μοντέλο 1 ανεπηρέαστο, έτσι ώστε να είναι εφικτή η συνέχιση της χρήσης του συνοδευτικού μοντέλου παλινδρόμησης (Πίνακας 4.1) ως βάση για την επιλογή τιμών παραμέτρων, ο τρόπος που επιλέχθηκε να γίνει η επέκταση του ήταν μέσω της προσθήκης σε αυτό ενός *derot* διαμερίσματος, παρόμοιο με αυτό που χρησιμοποιείται στο Μοντέλο 2. Το διαμέρισμα αυτό χρησιμοποιείται έτσι ώστε να προσομοιώνονται η μετατροπή της δοθείσας δόσης PRED στην βιοδιαθέσιμη ποσότητά της, ο μεταβολισμός πρώτης διόδου του φαρμάκου και η πρόσληψη, μέσω μιας διαδικασίας πρώτης τάξης, του ποσοστού του PRED το οποίο έχει μετατραπεί σε RPEDNL. Μία υπόθεση η οποία έγινε σχετικά με την μετατροπή του PRED σε PREDNL είναι ότι το μεγαλύτερο ποσοστό αυτής γίνεται μέσω του μεταβολισμού πρώτης διόδου και έτσι ο μεταγενέστερος αντιστρέψιμος μεταβολισμός

μεταξύ PRED και PREDNL δεν μοντελοποιείται. Η συγκεκριμένη υπόθεση συναντάται συχνά στη βιβλιογραφία όπως στην [196] (Μοντέλο 1) αλλά και στην [197] όπου μοντελοποιείται η φαρμακοκινητική το από στόματος χορηγούμενου PRED. Με βάση τα παραπάνω, η δομή του προτεινόμενου μοντέλου δίνεται στην ακόλουθη εικόνα.



Εικόνα 4.7 Δομή του προτεινόμενου μοντέλου για την προσομοίωση της φαρμακοκινητικής του PRED (Central Compartment: Κεντρικό διαμέρισμα, Peripheral Compartment: Περιφερικό διαμέρισμα)

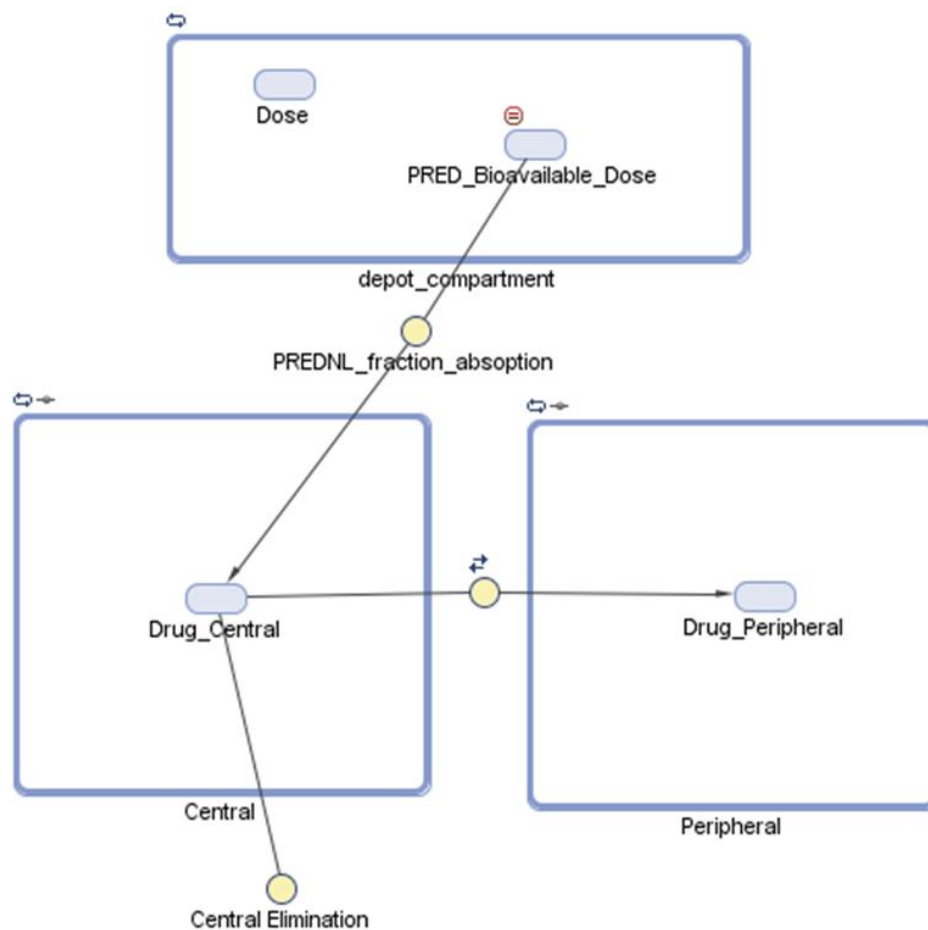
Παρόμοια με τα βήματα που παρουσιάστηκαν για το Μοντέλο 2, ακολούθως δίνονται τα βήματα που ακολουθούνται από το προτεινόμενο μοντέλο για την προσομοίωση της φαρμακοκινητικής του PRED:

1. Η από στόματος χορηγούμενη δόση PRED περνάει στο depot διαμέρισμα
2. Με βάση της βιοδιαθεσιμότητα του PRED μόνο ένα ποσοστό του φαρμάκου καταλήγει στο σύστημα, έτσι παρόμοια με το Μοντέλο 2, $A_{\alpha PRED}(0) = Dose_{PRED} * F_{PRED}$
3. Επειδή το PRED επιδέχεται ενός άμεσου ηπατικού μεταβολισμού πρώτης διόδου (first-pass conversion in liver), ένα ποσοστό του βιοδιαθέσιμου φαρμάκου

καταλήγει στη συστημακή κυκλοφορία ως PRED (f_{PRED}) και ένα ποσοστό ως PREDNL (f_{PREDNL}). Ωστόσο, για λόγους που εξηγήθηκαν προηγουμένως, μοντελοποιείται μόνο η απορρόφηση του PREDNL.

4. Μετά τον μεταβολισμό πρώτης διόδου, το PREDNL μεταφέρεται στο Κεντρικό Διαμέρισμα (Central Compartment) μέσω μιας διαδικασίας πρώτης τάξης με ρυθμό k_a
5. Η αποβολή (Clearance) και η δια-διαμερισματική αποβολή (Intercompartmental Clearance) του φαρμάκου μοντελοποιούνται όπως στο Μοντέλο 1.

Το προτεινόμενο μοντέλο υλοποιήθηκε επίσης στο SimBiology του Matlab με την ακόλουθη δομή:



Εικόνα 4.8 Δομή του προτεινόμενου μοντέλου στο περιβάλλον SimBiology (Central Compartment: Κεντρικό διαμέρισμα, Peripheral Compartment: Περιφερικό διαμέρισμα, Depot compartment: Αποθηκευτικό διαμέρισμα)

4.3.4 Εκπαίδευση και αξιολόγηση του προτεινόμενου μοντέλου

Προκειμένου να διασφαλιστεί ότι το προτεινόμενο μοντέλο, με δεδομένες τις τροποποιήσεις στη δομή του σε σχέση με το Μοντέλο 1, είναι ικανό να προσομοιώσει την φαρμακοκινητική του PREDNL μετά την από στόματος χορήγηση PRED, ακολουθήθηκε η

διαδικασία εκτίμησης των παραμέτρων του χρησιμοποιώντας δεδομένα από παιδιά στα οποία χορηγήθηκε η σχετική θεραπεία. Δυστυχώς, όπως ήδη έχει σημειωθεί, τέτοια δεδομένα δεν ήταν εφικτό να βρεθούν στην βιβλιογραφία για την περίπτωση παιδιών που πάσχουν από ALL. Ωστόσο, χρησιμοποιήθηκαν δεδομένα από παιδιά ασθενείς που αντιμετωπίζουν νεφρωσικό σύνδρομο (Nephrotic syndrome) και τα οποία έλαβαν αυτή τη θεραπεία. Τα δεδομένα αυτά λήφθηκαν από την εργασία [201]. Δεδομένου ότι μπορεί να θεωρηθεί ότι η συγκεκριμένη ασθένεια μπορεί να επηρεάσει την φαρμακοκινητική του φαρμάκου, χρησιμοποιήθηκαν οι μετρήσεις που λήφθηκαν κατά την ύφεση (remission) της ασθένειας, παρόλο που στην ίδια εργασία δείχνεται πως τα επίπεδα του PREDNL δεν παρουσιάζουν σημαντικές διαφορές μεταξύ της ενεργούς και αυτή της ύφεσης φάσεων. Στην ακόλουθη εικόνα δίδεται η μέση συγκέντρωση του ελεύθερου PREDNL στο αίμα, μετά την χορήγηση 55mg PRED (μέση δόση).

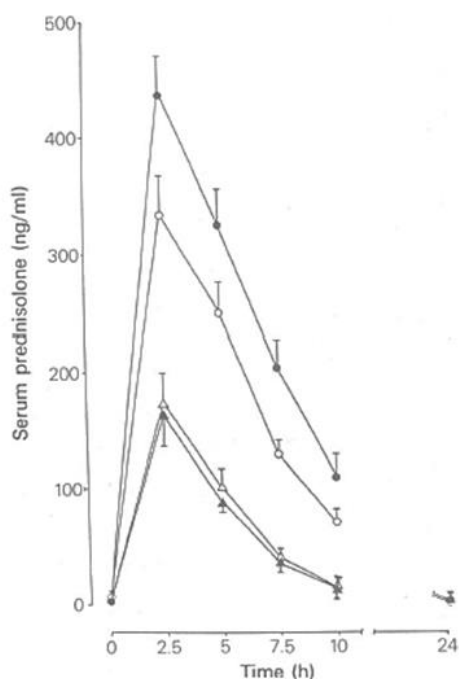


Figure 1 Levels of free (Δ , \blacktriangle) and total (\circ , \bullet) prednisolone levels during the active phase of the nephrotic syndrome (open symbols) and in remission (closed symbols). Values represent the means \pm s.e. mean.

Εικόνα 4.9 Δεδομένα που δίνονται στην [201] για τα επίπεδα (μέσες τιμές) της ελεύθερης Πρεδνιζολόνης μετά από χορήγηση 55mg Πρεδνιζόνης σε παιδιά με νεφρωτικό σύνδρομο σε φάση ύφεσης.

Η εκπαίδευση του μοντέλου στα παραπάνω δεδομένα, τα οποία ψηφιοποιήθηκαν με τη χρήση του λογισμικού Get Data Graph Digitizes (<http://www.getdata-graph-digitizer.com/>) έγινε μέσω της σχετικής λειτουργίας του SimBiology χρησιμοποιώντας την συνάρτηση NLINFIT για μεμονωμένη προσαρμογή (individual fit). Για τον υπολογισμό κάποιων αρχικών προσεγγίσεων των φαρμακοκινητικών παραμέτρων, χρησιμοποιήθηκε το

μοντέλο παλινδρόμησης όπως παρουσιάζει ο Πίνακας 4.1. Ωστόσο επειδή στην [201] το ύψος των ασθενών δεν γίνεται διαθέσιμο, το μέσο βάρος (32.3 kg) των ασθενών αυτών χρησιμοποιήθηκε ώστε να βρεθεί ένας ασθενής με παρόμοιο βάρος στην [196] και έτσι να ληφθεί μία, έστω κατά προσέγγιση, εκτίμηση για την μέση τιμή της BSA στην [201]. Επιλέχθηκε ασθενής με βάρος 36 kg και ύψος 131 cm. Επομένως η τιμή της BSA υπολογίστηκε στο 1.09 m² μέσω της φόρμουλας Gehan & George [202]. Θα πρέπει να σημειωθεί ωστόσο, πως οι προσεγγίσεις αυτές χρησιμοποιούνται μόνο προκειμένου να ληφθούν ρεαλιστικές αρχικές τιμές για την διαδικασία εκτίμησης παραμέτρων. Οι τιμές αυτές δίνονται στον ακόλουθο πίνακα.

Πίνακας 4.6 Αρχικές εκτιμήσεις για τις τιμές των παραμέτρων του προτεινόμενου μοντέλου με βάση το μοντέλο παλινδρόμησης που δίνεται στην [196]

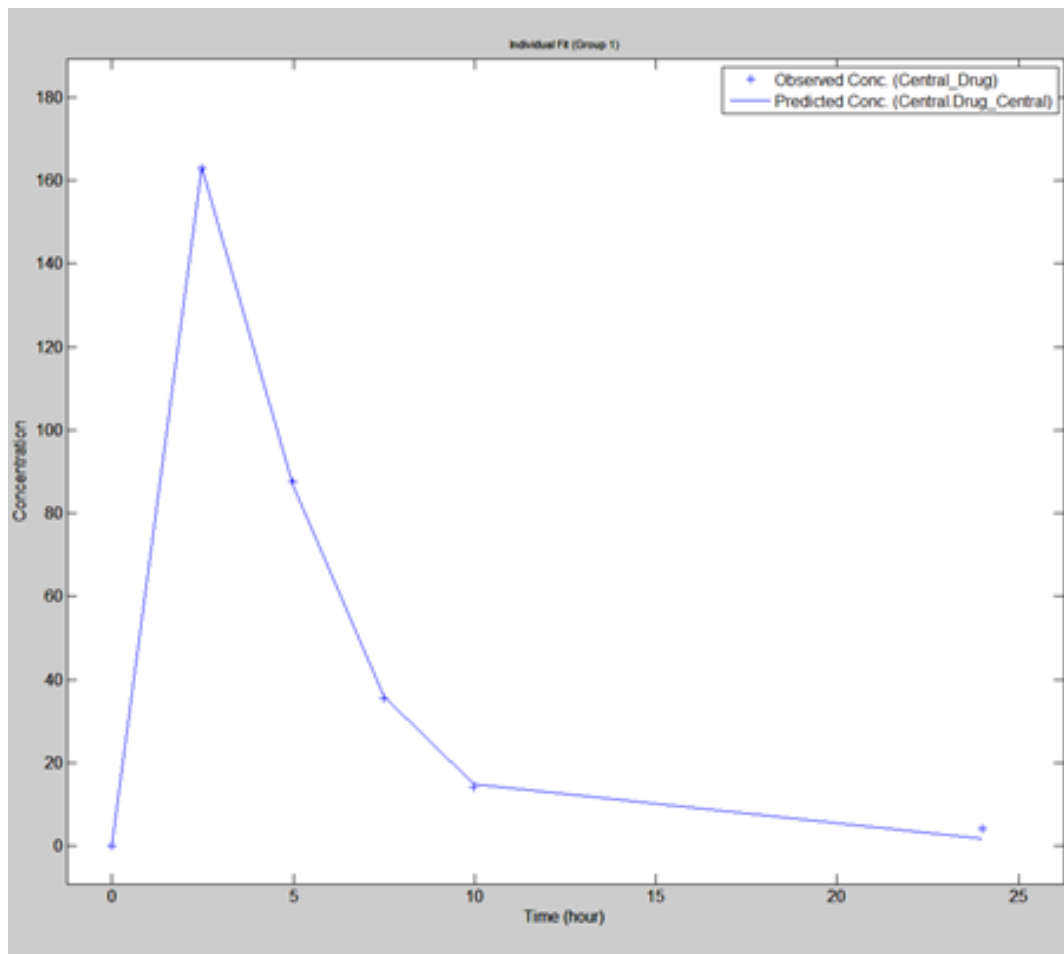
Παράμετρος	Αρχική Εκτίμηση
CL	33,463
CL_d	16,895 (χρήση του διαμέσου που δίνεται στην [196])
f_{PREDNL}	0.86
k_a	1.23 (μέση τιμή των παραμέτρων πρώτης τάξης στα Μοντέλα 1 και 2)
F	0.75
Vc	50.0

Μετά την εκτέλεση του αλγορίθμου εκτίμησης παραμέτρων λήφθηκαν οι εκτιμήσεις που δίνονται στον ακόλουθο πίνακα:

Πίνακας 4.7 Τελικές εκτιμήσεις παραμέτρων φαρμακοκινητικό μοντέλου

Παράμετρος	Εκτίμηση
CL	36.57618325
CL_d	3.94095906
f_{PREDNL}	0.61111965
k_a	0.89376099
F	0.6294353
Vc	75.65683239

Τέλος τα αποτελέσματα της εκτέλεσης του προτεινόμενου μοντέλου, συγκρινόμενα με τα διαθέσιμα δεδομένα δίνονται στην ακόλουθη εικόνα



Εικόνα 4.10 Αποτελέσματα εκτέλεσης του προτεινόμενου φαρμακοκινητικού μοντέλου, συγκρινόμενα με τα διαθέσιμα δεδομένα

Όπως είναι εμφανές, το προτεινόμενο μοντέλο, είναι ικανό να προσομοιώσει την φαρμακοκινητική του φαρμάκου PREDNL μετά την από στόματος χορήγηση PRED.

5. Προσαρμογή Ογκοπροσομοιωτή και Πρόβλεψη τιμών παραμέτρων του με τη χρήση δεδομένων έκφρασης γονιδίων

5.1 Εισαγωγή

Κοινός μακροπρόθεσμος στόχος των αναδυόμενων διεπιστημονικών τομέων της In Silico Ογκολογίας και Ιατρικής [123], [203] και της Πρωτοβουλίας για τον Εικονικό Φυσιολογικό Άνθρωπο (Virtual Physiological Human – VPH, <http://www.vph-institute.org/>), αποτελεί η ανάπτυξη υπολογιστικών μοντέλων που θα μπορούσαν να συμβάλουν στην εξατομίκευση (personalization) της θεραπείας των ασθενειών, κυρίως διαμέσου των πιθανών προβλεπτικών τους δυνατοτήτων. Μεταξύ διαφόρων άλλων σημαντικών συμβολών που αναφέρονται στην σχετική βιβλιογραφία, ο Ογκοπροσομοιωτής γενικά, όπως περιγράφεται στο Κεφάλαιο 2, αλλά και ο Ογκοπροσομοιωτής Λευχαιμίας (ALL Oncosimulator) ειδικά, αποτελούν μοντέλα τύπου VPH, με τον δεύτερο να στοχεύει ειδικά στην προσομοίωση της εξέλιξης της Οξείας Λεμφοβλαστικής Λευχαιμίας (ALL) και της απόκρισης αυτής σε θεραπεία. Στο παρόν Κεφάλαιο της διατριβής, παρουσιάζεται ένα σύνολο επιπλέον μεθοδολογιών που αναπτύχθηκαν και οι οποίες άπτονται των τομέων της υπολογιστικής βελτιστοποίησης (computational optimization) και της μηχανικής μάθησης (machine learning-ML), με στόχο την επέκταση των προβλεπτικών δυνατοτήτων του Ογκοπροσομοιωτή και την υποστήριξη της αξιολόγησής τους.

5.2 Πρόβλεψη τιμών Ογκοπροσομοιωτή και εξατομίκευση προσομοίωσης-προβλέψεων

Ο Ογκοπροσομοιωτής, όπως κάθε μοντέλο του τύπου αυτού, πρέπει να παραμετροποιηθεί όσο ακριβέστερα είναι δυνατό προκειμένου τα αποτελέσματα προσομοίωσης του να προσεγγίζουν την κλινική πραγματικότητα και να αποκτήσουν περιγραφική (descriptive) και πιθανόν προβλεπτική (predictive) αξία. Τα ανωτέρω υποστηρίζονται από ιδιότητες του μοντέλου όπως η παρατηρούμενη ευαισθησία (sensitivity) της απόκρισής του σε σχέση με την επιλογή των παραμέτρων εισόδου [204], [205].

Αναφερόμενοι σε παρελθοντικές/αναδρομικές (retrospective) περιπτώσεις ασθενών για τις οποίες είναι διαθέσιμα δεδομένα πριν, κατά τη διάρκεια αλλά και μετά τη θεραπεία, η εκτίμηση των τιμών παραμέτρων του μοντέλου μπορεί να επιτευχθεί με τη χρήση

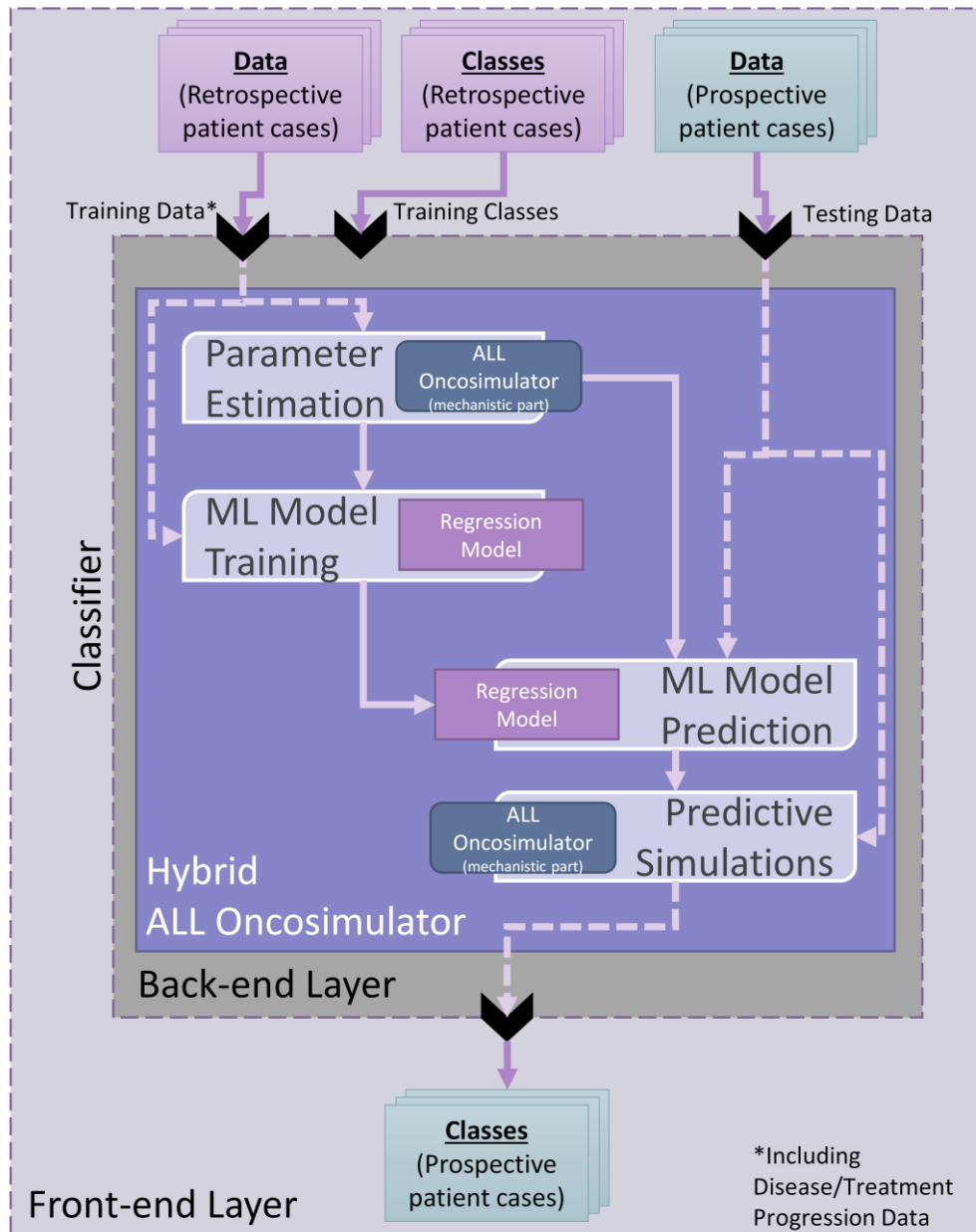
μεθόδων προσαρμογής/εκτίμησης παραμέτρων. Για μια νέα περίπτωση ασθενούς, ωστόσο, θα πρέπει να ακολουθηθεί μία εναλλακτική διαδικασία, λόγω της αποκλειστικής διαθεσιμότητας πληροφοριών, αποτελεσμάτων ιατρικών εξετάσεων και άλλων μορφών δεδομένων, πριν την έναρξη της θεραπείας. Επομένως, προκύπτει ξεκάθαρα η ανάγκη για την προσθήκη επιπλέον δυνατοτήτων-υπολογιστικών υπό-μονάδων (components) στον Ογκοπροσομοιωτή Λευχαιμίας. Αρχικά αυτές οι υπό-μονάδες θα πρέπει να είναι ικανές να ανιχνεύσουν το πλέον κατάλληλο σύνολο παραμέτρων εισόδου του Ογκοπροσομοιωτή Λευχαιμίας για τις αναδρομικές περιπτώσεις ασθενών (δηλαδή να βελτιστοποιούν το σύνολο παραμέτρων εισόδου έτσι ώστε το αποτέλεσμα προσομοίωσης του μοντέλου να συμφωνεί αρκούντως με την κατάσταση της ασθένειας του ασθενούς μετά τη θεραπεία). Επιπλέον, όμως, θα πρέπει να είναι ικανές να προβλέψουν, όσον το δυνατόν με μεγαλύτερη ακρίβεια, το σύνολο παραμέτρων εισόδου που θα οδηγήσει σε ορθή πρόβλεψη της απόκρισης στη θεραπεία ενός νέο-εισαγόμενου ασθενούς, πριν όμως αυτή η θεραπεία χορηγηθεί.

Έτσι, πρώτον, θα πρέπει να μελετηθούν οι μέθοδοι υπολογιστικής βελτιστοποίησης για την προσαρμογή/ εκτίμηση παραμέτρων της εισόδου του Ογκοπροσομοιωτή για τις αναδρομικές περιπτώσεις ασθενών. Δεύτερον, θα πρέπει να αναπτυχθούν μέθοδοι μηχανικής μάθησης που θα επιχειρήσουν να εξάγουν/μάθουν τη σχέση μεταξύ εξατομικευμένων δεδομένων ασθενών και τις βέλτιστης παραμετροποίησης του Ογκοπροσομοιωτή για κάθε ξεχωριστή περίπτωση του ασθενούς. Ο συνδυασμός αυτών των μεθόδων με τον Ογκοπροσομοιωτή, οδηγεί στην δημιουργία ενός υβριδικού υπολογιστικού μοντέλου που αποτελείται από ένα μηχανιστικό μέρος και από διάφορες υπολογιστικές υπο-μονάδες βελτιστοποίησης και μηχανικής μάθησης. Ο συνδυασμός αυτός ονομάζεται στα πλαίσια της διατριβής ως «Υβριδικός Ογκοπροσομοιωτής Λευχαιμίας» (Hybrid ALL Oncosimulator).

Στα πλαίσια της διατριβής, το συγκεκριμένο μοντέλο χρησιμοποιείται προκειμένου να προβλεφθεί το αποτέλεσμα της θεραπείας προ-φάσης (pre-phase) σε παιδιά με ALL. Το βασικό κριτήριο για την αξιολόγηση του μοντέλου ήταν η ακρίβειά του στην ταξινόμηση των ασθενών μεταξύ των κατηγοριών απόκρισης στο φάρμακο Πρεδνιζόνη (Prednisone). Για την αξιολόγηση αυτή επιλέχθηκε η μεθοδολογία αντεπικύρωσης (cross-validation), με δεδομένο ότι έγιναν διαθέσιμες μόνο αναδρομικές περιπτώσεις ασθενών, η οποία παρουσιάζεται λεπτομερώς στις ακόλουθες υπό-ενότητες. Λαμβάνοντας υπόψη την προγνωστική αξία της ταξινόμησης αυτής στην κλινική πράξη (Ενότητα 1.4.3), η ορθή

πρόβλεψη της κατηγορίας απόκρισης στο Prednisone, πριν την έναρξης της θεραπείας, θα διευκόλυνε την λήψη βέλτιστων επιλογών θεραπείας από τον κλινικό σαφώς νωρίτερα.

Παρά την αναμενόμενη πολυπλοκότητα του προτεινόμενου συνδυαστικού συστήματος, ο τελικός χρήστης αναμένεται να αλληλοεπιδρά με αυτό όπως θα αλληλοεπιδρούσε με οποιονδήποτε ταξινομητή (classifier). Μία απλουστευμένη παρουσίαση του μοντέλου, η οποία υποδεικνύει και τον τρόπο επικοινωνίας του χρήστη με το σύστημα δίνεται στην ακόλουθη εικόνα.



Εικόνα 5.1 Η περιληπτική δομή του Υβριδικού Ογκοπροσομοιωτή Λευχαιμίας και τα επίπεδα διεπαφής με το χρήστη-προσκήνιο (front-end) και παρασκήνιο (back-end).

Όπως φαίνεται στην Εικόνα 5.1, στο επίπεδο προσκήνιο (Front-end Layer) ο Hybrid ALL Oncosimulator αλληλεπιδρά με το χρήστη όπως ένας κοινός ταξινομητής.

Δηλαδή εισάγονται τα δεδομένα και οι κλάσεις των αναδρομικών περιπτώσεων ασθενών (Retrospective patient cases' Data and Classes) και με βάση τα δεδομένα των νέων περιπτώσεων ασθενών (Prospective patient cases' Data) εξάγεται η πρόβλεψη ταξινόμησης των νέων ασθενών (Prospective patient cases' classes). Στο επίπεδο παρασκηνίου (Backend Layer), τα δεδομένα που διατίθενται για τις αναδρομικές περιπτώσεις ασθενών, τα οποία θα πρέπει να περιέχουν δεδομένα εξέλιξης της ασθένειας και της θεραπείας (Disease/Treatment Progression Data), χρησιμοποιούνται για την εκτίμηση των παραμέτρων του Ογκοπροσομοιωτή και για την εκπαίδευση ενός μοντέλου παλινδρόμησης (regression model), το οποίο χρησιμοποιείται κατόπιν για την πρόβλεψη των τιμών των παραμέτρων αυτών για τις περιπτώσεις των νέων ασθενών (prospective patient cases). Οι προβλεφθείσες τιμές, τέλος, εισάγονται στον Ογκοπροσομοιωτή για την εξαγωγή της τελικής ταξινόμησης μέσω πολλαπλών προσομοιώσεων όπως θα περιγραφεί λεπτομερώς στις ακόλουθες υπό-ενότητες.

5.3 Διαθέσιμα δεδομένα και προετοιμασία τους

Τα δεδομένα που χρησιμοποιήθηκαν στο παρόν κεφάλαιο της διατριβής προέρχονται από έναν πληθυσμό 191 ασθενών με Οξεία Λεμφοβλαστικής Λευχαιμία παιδικής ηλικίας που συμμετείχαν στην σειρά κλινικών δοκιμών ALL-BFM 2000 και έγιναν διαθέσιμα σε ανωνυμοποιημένη μορφή στα πλαίσια του ερευνητικού προγράμματος P-medicine (FP7-ICT-2009-6-270089). Για τη συγκεκριμένη ομάδα ασθενών, σε συνδυασμό με τις κλινικές παραμέτρους που συλλέγονται στο κλινικό περιβάλλον και αφορούν την εξέλιξη της ασθένειας (μετρήσεις λευχαιμικών βλαστών-leukemic blast counts), μετρήθηκε επίσης η έκφραση του συνόλου του γονιδιώματος (whole genome expression). Το μοριακό μέρος αυτό του συνόλου δεδομένων αποτελείται από log₂ κανονικοποιημένους μέσους λόγους R/G μετρήσεων για 39778 δείκτες (probes) οι οποίοι είχαν σημασιολογικά περιγραφεί (annotated) μέσω IMAGE CloneIDs μοναδικών κλειδιών. Λεπτομέρειες σχετικά με την διαδικασία συγκέντρωσης και εξαγωγής των δεδομένων έκφρασης γονιδίων δίνονται στην [206].

Όλοι οι ασθενείς έλαβαν θεραπεία προ-φάσης (pre-phase) η οποία αποτελείται από 7 ημέρες από στόματος (orally) μονοθεραπεία με Πρεδνιζόνη (60 mg/m² ανά ημέρα) και μία δόση ενδοραχιαίας (intrathecal) Μεθοτρεξάτης (Methotrexate) στην πρώτη ημέρα της θεραπείας [207]. Κριτήρια επιλογής (eligibility criteria) ασθενών από το σύνολο δεδομένων αποτέλεσαν εκτός από την διαθεσιμότητα των δεδομένων έκφρασης γονιδίων, η διαθεσιμότητα μετρήσεων των λευχαιμικών βλαστών πριν τη θεραπεία αλλά και στην 8^η

μέρα αυτής (μετά δηλαδή την ολοκλήρωση της προ-φάσης). Επίσης, όπως προτείνεται στην [206], οι ασθενείς που εμφανίζουν τις χρωμοσωμικές αναδιατάξεις (re-arrangements) BCR-ABL, MLL-AF4 και TEL-AML1 όπως επίσης και οι ασθενείς με δείκτη (index) DNA διαφορετικό του 1, εξαιρέθηκαν από την περαιτέρω ανάλυση, καταλήγοντας σε 87 περιπτώσεις ασθενών που μπορούσαν να αξιοποιηθούν. Στον ακόλουθο πίνακα, δίνονται διάφορες στατιστικές του τελικώς διαθέσιμου συνόλου δεδομένων.

Πίνακας 5.1 Στατιστικές πληροφορίες των διαθέσιμων και αξιοποιήσιμων δεδομένων

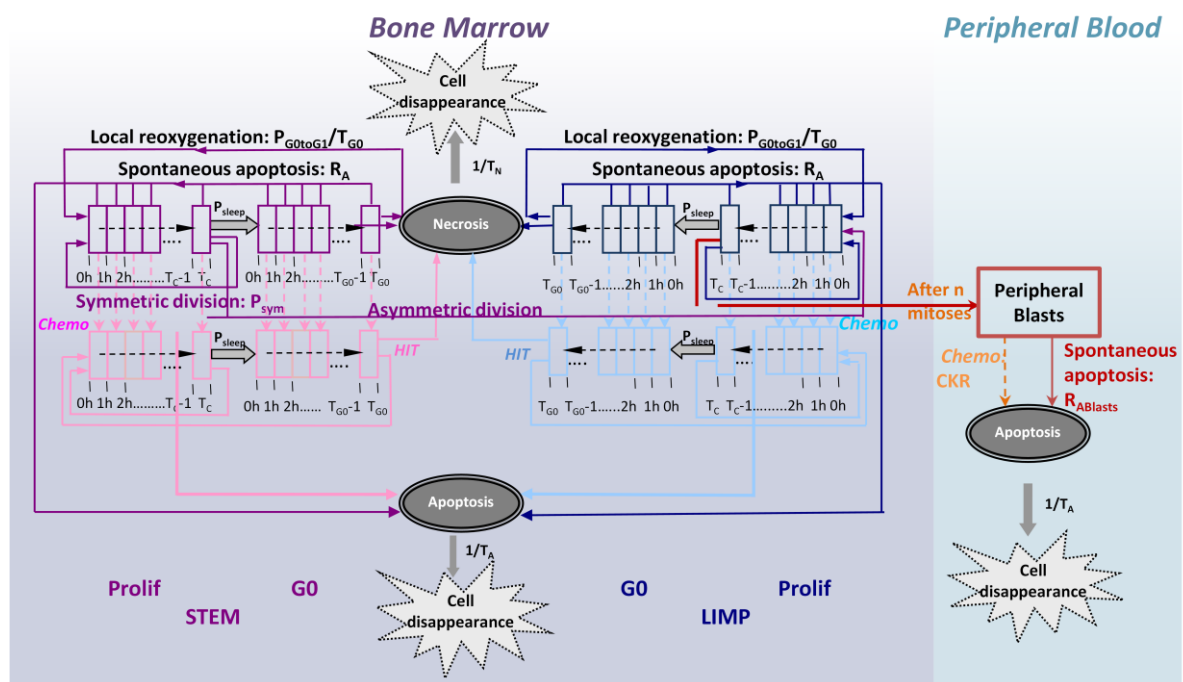
Παράγοντας		
Ηλικία	<10 χρ. (0.54%)	>10 χρ. (0.46%)
Φύλο	Θήλυ (41%)	Άρρεν (59%)
Αριθμός λευκοκυττάρων κατά την διάγνωση (λεμφοβλάστες/μl)	<50.000 (41%)	>50.000 (59%)
Απόκριση στην Πρεδνιζόνη	Καλή (44%)	Κακή (56%)
Ανοσοφαινότυπος	B-σειράς ALL (78%)	T-σειράς ALL (22%)

5.4 Ο Ογκοπροσομοιωτής Λευχαιμίας

Ο Ογκοπροσομοιωτής Λευχαιμίας (ALL Oncosimulator) αποτελεί μία σημαντική τροποποίηση των αλγορίθμων και μοντέλων (Ογκοπροσομοιωτών) που έχουν αναπτυχθεί από την ομάδα για την In Silico Ογκολογία και την In Silico Ιατρική (<http://in-silico-oncology.iccs.ntua.gr/>) όπως αυτά δίνονται για παράδειγμα στις εργασίες [117], [123]–[126], [204] ενώ βασίζεται στην γενικότερη προσέγγιση του Ογκοπροσομοιωτή μη χωρικής προσέγγισης (όπως παρουσιάστηκε στην Ενότητα 2.5.2). Το συγκεκριμένο μοντέλο αναπτύχθηκε στα πλαίσια του προγράμματος p-medicine (<http://www.p-medicine.eu/>) ενώ παρουσιάστηκε πρωτότυπα στην [130]. Οι απαραίτητες για τον ορισμό του Ογκοπροσομοιωτή Λευχαιμίας αλλαγές, καθορίστηκαν κυρίως από την ειδική φύση της Λευχαιμίας η οποία αποτελεί ένα μη-συμπαγή όγκο (non-solid tumour).

Όπως οι περισσότεροι Ογκοπροσομοιωτές που αναφέρθηκαν προηγουμένως έτσι και ο Ογκοπροσομοιωτής Λευχαιμίας βασίζεται στην εκτεταμένη χρήση των κυτταρικών αυτομάτων. Σε σχέση με τους Ογκοπροσομοιωτές χωρικής προσέγγισης, ο Ογκοπροσομοιωτής Λευχαιμίας, βασιζόμενος στη μη-χωρική προσέγγιση, παρουσιάζει νέα χαρακτηριστικά όπως το ότι θεωρεί μεγαλύτερο αριθμό διαμερισμάτων (compartments) για τα πολλαπλασιαζόμενα (proliferating) και τα κύτταρα που βρίσκονται σε φάση ύπνωσης (dormant cells). Στους Ογκοπροσομοιωτές χωρικής προσέγγισης (π.χ. [1,7,13–16]) τα εντός

του κυτταρικού κύκλου κύτταρα κατανέμονται σε 4 διαφορετικά διαμερίσματα (αντίστοιχα των φάσεων G1,S,G2,M) ενώ τα σε ύπνωση κύτταρα σε μία ξεχωριστή κατηγορία κυττάρων (G0). Ο Ογκοπροσομοιωτής Λευχαιμίας θεωρεί έναν αριθμό διαμερισμάτων για τα πολλαπλασιαζόμενα και για τα σε ύπνωση κύτταρα, ίσο με τον διακριτή διάρκεια του ενεργού κυτταρικού κύκλου ή της φάση ύπνωσης αντίστοιχα, εκφρασμένη σε ώρες. Συγκεκριμένα κάθε διαμέρισμα/κυτταρική κλάση αντιστοιχεί σε ένα χρονικό διάστημα ίσο με μία ώρα εντός της αντίστοιχης φάσης του κυτταρικού κύκλου όπως δίνεται στην Εικόνα 5.2.



Εικόνα 5.2 Γενικό κυτταροκινητικό μοντέλο του Ογκοπροσομοιωτή Λευχαιμίας για τα διαμερίσματα του μυελού των οστών (bone marrow) και του περιφερικού αίματος (peripheral blood) [130]. STEM: στελεχειαίο κύτταρο/stem cell. LIMP: Limited Mitotic Potential cancer cell. Prolif: Πολλαπλασιαζόμενα κύτταρα/Proliferating cell. G0: Κύτταρα σε ύπνωση/dormant cell. Chemo: Χημειοθεραπεία. HIT: Κύτταρα στα οποία έχει εκκινήσει η διαδικασία κυτταρικού θανάτου λόγω της θεραπείας

Παρόμοια με άλλους Ογκοπροσομοιωτές, οι κανόνες πολλαπλασιασμού των λευχαιμικών κυττάρων βασίζονται στην καλώς τεκμηριωμένη υπόθεση των βλαστικών/στελεχειαίων καρκινικών κυττάρων [145]–[147]. Επίσης, ορίζονται δύο βασικά κυτταρικά διαμερίσματα αυτά του Μυελού των οστών (bone marrow) και του Περιφερικού αίματος (peripheral blood). Το μοντέλο θεωρεί ότι τα λευχαιμικά βλαστικά κύτταρα βρίσκονται στον Μυελό των οστών και έχουν την δυνατότητα απεριόριστης ανανέωσης αλλά και έναρξης της διαδικασίας διαφοροποίησης (Εικόνα 5.2). Για το συγκεκριμένο διαμέρισμα θεωρούνται τρεις επιπλέον κατηγορίες κυττάρων:

- Κύτταρα πεπερασμένης δυνατότητας πολλαπλασιασμού (Limited Mitotic Potential, LIMP)
- Τα αποπτωτικά κύτταρα (apoptotic cells)
- Νεκρωτικά κύτταρα (necrotic cells)

Τα βλαστικά κύτταρα όπως και τα LIMP Κύτταρα μπορούν είτε να πολλαπλασιάζονται είτε να βρίσκονται σε ύπνωση. Το περιφερικό αίμα, στο οποίο τα λευχαιμικά κύτταρα κυκλοφορούν (circulate), αποτελεί το δεύτερο βασικό κυτταρικό διαμέρισμα του μοντέλου. Ως μία πρώτη προσέγγιση, τα λευχαιμικά κύτταρα στο περιφερικό αίμα θεωρούνται ότι βρίσκονται σε ύπνωση (εκτός κυτταρικού κύκλου) και έχουν οδηγηθεί μετά από έναν αριθμό μιτώσεων (παράμετρος τους μοντέλου) σε κατάσταση όμοια αυτής της διαφοροποίησης.

Όπως εν γένει ο Ογκοπροσομοιωτής (όπως παρουσιάστηκαν στην Ενότητα 2.5), το μοντέλο προσομοιώνει μια πληθώρα βιο-μηχανισμών από το κυτταρικό και το υπέρ-κυτταρικό επίπεδο:

1. Την πρόοδο των πολλαπλασιαζόμενων κυττάρων εντός του ενεργού κυτταρικού κύκλου, την έξοδο αυτών προς την φάση ύπνωσης G0 αλλά και την επαν-είσοδό τους στο κυτταρικό κύκλο
2. Την συμμετρική και την ασύμμετρη διαίρεση των βλαστικών/στελεχιαίων κυττάρων (με αποτέλεσμα την δημιουργία νέων βλαστικών κυττάρων και LIMP κυττάρων).
3. Την ωρίμανση των LIMP κυττάρων μετά την πραγματοποίηση ενός αριθμού διαιρέσεων, την παύση (arrest) του κυτταρικού τους κύκλου και την είσοδό τους στο περιφερικό αίμα
4. Την απώλεια κυττάρων κυρίως μέσω της απόπτωσης, είτε αυθόρμητα (spontaneous apoptosis) είτε λόγω θεραπείας και στα δύο διαμερίσματα (μυελό των οστών και περιφερικό αίμα).

Οι κανόνες που διέπουν τις μεταβάσεις αυτές δίνονται στην Εικόνα 5.2 ενώ τις παραμέτρους που αφορούν τις συγκεκριμένες διαδικασίες παρουσιάζει ο Πίνακας 5.2.

Πίνακας 5.2 Παράμετροι εισόδου του Ογκοπροσομοιωτή Λευχαιμίας (μηχανιστικό μέρος) και εύρος τιμών τους (κατά την εκτίμηση παραμέτρων)

Παράμετρος	Περιγραφή	Εύρος τιμών κατά την εκτίμηση παραμέτρων	Αναφορές, τρόπος απόκτησης τιμής
T_C	Διάρκεια κυτταρικού κύκλου	24-200 h	[151], [208]–[210]
T_{G0}	Διάρκεια της φάσης ύπνωσης (G0)	0-120 h	(εκτιμήθηκε) *
T_N	Χρόνος που απαιτείται για την ολοκλήρωση της νέκρωσης και για την απομάκρυνση των προϊόντων της από τον μυελό των οστών	100-140 h	[211], (εκτιμήθηκε) *
T_A	Χρόνος που απαιτείται για την ολοκλήρωση της απόπτωσης και για την απομάκρυνση των προϊόντων της	6 h	[162]
N_{LIMP}	Αριθμός μιτώσεων που πραγματοποιούνται από ένα LIMP [†] κύτταρο πριν την παύση του κυτταρικού κύκλου	7	[130],(παραδοχή) [#]
R_A	Ποσοστό κυττάρων (στελεχιαία και LIMP) που πεθαίνουν μέσω Απόπτωσης ανά ώρα	0.0001-0.1 h ⁻¹	(εκτιμήθηκε) *
$R_{A(Peripheral\ Blasts)}$	Ποσοστό κυττάρων στο περιφερικό αίμα που πεθαίνουν μέσω Απόπτωσης ανά ώρα	τέθηκε ίσο με R_A	
P_{G0toG1}	Ποσοστό κυττάρων που βρίσκονταν σε ύπνωση (στελεχιαία και LIMP) που μόλις επέστρεψαν στον κυτταρικό κύκλο	0.005-0.9	(εκτιμήθηκε) *
P_{sleep}	Ποσοστό κυττάρων που μεταβαίνουν σε ύπνωση μετά την μίτωση	0.001-0.3	(εκτιμήθηκε) *
P_{sym}	Ποσοστό στελεχιαίων κυττάρων που πραγματοποιούν συμμετρική διαίρεση	0.2-0.8	(εκτιμήθηκε) *
CKR_{PRED}	Ποσοστό κυττάρων που πεθαίνουν λόγω της θεραπείας με Πρεδνιζόνη (Cell kill rate of Prednisone)	0-0.8	(εκτιμήθηκε) *
CKR_{MTX}	Ποσοστό κυττάρων που πεθαίνουν λόγω της θεραπείας με Μεθοτρεξάτη (Cell kill rate of Methotrexate)	0.2	(παραδοχή) [#]
[†] LIMP: •Κύτταρα πεπερασμένης δυνατότητας πολλαπλασιασμού * Η ακριβής τιμή της παραμέτρου εκτιμήθηκε κατά την διαδικασία εκτίμησης παραμέτρων [#] Η τιμή της παραμέτρου επιλέχθηκε μέσω υπόθεσης/παραδοχής			

Σχετικά με την προσομοίωση της θεραπείας της ALL, τα κύτταρα του μυελού των οστών που προσβάλλονται θανάσιμα από ένα φάρμακο, εισάγονται σε έναν εκφυλισμένο (rudimentary) κυτταρικό κύκλο που οδηγεί στον αποπτωτικό θάνατο διαμέσου συγκεκριμένης φάσης. Η προσβολή του φαρμάκου εφαρμόζεται σε ένα ποσοστό των

καρκινικών κυττάρων το οποίο καθορίζεται από την ευαισθησία των κυττάρων αυτών στο συγκεκριμένο χορηγούμενο φάρμακο. Επί αυτού ορίζεται μία συγκεκριμένη παράμετρος του Ογκοπροσομοιωτή Λευχαιμίας, η CKR (Cell Kill Rate), για κάθε χορηγούμενο φάρμακο. Στα πλαίσια της παρούσας διατριβής, ορίζονται δύο τέτοιες παράμετροι, μία για το φάρμακο Πρεδνιζόνη (Prednisone, CKR_{PRED}) και μία για το φάρμακο Μεθοτρεξάτη (Methotrexate, CKR_{MTX}). Η ακριβής φάση του κυτταρικού κύκλου από την οποία τα κύτταρα εξέρχονται αυτού και εισέρχονται στην διαδικασία απόπτωσης, καθορίζεται από τον μηχανισμό δράσης του εκάστοτε φαρμάκου. Η Μεθοτρεξάτη, η οποία αποτελεί ανάλογο (έχει παρόμοια δομή) του φυλλικού οξέος, παρουσιάζει δράση στην S φάση [212], θεωρείται πως απορροφάται σε όλες τις φάσεις του κυτταρικού κύκλου ενώ ο αποπτωτικός θάνατος λόγω της χορήγησής της λαμβάνει χώρα κατά την S φάση. Η Πρεδνιζόνη, η οποία αποτελεί μη ειδικό του κυτταρικού κύκλου φάρμακο (cell cycle non-specific drug), θεωρείται πως επηρεάζει τα κύτταρα σε όλες τις φάσεις του κυτταρικού κύκλου όπως και στη G0 ενώ ο κυτταρικός θάνατος λόγω αυτής πραγματοποιείται στο τέλος της G1 φάσης [213]. Τέλος, με βάση την τρέχουσα υλοποίηση του Ογκοπροσομοιωτή Λευχαιμίας, μόνο η Πρεδνιζόνη θεωρείται πως έχει κυτταροτοξική επιρροή επί των λευχαιμικών κυττάρων που βρίσκονται στο περιφερικό αίμα.

Ο Ογκοπροσομοιωτής Λευχαιμίας (μηχανιστικό μοντέλο) έχει υλοποιηθεί από την ομάδα για την In Silico Ογκολογία και την In Silico Ιατρική με τη χρήση της γλώσσας προγραμματισμού C++. Για τις ανάγκες της παρούσας διατριβής το μοντέλο χρησιμοποιήθηκε και σε εκτελέσιμη μορφή (.exe) αλλά και ως βιβλιοθήκη δυναμικής σύνδεσης (Dynamic-link library, .dll). Η .dll μορφή επιλέχθηκε συγκεκριμένα για την αποδοτική ενσωμάτωση του μοντέλου στο περιβάλλον της γλώσσας R [214] στην οποία έγινε η ανάπτυξη του Υβριδικού Ογκοπροσομοιωτή Λευχαιμίας όπως αυτός παρουσιάζεται στις ακόλουθες ενότητες. Για την ανάγκες της ενσωμάτωσης αυτής χρησιμοποιήθηκε η διεπαφή .C της γλώσσας R.

5.5 Υπό-μονάδες μηχανικής μάθησης του Υβριδικού

Ογκοπροσομοιωτή

Όπως ήδη συζητήθηκε σε προηγούμενες υπό-ενότητες, η δυνατότητα του μηχανιστικού σκέλους του Ογκοπροσομοιωτή να προσομοιώσει με ακρίβεια ή να προβλέψει την εξέλιξη της ασθένειας και του αποτελέσματος της θεραπείας αυτής για μία δοθείσα περίπτωση ασθενούς καθορίζεται σημαντικά από την επιλογή των τιμών των

παραμέτρων του. Επομένως, ένας συνδυασμός μεθόδων βελτιστοποίησης και μηχανικής μάθησης με το μηχανιστικό σκέλος του Ογκοπροσομοιωτή Λευχαιμίας προκειμένου να δημιουργηθεί ο Υβριδικός Ογκοπροσομοιωτής Λευχαιμίας προτείνεται και αξιολογείται στην παρούσα διατριβή. Οι προστιθέμενες επιπλέον υπό-μονάδες καλύπτουν τις ανάγκες για:

1. Την αυτόματη και ακριβή εκτίμηση παραμέτρων για ένα σύνολο αναδρομικών περιπτώσεων ασθενών
2. Την δημιουργία μια διαδικασίας που θα επιτρέψει την πρόβλεψη των πλέον κατάλληλων τιμών παραμέτρων για νέες περιπτώσεις ασθενών, με βάση την γνώση που έχει αποκτηθεί από αναδρομικές περιπτώσεις.

Προφανώς, όσο πιο ακριβής επιτευχθεί να είναι αυτή η πρόβλεψη παραμέτρων, τόσο ακριβέστερη θα είναι η τελική πρόβλεψη του αποτελέσματος της θεραπείας (π.χ. Κατηγορίας απόκρισης στην Πρεδνιζόνη) μέσω προσομοιώσεων του ALL Ογκοπροσομοιωτή.

Στις ακόλουθες υπό-ενότητες παρουσιάζονται λεπτομερώς οι επιλογές που έγιναν για την υλοποίηση αυτών των υπό-μονάδων.

5.5.1 Εκτίμηση παραμέτρων μηχανιστικού μοντέλου

Όπως μπορεί να εξαχθεί από την παρουσίαση του Ογκοπροσομοιωτή Λευχαιμίας (Ενότητα 5.4), δύο βασικά σύνολα παραμέτρων πρέπει να εκτιμηθούν [130]. Η πρώτη περιλαμβάνει τις παραμέτρους που αφορούν στις ιδιότητες αύξησης του όγκου (παραμέτροι T_c έως P_{sym} , Πίνακας 5.2), ενώ η δεύτερη περιλαμβάνει τις σχετικές με την χημειοευαισθησία του κυττάρου παραμέτρους (CKRs).

Εν γένει, ο στόχος της διαδικασίας εκτίμησης παραμέτρων είναι να ελαχιστοποιηθεί η διαφορά μεταξύ της προσομοιωμένης εξέλιξης του όγκου (παρουσία είτε απουσία θεραπείας) και της κλινικά παρατηρούμενης εξέλιξης της ασθένειας για τις αναδρομικές περιπτώσεις ασθενών. Ωστόσο, για τις σχετικές με την ελεύθερη ανάπτυξη του όγκου παραμέτρους, η εκτίμησή τους απευθείας από τα δεδομένα είναι συνήθως μη εφικτή/πρακτική. Αυτό είναι αποτέλεσμα του ότι συνήθως τα διαθέσιμα κλινικά σύνολα δεδομένων δεν περιλαμβάνουν συγκεκριμένους και εξατομικευμένους για κάθε ασθενή δείκτες πολλαπλασιασμού ή μετρήσεις του μεγέθους του όγκου (στην λευχαιμία του αριθμού των λεμφοβλαστών) για μία αρκούντως μεγάλη περίοδο πριν την έναρξη της θεραπείας. Σε αυτή την περίπτωση, περιγραφικά χαρακτηριστικά της αύξησης του όγκου όπως αυτά αναφέρονται στην βιβλιογραφία, μπορούν να χρησιμοποιηθούν προκειμένου να

εκτιμηθούν τα σύνολα παραμέτρων που είναι ικανά να οδηγήσουν στην προσομοίωση ενός όγκου με συγκεκριμένα χαρακτηριστικά [130]. Για παράδειγμα, ο χρόνος διπλασιασμού (Doubling Time, T_d) αλλά και οι δείκτες (indexes) ή τα κλάσματα (fractions) αύξησης (growth), απόπτωσης (apoptotic) και νέκρωσης (necrotic) του υπό εξέταση όγκου μπορούν να χρησιμοποιηθούν.

Τα όσα περιεγράφηκαν παραπάνω ισχύουν και για το διαθέσιμο στα πλαίσια της διατριβής σύνολο δεδομένων. Επομένως, εύλογα σύνολα παραμέτρων που οδηγούν σε 6 διαφορετικές ρεαλιστικές τιμές για το χρόνο διπλασιασμού (T_d) της ALL (7-42 μέρες) [150], [162], [215] εκτιμήθηκαν ελαχιστοποιώντας την ακόλουθη συνάρτηση κόστους (objective function):

$$T_{d_{objective_{fun_i}}} (T_C, T_{G0}, T_N, R_A, R_{A(Peripheral\ Blasts)}, P_{sleep}, P_{sym}, P_{G0toG1}) \\ = \left| T_{d_{simulated}} (T_C, T_{G0}, T_N, R_A, R_{A(Peripheral\ Blasts)}, P_{sleep}, P_{sym}, P_{G0toG1}) - T_{d_{objective_i}} \right|, i \\ = 1, \dots, 6$$

όπου $T_{d_{simulated}} (T_C, T_{G0}, T_N, R_A, R_{ABlast}, P_{sleep}, P_{sym}, P_{G0toG1})$ ο χρόνος διπλασιασμού του λευχαιμικού όγκου όπως εκτιμάται από τον Ογκοπροσομοιωτή Λευχαιμίας για συγκεκριμένες δοθείσες τιμές των παραμέτρων του και $T_{d_{objective_i}}$, ο χρόνος διπλασιασμού που τέθηκε ως στόχος (7,14,21,28,35,42 ημέρες αντίστοιχα). Η επιλογή της ανωτέρω παρουσιαζόμενης συνάρτησης κόστους καθορίστηκε από την ανάγκη για την απόκτηση παραμετροποιήσεων του Ογκοπροσομοιωτή που θα οδηγήσουν σε προσομοιωμένους όγκους με χρόνο διπλασιασμού όσο το δυνατόν εγγύτερο σε αυτούς που τέθηκαν ως στόχος. Επιπροσθέτως, η αμεσότητα της σύγκρισης μεταξύ προσομοιωμένου και στοχευόμενου χρόνου διπλασιασμού καθιστά την απλή απόλυτη διαφορά, η οποία συχνά χρησιμοποιείται σε αντίστοιχες εργασίες εκτίμησης παραμέτρων, αρκετή προκειμένου ο αλγόριθμος βελτιστοποίησης να συγκλίνει.

Η διαδικασία βελτιστοποίησης υπόκειται επίσης σε συγκεκριμένους περιορισμούς. Πρώτον, στο κριτήριο για αυτοσυντηρούμενη ανάπτυξη του μη θεραπευόμενου όγκου [204] (self-sustained untreated tumour) όπως παρουσιάστηκε κατά την παρουσίαση του Ογκοπροσομοιωτή (Ενότητα 2.5) και όπως ήδη υλοποιείται σε αυτόν και δεύτερον, σε συγκεκριμένους περιορισμούς για την δυναμική του όγκου (tumour dynamics) και για τα χαρακτηριστικά σύνθεσης (constitution characteristics) όπως δίνονται στον ακόλουθο πίνακα

Πίνακας 5.3 Περιορισμοί που τέθηκαν για τις ιδιότητες αύξησης της ALL κατά την διαδικασία εκτίμησης παραμέτρων

Περιορισμός	Αναφορές/Υποστήριξη
$0.7 \leq Growth\ Fraction_{Bone\ Marrow} \leq 1$	[142]
$0.001 \leq Apoptotic\ Fraction \leq 0.072$	[151]
$Necrotic\ Fraction \leq 0.02$	Επιτρέπεται ένα σχετικά μικρό ποσοστό νεκρωτικών κυττάρων

Η υλοποίηση της διαδικασίας βελτιστοποίησης έγινε μέσω της εκτίμησης αληθοφανών συνόλων τιμών παραμέτρων για τον Ογκοπροσομοιωτή Λευχαιμίας (στη λειτουργία ελεύθερης ανάπτυξης) χρησιμοποιώντας τον αλγόριθμο ολικής στοχαστικής βελτιστοποίησης Διαφοροποιημένης Εξέλιξης (Differential Evolution) [216] όπως υλοποιείται στο DEoptim πακέτο της R [217]. Ο Πίνακας 5.2 παρουσιάζει τις υπό-εκτίμηση παραμέτρους και τα διαστήματα που τέθηκαν για αυτές κατά τη διάρκεια της εκτίμησης παραμέτρων παρουσιάζονται (παράμετροι Tc έως Psym).

Ακολουθως, οι σχετικές τιμές CKR για κάθε ασθενή του διαθέσιμου συνόλου δεδομένων και για κάθε T_d θα πρέπει να εκτιμηθούν. Αυτό μπορεί να επιτευχθεί ελαχιστοποιώντας την διαφορά μεταξύ της προβλεφθείσας από τον Ογκοπροσομοιωτή τιμής για τον αριθμό των λευχαιμικών κυτταρικών στο περιφερικό αίμα την ημέρα 8 μετά την έναρξη της θεραπείας και της πραγματικής τιμής που παρατηρείται στο σύνολο δεδομένων. Στα πλαίσια της διατριβής, και ελλείψει σημαντικών επιπλέον γνώσεων, η τιμή της CKR για το φάρμακο Μεθοτρεξάτη (Methotrexate, CKR_{MTX}) τέθηκε στην τιμή 0.2 [130]. Η τιμή αυτή επιλέχθηκε έτσι ώστε η κυτταροτοξική επίδραση του φαρμάκου να είναι σχετικά χαμηλή με δεδομένο ότι η δοθείσα δόση θεωρείται χαμηλή. Με δεδομένη την επιλογή της τιμής για την CKR_{MTX} , η CKR για το φάρμακο Πρεδνιζόνη (CKR_{PRED}) φράσσεται εντός του διαστήματος [0–0.8]. Με δεδομένη την παρούσα υλοποίηση του Ογκοπροσομοιωτή Λευχαιμίας, όπου οι τιμές των CKR παραμένουν σταθερές καθ'όλη τη διάρκεια της φάση θεραπείας, τη συγχορήγηση των MTX και PRED την πρώτη ημέρα της θεραπείας και με δεδομένο ότι τα δύο φάρμακα επενεργούν και στις ίδιες κατηγορίες κυττάρων, η συνολική τιμή του CKR (άθροισμα των δύο CKR στη συγκεκριμένη περίπτωση) δεν θα πρέπει να ξεπερνάει την τιμή 1 (δηλαδή στην αφαίρεση του 100% των κυττάρων στα οποία επενεργούν τα φάρμακα).

Η εκτίμηση της εξατομικευμένης τιμής για την CKR_{PRED} για κάθε T_d , πραγματοποιήθηκε με τη χρήση της συνάρτησης *optimize* του πακέτου *stats* [214] στη γλώσσα R μέσω της ελαχιστοποίησης της ακόλουθης συνάρτησης κόστους

$$\begin{aligned}
& CKR_{PRED_{objective}} \text{fun}_{i,j} (CKR_{PRED_{i,j}}) \\
&= \left| PBblasts_{Day8_{predicted}} \left(T_{c_i}, T_{G0_i}, T_{N_i}, R_{A_i}, R_{A(Peripheral\ Blasts)_i}, P_{sleep_i}, P_{sym_i}, P_{GotoG1_i}, CKR_{PRED_{i,j}}, PBblasts_{Day0_{observed}}(Patient_j) \right) \right. \\
&\quad \left. - PBblasts_{Day8_{observed}}(Patient_j) \right|
\end{aligned}$$

όπου

$i = 1, 2, \dots, 6$ (τα σύνολα παραμέτρων για τους διαφορετικούς χρόνους διπλασιασμού)

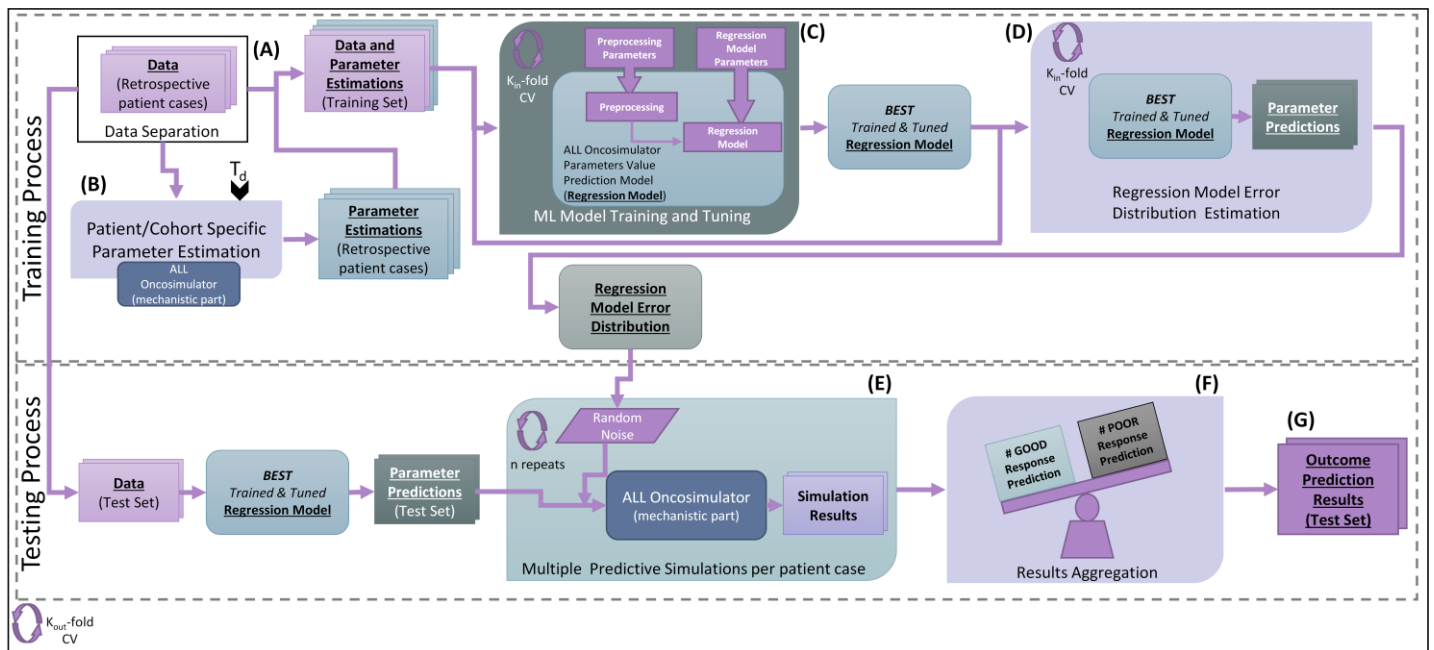
$j = 1, 2, \dots, 87$ (οι διαφορετικές περιπτώσεις ασθενών που μελετήθηκαν)

Οι λόγοι που καθόρισαν την επιλογή της συνάρτησης κόστους είναι παρόμοιοι με αυτούς που ίσχυσαν για την αντίστοιχη συνάρτηση κόστους για τους χρόνους διπλασιασμού. Τέλος θα πρέπει να σημειωθεί ότι η ανωτέρω περιγραφείσα διαδικασία, εκτελέστηκε ανεξάρτητα για κάθε περίπτωση ασθενούς. Επομένως δεν αναμένονται συσχετίσεις στην τιμή της CKR_{PRED} μεταξύ των ασθενών.

5.5.2 Δόμηση, υλοποίηση και εκπαίδευση συστήματος πρόβλεψης αποτελέσματος θεραπείας

Ακολουθώντας της παρουσίασης της διαδικασίας προσαρμογής του μηχανιστικού τμήματος του Ογκοπροσομοιωτή στα διαθέσιμα αναδρομικά δεδομένα, τα κομμάτια του Υβριδικού Ογκοπροσομοιωτή που οδηγούν στην πρόβλεψη του αποτελέσματος θεραπείας για μία νέα περίπτωση ασθενούς ορίζονται στην παρούσα ενότητα.

Το συνολικό σύστημα χωρίζεται σε 7 βασικά σύνολα βημάτων, συμπεριλαμβανομένων αυτών που επιτρέπουν την αξιολόγησή του μέσω της διαδικασίας αντεπικύρωσης (cross-validation). Τα βήματα αυτά παρουσιάζονται στην Εικόνα 5.3.



Εικόνα 5.3 Λεπτομερής δομή του Υβριδικού Ογκοπροσομοιωτή Λευχαιμίας (Hybrid ALL Oncosimulator) και βήματα για την αξιολόγηση του μέσω αντεπικύρωσης (cross-validation).

Το σύστημα υλοποιήθηκε με τη χρήση της γλώσσας προγραμματισμού R (εκδόσεις 3.2.1-3.3.1) και διαφόρων επιπλέον πακέτων αυτής, τα οποία σημειώνονται λεπτομερώς στη συνέχεια του κειμένου. Τα βήματα που παρουσιάζονται μπορούν να εκτελεστούν είτε μία είτε πολλές φορές (με μία λογική εξωτερικής αντεπικύρωσης) προκειμένου να εκτιμηθεί η μέση ακρίβεια πρόβλεψης. Στα πλαίσια της παρούσας παρουσίασης, πραγματοποιήθηκαν 5 εξωτερικές επαναλήψεις αντεπικύρωσης (cross-validation CVs, k_{out} -CV στην Εικόνα 5.3).

Στο βήμα (A), το σύνολο δεδομένων χωρίζεται τυχαία σε σύνολο εκπαίδευσης (train) και σε σύνολο δοκιμής (test) για τις ανάγκες της εξωτερικής αντεπικύρωσης. Το βήμα αυτό είναι απαραίτητο στην περίπτωση όπου η αξιολόγηση του συστήματος θα πρέπει να γίνει απουσία νέων περιπτώσεων ασθενών (prospective data), όπως συμβαίνει και στην περίπτωση της παρούσας διατριβής. Σε αντίθετη περίπτωση, το βήμα αυτό μπορεί να απαλειφθεί και τα δεδομένα νέων περιπτώσεων να τροφοδοτηθούν απευθείας στα σχετικά με την πρόβλεψη βήματα του συστήματος. Αυτή η προσέγγιση διαχωρισμού του συνόλου δεδομένων ακολουθείται ευρέως στη βιβλιογραφία για την αξιολόγηση ταξινομητών. Στα πλαίσια της διατριβής το βήμα αυτό υλοποιήθηκε με τη χρήση της συνάρτησης `createDataPartition` function (συγκεκριμένα την `createFolds` λειτουργία) του πακέτου `caret` στην γλώσσα R. Η συγκεκριμένη συνάρτηση χωρίζει το διαθέσιμο σύνολο δεδομένων σε μη αλληλεπικαλυπτόμενα μέρη, συγκεκριμένα το σύνολο Εκπαίδευσης (Train) και το σύνολο Δοκιμής (Test). Ο διαχωρισμός εξισορροπείται (is balanced) με βάση τις παρατηρούμενες κλάσεις των ασθενών στο σύνολο δεδομένων, συγκεκριμένα στην

παρούσα προσπάθεια η δυαδική κατηγορία απόκρισης στο φάρμακο Πρεδνιζόνη (δηλαδή οι αναλογίες των κλάσεων διατηρούνται στα σύνολα Εκπαίδευσης και Δοκιμής). Για τις ανάγκες της παρούσας αξιολόγησης του Υβριδικού Ογκοπροσομοιωτή Λευχαιμίας, δημιουργήθηκαν 5 μέρη(folds) του αξιοποιήσιμου μέρους του συνόλου δεδομένων. Συγκεκριμένα, το σύνολο δεδομένων χωρίστηκε σε 5 εξισορροπημένα μέρη και κάθε μέρος χρησιμοποιήθηκε ως το σύνολο Δοκιμής για κάθε εκτέλεση του συνολικού συστήματος (εξωτερική αντεπικύρωση, external cross-validation) ενώ τα υπόλοιπα 4 μέρη συνδυασμένα αποτέλεσαν τα σύνολα εκπαίδευσης.

Το παραπάνω βήμα ακολουθείται από το βήμα εκτίμησης παραμέτρων (Βήμα Β, Step B) για το σύνολο Εκπαίδευσης, όπως αυτό περιγράφηκε στην προηγούμενη ενότητα. Πριν την εκτέλεση του συγκεκριμένου βήματος, τα κλινικά δεδομένα όπως επίσης και η επιλογή των σεναρίων Td θα πρέπει να εισαχθούν στο σύστημα. Δηλαδή, θα πρέπει να επιλεχθούν οι χρόνοι διπλασιασμού που ο προσομοιωμένος όγκος θα παρουσιάζει. Στην περίπτωση που είναι επιθυμητή η επιλογή από ένα σύνολο προ-καθορισμένων χρόνων διπλασιασμού (όπως αυτό που παρουσιάστηκε στην προηγούμενη υπό-ενότητα), το μηχανιστικό μέρος του μοντέλου παραμετροποιείται άμεσα, με δεδομένο ότι οι σχετικές παράμετροι του Ογκοπροσομοιωτή (μπορούν να) έχουν εκτιμηθεί από πριν. Σε αντίθετη περίπτωση, οι παράμετροι που οδηγούν στον επιλεχθέντα χρόνο διπλασιασμού (όπως επίσης και οι σχετικές με την χημειοευαισθησία παράμετροι) θα πρέπει πρώτα να εκτιμηθούν (όπως παρουσιάζεται στην προηγούμενη υπό-ενότητα).

Στο επόμενο βήμα (Βήμα Γ, Step C) ένα μοντέλο με στόχο την πρόβλεψη των τιμών παραμέτρων του Ογκοπροσομοιωτή Λευχαιμίας (μοντέλο παλινδρόμησης, regression model) εκπαιδεύεται με βάση το σύνολο εκπαίδευσης. Οι παράμετροι του μοντέλου αυτού (μοντέλου παλινδρόμησης) βελτιστοποιούνται με τη χρήση μια εσωτερικής k-fold διαδικασίας αντεπικύρωσης (k_{in} -CV στην Εικόνα 5.3) με τη χρήση του πακέτου caret [218] στη γλώσσα R. Με βάση προηγούμενες δοκιμές και εμπειρία [130], [219], ο αλγόριθμος «τυχαία δάση» (Random Forests) [220] επιλέχθηκε προς αξιολόγηση. Επίσης δοκιμάστηκε ο κοινός αλλά ευρέως διαδεδομένος αλγόριθμος k-Κοντινότερων Γειτόνων με βάρη (Weighted k-Nearest Neighbours, k-NN) [221], ως ένα παράδειγμα απλούστερου αλγορίθμου σε σχέση με τον Random Forest. Και οι δύο αλγόριθμοι προσφέρονται από το πακέτο caret (και έχουν υλοποιηθεί στα πακέτα *Random Forests* [222] και *kkn* [223]).

Ο αλγόριθμος k-Κοντινότερων Γειτόνων είναι μία μη-παραμετρική μέθοδος (η δομή του μοντέλου δεν είναι προκαθορισμένη αλλά καθορίζεται από τα δεδομένα) μηχανικής μάθησης που χρησιμοποιείται για τις διεργασίες της ταξινόμησης αλλά και της

παλινδρόμησης [224]. Και στις δύο περιπτώσεις λαμβάνονται υπόψη οι k κοντινότερες ήδη γνωστές περιπτώσεις/δείγματα (από το σύνολο εκπαίδευσης) στο χώρο παραμέτρων. Η έξοδος του αλγόριθμου είναι:

- Στην περίπτωση της ταξινόμησης, η κλάση στην οποία προβλέπεται (μέσω του αλγορίθμου) να ανήκει το νέα δείγμα. Το νέο δείγμα ταξινομείται μέσω πλειοψηφικής ψηφοφορίας σε σχέση με τους κοντινότερους γείτονές του και λαμβάνει την κλάση η οποία πλειοψηφεί μεταξύ αυτών.
- Στην περίπτωση της παλινδρόμησης, η τιμή της εξαρτημένης μεταβλητής για το συγκεκριμένο δείγμα προκύπτει ως η μέση τιμή των αντίστοιχων τιμών για του κοντινότερους γείτονες.

Τα δείγματα εκπαίδευσης ουσιαστικά αποτελούν διανύσματα στον πολυδιάστατο χώρο παραμέτρων όπου το καθένα ανήκει σε μία από τις προκαθορισμένες κλάσεις και η φάση εκπαίδευσης του αλγορίθμου περιέχει μόνο την αποθήκευση αυτών. Για τη διαδικασία ταξινόμησης ή υπολογισμού της εξαρτημένης μεταβλητής ενός νέου δείγματος, υπολογίζεται εκ νέου η απόσταση του δείγματος αυτού (στο χώρο παραμέτρων) από τα δείγματα εκπαίδευσης και του ανατίθεται η πιο συχνά εμφανιζόμενη κλάση (ή η μέση τιμή της εξαρτημένης μεταβλητής) μεταξύ των k κοντινότερων δειγμάτων. Η τιμή της παραμέτρου k καθορίζεται από το χρήστη.

Για τον υπολογισμό της απόστασης μεταξύ του νέου και των γνωστών δειγμάτων χρησιμοποιούνται μετρικές απόστασης. Η πιο συχνά χρησιμοποιούμενη απόσταση για την περίπτωση όπου οι μεταβλητές-χαρακτηριστικά των δειγμάτων είναι συνεχής είναι η Ευκλείδεια απόσταση η οποία υπολογίζεται ως εξής:

$$d(x_i, x_j) = \left(\sum_{s=1}^p (x_{is} - x_{js})^2 \right)^{\frac{1}{2}}$$

όπου x_i, x_j δύο δείγματα ενώ $s = 1, \dots, p$ ο αριθμός των μεταβλητών/χαρακτηριστικών των δειγμάτων.

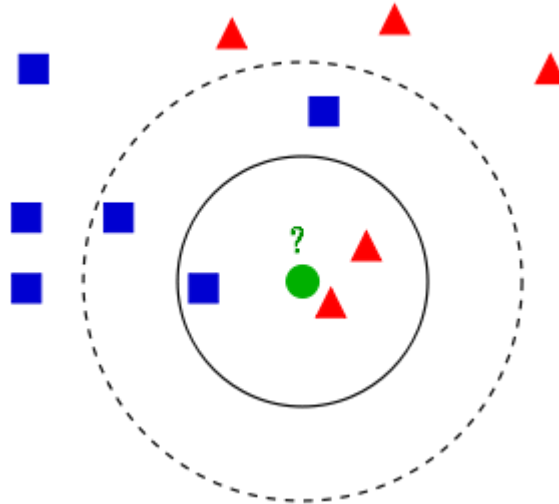
Η Ευκλείδεια απόσταση γενικεύεται στην απόσταση Minkowski η οποία ορίζεται ως:

$$d(x_i, x_j) = \left(\sum_{s=1}^p |x_{is} - x_{js}|^q \right)^{\frac{1}{q}}$$

όπου για $q = 2$ λαμβάνουμε την Ευκλείδεια απόσταση και για $q = 1$ την απόλυτη απόσταση.

Στην περίπτωση που τα χαρακτηριστικά είναι διακριτά, χρησιμοποιούνται διαφορετικές μετρικές όπως η απόσταση Hamming.

Ένα παράδειγμα ταξινόμησης μέσω του αλγορίθμου k-NN δίνεται στην ακόλουθη εικόνα.



Εικόνα 5.4 Παράδειγμα ταξινόμησης μέσω του k-NN για k=3 και για k=5

Στο συγκεκριμένο παράδειγμα χρησιμοποιείται η Ευκλείδεια απόσταση σε χώρο παραμέτρων δύο διαστάσεων. Όπως φαίνεται στην εικόνα, στην περίπτωση που η τιμή της παραμέτρου k τεθεί ίση με 3, το νέο δείγμα (πράσινο) θα ταξινομηθεί στην κόκκινη κλάση, ενώ τεθεί ίση με 5 θα ταξινομηθεί στην μπλε κλάση.

Ένα πιθανό μειονέκτημα της πλειοψηφικής διαδικασίας που παρουσιάστηκε εμφανίζεται ότι η κατανομή της κλάσης είναι ασύμμετρη (skewed), δηλαδή δείγματα που ανήκουν σε μία κλάση εμφανίζονται συχνότερα σε σχέση με αυτά της άλλης. Κάτι τέτοιο μπορεί να οδηγήσει σε μη επιθυμητή συνεχή επιλογή της συχνότερης κλάσης. Αντίστοιχα σε ένα πρόβλημα παλινδρόμησης, δείγματα που διαφέρουν σημαντικά, δηλαδή βρίσκονται σε μεγάλη απόσταση, από το νέο δείγμα, επηρεάζουν ισόποσα την προβλεφθείσα τιμή της εξαρτημένης μεταβλητής. Η προτεινόμενη λύση για την απάλειψη του προβλήματος, είναι η χρήση κατάλληλων βαρών έτσι ώστε να λαμβάνεται υπόψη η απόσταση των k κοντινότερων γειτόνων από το νέο δείγμα. Κατά την διαδικασία ψηφοφορίας (ή κατά τον υπολογισμό της μέσης τιμής για την εξαρτημένη μεταβλητή), η κλάση η εξαρτημένη μεταβλητή των k κοντινότερων δειγμάτων εκπαίδευσης, πολλαπλασιάζεται με ένα βάρος (weight, w) ανάλογα του αντίστροφου της απόστασης από το νέο δείγμα. Στην περίπτωση που δεν χρησιμοποιούνται βάρη, ο k-NN λειτουργεί παρόμοια με το απέδιδο βάρη ίσα με $1/k$ σε κάθε ένα από τα k κοντινότερα δείγματα και 0 στα υπόλοιπα. Σε περίπτωση που είναι επιθυμητή η χρήση βαρών, συχνά χρησιμοποιείται μία συνάρτηση πυρήνα για τον

υπολογισμό των βαρών αυτών. Μεταξύ άλλων επιλογές για την συνάρτηση πυρήνα αποτελούν ο τετραγωνικός (rectangular, ισοδύναμος με την περίπτωση χωρίς βάρη), ο τριγωνικός, ο *erpanechnikov* (beta(2,2)), ο διβαρής(biweight, beta(3,3)), ο τριβαρής(triweight, beta(3,3)), ο gaussian και ο optimal [225]. Για παράδειγμα, αν επιλεγθεί ο Gaussian kernel και για την περίπτωση μονοδιάστατου συνόλου δεδομένων, ο υπολογισμός των βαρών έχει ως ακολούθως

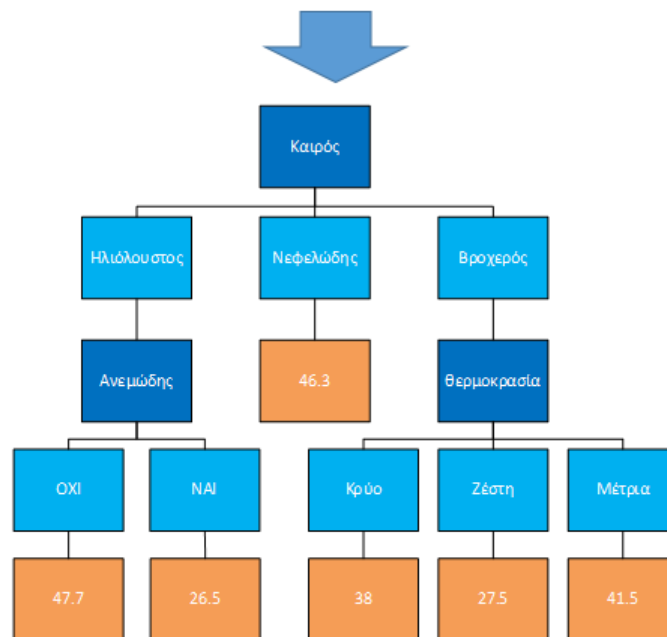
$$w_i = Kernel_{\lambda}(|x_i - x_q|) = \exp\left(-\frac{(x_i - x_q)^2}{\lambda}\right)$$

όπου i το i -οστό δείγμα από το σύνολο εκπαίδευσης, q το νέο δείγμα προς ταξινόμηση/παλινδρόμηση x_i, x_q οι αντίστοιχες τιμές χαρακτηριστικών τους και λ η παράμετρος που καθορίζει τον ρυθμό με τον οποίο μειώνεται η τιμή του βάρους όσο αλλάζει η απόσταση των δειγμάτων.

Ο αλγόριθμος των Τυχαίων Δασών (Random Forests) βασίζεται στη μέθοδο των Δέντρων Απόφασης (Decision Trees) [226] και στις αντίστοιχους αλγορίθμους για την μάθηση αυτών. Ανάλογα με τον τύπο της πληροφορίας για τον οποίο τίθεται ως στόχος η πρόβλεψη, τα Δέντρα Απόφασης χωρίζονται σε δύο βασικές κατηγορίες. Τα Δέντρα Ταξινόμησης (Classification Tree) και τα Δέντρα Παλινδρόμησης (Regression Trees). Στην πρώτη περίπτωση εκπαιδεύεται ένα δέντρο προκειμένου να μπορεί να προβλέψει την κλάση στην οποία ανήκουν τα δεδομένα, ενώ στη δεύτερη επιχειρείται η πρόβλεψη μίας συνεχούς μεταβλητής (εξαρτημένη μεταβλητή) με βάση τις τιμές των υπόλοιπων χαρακτηριστικών.

Για την κατασκευή ενός δέντρου απόφασης ανατίθεται αρχικά στη ρίζα του το σύνολο των δειγμάτων εκπαίδευσης. Κάθε ενδιάμεσος κόμβος περιέχει υποσύνολο των δειγμάτων το οποίο μέσω της εφαρμογής ενός κατάλληλου ελέγχου διαχωρίζεται σε δυο ή περισσότερα μικρότερα υποσύνολα στο επόμενο επίπεδο. Ο έλεγχος συνήθως αφορά ένα υποσύνολο των χαρακτηριστικών των δειγμάτων εκπαίδευσης. Η επιλογή του καλύτερου διαχωρισμού γίνεται σύμφωνα με ένα κατάλληλο μέτρο όπως ο δείκτης νοθείας Gini (Gini Impurity index)[226] ή το σφάλμα ταξινόμησης για τα δέντρα Ταξινόμησης και μετρήσεις όπως το Μέσο Τετραγωνικό Σφάλμα για τα Δέντρα Παλινδρόμησης. Ένα απλό παράδειγμα δέντρου παλινδρόμησης δίνεται στην ακόλουθη εικόνα.

Ανεξάρτητες Μεταβλητές				Εξαρτημένη Μεταβλητή
Καιρός	Θερμοκρασία	Υγρασία	Ανεμώδης	Ώρες Δραστηριότητας
Βροχερός	Ζέστη	Υψηλή	ΟΧΙ	26
Βροχερός	Ζέστη	Υψηλή	ΝΑΙ	30
Νεφελώδης	Ζέστη	Υψηλή	ΟΧΙ	48
Ηλιόλουστος	Μέτρια	Υψηλή	ΟΧΙ	46
Ηλιόλουστος	Κρύο	Κανονική	ΟΧΙ	62
Ηλιόλουστος	Κρύο	Κανονική	ΝΑΙ	23
Νεφελώδης	Κρύο	Κανονική	ΝΑΙ	43
Βροχερός	Μέτρια	Υψηλή	ΟΧΙ	36
Βροχερός	Κρύο	Κανονική	ΟΧΙ	38
Ηλιόλουστος	Μέτρια	Κανονική	ΟΧΙ	48
Βροχερός	Μέτρια	Κανονική	ΝΑΙ	48
Νεφελώδης	Μέτρια	Υψηλή	ΝΑΙ	62
Νεφελώδης	Ζέστη	Κανονική	ΟΧΙ	44
Ηλιόλουστος	Μέτρια	Υψηλή	ΝΑΙ	30

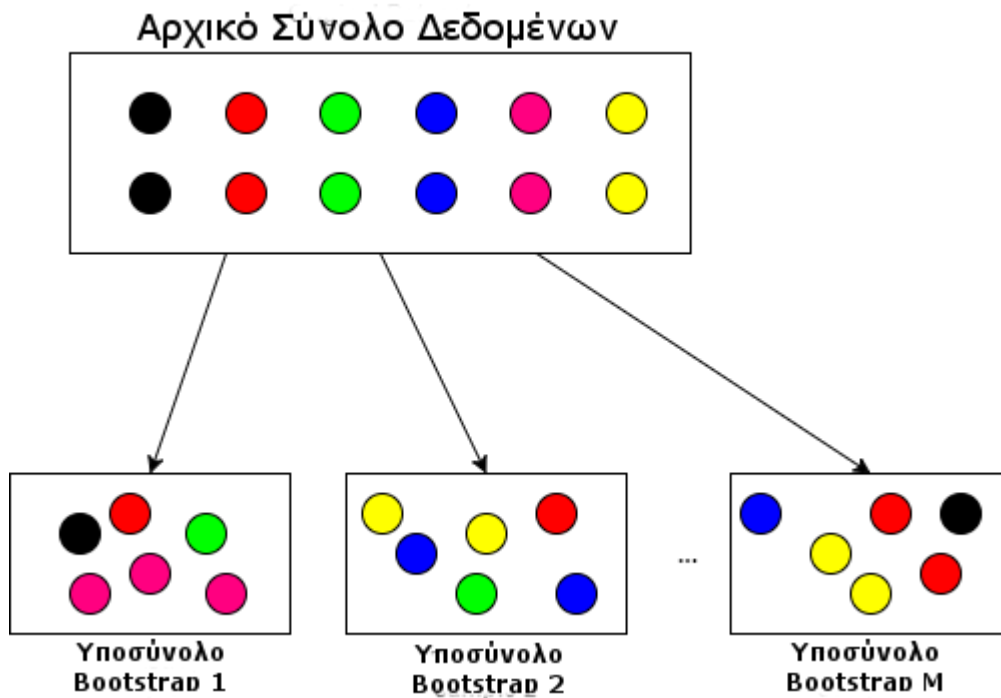


Εικόνα 5.5 Παράδειγμα δέντρου παλινδρόμησης με κατηγορικές ανεξάρτητες μεταβλητές. Η εξαρτημένη μεταβλητή περιέχει τις ώρες (σε εβδομαδιαία βάση) δραστηριότητας στην ύπαιθρο ανάλογα με τις μεταβλητές του καιρού.

Ο αλγόριθμος των τυχαίων δασών (Random Forests) βασίζεται στην προσέγγιση της συλλογικής μάθησης (ensemble learning) [227] και ουσιαστικά αποτελεί στον συνδυασμό πολλών δέντρων απόφασης προκειμένου να προκύψει ένα τελικό συλλογικό/συνδυαστικό μοντέλο. Για την αποδοτική μάθηση και συνδυασμό πολλών δέντρων με τη χρήση του ίδιου συνόλου εκπαίδευσης, ο αλγόριθμος Random Forest βασίζεται στην μεθοδολογία Bootstrap.

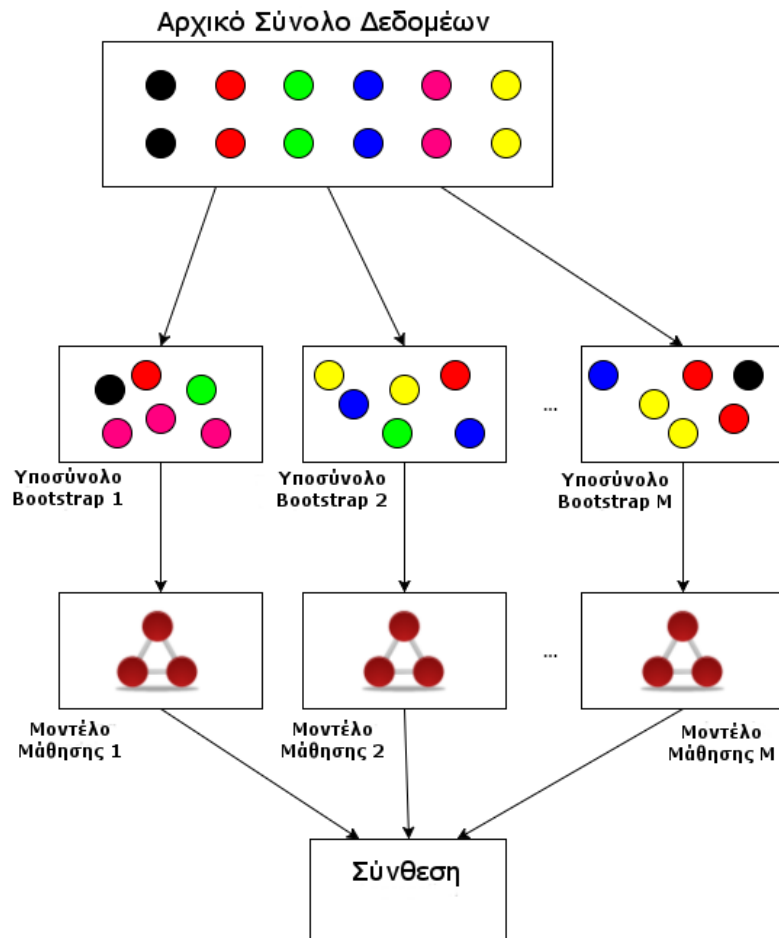
Η μεθοδολογία Bootstrap (ή bagging) είναι μία αποδοτική τεχνική που έχει ως στόχο την μείωση της υπερ-προσαρμογής (overfitting) των μοντέλων και τη βελτίωση των αποτελεσμάτων μηχανικής μάθησης σε μικρά ή ασταθή σύνολα δεδομένων. Η μέθοδος bagging αρχικά δημιουργεί M υποσύνολα (subsets) του συνόλου εκπαίδευσης με n δείγματα

ανά υποσύνολο. Τα n αυτά δείγματα λαμβάνονται ομοιόμορφα με αντικατάσταση, όπως αναπαρίσταται στην ακόλουθη εικόνα.



Εικόνα 5.6 Παράδειγμα δημιουργίας υποσυνόλων Bootstrap

Εν συνεχεία, όπως παρουσιάζεται στην ακόλουθη εικόνα, M διαφορετικά μοντέλα μάθησης (ensembles) δημιουργούνται για κάθε υποσύνολο. Οι έξοδοι των διαφορετικών μοντέλων μάθησης συνθέτουν την τελική πρόβλεψη μέσω διαδικασιών ψηφοφορίας ή υπολογισμού του μέσου όρου των προβλέψεων.



Εικόνα 5.7 Δημιουργία και σύνθεση πολλαπλών μοντέλων μάθησης για διαφορετικά υποσύνολα bootstrap

Η παραπάνω περιγραφή αφορούσε τον αλγόριθμο bagging γενικά για τα δέντρα απόφασης. Τα τυχαία δάση (Random forests) διαφέρουν από το συγκεκριμένο γενικό σχήμα στο ότι χρησιμοποιούν έναν τροποποιημένο αλγόριθμο μάθησης των δέντρων αποφάσεων ο οποίος σε κάθε φάση δημιουργίας διακλάδωσης, επιλέγει ένα τυχαίο σύνολο χαρακτηριστικών από το σύνολο των χαρακτηριστικών του συνόλου εκπαίδευσης βάσει του οποίου αποφασίζεται η διακλάδωση. Αυτή η διαδικασία αναφέρεται ως bagging επί των χαρακτηριστικών (feature bagging) και αποσκοπεί στη μείωση της συσχέτισης μεταξύ των διαφόρων δέντρων που αναπτύσσονται. Ο αριθμός των χαρακτηριστικών/μεταβλητών που επιλέγονται τυχαία αναφέρεται συνήθως ως η παράμετρος $mtry$.

Στα πλαίσια της παρούσας προσπάθειας, οι τύποι δεδομένων που χρησιμοποιήθηκαν ως χαρακτηριστικά για την εκπαίδευση του μοντέλου παλινδρόμησης αποτέλεσαν τα δεδομένα έκφρασης γονιδίων και η μέτρηση των λευχαιμικών κυττάρων στο περιφερικό αίμα την ημέρα 0, δηλαδή πριν την έναρξη της θεραπείας (initial peripheral blood blast count at Day 0), για κάθε ασθενή. Επίσης, η παράμετρος CKR_{PRED} επιλέχθηκε προκειμένου να προβλεφθεί (εξαρτημένη μεταβλητή, response variable). Επομένως, οι εκτιμήσεις του

CKR_{PRED} για τους ασθενείς που εντάσσονται στο σύνολο Εκπαίδευσης τροφοδοτούνται στο μοντέλο παλινδρόμησης. Θα πρέπει να τονιστεί ότι για κάθε επανάληψη της εξωτερικής αντεπικύρωσης (external-CV fold), η τιμή της CKR_{PRED} για του ασθενείς που εντάσσονται στο σύνολο δοκιμής είτε δεν εκτιμήθηκαν είτε αποκρύφθηκαν από το υπόλοιπο σύστημα.

Επιπροσθέτως, κάθε φορά που οι διαδικασίες εκπαίδευσης ή πρόβλεψης του αλγορίθμου παλινδρόμησης καλούνται, εκτελείται μία ακολουθία βημάτων προεπεξεργασίας δεδομένων.

Αρχικά, τα δεδομένα έκφρασης γονιδίων που έγιναν διαθέσιμα αποτελούνται από μετρήσεις διαφορετικών ανιχνευτών (probes), ωστόσο πολλοί από αυτούς αναφέρονται σε ίδια γονίδια. Επομένως, προκειμένου να καταστεί το σύνολο δεδομένων συμβατό με βασιζόμενη σε βιολογικά μονοπάτια ανάλυση (pathway-based analysis), να ενισχυθεί η βιολογική του ερμηνευσιμότητα και να καταστούν τα επόμενα βήματα της ανάλυσης ανεξάρτητα από την πλατφόρμα μέτρησης έκφρασης γονιδίων όπως προτείνεται στην [228], το σύνολο δεδομένων θα πρέπει να μετασχηματιστεί (collapse) από το επίπεδο των ανιχνευτών (probe level, το οποίο στο παρόν σύνολο δεδομένων είναι χαρακτηρισμένο μέσω IMAGE CloneIDs) στο επίπεδο των γονιδίων (π.χ. EntrezIDs). Προκειμένου να εκτελεστεί το συγκεκριμένο βήμα χρησιμοποιήθηκε η συνάρτηση *collapseRows* [228] του πακέτου WGCNA [229]. Μεταξύ των διαφόρων επιλογών για τον υπολογισμό της έκφρασης στο επίπεδο των γονιδίων, επιλέχθηκε η μέση (Average) έκφραση των ανιχνευτών. Ειδικότερα, υπολογίστηκε η μέση ένταση των ανιχνευτών που αναφέρονται στο ίδιο γονίδιο. Για την υλοποίηση του μετασχηματισμού (collapsing) είναι αναγκαία η τροφοδότηση στην συνάρτηση *CollapseRows* ενός πίνακα αντιστοίχισης μεταξύ των CloneIDs αριθμών αναφοράς και των αντίστοιχων σε EntrezIDs. Η αντιστοίχιση αυτή δημιουργήθηκε μέσω της διαθέσιμης στο διαδίκτυο πλατφόρμας SOURCE η οποία αρχικώς δημιουργήθηκε από το Τμήμα Γενετικής του Πανεπιστημίου του Stanford (<http://source-search.princeton.edu/cgi-bin/source/sourceBatchSearch>).

Στη συνέχεια, ακολουθείται μία διαδικασία φιλτραρίσματος γονιδίων (gene filtering) με βάση το ποσοστό των ελλিপών τιμών (missing values). Παρόμοια με την πλειοψηφία των συνόλων δεδομένων έκφρασης γονιδίων, το διαθέσιμο σύνολο δεδομένων εμφανίζει έναν αριθμό ελλিপών τιμών στις μετρήσεις των ανιχνευτών/γονιδίων. Παρόλο που το προηγούμενο βήμα μετασχηματισμού των ανιχνευτών σε γονίδια είναι ικανό να μειώσει τον αριθμό των ελλিপών τιμών (η συνάρτηση *CollapseRows* αγνοεί ανιχνευτές που εμφανίζουν ελλιπείς τιμές σε ποσοστό μεγαλύτερο του 90% των περιπτώσεων), ένα σημαντικός αριθμός αυτού του είδους των τιμών πιθανόν να συνεχίζει να υπάρχει. Επομένως, τα

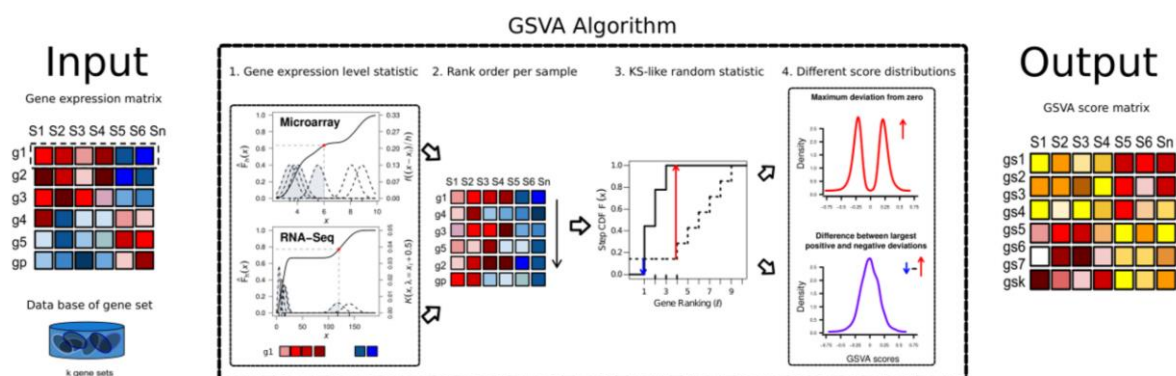
γονίδια που εμφανίζουν υψηλή συχνότητα ελλιπών τιμών (σε άνω του 20% των περιπτώσεων) φιλτραρίστηκαν και απομακρύνθηκαν από το σύνολο δεδομένων. Τα γονίδια αυτά μπορεί να θεωρηθεί ότι μετρήθηκαν ελλιπώς. Το βήμα αυτό υλοποιήθηκε με τη χρήση της συνάρτησης *goodGenes* του πακέτου WGCNA επιλέγοντας το ελάχιστο ποσοστό μη-ελλιπών δειγμάτων για να ένα γονίδιο προκειμένου να χαρακτηριστεί επαρκώς μετρημένο να είναι στο 80%.

Οι υπολειπόμενες ανεπιτυχώς μετρημένες τιμές έκφρασης γονιδίων συμπληρώθηκαν μέσω διαδικασίας απόδοσης/συμπλήρωσης τιμής (imputation). Το βήμα αυτό θεωρείται προτιμότερο σε σχέση με την εξαίρεση όλων των γονιδίων που εμφανίζουν ελλιπείς τιμές, με δεδομένη την εμπειρία που έχει συσσωρευθεί στην βιβλιογραφία [230]–[234] η οποία δείχνει ότι μία ανεπιτυχής διαχείριση των ελλιπών τιμών μπορεί να επηρεάσει τα περαιτέρω βήματα ανάλυσης. Εξαίρεση αποτέλεσαν τα γονίδια που φιλτραρίστηκαν στο προηγούμενο βήμα μιας και έχει βρεθεί ότι η ακρίβεια των μεθόδων απόδοσης τιμής μειώνεται σημαντικά για αυτές τις περιπτώσεις [231], [234]. Ο αλγόριθμος που επιλέχθηκε για την διαδικασία απόδοσης τιμής ήταν ο αλγόριθμος Κοντινότερων Γειτόνων (Nearest Neighbours algorithms, knn-imputation) [235], ο οποίος αναγνωρίζεται ως ο πλέον ευρέως και συχνότερα χρησιμοποιούμενος αλγόριθμος απόδοσης τιμών (imputation algorithm). Η απόδοση τιμών μέσω knn υλοποιήθηκε με τη χρήση του πακέτου *Impute* της πλατφόρμας Bioconductor (<http://www.bioconductor.org/>) στην γλώσσα R. Η τιμή για την παράμετρο *k* (οι γείτονες που θα χρησιμοποιηθούν για τον υπολογισμό της ελλιπούς τιμής) που τέθηκε ήταν 20. Η τιμή αυτή είναι η συχνότερα χρησιμοποιούμενη τιμή για την παράμετρο αυτή όταν ο αλγόριθμος *k*-knn χρησιμοποιείται για την απόδοση τιμών [231].

Ακολούθως, προκειμένου να πραγματοποιηθεί μία βασιζόμενη σε βιολογικά μονοπάτια ανάλυση της έκφρασης γονιδίων (pathway-based gene expression analysis), ακολουθήθηκε μία διαδικασία σύνοψης (aggregation) των εκφράσεων γονιδίων σε ενεργοποίηση των βιολογικών μονοπατιών της βάσης KEGG. Όπως αναφέρεται στη βιβλιογραφία [236]–[238], ο μετασχηματισμός των δεδομένων έκφρασης γονιδίων από το χώρο των γονιδίων στο χώρο των βιολογικών μονοπατιών (pathways) αναμένεται να οδηγήσει σε αυξημένη ευρωστία (robustness) των αποτελεσμάτων ανάλυσης μοριακών δεδομένων, σε αντίθεση με την υπογραφές γονιδίων (gene signatures) οι οποίες συχνά παρουσιάζονται ασταθείς. Επιπλέον, η σύνοψη της έκφρασης γονιδίων σε έναν διαφορετικό κοινώς χρησιμοποιούμενο χώρο, δηλαδή τον χώρο των βιολογικών μονοπατιών, αναμένεται να μειώσει την εγγενή (intrinsic) τεχνολογική και βιολογική διακύμανση μεταξύ των διαφόρων δειγμάτων δεδομένων. Η μέθοδος που επιλέχθηκε για την υλοποίηση της συγκεκριμένης

διαδικασίας ήταν η Ανάλυση Διακύμανσης Συνόλου Γονιδίων (Gene Set Variation Analysis ή GSVA) [239] όπως υλοποιείται στο ομώνυμο πακέτο της πλατφόρμας Bioconductor.

Στην ακόλουθη εικόνα δίνεται μία συνοπτική περιγραφή των βημάτων που ακολουθούνται από την GSVA προκειμένου να μετασχηματιστεί η έκφραση γονιδίων ανά δείγμα (sample-wise) σε βαθμό εμπλουτισμού συνόλου γονιδίων (gene set enrichment) ανά δείγμα.



Εικόνα 5.8 Τα βήματα εκτέλεσης του αλγορίθμου GSVA

Η μέθοδος GSVA λαμβάνει ως είσοδο ένα πίνακα $X = \{x_{ij}\}_{p \times n}$ με εκφράσεις γονιδίων στη μορφή \log_2 τιμών έκφρασης από microarray (ή RNA-seq μετρήσεις) για p γονίδια και n δείγματα και μία βάση από σύνολα γονιδίων $\Gamma = \{\gamma_1, \dots, \gamma_m\}$ (δηλαδή ομάδες γονιδίων με κάποια συσχέτιση μεταξύ τους, όπως λ.χ. η συμμετοχή στο ίδιο βιολογικό μονοπάτι). Με x_{ij} συμβολίζεται η τιμή έκφρασης τους i -οστού γονιδίου για το j -οστό δείγμα ενώ με $\gamma_k \subset \{1, \dots, p\}$ ορίζεται το υποσύνολο των γραμμών του X που απαρτίζουν ένα βιολογικό μονοπάτι ή μία άλλη λειτουργική μονάδα.

Η μέθοδος ξεκινάει με την αξιολόγηση του αν ένα γονίδιο i είναι υπερ- ή υπό-εκφρασμένο σε ένα δείγμα j στα πλαίσια της κατανομής του δείγματος πληθυσμού. Προκειμένου οι τιμές εκφράσεων να βρεθούν σε ίδιο κλίμακα εφαρμόζεται η ακόλουθη στατιστική. Για κάθε προφίλ έκφραση γονιδίου $x_i = \{x_{i1}, \dots, x_{in}\}$ μία μη παραμετρική εκτίμηση της αθροιστικής συνάρτησης κατανομής με πυρήνα (kernel estimation of cumulative density function, cdf). Στην περίπτωση των microarrays, χρησιμοποιείται Gaussian πυρήνας:

$$\widehat{F}_{h_i}(x_{ij}) = \frac{1}{n} \sum_{k=1}^n \int_{-\infty}^{(x_{ij}-x_{ik})/h_i} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

όπου h_i η παράμετρος εύρους ζώνης που καθορίζει την ανάλυση της εκτίμησης πυρήνα για κάθε γονίδιο και τίθεται $h_i = s_i/4$, όπου s_i η τυπική απόκλιση του i -οστού

γονιδίου. Αν θεωρήσουμε ότι $\widehat{F}_{h_i}(x_{ij}) = z_{ij}$ προκειμένου να μειωθεί η επιρροή από πιθανές ακραίες τιμές, μετατρέπεται η στατιστική z_{ij} σε τάξεις (ranks) $z_{(i)j}$ και προκειμένου αυτές να κανονικοποιηθούν και να είναι συμμετρικές γύρω από το 0 εφαρμόζεται η $r_{ij} = \left| \frac{p}{2} - z_{(i)j} \right|$. Ο υπολογισμός του enrichment score γίνεται παρόμοια με την ευρέως γνωστή GSEA μέθοδο με τη χρήση της ομοιάζουσας με Kolmogorov-Smirnov τυχαίο περίπατο (random walk) στατιστική:

$$v_{jk}(l) = \frac{\sum_{i=1}^l |r_{ij}|^\tau I(g(i) \in \gamma_k)}{\sum_{i=1}^p |r_{ij}|^\tau I(g(i) \in \gamma_k)} - \frac{\sum_{i=1}^l I(g(i) \notin \gamma_k)}{p - |\gamma_k|}$$

όπου τ η παράμετρος που περιγράφει το βάρος του άκρου στον τυχαίο περίπατο (συνήθως $\tau = 1$), γ_k το k -οστό σύνολο γονιδίων και $I(g(i) \in \gamma_k)$ η συνάρτηση που δείχνει αν το i -οστό γονίδιο ανήκει στο σύνολο γ_k .

Για την μετατροπή της στατιστικής αυτής σε enrichment score (GSVA score) προσφέρονται δύο μέθοδοι:

1. Η μέθοδος της μέγιστης απόκλισης (maximum deviation) η οποία δίνεται από

$$\text{την } ES_{jk}^{max} = v_{ij} \left[\arg \max_{l=1, \dots, p} |v_{jk}(l)| \right]$$

2. Η κανονικοποιημένη enrichment score που δίνεται από την $ES_{jk}^{diff} =$

$$\max_{l=1, \dots, p} (0, v_{jk}(l)) - \min_{l=1, \dots, p} (0, v_{jk}(l))$$

Τα σύνολα γονιδίων (gene sets) που επιλέχθηκαν προκειμένου να πραγματοποιηθεί η συσσωμάτωση των εκφράσεων γονιδίων ήταν αυτά που απαρτίζουν τα βιολογικά μονοπάτια της βάσης KEGG (<http://www.genome.jp/kegg/pathway.html>) τα οποία κατά το χρόνο εκπόνησης της διατριβής ήταν 186 στον αριθμό. Τα σύνολα γονιδίων ανακτήθηκαν από τη βάση MsigDB του Broad Institute (<http://www.broadinstitute.org/gsea/msigdb>) στη μορφή .gmt αρχείων (CP:KEGG: KEGG gene sets) και εισήχθησαν στο περιβάλλον της R με τη χρήση της συνάρτησης getGmt του πακέτου *GSEABase* [240]. Σε συνέχεια της εφαρμογής της μεθόδου GSVA, το μοριακό κομμάτι του συνόλου δεδομένων αποτελείται πλέον από 186 χαρακτηριστικά για κάθε ασθενή που αντιστοιχούν στην ενεργοποίηση των βιολογικών μονοπατιών.

Με το σύνολο δεδομένων στην τελική του πλέον μορφή, δύο επιπλέον βήματα προεπεξεργασίας πραγματοποιούνται. Πρώτον οι τιμές των χαρακτηριστικών κανονικοποιούνται με βάση την μέση τιμή τους (mean-centered). Δεύτερον, ακολουθήθηκε ένα βήμα μείωσης του επιπέδου συσχέτισης των διαφόρων μεταβλητών πρόβλεψης

(predictors) μεταξύ τους, αφαιρώντας τις έντονα συσχετισμένες μεταβλητές με τη χρήση της συνάρτησης findCorellation του πακέτου caret.

Θα πρέπει ιδιαίτερα να σημειωθεί πως για τα ανωτέρω παρουσιαζόμενα βήματα, ειδικές εκδόσεις του φιλτραρίσματος γονιδίων, της απόδοσης τιμών, της κανονικοποίησης και της μείωσης της συσχέτισης μεταξύ των χαρακτηριστικών υλοποιήθηκαν ειδικά για το βήμα πρόβλεψης του αλγορίθμου παλινδρόμησης. Αυτό αποσκοπεί στο να βασιστεί η προ-επεξεργασία των δεδομένων δοκιμής στη γνώση που αποκτήθηκε από την προ-επεξεργασία των δεδομένων εκπαίδευσης μιας και μόνο αυτή είναι διαθέσιμη κατά τη φάση πρόβλεψης.

Κατά τη διάρκεια των ανωτέρω παρουσιαζόμενων διαδικασιών εκπαίδευσης και προσαρμογής, πραγματοποιείται βελτιστοποίηση ενός αριθμού παραμέτρων που αφορούν τη διαδικασία προ-επεξεργασίας δεδομένων όπως επίσης και το μοντέλο παλινδρόμησης. Ο Πίνακας 5.4 παρουσιάζει τις παραμέτρους αυτές, μία σύντομη περιγραφή τους όπως επίσης και τις τιμές που δοκιμάστηκαν κατά τη διάρκεια της βελτιστοποίησης. Η επιλογή του καλύτερου συνόλου τιμών για τις παραμέτρους αυτές πραγματοποιείται με βάση την σχετική με την ρίζα της μέσης τετραγωνικής απόκλισης απόδοση (root mean square error (RMSE) performance) στην πρόβλεψη της τιμής της CKR_{PRED} που παρουσιάζει το μοντέλο παλινδρόμησης .

Πίνακας 5.4 Παράμετροι της διαδικασίας προ-επεξεργασίας και των μοντέλων παλινδρόμησης που βελτιστοποιήθηκαν μέσω εσωτερικής αντεπικύρωσης (internal cross-validation)

Όνομα παραμέτρου	Περιγραφή	Τιμές
Παράμετροι προ-επεξεργασίας		
Μέθοδος μετασχηματισμού (Collapse method)	Μέθοδος που χρησιμοποιήθηκε για τον μετασχηματισμό της έκφρασης γονιδίων από το επίπεδο των ανιχνευτών (probes) στο επίπεδο γονιδίων	Μέση τιμή (Average)
Ελάχιστο ποσοστό μη-ελλiptών τιμών (Good genes minimum fraction)	Ελάχιστο ποσοστό μη-ελλiptών τιμών προκειμένου η έκφραση ενός γονιδίου να θεωρηθεί ότι μετρήθηκε επαρκώς	0.8
Παράμετρος k αλγορίθμου knn για την συμπλήρωση τιμών	Αριθμός γειτόνων που χρησιμοποιήθηκαν για την συμπλήρωση τιμών	20

Όριο συσχέτισης (Correlation cutoff)	Η τιμή ορίου για την ανά-δύο απόλυτη συσχέτιση μεταξύ μεταβλητών-χαρακτηριστικών	0.5, 0.7, 0.9, 1.0
Παράμετροι Random Forest		
mtry	Αριθμός χαρακτηριστικών που επιλέγονται τυχαία για κάθε διακλάδωση	Δοκιμάστηκαν 5 διαφορετικές τιμές της mtry όπως παράγονται από την συνάρτηση <i>var_seq</i> του πακέτου <i>caret</i> [218] στην R. Οι τιμές αυτές εξαρτώνται από τον αριθμό των χαρακτηριστικών του συνόλου δεδομένων.
Παράμετροι αλγορίθμου k-Κοντινότερων Γειτόνων με βάρη		
k	Αριθμός γειτόνων	5,7,9
Απόσταση (distance)	Παράμετρος της απόστασης Minkowski	0.5, 1.0, 2.0, 3.0
Πυρήνας (kernel)	Συνάρτηση πυρήνα που χρησιμοποιήθηκε για τον υπολογισμό των βαρών των γειτόνων ανάλογα με την απόστασή τους	optimal [225]

Συνεχίζοντας την περιγραφή του συστήματος, στο βήμα D (Step D) ακολουθείται μία νέα σειρά εκπαιδεύσεων του μοντέλου παλινδρόμησης και ανάλυσης της απόδοσης του μέσω αντεπικύρωσης (cross-validation). Σε αυτά τα πλαίσια πραγματοποιείται επιπλέον διαχωρισμός του συνόλου εκπαίδευσης της τρέχουσας εξωτερικής αντεπικύρωσης (external cross-validation) προκειμένου να αναλυθεί η συμπεριφορά του σφάλματος του βελτιστοποιημένου (σε επίπεδο παραμέτρων) μοντέλου παλινδρόμησης. Επιπλέον, μέσω αυτής της διαδικασίας, αποκτάται μία εκτίμηση της κατανομής του αναμενόμενου σφάλματος του μοντέλου παλινδρόμησης παρατηρώντας τις αποκλίσεις (residuals) των προβλέψεων του από την πραγματική τιμή του CKR_{PRED} .

Για κάθε επανάληψη της εξωτερικής αντεπικύρωσης, ακολουθείται μία επιπλέον διαδικασία 5-fold αντεπικύρωσης. Αρχικά το σύνολο εκπαίδευσης χωρίζεται περαιτέρω σε Σύνολο Εκπαίδευσης Κατανομής Σφάλματος (Error Distribution Train Set) και σε Σύνολο Δοκιμής Κατανομής Σφάλματος (Error Distribution Test Set). Ακολούθως το μοντέλο παλινδρόμησης εκπαιδεύεται χρησιμοποιώντας τις βέλτιστες παραμέτρους (όπως εκτιμήθηκαν στο βήμα B) και με βάση αυτό εξάγονται προβλέψεις για τις τιμές του CKR_{PRED} για το Σύνολο Δοκιμής Κατανομής Σφάλματος. Τέλος, οι αποκλίσεις των προβλέψεων αυτών από τις εκτιμηθείσες από τα πραγματικά δεδομένα τιμές CKR_{PRED} υπολογίζονται για κάθε περίπτωση ασθενούς που εντάσσεται στο Σύνολο Δοκιμής Κατανομής Σφάλματος. Οι αποκλίσεις όλων των επαναλήψεων συνδυάζονται και

χρησιμοποιούνται, με τρεις πιθανούς τρόπους, προκειμένου να εκτιμηθεί η κατανομή του σφάλματος του μοντέλου παλινδρόμησης. Ο πρώτος τρόπος βασίζεται στον υπολογισμό του ιστογράμματος των αποκλίσεων (αποθηκεύοντας τις τιμές των κεντρικών τιμών/midpoints και τις πιθανότητες αυτών) με τη χρήση της συνάρτησης *hist* στην γλώσσα R. Ο δεύτερος τρόπος αφορά στην προσαρμογή μια κανονικής κατανομής επί των δεδομένων απόκλισης (με τη χρήση της συνάρτησης *fitdist* του πακέτου *fitdistplus* [241] στην γλώσσα R). Ο τρίτος κάνει χρήση του υπολογισμού εκτιμήσεων πυκνότητας πυρήνα (kernel density estimates) με τη χρήση της συνάρτησης *density* στη γλώσσα R (χρησιμοποιώντας τις προκαθορισμένες τιμές παραμέτρων). Ο τρόπος που αυτή η εκτιμηθείσα κατανομή χρησιμοποιείται στα επόμενα βήματα του συστήματος όπως και η επιρροή της επιλογής της μεθόδου στα τελικά συστήματα συζητείται στη συνέχεια του κειμένου.

Στο βήμα E (Step E), το τελικώς εκπαιδευμένο και βελτιστοποιημένο μοντέλο παλινδρόμησης χρησιμοποιείται για την εξαγωγή των προβλέψεων για την τιμή της CKR_{PRED} , για τις περιπτώσεις ασθενών που εντάσσονται στο Σύνολο Δοκιμής (εξωτερικής αντεπικύρωσης) ή γενικότερα για νέες περιπτώσεις ασθενών. Οι εξαχθείσες προβλέψεις, χρησιμοποιούνται για την πραγματοποίηση ενός αριθμού προσομοιώσεων μέσω του μηχανιστικού μέρους του Ογκοπροσομοιωτή Λευχαιμίας για κάθε περίπτωση ασθενούς (*n repeats* στην Εικόνα 5.3). Στα πλαίσια των δοκιμών για την παρούσα διατριβή ο αριθμός των προσομοιώσεων επιλέχθηκε να είναι 500. Σε κάθε προσομοίωση, μία διαφορετική τιμή προσθετικού θορύβου, ο οποίος εξάγεται ως δείγμα από την προαναφερθείσα Κατανομή Σφάλματος, προστίθεται στην προβλεφθείσα για τον συγκεκριμένο ασθενή τιμή της CKR_{PRED} . Έχοντας αυτή την τιμή της παραμέτρου ως είσοδο, εξάγεται μία πρόβλεψη για τον αριθμό των λεμφοβλαστών στο περιφερικό αίμα την ημέρα 8 μέσω του Ογκοπροσομοιωτή.

Μετά την εκτέλεση των προσομοιώσεων, στο βήμα F (Step F), κάθε ασθενής ταξινομείται τελικά στην κατηγορία απόκρισης στην Πρεδνιζόνη στην οποία προβλέφθηκε να ανήκει στην πλειοψηφία των προσομοιώσεων που τον αφορούσαν (με βάση το κριτήριο των 1000 λεμφοβλαστών/μl). Στην ακραία περίπτωση όπου ένα ασθενής προβλέφθηκε να ανήκει ίσες φορές στις δύο κατηγορίες, εκτελείται μία τελευταία προσομοίωση και το αποτέλεσμα αυτής λαμβάνεται ως η τελική πρόβλεψη.

Η εκτέλεση πολλαπλών προσομοιώσεων θεωρείται ιδιαίτερα σημαντική με δεδομένη την *a priori* γνώση ότι η CKR_{PRED} προβλέπεται με σχετικό σφάλμα από το μοντέλο παλινδρόμησης. Επομένως, αντί να υιοθετηθεί άνευ όρων η προβλεφθείσα τιμή της CKR_{PRED} και μέσω αυτής να εξαχθεί η πρόβλεψη κατηγορίας ασθενούς, ένα αριθμός

προσομοιώσεων με ελαφρώς τροποποιημένες παραμέτρους εισόδου, αναμένεται να είναι ικανή να εμφανίζει πιθανές τάσεις για την κατηγοριοποίηση της απόκρισης του ασθενούς, Επιπλέον, μία ποσοτικοποίηση της πεποίθησης (confidence) της εξαχθείσας ταξινόμησης είναι εφικτή να εξαχθεί παρατηρώντας το ποσοστό των προσομοιώσεων για κάθε ασθενή που έδωσαν την ίδια με την τελική πρόβλεψη. Αυτό το χαρακτηριστικό θεωρείται κρίσιμο για κάθε σύστημα που στοχεύει σε μελλοντική κλινική χρήση, όπως ο Ογκοπροσομοιωτής Λευχαιμίας με δεδομένο ότι μπορεί να προσφέρει τη δυνατότητα στον χρήστη να κρίνει αν η τελική πρόβλεψη θα πρέπει ή όχι να υιοθετηθεί.

Τέλος, στο βήμα G (Step G) τα αποτελέσματα ταξινόμησης για κάθε επανάληψη της εξωτερικής αντεπικύρωσης αλλά και για την συνολική διαδικασία συλλέγονται και ενοποιούνται.

5.6 Αποτελέσματα

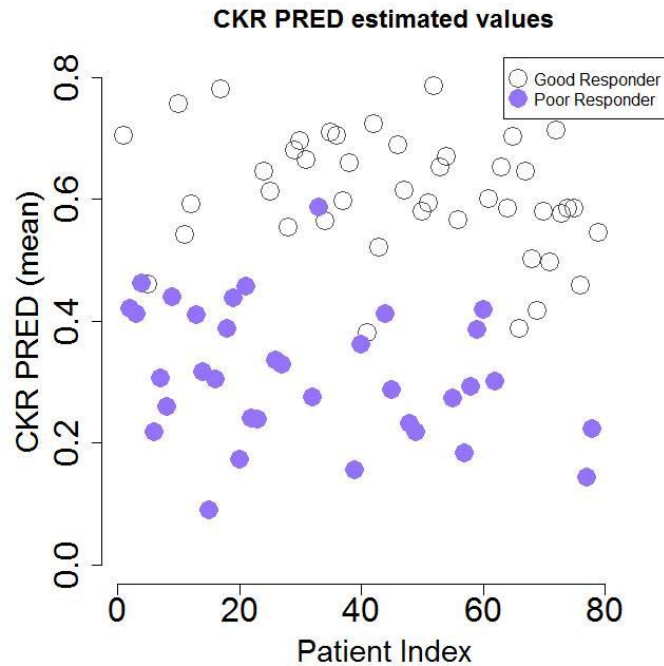
5.6.1 Αποτελέσματα εκτίμησης παραμέτρων

Όσον αφορά την εκτίμηση των παραμέτρων ελεύθερης ανάπτυξης του όγκου για τα 6 σενάρια T_d , αυτή επετεύχθη με τιμές για την σχετική συνάρτηση κόστους (Στοχευόμενο T_d μείον προσομοιωμένο από τον Ογκοπροσομοιωτή T_d) που κυμαίνονταν από 0.00018 μέρες έως 0.00626 μέρες. Αυτό είναι ενδεικτικό μίας σχεδόν πλήρως επιτυχημένης πρόβλεψης παραμέτρων. Ο Πίνακας 5.5 παρουσιάζει τις εκτιμηθείσες τιμές των παραμέτρων.

Πίνακας 5.5 Αποτελέσματα εκτίμησης παραμέτρων για τις σχετικές με την ελεύθερη ανάπτυξη του όγκου του μηχανιστικού μέρους του Ογκοπροσομοιωτή Λευχαιμίας για 6 διαφορετικά σενάρια χρόνου διπλασιασμού (doubling times (T_d)).

Όνομα Παραμέτρου	Εκτιμηθείσες τιμές					
	$T_d=7$ d	$T_d=14$ d	$T_d= 21$ d	$T_d= 28$ d	$T_d= 35$ d	$T_d= 42$ d
T_c	82 h	99 h	109 h	135 h	158 h	182 h
T_{G0}	82 h	82 h	80 h	78 h	108 h	97 h
T_N	106 h	119 h	126 h	113 h	133 h	120 h
R_A = $R_{A(Peripheral}$ Blasts)	0.001508683 h^{-1}	0.001280446 h^{-1}	0.0005444389 h^{-1}	0.0008812038 h^{-1}	0.0006393167 h^{-1}	0.000464761 h^{-1}
P_{sleep}	0.09283396	0.1712635	0.1044412	0.1523759	0.0735677	0.1194185
P_{sym}	0.6881355	0.4957694	0.2705583	0.3605559	0.3041863	0.3208003
P_{G0toG1}	0.5676784	0.7851885	0.8349451	0.7924908	0.6382258	0.4995043

Τα αποτελέσματα της συνολικής διαδικασίας προσαρμογής του μοντέλου, συνοψίζονται στην μέση εκτίμηση της CKR_{PRED} όπως δίδεται στην Εικόνα 5.9.



Εικόνα 5.9 Διάγραμμα διασποράς (scatter plot) της μέσης εκτιμηθείσας CKR_{PRED} (διαφορετικά σενάρια χρόνου διπλασιασμού) για τους ασθενείς που εντάχθηκαν στην μελέτη. Οι ασθενείς που κατατάσσονται στην κατηγορία κακής απόκρισης στην Πρεδνιζόνη τείνουν να εμφανίζουν χαμηλότερη τιμή χημειοευαισθησίας σε σχέση με τους ασθενείς που κατατάσσονται στην κατηγορία καλής απόκρισης στην Πρεδνιζόνη

Για κάθε ασθενή, 6 διαφορετικές τιμές της CKR_{PRED} εκτιμήθηκαν, μία για κάθε σενάριο χρόνου διπλασιασμού. Θα πρέπει να σημειωθεί ότι για ένα αριθμό ασθενών, κάποια σενάρια δεν οδήγησαν στην εκτίμηση έγκυρης τιμής για την CKR_{PRED} . Οι ασθενείς αυτοί χαρακτηρίζοντας από υψηλότερο, αντί για χαμηλότερο, αριθμό λεμφοβλαστών την ημέρα 8 της θεραπείας σε σύγκριση με την αντίστοιχη μέτρηση στην έναρξη αυτής. Η συμπεριφορά αυτή θα μπορούσε να εξηγηθεί στα πλαίσια μίας ιδιαίτερα γρήγορης αύξησης του όγκου με πιθανόν ταυτόχρονη χαμηλή χημειοευαισθησία. Έτσι, μικροί ρυθμοί αύξησης, που αντιστοιχούν σε σενάρια με μεγάλο χρόνο διπλασιασμού, εμφανίζονται ανεπαρκείς προκειμένου να προσομοιώσουν την παρατηρούμενη συμπεριφορά. Ειδικότερα, κατά τη διάρκεια των σχετικών δοκιμών για αυτούς τους ασθενείς και για μεγάλες τιμές του χρόνου διπλασιασμού, ο αριθμός των λεμφοβλαστών που επέστρεφε η εκτέλεση του Ογκοπροσομοιωτή Λευχαιμίας ήταν πάντα μικρότερος του παρατηρούμενου στα κλινικά δεδομένα, ακόμα και απουσία θεραπείας. Έτσι, οι ασθενείς αυτοί εξαιρέθηκαν σε αυτή τη φάση από την περαιτέρω μελέτη.

Για κάθε ασθενή, για το διάστημα τιμών των παραμέτρων του μοντέλου και των υπό-εξέταση σεναρίων χρόνου διπλασιασμού, η διακύμανση (variance) της CKR_{PRED} για τα έξι σενάρια Td δεν βρέθηκε να είναι σημαντική (μέση variance της CKR_{PRED} : 1.5036×10^{-4}). Η παρατήρηση αυτή μπορεί να εξηγηθεί από το γεγονός ότι η διάρκεια της θεραπείας

προ-φάσης (7 ημέρες) είναι πιθανότατα ιδιαίτερα μικρή προκειμένου η επιρροή των διαφόρων χρόνων διπλασιασμού στην προσομοίωση της θεραπείας να γίνουν εμφανή. Επομένως, στα πλαίσια της παρούσας δοκιμής μελετήθηκε περαιτέρω μόνο το *Td* σενάριο των 7 ημερών.

Όπως είναι εμφανές στην Εικόνα 5.9, οι ασθενείς που κατατάσσονται κλινικά στην κατηγορία καλής απόκρισης στην Πρεδνιζόνη (Prednisone Good Responders) εμφανίζουν υψηλότερη μέση τιμή της CKR_{PRED} σε σχέση με την κατηγορία κακή απόκρισης στο φάρμακο αυτό (Prednisone Poor Responders). Το εύρημα αυτό υποστηρίζει περαιτέρω την ορθότητα της διαδικασίας εκτίμησης παραμέτρων/προσαρμογής αλλά και του Ογκοπροσομοιωτή Λευχαιμίας συνολικά. Ωστόσο, για το διάστημα τιμών της CKR_{PRED} [0.4-0.6] παρατηρείται μία αλληλοεπικάλυψη μεταξύ των κατηγοριών. Για τις περιπτώσεις αυτές, η πρόβλεψη του αποτελέσματος της θεραπείας αναμένεται να είναι πιο δύσκολη. Θα πρέπει να σημειωθεί, ωστόσο, ότι η παρατηρούμενη αλληλοεπικάλυψη επί του συγκεκριμένου χαρακτηριστικού δεν αναμένεται να αφαιρέσει την ικανότητα του συστήματος στο να ταξινομήσει τους ασθενείς μεταξύ των δύο κατηγοριών με δεδομένο ότι το συνολικό σύστημα βασίζεται εκτός από την πρόβλεψη της τιμής της CKR και στον αρχικό αριθμό βλαστών (initial blast cell count) αλλά και στις βιολογικές αρχές (biological rules) όπως αυτές υλοποιούνται από το μηχανιστικό μοντέλο. Αυτές οι δύο πηγές πληροφορίας αναμένεται να συμβάλουν σημαντικά στον επιτυχή διαχωρισμό των ασθενών.

5.6.2 Αποτελέσματα πρόβλεψης κατηγορίας απόκρισης στην Πρεδνιζόνη

Τα αποτελέσματα εκτέλεσης του συστήματος παρουσιάζονται στον ακόλουθο πίνακα.

Πίνακας 5.6 Αποτελέσματα αξιολόγησης του Υβριδικού Ογκοπροσομοιωτής Λευχαιμίας, μέσω εξωτερικής αντεπικύρωσης (cross-validation, CV) στο να ταξινομεί ασθενείς μεταξύ των κατηγοριών Καλής και Κακής απόκρισης στην Πρεδνιζόνη

Μέθοδος εκτίμησης του σφάλματος του Μοντέλου Παλινδρόμησης	Κατώφλι Πεποίθησης	Απόδοση Ταξινόμησης (Ευαισθησία-Sensitivity, Ειδικότητα-Specificity) ^{*,**} [0-1]	Αριθμός ασθενών από το σύνολο δοκιμής που ταξινομήθηκαν με επαρκή πεποίθηση ^{**}	Ποσοστό πεποίθησης για τις σωστά ταξινομημένες περιπτώσεις ^{**} [0-1]	Ποσοστό πεποίθησης για τις λάθος ταξινομημένες περιπτώσεις ^{**} [0-1]	CKR RMSE στο εσωτερικό CV ^{**}	CKR R ² στο εσωτερικό CV ^{**}
Αλγόριθμος k-Κοντινότερων Γειτόνων με βάρη							
Εκτίμηση Πυκνότητας Πυρήνα	-	0.682 (0.67, 0.695)	all	0.759	0.703	0.143	0.295
	0.7	0.709 (0.646,0.753)	9.8				
	0.75	0.749 (0.716,0.768)	7.6				
	0.8	0.736 (0.634,0.753)	5.6				

	0.9	0.95 (0.6,0.933)	2.2				
Ιστόγραμμα	-	0.623 (0.58,0.65)	all	0.778	0.745		
	0.7	0.694 (0.607,0.724)	10.2				
	0.75	0.742 (0.598,0.805)	8.4				
	0.8	0.753 (0.45, 0.766)	5.2				
	0.9	0.533 (0.6, 0.933)	3.6				
Προσαρμογή Κανονικής Κατανομής	-	0.622 (0.558,0.686)	all	0.754	0.649		
	0.7	0.747 (0.633, 0.696)	7.6				
	0.75	0.779 (0.7, 0.712)	7				
	0.8	0.893 (0.8, 0.81)	5.2				
	0.9	0.76 (0.6, 0.55)	2				
Τυχαία Δάση (Random Forests)							
Εκτίμηση Πυκνότητας Πυρήνα	-	0.558 (0.508, 0.609)	all	0.681	0.631		
	0.7	0.648 (0.582, 0.62)	5.4				
	0.75	0.703 (0.533, 0.6)	3.8				
	0.8	0.667 (0.2, 0.566)	2				
	0.9	0.5 (- ⁺ , 0.5)	1.2				
Ιστόγραμμα	-	0.597 (0.52,0.694)	all	0.662	0.647	0.154	0.193
	0.7	0.5667 (0.285, 0.5)	5				
	0.75	0.58 (0.333, 0.5)	4				
	0.8	0.587 (0.3, 0.52)	3				
	0.9	0.32 (- ⁺ , 0.5)	1.8				
Προσαρμογή Κανονικής Κατανομής	-	0.658 (0.625, 0.683)	all	0.689	0.679		
	0.7	0.7405 (0.733, 0.786)	5.6				
	0.75	0.7843(0.73,0. 82)	4.2				
	0.8	0.75 (0.6, 0.85)	3.4				
	0.9	0.5667 (- ⁺ , 0.66)	1.4				
<p>*Η κακή απόκριση στην Πρεδνιζόνη επιλέχθηκε ως η «θετική» κλάση **Μέση τιμή για τις 5 επαναλήψεις εξωτερικής αντεπικύρωσης ⁺ Για κάποιες τιμές κατωφλίου πεποίθησης η «θετική» κλάση απουσίαζε από τις εναπομείνουσες περιπτώσεις ασθενών. Σε αυτή την περίπτωση, η μετρική ευαισθησίας (sensitivity) δεν εφαρμόζεται</p>							

Για κάθε αλγόριθμο παλινδρόμησης (regression algorithm) και για κάθε μέθοδο εκτίμησης της κατανομής σφάλματος αυτού (regression model error distribution estimation method), εκτελέστηκαν 5 εξωτερικές επαναλήψεις αντεπικύρωσης. Ο Πίνακας 5.6 παρουσιάζει τη μέση απόδοση του συστήματος για αυτές τις 5 επαναλήψεις. Όπως παρατηρείται, η ακρίβεια ταξινόμησης για το πλήρες Σύνολο Δοκιμής (16 ασθενείς για κάθε επανάληψη της εξωτερικής αντεπικύρωσης) κυμαίνεται στο διάστημα [0.62-0.68] για τον αλγόριθμο weighted k nearest neighbours και στο διάστημα [0.55-0.65] για τον αλγόριθμο των τυχαίων δασών (random forests). Από αυτό εξάγεται ότι για μία νέα

περιπτώσεις ασθενών, το πλήρως βελτιστοποιημένο σύστημα (έχοντας επιλέξει τον καλύτερο μοντέλο παλινδρόμησης και την καλύτερη μέθοδο εκτίμησης της κατανομής του σφάλματος αυτού) αναμένεται να ταξινομεί σωστά περίπου το 70% των ασθενών.

Η εκτέλεση των 5-fold εξωτερικής αντεπικύρωσης του όλου συστήματος (συμπεριλαμβανομένων δηλαδή της προ-επεξεργασίας δεδομένων, της εκπαίδευσης και βελτιστοποίησης του μοντέλου παλινδρόμησης, της εκτίμησης της κατανομής σφάλματος και της πρόβλεψης της κατηγορίας απόκρισης στην Πρεδνιζόνη μέσω πολλαπλών εκτελέσεων του Ογκοπροσομοιωτή λευχαιμίας) απαιτήσε κατά μέσο όρο 189 λεπτά σε προσωπικό υπολογιστή με 4 επεξεργαστικούς πυρήνες συγχρονισμένους στα 3.5 GHz και 16 GB μνήμης RAM. Παρόλου που η εκπαίδευση του συστήματος μπορεί να απαιτεί αρκετό χρόνο, η διαδικασία πρόβλεψης της κατηγορίας απόκρισης ενός νέου ασθενούς (η οποία περιλαμβάνει τα βήμα της προ-επεξεργασίας των εξατομικευμένων δεδομένων του, την εξατομικευμένη πρόβλεψη της τιμής της CKR_{PRED}, την εκτέλεση 500 προσομοιώσεων με τον Ογκοπροσομοιωτή Λευχαιμίας και την ενσωμάτωση των αποτελεσμάτων σε μία τελική απόφαση ταξινόμησης) διαρκεί λιγότερο από ένα λεπτό της ώρας. Θα πρέπει να σημειωθεί ωστόσο ότι ο χρόνοι αυτοί, και ειδικότερα ο χρόνος που απαιτείται για την διαδικασία αντεπικύρωσης, μπορεί να αυξηθούν σημαντικά στην περίπτωση που ένα μεγαλύτερο σύνολο δεδομένων (συγκεκριμένα μεγαλύτερος αριθμός ασθενών) γίνει διαθέσιμο στο μέλλον. Ωστόσο, και οι διαδικασία της αντεπικύρωσης αλλά και οι πολλαπλές προσομοιώσεις του Ογκοπροσομοιωτή Λευχαιμίας μπορούν παράλληλα με δεδομένο ότι δεν υπάρχουν αλληλοεξαρτήσεις μεταξύ των διαφόρων επαναλήψεων εκτέλεσης. Ένα τέτοιο βήμα παραλληλοποίησης της διαδικασίας μπορεί να μειώσει σημαντικά το χρόνο που απαιτείται και για την εκτέλεση και των δύο διαδικασιών.

Επί του παρόντος ο Υβριδικός Ογκοπροσομοιωτής Λευχαιμίας εκτελείται στο περιβάλλον της γλώσσας R. Επομένως, κάθε υπολογιστική υποδομή που υποστηρίζει το συγκεκριμένο περιβάλλον, συμπεριλαμβανομένων προσωπικών υπολογιστών, εικονικών μηχανών στο νέφος (virtual machines at the cloud) και εξυπηρετητές (server, π.χ. R server) μπορούν να χρησιμοποιηθούν για την εκτέλεση του συστήματος. Σχετικά με το υποκείμενο μηχανιστικό μοντέλο (Ογκοπροσομοιωτής Λευχαιμίας), αυτό έχει υλοποιηθεί με τη χρήση της γλώσσας C++ και έτσι όλα τα ευρέως χρησιμοποιούμενα λειτουργικά συστήματα είναι ικανά να υποστηρίξουν την εκτέλεσή του είτε ως εκτελέσιμο αρχείο (executable file) είτε ως δυναμική βιβλιοθήκη (dynamic library).

Συνεχίζοντας την ανάλυση των αποτελεσμάτων, για την πλειοψηφία των συνδυασμών, το σύστημα βρέθηκε να αποκρίνεται με μεγαλύτερη πεποίθηση όταν το

αποτέλεσμα ταξινόμησης ήταν σωστό. Επομένως, η ακόλουθη υπόθεση μπορεί να διατυπωθεί:

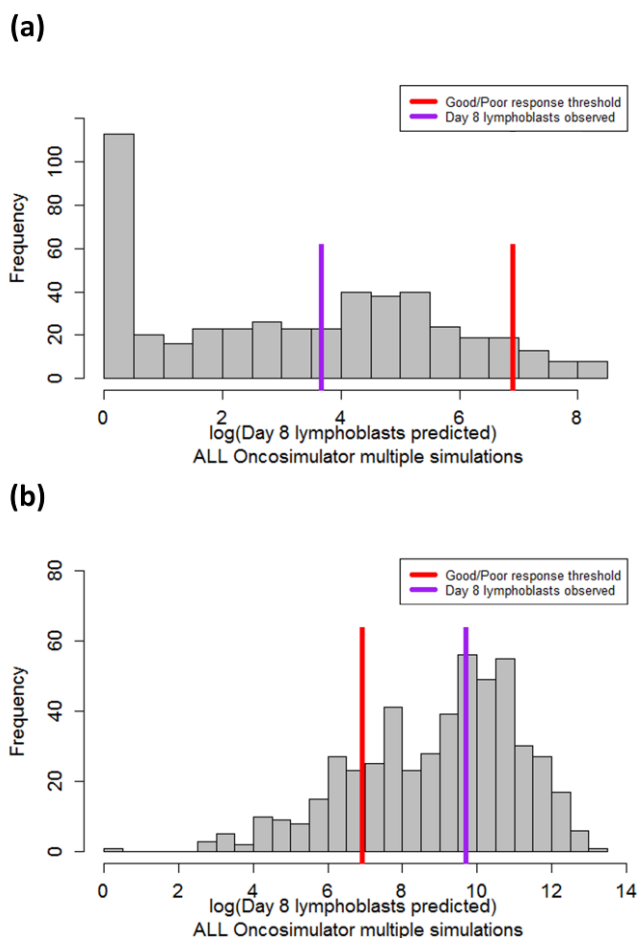
«Αν μόνο οι περιπτώσεις ασθενών του Συνόλου Δοκιμής για τις οποίες ο Υβριδικός Ογκοπροσομοιωτής Λευχαιμίας αποκρίνεται με σχετικά υψηλότερη πεποίθηση ληφθούν υπόψη, η ακρίβεια ταξινόμησης μπορεί να αυξηθεί και πιθανές παραπλανητικές προβλέψεις μπορούν να αποφευχθούν»

Η συγκεκριμένη υπόθεση δοκιμάστηκε θέτοντας 4 διαφορετικά όρια πεποίθησης και επαν-υπολογίζοντας την ακρίβεια ταξινόμησης μόνο για τις περιπτώσεις για τις οποίες η ταξινόμησή τους προβλέφθηκε με μεγαλύτερη από το αντίστοιχο όριο πεποίθησης. Όπως παρουσιάζει ο Πίνακας 5.6, θέτοντας το όριο πεποίθησης στην τιμή 0.9, η απόδοση του συστήματος φτάνει μέχρι και στο 0.95 όταν ο weighted-kNN και η Kernel Density Estimation μέθοδοι χρησιμοποιούνται. Ωστόσο, αυτή η αύξηση της ακρίβειας συνοδεύεται από την επίπτωση της άρνησης του συστήματος να ταξινομήσει τελικώς την συντριπτική πλειοψηφία των ασθενών. Συγκεκριμένα μόνο 2.2 από τους 16 ασθενείς του Συνόλου Δοκιμής ταξινομούνται κατά μέσο όρο. Για χαμηλότερες τιμές του ορίου, «έμπιστα» αποτελέσματα ταξινόμησης επιστρέφονται για μεγαλύτερο αριθμό ασθενών και εν γένει με υψηλότερη ακρίβεια, σε σχέση με την περίπτωση όπου δεν χρησιμοποιείται όριο, και ειδικότερα όταν ο επιλέγεται ο weighted-kNN.

Θα πρέπει να σημειωθεί ότι η αρχιτεκτονική του Υβριδικού Ογκοπροσομοιωτή Λευχαιμίας, αποτελεί πρόταση για τον τρόπο με τον οποίο ένα μοντέλο τύπου VPH μπορεί να καταστεί ικανό να προβλέψει το αποτέλεσμα της θεραπείας σε ένα πραγματικό κλινικό σενάριο. Το χαρακτηριστικό της χρήσης ορίου πεποίθησης (όπως περιγράφηκε προηγουμένως) μπορεί να επιτρέψει στο χρήστη (πιθανόν κλινικό) να επιλέξει το βαθμό πεποίθησης βάσει του οποίου θα εμπιστευτεί τις προβλέψεις ενός συστήματος αυτής της φύσης. Μία τέτοια στρατηγική πιστεύεται ότι μπορεί να υποστηρίξει σημαντικά τις σχετικές με τη θεραπεία αποφάσεις όταν αυτές βασίζονται και σε προβλέψεις μοντέλων.

Κεντρική ιδιότητα του συστήματος η οποία επιτρέπει την καθοδήγηση αυτή αποτελεί η δυνατότητα εκτέλεσης πολλαπλών προσομοιώσεων του Ογκοπροσομοιωτή Λευχαιμίας. Προκειμένου να παρουσιαστεί ο τρόπος με τον οποίο οι προσομοιώσεις αυτές αξιοποιούνται κατά τη λήψη της τελικής απόφασης ταξινόμησης, στην Εικόνα 5.10 δίνονται τα ιστογράμματα των λεμφοβλαστών στο περιφερικό αίμα στο τέλος της θεραπείας προφάσης (σε λογαριθμική κλίμακα), όπως προβλέφθηκαν από τις πολλαπλές εκτελέσεις του

Ογκοπροσομοιωτή Λευχαιμίας, για δύο ενδεικτικές περιπτώσεις ασθενών. Και οι δύο περιπτώσεις ταξινομήθηκαν σωστά με τη χρήση των weighted-kNN και kernel density estimation.



Εικόνα 5.10 Ιστογράμματα πολλαπλών εκτελέσεων/προσομοιώσεων του Υβριδικού Ογκοπροσομοιωτή Λευχαιμίας για δύο ενδεικτικές περιπτώσεις ασθενών. Και οι δύο περιπτώσεις ταξινομήθηκαν σωστά από το σύστημα. Το κατώφλι Καλής/Κακής απόκρισης στην Πρεδνιζόνη (Good/Poor response threshold) σημειώνεται με κόκκινη γραμμή ενώ ο αριθμός των λεμφοβλαστών την 8^η ημέρα της θεραπείας του κάθε ασθενή με μωβ γραμμή. (a) Οι πλειοψηφία των προσομοιώσεων καταλήγουν σε πρόβλεψη για τον αριθμό των λεμφοβλαστών την 8^η ημέρα χαμηλότερο του κατωφλίου για ασθενή που κατατάσσεται στην κατηγορία Καλής απόκρισης. (b) Οι προσομοιώσεις επιστρέφουν συχνότερα τιμές υψηλότερες του κατωφλίου για ασθενή που κατατάσσεται στην κατηγορία Κακής απόκρισης.

Η πρώτη περίπτωση (Εικόνα 5.10(a)),αφορά σε ασθενή με καλή απόκριση στην Πρεδνιζόνη (Prednisone) ενώ η δεύτερη (Εικόνα 5.10(b)) σε ασθενή με κακή απόκριση. Και στις δύο εικόνες το όριο των 1000 λεμφοβλαστών/μλ σημειώνεται με κόκκινο ενώ ο πραγματικός αριθμός αυτών στο περιφερικό αίμα των ασθενών στο τέλος της θεραπείας προ-φάσης σημειώνεται με μωβ. Όπως είναι εμφανές, η πλειοψηφία των προσομοιώσεων για τον ασθενή καλής απόκρισης καταλήγουν σε πρόβλεψη για τον αριθμό των βλαστών

χαμηλότερη του ανωτέρω αναφερόμενου ορίου, ενώ το αντίθετο ισχύει για την περίπτωση του ασθενούς κακής απόκρισης.

Σχετικά με την σύγκριση των δύο αλγορίθμων παλινδρόμησης που δοκιμάστηκαν, ο weighted k-NN αποδείχθηκε να επιτυγχάνει καλύτερα αποτελέσματα σε σχέση με τον αλγόριθμο των τυχαίων δασών (Random Forests). Αυτό αφορά όχι μόνο την ακρίβεια ταξινόμησης αλλά και στον αριθμό των ασθενών για τους οποίους επιστρέφεται αποτέλεσμα με υψηλή πεποίθηση όταν τεθεί σχετικό όριο. Αυτό μπορεί να εξηγηθεί όχι μόνο λόγω της αυξημένης ακρίβειας που παρουσιάζει στην πρόβλεψη της τιμής της CKR_{PRED} (Πίνακας 5.6) αλλά και στην σημαντικά αυξημένη τιμή της R^2 (ποσοστό της διακύμανση στην εξαρτημένη μεταβλητή που προβλέπεται από τις ανεξάρτητες μεταβλητές-χαρακτηριστικά).

Τα αποτελέσματα που παρουσιάστηκαν εμφανίζονται ελπιδοφόρα σχετικά με την ορθότητα του προτεινόμενου συνδυασμού μεθόδων ως ένα ολοκληρωμένο σύστημα (Υβριδικός Ογκοπροσομοιωτής Λευχαιμίας) αλλά και τη δυνατότητα να μπορεί να θεωρηθεί ως συνεισφορά στην δημιουργία θεμελιωδών αρχών για την εφαρμογή μοντέλων τύπου VPH σε διαδικασίες με στόχο την πρόβλεψη.

5.7 Συμπεράσματα, γενική αξιολόγηση και μελλοντικοί στόχοι

Στα πλαίσια του παρόντος κεφαλαίου της Διατριβής, παρουσιάστηκε ένας συνδυασμός ενός μηχανιστικού μοντέλου τύπου VPH το οποίο προσομοιώνει την εξέλιξη της Οξείας Λεμφοβλαστικής Λευχαιμίας (ALL) και τη θεραπεία αυτής (Ογκοπροσομοιωτής Λευχαιμίας, ALL Oncosimulator) με μεθόδους υπολογιστικής βελτιστοποίησης και μηχανικής μάθησης. Οι μέθοδοι υπολογιστικής βελτιστοποίησης χρησιμοποιήθηκαν με σκοπό την εκτίμηση παραμέτρων ενώ οι μέθοδοι μηχανικής μάθησης με σκοπό την πρόβλεψη των τιμών αυτών. Το σύστημα που δημιουργήθηκε, το οποίο ονομάστηκε Υβριδικός Ογκοπροσομοιωτής Λευχαιμίας (Hybrid ALL Oncosimulator), δοκιμάστηκε στο να προβλέψει την κατηγορία απόκρισης στο φάρμακο Πρεδνιζόνη (Prednisone) μίας νέας περίπτωσης ασθενούς με ALL. Τα αποτελέσματα αντεπικύρωσης (cross-validation) έδειξαν ότι τα προτεινόμενα σύστημα αναμένεται να ταξινομή σωστά περίπου το 70% των ασθενών. Επιπλέον η απόδοση/ορθότητα (accuracy) πρόβλεψης μπορεί να ανέλθει στο 95% όταν απαιτείται η ακρίβεια (precision) για την διαδικασία ταξινόμησης μέσω πολλαπλών προσομοιώσεων να είναι υψηλή (όριο πεποίθησης ταξινόμησης στο 0.9) προκειμένου μόνο «έμπιστες» αποφάσεις ταξινόμησης να επιστρέφονται και πιθανές παραπλανητικές

προβλέψεις να απορρίπτονται. Η υιοθέτηση ήδη αποδεκτών στην κλινική πράξη κριτηρίων ταξινόμησης και διαστρωμάτωσης (stratification) θεωρείται προτιμότερη σε σύγκριση με την υιοθέτηση μίας όχι και τόσο σίγουρης πρόβλεψης από το σύστημα. Επομένως το χαρακτηριστικό αυτό του συστήματος θεωρείτε κρίσιμο στα πλαίσια του στόχου της μελλοντικής εφαρμογής αυτού του είδους των συστημάτων στην κλινική πράξη. Στην παρούσα εργασία, ωστόσο, η αύξηση της απόδοσης στην ταξινόμηση συνοδεύεται και από σημαντική μείωση του αριθμού των ασθενών για τους οποίους επιστρέφεται «έμπιστη» πρόβλεψη (μόνο 2.2 ασθενείς από τους 16 για κάθε fold της διαδικασίας αντεπικύρωσης τελικώς ταξινομήθηκαν όταν το όριο πεποίθησης τέθηκε στο 0.9). Μεταξύ άλλων πιθανών αιτιών που μπορεί να σχετίζονται με τη σχεδίαση του συστήματος (λεπτομερής συζήτηση επί αυτού γίνεται παρακάτω), η έλλειψη ενός αρκούντως μεγάλου συνόλου ασθενών για την πραγματοποίηση της παρούσας ανάλυσης μπορεί να ευθύνεται για την σημαντική μείωση της πληρότητας των διαφορετικών κλάσεων απόκρισης στην Πρεδνιζόνη. Επομένως, πιο πλήρη και τυποποιημένα (standardized) σύνολα δεδομένων μεγάλης κλίμακας θα χρειαστούν στο μέλλον προκειμένου να υποστηριχθεί ο συγκεκριμένος δι-επιστημονικός τομέας έρευνας.

Δεχόμενοι την σχετική επιτυχία του προτεινόμενου Υβριδικού Ογκοπροσομοιωτή Λευχαιμίας στο να προβλέψει το αποτέλεσμα της θεραπείας προ-φάσης στην ALL ως μία απόδειξη της ορθότητας της γενικότερης ιδέας (proof of concept), τα βασικά βήματα που διαμορφώθηκαν στην παρούσα εργασία θα μπορούσαν να θεωρηθούν ως μια καλή βάση για μελλοντικές βελτιώσεις του συστήματος. Αρχικά, η απόδοση (accuracy) πρόβλεψης αλλά και η ευρωστία (robustness) του συστήματος θα πρέπει να αυξηθούν όσο περισσότερο είναι εφικτό. Σε μελλοντικές προσπάθειες, επιπλέον έρευνα επί των μεθόδων παλινδρόμησης και επί των μεθοδολογιών προ-επεξεργασίας δεδομένων αλλά και η χρήση περισσότερων μορφών δεδομένων (για παράδειγμα επιπλέον τύπων -ομικών δεδομένων) θα μπορούσαν να οδηγήσουν σε αύξηση της απόδοσης του συστήματος. Επιπλέον, πιθανοί εναλλακτικοί τρόποι για την αξιοποίηση της δυνατότητας του συστήματος να εκτελεί πολλαπλές προσομοιώσεις θα πρέπει να μελετηθούν και όσον αφορά την λήψη τελικής απόφαση ταξινόμησης αλλά και τον αντίστοιχο υπολογισμό της πεποίθησης για αυτή.

Η αξιολόγηση των ταξινομητών μέσω της διαδικασίας αντεπικύρωσης εμφανίζεται συχνά την σχετική με τη μηχανική μάθηση βιβλιογραφία. Ωστόσο, ένα προβλεπτικό σύστημα, όπως αυτό που προτείνει η παρούσα διατριβή (μαζί με όποιες πιθανές μελλοντικές βελτιώσεις), θα πρέπει να επικυρωθεί διεξοδικά όσον αφορά τη δυνατότητα του να προβλέπει την απόκριση πραγματικά νέων περιπτώσεων ασθενών (prospective

patient cases) οι οποίοι θα απαρτίζουν το ανεξάρτητο σύνολο επικύρωσης (independent validation set). Επιπλέον, ειδικά για τον σχετικό με τη θεραπεία προ-φάσης Υβριδικό Ογκοπροσομοιωτή Λευχαιμίας, τα πραγματικά οφέλη της πρόβλεψης της κατηγορίας απόκρισης στην Πρεδνιζόνη για νεοεισερχόμενους ασθενείς, θα πρέπει να επιβεβαιωθούν σε σχέση με την αποδοτική τροποποίηση της θεραπείας τους.

Και τα δύο παραπάνω βήματα θα πρέπει να ενσωματωθούν σε μελλοντική κλινική δοκιμή η οποία θα περιλαμβάνει δύο επιπλέον διακριτές φάσης. Κατά τη διάρκεια της πρώτης φάσης, ο ήδη εκπαιδευμένος σε προηγούμενες περιπτώσεις ασθενών Υβριδικός Ογκοπροσομοιωτής Λευχαιμίας θα πρέπει να αποδείξει την δυνατότητα του στο να προβλέπει την κατηγορία απόκρισης. Σε συνέχεια της επιβεβαίωσης αυτής, στη δεύτερη φάση, ένα πρώτο σύνολο ασθενών θα μπορούσε να θεραπευτεί με βάση το τρέχον και καθιερωμένο στην κλινική πράξη πρωτόκολλο, συμπεριλαμβάνοντας την αξιολόγηση της απόκρισης στην Πρεδνιζόνη μέσω της θεραπείας προ-φάσης. Για ένα άλλο σύνολο ασθενών, η κατηγορία απόκρισης στην Πρεδνιζόνη, θα μπορούσε να προβλεφθεί μέσω του Υβριδικού Ογκοπροσομοιωτή Λευχαιμίας και οι αποφάσεις διαστρωμάτωσης που επηρεάζονται από την απόκριση στην Πρεδνιζόνη (βλ. Κεφάλαιο 1) να ληφθούν άμεσα και πριν τη θεραπεία προ-φάσης. Μετά την ολοκλήρωση της θεραπείας και των δύο συνόλων ασθενών, τα πιθανά πλεονεκτήματα της επιτάχυνσης της εξατομίκευσης θα μπορούσαν να εξαχθούν συγκρίνοντας τα ποσοστά επιτυχίας ελέγχου της ασθένειας μεταξύ των δύο συνόλων.

Έπειτα από μία τέτοια μελλοντική επικύρωση, επόμενο βήμα θα μπορούσε να αποτελέσει η μελέτη της δυνατότητας ενός παρόμοιου συστήματος στο να προβλέψει το αποτέλεσμα μιας μεγαλύτερης σε διάρκεια και πιο σύνθετης θεραπείας ή ενός συνδυασμού φάσεων θεραπείας. Για παράδειγμα, η ανίχνευση της Ελάχιστης Υπολειπόμενη Ασθένειας (Minimal Residual Disease-MRD) είναι ιδιαίτερα σημαντικής αξίας για την αξιολόγηση της απόκρισης στη θεραπεία και για την επιπλέον διαστρωμάτωση των ασθενών σε κατηγορίας κινδύνου (risk groups) [242]–[244]. Αυτό το βήμα απαιτεί από τη μία την μοντελοποίηση της χορήγησης επιπλέον φαρμάκων από το μηχανιστικό μέρος του μοντέλου (Ογκοπροσομοιωτής Λευχαιμίας). Από την άλλη απαιτεί πιθανόν την πρόβλεψη περισσότερων από μίας παραμέτρων μέσω των μεθόδων μηχανικής μάθησης. Αυτές οι παράμετροι θα μπορούσαν να είναι είτε σχετικές με την χημειοευαισθησία σε φάρμακα είτε σχετικές με την αύξηση του όγκου. Ειδικότερα, η πρόβλεψη των τελευταίων θεωρείται κρίσιμη μιας και ο χρόνος διπλασιασμού του προσομοιωμένου όγκου μπορεί να επηρεάσει σημαντικά τα αποτελέσματα προσομοίωσης μιας εκτεταμένης χρονικά φάσης θεραπείας.

Θα πρέπει να σημειωθεί πως η πρόβλεψη περισσότερων από μία παραμέτρους απαιτεί είτε τη χρήση πολλαπλών μοντέλων παλινδρόμησης, είτε ενός μοντέλου παλινδρόμησης με πολλαπλές εξαρτημένες μεταβλητές (multiple responses). Ωστόσο η διαδικασία αυτή θα μπορούσε να υποστηριχθεί μέσω της χρήσης επιπλέον μηχανιστικών μοντέλων που εστιάζουν στα σχετικά με την ALL φαινόμενα (π.χ. τα μοντέλα που παρουσιάζονται στις [245]–[247], στις οποίες συμπεριλαμβάνεται και το μοντέλο που παρουσιάστηκε στο Κεφάλαιο 3 της παρούσας διατριβής).

Τα αποτελέσματα που παρουσιάστηκαν στο παρόν κεφάλαιο της διατριβής, υποστηρίζουν περαιτέρω την ιδέα πως ένας αποδοτικός συνδυασμός διαφόρων ετερογενών μοντέλων (και σε επίπεδο επιπέδων βιο-πολυπλοκότητας και σε επίπεδο προσεγγίσεων μοντελοποίησης) μπορεί να οδηγήσει στην δημιουργία συστημάτων ικανών να υποστηρίξουν τις σχετικές με τη θεραπεία κλινικές αποφάσεις. Τέλος, προβλέπεται ότι το ιατρικό προσωπικό θα μπορεί εύκολα να αλληλοεπιδρά με το υποκείμενο σύνθετο σύστημα μέσω της χρήσης συστημάτων Υποστήριξης Κλινικής Απόφασης (Clinical Decision Support system) όπως προτείνεται στην [248].

6. Γενικές παρατηρήσεις, συμπεράσματα και συνολικοί μελλοντικοί στόχοι

Στα πλαίσια της παρούσας διδακτορικής διατριβής τέθηκε ως βασικός στόχος η μελέτη των επιστημονικών τομέων της Μηχανικής Μάθησης, της Βιοπληροφορικής και της Υπολογιστικής Συστημικής Βιολογίας και η αξιοποίηση αλλά και η ανάπτυξη μεθόδων και μοντέλων που άπτονται αυτών στο χώρο της In Silico Ογκολογίας. Βασική κατευθυντήρια γραμμή για την αναζήτηση μεθόδων και για την ανάπτυξη νέων μοντέλων αποτέλεσε η κατά το δυνατόν συνάφεια και συσχέτιση με μοντέλα και μεθόδους που αναπτύσσονται από την Ομάδα για την In Silico Ογκολογία και την In Silico Ιατρική του ΕΠΙΣΕΥ. Συγκεκριμένα επιλέχθηκε η περίπτωση της Οξείας Λεμφοβλαστικής Λευχαιμίας για την οποία έχει ήδη αναπτυχθεί αλλά και βρίσκεται υπό συνεχή εξέλιξη από την προαναφερθείσα ομάδα, προσομοιωτικό μοντέλο για την ανάπτυξη αλλά και την απόκριση στη θεραπεία. Το συγκεκριμένο μοντέλο, ο Ογκοπροσομοιωτής Λευχαιμίας (ALL Oncosimulator), εστιάζει στα φαινόμενα που άπτονται του κυτταρικού αλλά και ανώτερων επιπέδων.

Η συνεισφορά της παρούσας διατριβής έγκειται στην ανάπτυξη μοντέλων και μεθόδων που ως βασικούς στόχους έχουν:

- Την εστίαση σε φαινόμενα και πηγές δεδομένων που άπτονται του μοριακού επιπέδου, αφορούν την Οξεία Λεμφοβλαστική Λευχαιμία και σχετίζονται με κυτταρικά φαινόμενα που λαμβάνει υπόψη ο Ογκοπροσομοιωτής Λευχαιμίας.
- Την εξατομίκευση της παραμετροποίησης του Ογκοπροσομοιωτή Λευχαιμίας προκειμένου να υποστηριχθεί η ακριβέστερη μελέτη των ήδη γνωστής εξέλιξης και απόκρισης στη θεραπεία περιπτώσεων ασθενών αλλά και η πρόβλεψη αυτών για νέες περιπτώσεις ασθενών.

Στα πλαίσια αυτά αναπτύχθηκαν:

- Κεφάλαιο 3: Ένα μοντέλο Συστημικής Βιολογίας που προσομοιώνει το δίκτυο βιοχημικών αντιδράσεων που ρυθμίζει την εξέλιξη του κυτταρικού κύκλου και συγκεκριμένα του «Σημείου Περιορισμού» (Μετάβαση από την G1 φάση στην S φάση). Το συγκεκριμένο μοντέλο αναπτύχθηκε με βάση τα παρατηρούμενα στην Οξεία Λεμφοβλαστική Λευχαιμία φαινόμενα τα οποία αποδείχθηκε να είναι ικανό να προσομοιώσει. Επίσης, βασική παράμετρος

εξόδους του μοντέλου αποτελεί η διάρκεια του κυτταρικού κύκλου, η οποία αποτελεί βασική παράμετρο εισόδου του Ογκοπροσομοιωτή.

- Κεφάλαιο 4: Ένα μοντέλο φαρμακοκινητικής για την από στόματος χορήγηση του φαρμάκου Πρεδνιζόνη, το οποίο εντάσσεται στο βασικό κορμό της θεραπείας της Οξείας Λεμφοβλαστικής Λευχαιμίας σε παιδιά. Για την ανάπτυξη του συγκεκριμένου μοντέλου χρησιμοποιήθηκαν εργαλεία Συστημικής Βιολογίας και μέθοδοι βελτιστοποίησης προκειμένου να προσαρμοστεί σε διαθέσιμα από την βιβλιογραφία δεδομένα. Η εξατομικευμένη πρόβλεψη της φαρμακοκινητικής του συγκεκριμένου φαρμάκου που εξάγει το συγκεκριμένο μοντέλο, μπορεί να συσχετιστεί με την παραμετροποίηση της εισόδου του Ογκοπροσομοιωτή όταν αυτός προσομοιώνει φάση θεραπείας που εντάσσει το συγκεκριμένο φάρμακο.
- Κεφάλαιο 0: Μία μεθοδολογία προσαρμογής της παραμετροποίησης εισόδου Ογκοπροσομοιωτή Λευχαιμίας, μέσω μεθόδων υπολογιστικής βελτιστοποίησης, σε δεδομένα ασθενών για τους οποίους είναι ήδη γνωστή η εξέλιξη της ασθένειας τους και η απόκρισή τους στη θεραπεία και εξαγωγής γνώσης για την δημιουργία συστήματος εξατομικευμένης πρόβλεψης της απόκρισης στην ίδια θεραπεία για νέες περιπτώσεις ασθενών. Το σύστημα που αναπτύχθηκε ονομάστηκε Υβριδικός Ογκοπροσομοιωτής Λευχαιμίας με δεδομένο ότι αξιοποιεί μεθόδους Μηχανικής Μάθησης σε συνδυασμό με το μηχανιστικό μοντέλο (Ογκοπροσομοιωτής Λευχαιμίας), προκειμένου να προβλέψει την απόκριση στη θεραπεία. Επιπλέον, ο Υβριδικός Ογκοπροσομοιωτής που αναπτύχθηκε βασίστηκε στην αξιοποίηση δεδομένων έκφρασης γονιδίων για την εξατομίκευση και πρόβλεψη της εισόδου του μηχανιστικού μοντέλου (Ογκοπροσομοιωτής), το οποίο απαίτησε την εκτεταμένη χρήση μεθόδων Βιοπληροφορικής.

Στα επιμέρους αυτά κεφάλαια έχουν ήδη δοθεί συμπερασματικές σκέψεις και μελλοντικοί στόχοι. Συνολικά, θεωρείται πως οι συγκεκριμένες συνεισφορές της διατριβής κατάφεραν επαρκώς να δείξουν την αξία της εφαρμογής των επιστημονικών τομέων της Μηχανικής Μάθησης, της Βιοπληροφορικής και της Συστημικής Βιολογίας στην In Silico Ογκολογία ενώ αποτελούν και από μόνες τους σημαντικά επιτεύγματα. Η επιτυχής μοντελοποίηση και προσομοίωση της βιοχημικής ρύθμισης του κυτταρικού κύκλου προσαρμοσμένη στα μη κανονικά φαινόμενα που παρατηρούνται στην Οξεία

Λεμφοβλαστική Λευχαιμία θεωρείται κρίσιμη για την επιβεβαίωση των σχετικών υποθέσεων και την περαιτέρω μελέτη αυτών. Η μοντελοποίηση της φαρμακοκινητικής της από στόματος χορήγησης της Πρεδνιζόνης κρίνεται κρίσιμη μιας και το συγκεκριμένο μοντέλο απουσίαζε από την βιβλιογραφία. Τέλος, και πιο σημαντικά, η πρωτοπόρα προσπάθεια και η επαρκής επιτυχία πρόβλεψης της απόκρισης στη θεραπεία προ-φάσης (απόκριση στην Πρεδνιζόνη) παιδιών με Οξεία Λεμφοβλαστική Λευχαιμία μέσω του Υβριδικού Ογκοπροσομοιωτή Λευχαιμίας, θεωρείται πως αποτελεί κομβικό σημείο για τον χώρο της In Silico Ογκολογίας και την υποστηρίζει σημαντικά προς την επίτευξη του κεντρικού της στόχου, αυτόν της υποστήριξης των κλινικών στην βελτιστοποίηση της θεραπείας. Επίσης αποτελεί σημαντική επιβεβαίωση της ορθότητας της προσέγγισης αυτής.

Η περαιτέρω μελέτη και αναβάθμιση των επιμέρους μοντέλων και μεθόδων που αναπτύχθηκαν στην διατριβή αποτελεί αναγκαιότητα για την βελτίωση της ακρίβειας των εξαγόμενων αποτελεσμάτων, όπως λεπτομερώς παρουσιάζεται στις επιμέρους συμπερασματικές υπό-ενότητες. Σημαντικό μελλοντικό στόχο ωστόσο αποτελεί ο περαιτέρω συνδυασμός των συγκεκριμένων μεθόδων και μοντέλων αλλά και γενικότερα των επιστημονικών τομέων που μελετήθηκαν, σε συνδυασμό με άλλες προσπάθειες από την επιστημονική κοινότητα, προκειμένου να επιτευχθεί η κατά το δυνατόν ακριβέστερη πρόβλεψη της απόκρισης στη θεραπεία.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] D. Loukopoulos, M. Politi, Δ. Λουκόπουλος, and Π. Μαριάννα, *HEMATOLOGY LESSONS*. 2015.
- [2] S. Davidson, *Davidson's γενικές αρχές και κλινική πράξη της ιατρικής παθολογίας*, 4η έκδ. Αθήνα: Επιστημονικές Εκδόσεις Παρισιάνου, 2014.
- [3] R. A. Weinberg, "How cancer arises," *Sci. Am.*, vol. 275, no. 3, pp. 62–70, Sep. 1996.
- [4] B. G. Katzung, McGraw-Hill Companies, and Teton Data Systems, *Basic & clinical pharmacology*. New York: McGraw-Hill, 2001.
- [5] H. Lodish, A. Berk, S. L. Zipursky, P. Matsudaira, D. Baltimore, and J. Darnell, *Molecular Cell Biology*, 4th ed. W. H. Freeman, 2000.
- [6] G. G. Steel, *Basic clinical radiobiology*, 3rd ed. London : Arnold ; New York : Oxford University Press, 2002.
- [7] M. C. Perry, *The Chemotherapy Source Book*. Lippincott Williams & Wilkins, 2008.
- [8] R. J. Epstein, "Human Molecular Biology: An Introduction to the Molecular Basis of Health and Disease," *Cambridge Core*, Oct-2002. [Online]. Available: /core/books/human-molecular-biology/35AC5B24E00C17DE5BD6FC6B1B4B075F. [Accessed: 05-Jan-2018].
- [9] "Τελομερές," *Βικιπαίδεια*. 07-May-2017.
- [10] E. S. Hwang, "Replicative senescence and senescence-like state induced in cancer-derived cells," *Mech. Ageing Dev.*, vol. 123, no. 12, pp. 1681–1694, Nov. 2002.
- [11] K. Vermeulen, D. R. Van Bockstaele, and Z. N. Berneman, "The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer," *Cell Prolif.*, vol. 36, no. 3, pp. 131–149, Jun. 2003.
- [12] P. M. Henson, D. L. Bratton, and V. A. Fadok, "Apoptotic cell removal," *Curr. Biol. CB*, vol. 11, no. 19, pp. R795-805, Oct. 2001.
- [13] Y. Soini, P. Pääkkö, and V.-P. Lehto, "Histopathological Evaluation of Apoptosis in Cancer," *Am. J. Pathol.*, vol. 153, no. 4, pp. 1041–1053, Oct. 1998.
- [14] P. Nicotera, M. Leist, and E. Ferrando-May, "Intracellular ATP, a switch in the decision between apoptosis and necrosis," *Toxicol. Lett.*, vol. 102–103, pp. 139–142, Dec. 1998.
- [15] G. Majno and I. Joris, "Apoptosis, oncosis, and necrosis. An overview of cell death," *Am. J. Pathol.*, vol. 146, no. 1, pp. 3–15, Jan. 1995.
- [16] T. Shirai, "Commentary: oncosis and apoptosis: two faces of necrosis in a new proposal to clear up the confusion regarding cell death," *Toxicol. Pathol.*, vol. 27, no. 4, pp. 495–496, Aug. 1999.
- [17] M. Leist and P. Nicotera, "The shape of cell death," *Biochem. Biophys. Res. Commun.*, vol. 236, no. 1, pp. 1–9, Jul. 1997.
- [18] S. Huettnerbrenner *et al.*, "The evolution of cell death programs as prerequisites of multicellularity," *Mutat. Res. Mutat. Res.*, vol. 543, no. 3, pp. 235–249, Jun. 2003.
- [19] C.-H. Pui and W. E. Evans, "Treatment of acute lymphoblastic leukemia," *N. Engl. J. Med.*, vol. 354, no. 2, pp. 166–178, Jan. 2006.
- [20] R. A. al et, "A systematic literature review of the clinical and epidemiological burden of acute lymphoblastic leukaemia (ALL). - PubMed - NCBI." [Online]. Available: <https://www.ncbi.nlm.nih.gov/pubmed/15698386>. [Accessed: 21-Dec-2017].
- [21] R. W. Miller, "Down's Syndrome (Mongolism), Other Congenital Malformations and Cancers among the Sibs of Leukemic Children," *N. Engl. J. Med.*, vol. 268, no. 8, pp. 393–401, Feb. 1963.
-

- [22] “Relation Between Cancer and Congenital Defects: An Epidemiologic Evaluation | JNCI: Journal of the National Cancer Institute | Oxford Academic.” [Online]. Available: <https://academic.oup.com/jnci/article-abstract/40/5/1079/901277?redirectedFrom=PDF>. [Accessed: 21-Dec-2017].
- [23] M. Swift, “Fanconi’s Anaemia in the Genetics of Neoplasia,” *Nature*, vol. 230, no. 5293, p. 370, Apr. 1971.
- [24] M. F. Greaves, A. T. Maia, J. L. Wiemels, and A. M. Ford, “Leukemia in twins: lessons in natural history,” *Blood*, vol. 102, no. 7, pp. 2321–2333, Oct. 2003.
- [25] J. M. Bennett *et al.*, “The morphological classification of acute lymphoblastic leukaemia: concordance among observers and clinical correlations,” *Br. J. Haematol.*, vol. 47, no. 4, pp. 553–561, Apr. 1981.
- [26] L. Sen and L. Borella, “Clinical importance of lymphoblasts with T markers in childhood acute leukemia,” *N. Engl. J. Med.*, vol. 292, no. 16, pp. 828–832, Apr. 1975.
- [27] J. Kersey, M. Nesbit, H. Hallgren, A. Sabad, E. Yunis, and K. Gajl-Peczalska, “Evidence for origin of certain childhood acute lymphoblastic leukemias and lymphomas in thymus-derived lymphocytes,” *Cancer*, vol. 36, no. 4, pp. 1348–1352, Oct. 1975.
- [28] M. F. Greaves, “Differentiation-linked leukemogenesis in lymphocytes,” *Science*, vol. 234, no. 4777, pp. 697–704, Nov. 1986.
- [29] C. A. Hurwitz *et al.*, “Asynchronous antigen expression in B lineage acute lymphoblastic leukemia,” *Blood*, vol. 72, no. 1, pp. 299–307, Jul. 1988.
- [30] M. C. Bene *et al.*, “Proposals for the immunological classification of acute leukemias. European Group for the Immunological Characterization of Leukemias (EGIL),” *Leukemia*, vol. 9, no. 10, pp. 1783–1786, Oct. 1995.
- [31] S. SH *et al.*, *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*. .
- [32] H. Inaba and C.-H. Pui, “Glucocorticoid use in acute lymphoblastic leukaemia,” *Lancet Oncol.*, vol. 11, no. 11, pp. 1096–1106, Nov. 2010.
- [33] D. Pinkel, “Five-year follow-up of ‘total therapy’ of childhood lymphocytic leukemia,” *JAMA*, vol. 216, no. 4, pp. 648–652, Apr. 1971.
- [34] H. Riehm, H. Gadner, G. Henze, H.-J. Langermann, and E. Odenwald, “The Berlin Childhood Acute Lymphoblastic Leukemia Therapy Study, 1970-1976,” *J. Pediatr. Hematol. Oncol.*, vol. 2, Jan. 1980.
- [35] L. K. Jones and V. Saha, “Philadelphia positive acute lymphoblastic leukaemia of childhood,” *Br. J. Haematol.*, vol. 130, no. 4, pp. 489–500, Aug. 2005.
- [36] R. Pieters *et al.*, “A treatment protocol for infants younger than 1 year with acute lymphoblastic leukaemia (Interfant-99): an observational study and a multicentre randomised trial,” *Lancet Lond. Engl.*, vol. 370, no. 9583, pp. 240–250, Jul. 2007.
- [37] M. Dördelmann *et al.*, “Prednisone Response Is the Strongest Predictor of Treatment Outcome in Infant Acute Lymphoblastic Leukemia,” *Blood*, vol. 94, no. 4, pp. 1209–1217, Aug. 1999.
- [38] S. J. Russell, P. Norvig, and E. Davis, *Artificial intelligence: a modern approach*. Upper Saddle River, NJ: Prentice Hall, 2010.
- [39] X. Χαλκίδη and M. Βαζιργιάννης, *Εξόρυξη Γνώσης από Βάσεις Δεδομένων και τον Παγκόσμιο Ιστό*, 2nd ed. Τυπωθήτω, 2005.
- [40] “History of biology,” *Wikipedia*. 12-Nov-2017.
- [41] “Online Education Kit: Timeline: From Darwin and Mendel to the Human Genome Project,” *National Human Genome Research Institute (NHGRI)*. [Online]. Available: <https://www.genome.gov/25019887/Online-Education-Kit-Timeline-From-Darwin-and-Mendel-to-the-Human-Genome-Project>. [Accessed: 01-Dec-2017].
- [42] “Η καταγωγή των ειδών,” *Βικιπαίδεια*. 18-Nov-2017.

- [43] “Genetics,” *Wikipedia*. 25-Nov-2017.
- [44] “Mendelian inheritance,” *Wikipedia*. 22-Oct-2017.
- [45] “Polymerase chain reaction,” *Wikipedia*. 30-Nov-2017.
- [46] “Genetic screen,” *Wikipedia*. 19-Oct-2017.
- [47] “Genetic linkage,” *Wikipedia*. 27-Sep-2017.
- [48] “Bioinformatics,” *Wikipedia*. 30-Nov-2017.
- [49] “DNA sequencing,” *Wikipedia*. 30-Nov-2017.
- [50] “Pattern recognition,” *Wikipedia*. 05-Nov-2017.
- [51] “Data mining,” *Wikipedia*. 21-Nov-2017.
- [52] “Machine learning,” *Wikipedia*. 01-Dec-2017.
- [53] “Biological data visualization,” *Wikipedia*. 26-Oct-2017.
- [54] “Sequence assembly,” *Wikipedia*. 30-Nov-2017.
- [55] “Sequence alignment,” *Wikipedia*. 09-Nov-2017.
- [56] “Structural alignment,” *Wikipedia*. 15-Oct-2017.
- [57] “Protein structure prediction,” *Wikipedia*. 25-Oct-2017.
- [58] “Protein–protein interaction,” *Wikipedia*. 08-Oct-2017.
- [59] “Genome-wide association study,” *Wikipedia*. 17-Oct-2017.
- [60] H. Kitano, “Systems biology: a brief overview,” *Science*, vol. 295, no. 5560, pp. 1662–1664, Mar. 2002.
- [61] M. Ullah and O. Wolkenhauer, “Stochastic approaches in systems biology,” *Wiley Interdiscip. Rev. Syst. Biol. Med.*, vol. 2, no. 4, pp. 385–397, Aug. 2010.
- [62] “Avogadro constant,” *Wikipedia*. 04-Nov-2017.
- [63] F. Schlögl, “Chemical reaction models for non-equilibrium phase transitions,” *Z. Für Phys.*, vol. 253, no. 2, pp. 147–161, Apr. 1972.
- [64] “SBML,” *Wikipedia*. 30-Jun-2017.
- [65] M. Hucka *et al.*, “The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models,” *Bioinforma. Oxf. Engl.*, vol. 19, no. 4, pp. 524–531, Mar. 2003.
- [66] “Main Page - SBML.caltech.edu.” [Online]. Available: http://sbml.org/Main_Page. [Accessed: 01-Dec-2017].
- [67] “Systems Biology Graphical Notation (SBGN).” [Online]. Available: <http://sbgn.github.io/sbgn/>. [Accessed: 01-Dec-2017].
- [68] “XML,” *Wikipedia*. 17-Nov-2017.
- [69] “Systems Biology Graphical Notation,” *Wikipedia*. 15-Sep-2017.
- [70] S. Moodie, N. Le Novère, E. Demir, H. Mi, and A. Villéger, “Systems Biology Graphical Notation: Process Description language Level 1 Version 1.3,” *J. Integr. Bioinforma.*, vol. 12, no. 2, p. 263, Sep. 2015.
- [71] A. Sorokin *et al.*, “Systems Biology Graphical Notation: Entity Relationship language Level 1 Version 2,” *J. Integr. Bioinforma.*, vol. 12, no. 2, p. 264, Sep. 2015.
- [72] H. Mi *et al.*, “Systems Biology Graphical Notation: Activity Flow language Level 1 Version 1.2,” *J. Integr. Bioinforma.*, vol. 12, no. 2, p. 265, Sep. 2015.
- [73] D. T. Gillespie, “Stochastic simulation of chemical kinetics,” *Annu. Rev. Phys. Chem.*, vol. 58, pp. 35–55, 2007.
- [74] D. J. Wilkinson, “Stochastic modelling for quantitative description of heterogeneous biological systems,” *Nat. Rev. Genet.*, vol. 10, no. 2, pp. 122–133, 2009.
- [75] “Monte Carlo method,” *Wikipedia*. 15-Nov-2017.
- [76] D. T. Gillespie, “Exact stochastic simulation of coupled chemical reactions,” *J. Phys. Chem.*, vol. 81, no. 25, pp. 2340–2361, Dec. 1977.

- [77] D. T. Gillespie, “A general method for numerically simulating the stochastic time evolution of coupled chemical reactions,” *J. Comput. Phys.*, vol. 22, no. 4, pp. 403–434, Dec. 1976.
- [78] M. A. Gibson and J. Bruck, “Efficient Exact Stochastic Simulation of Chemical Systems with Many Species and Many Channels,” *J. Phys. Chem. A*, vol. 104, no. 9, pp. 1876–1889, Mar. 2000.
- [79] J. Pahle, “Biochemical simulations: stochastic, approximate stochastic and hybrid approaches,” *Brief. Bioinform.*, vol. 10, no. 1, pp. 53–64, Jan. 2009.
- [80] D. T. Gillespie, “Approximate accelerated stochastic simulation of chemically reacting systems,” *J. Chem. Phys.*, vol. 115, no. 4, pp. 1716–1733, Jul. 2001.
- [81] D. T. Gillespie, “The chemical Langevin equation,” *J. Chem. Phys.*, vol. 113, no. 1, pp. 297–306, Jun. 2000.
- [82] “Stochastic differential equation,” *Wikipedia*. 14-Nov-2017.
- [83] S. Sahle, P. Mendes, S. Hoops, and U. Kummer, “A new strategy for assessing sensitivities in biochemical models,” *Philos. Transact. A Math. Phys. Eng. Sci.*, vol. 366, no. 1880, pp. 3619–3631, Oct. 2008.
- [84] B. Ingalls, “Sensitivity analysis: from model parameters to system behaviour,” *Essays Biochem.*, vol. 45, pp. 177–193, 2008.
- [85] A. de la Fuente, J. L. Snoep, H. V. Westerhoff, and P. Mendes, “Metabolic control in integrated biochemical systems,” *Eur. J. Biochem.*, vol. 269, no. 18, pp. 4399–4408, 2002.
- [86] “Effector (biology),” *Wikipedia*. 07-Nov-2017.
- [87] “Lyapunov exponent,” *Wikipedia*. 23-Nov-2017.
- [88] “COPASI: Support/User Manual/Methods/Lyapunov Exponents Calculation.” [Online]. Available: http://copasi.org/Support/User_Manual/Methods/Lyapunov_Exponents_Calculation/. [Accessed: 01-Dec-2017].
- [89] I. Surovtsova, N. Simus, T. Lorenz, A. König, S. Sahle, and U. Kummer, “Accessible methods for the dynamic time-scale decomposition of biochemical systems,” *Bioinforma. Oxf. Engl.*, vol. 25, no. 21, pp. 2816–2823, Nov. 2009.
- [90] “COPASI: Support/User Manual/Methods/Time Scale Separation Methods.” [Online]. Available: http://copasi.org/Support/User_Manual/Methods/Time_Scale_Separation_Methods/. [Accessed: 01-Dec-2017].
- [91] “COPASI: Support/User Manual/Methods/Steady State Calculation.” [Online]. Available: http://copasi.org/Support/User_Manual/Methods/Steady_State_Calculation/. [Accessed: 01-Dec-2017].
- [92] M. Quach, N. Brunel, and F. d’Alché-Buc, “Estimating parameters and hidden variables in non-linear state-space models based on ODEs for biological networks inference,” *Bioinforma. Oxf. Engl.*, vol. 23, no. 23, pp. 3209–3216, Dec. 2007.
- [93] “Global optimization,” *Wikipedia*. 10-Nov-2017.
- [94] S. M. Baker, K. Schallau, and B. H. Junker, “Comparison of different algorithms for simultaneous estimation of multiple parameters in kinetic metabolic models,” *J. Integr. Bioinforma.*, vol. 7, no. 3, Mar. 2010.
- [95] C. G. Moles, P. Mendes, and J. R. Banga, “Parameter estimation in biochemical pathways: a comparison of global optimization methods,” *Genome Res.*, vol. 13, no. 11, pp. 2467–2474, Nov. 2003.
- [96] “COPASI: Support/User Manual/Methods/Optimization Methods.” [Online]. Available: http://copasi.org/Support/User_Manual/Methods/Optimization_Methods/. [Accessed: 01-Dec-2017].

- [97] “Parameter Estimation and Model Selection in Computational Biology.” [Online]. Available: <http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1000696>. [Accessed: 01-Dec-2017].
- [98] “Likelihood function,” *Wikipedia*. 20-Nov-2017.
- [99] “Bayesian statistics,” *Wikipedia*. 01-Nov-2017.
- [100] “Bayesian inference,” *Wikipedia*. 30-Nov-2017.
- [101] “Kalman filter,” *Wikipedia*. 20-Nov-2017.
- [102] “Particle filter,” *Wikipedia*. 15-Nov-2017.
- [103] T. Tian, S. Xu, J. Gao, and K. Burrage, “Simulated maximum likelihood method for estimating kinetic rates in gene expression,” *Bioinforma. Oxf. Engl.*, vol. 23, no. 1, pp. 84–91, Jan. 2007.
- [104] X. Sun, L. Jin, and M. Xiong, “Extended Kalman Filter for Estimation of Parameters in Nonlinear State-Space Models of Biochemical Networks,” *PLOS ONE*, vol. 3, no. 11, p. e3758, 2008.
- [105] *Mathematical Biology II - Spatial Models and Biomedical* / J.D. Murray / Springer. .
- [106] K. R. Swanson, E. C. Alvord, and J. D. Murray, “Virtual brain tumours (gliomas) enhance the reality of medical imaging and highlight inadequacies of current therapy,” *Br. J. Cancer*, vol. 86, no. 1, pp. 14–18, Jan. 2002.
- [107] C. J. W. Breward, H. M. Byrne, and C. E. Lewis, “A multiphase model describing vascular tumour growth,” *Bull. Math. Biol.*, vol. 65, no. 4, pp. 609–640, Jul. 2003.
- [108] O. Clatz *et al.*, “Realistic Simulation of the 3D Growth of Brain Tumors in MR Images Coupling Diffusion with Biomechanical Deformation,” *Ieee Trans. Med. Imaging*, vol. 24, no. 10, p. 1334, Oct. 2005.
- [109] H. B. Frieboes, X. Zheng, C.-H. Sun, B. Tromberg, R. Gatenby, and V. Cristini, “An integrated computational/experimental model of tumor invasion,” *Cancer Res.*, vol. 66, no. 3, pp. 1597–1604, Feb. 2006.
- [110] C. Guiot, P. P. Delsanto, A. Carpinteri, N. Pugno, Y. Mansury, and T. S. Deisboeck, “The dynamic evolution of the power exponent in a universal growth model of tumors,” *J. Theor. Biol.*, vol. 240, no. 3, pp. 459–463, Jun. 2006.
- [111] H. Enderling, M. A. J. Chaplain, A. R. A. Anderson, and J. S. Vaidya, “A mathematical model of breast cancer development, local treatment and recurrence,” *J. Theor. Biol.*, vol. 246, no. 2, pp. 245–259, May 2007.
- [112] W. Düchting and T. Vogelsaenger, “Three-dimensional pattern generation applied to spheroidal tumor growth in a nutrient medium,” *Int. J. Biomed. Comput.*, vol. 12, no. 5, pp. 377–392, Sep. 1981.
- [113] T. Ginsberg, W. Ulmer, and W. Duechting, “Computer simulation of fractionated radiotherapy: Further results and their relevance to percutaneous irradiation and brachytherapy,” *Strahlenther. Onkol.*, vol. 169, no. 5, pp. 304–310, 1993.
- [114] G. S. Stamatakos, D. D. Dionysiou, E. I. Zacharakis, N. A. Mouravliansky, K. S. Nikita, and N. K. Uzunoglu, “In silico radiation oncology: combining novel simulation algorithms with current visualization techniques,” *Proc. IEEE*, vol. 90, no. 11, pp. 1764–1777, 2002.
- [115] G. S. Stamatakos, V. P. Antipas, N. K. Uzunoglu, and R. G. Dale, “A four-dimensional computer simulation model of the in vivo response to radiotherapy of glioblastoma multiforme: studies on the effect of clonogenic cell density,” *Br. J. Radiol.*, vol. 79, no. 941, pp. 389–400, May 2006.
- [116] G. S. Stamatakos, V. P. Antipas, and N. K. Uzunoglu, “A spatiotemporal, patient individualized simulation model of solid tumor response to chemotherapy in vivo: the paradigm of glioblastoma multiforme treated by temozolomide,” *IEEE Trans. Biomed. Eng.*, vol. 53, no. 8, pp. 1467–1477, 2006.

- [117] G. S. Stamatakos *et al.*, “The ‘Oncosimulator’: a multilevel, clinically oriented simulation system of tumor growth and organism response to therapeutic schemes. Towards the clinical evaluation of in silico oncology,” *Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.*, vol. 2007, pp. 6629–6632, 2007.
- [118] D. D. Dionysiou, G. S. Stamatakos, N. K. Uzunoglu, K. S. Nikita, and A. Marioli, “A four-dimensional simulation model of tumour response to radiotherapy in vivo: parametric validation considering radiosensitivity, genetic profile and fractionation,” *J. Theor. Biol.*, vol. 230, no. 1, pp. 1–20, 2004.
- [119] D. D. Dionysiou, G. S. Stamatakos, N. K. Uzunoglu, and K. S. Nikita, “A computer simulation of in vivo tumour growth and response to radiotherapy: New algorithms and parametric results,” *Comput. Biol. Med.*, vol. 36, no. 5, pp. 448–464, 2006.
- [120] D. D. Dionysiou and G. S. Stamatakos, “Applying a 4D multiscale in vivo tumor growth model to the exploration of radiotherapy scheduling: The effects of weekend treatment gaps and p53 gene status on the response of fast growing solid tumors,” *Cancer Inform.*, vol. 2, pp. 113–121, Feb. 2007.
- [121] T. S. Deisboeck, L. Zhang, J. Yoon, and J. Costa, “In silico cancer modeling: is it ready for primetime?,” *Nat. Clin. Pract. Oncol.*, vol. 6, no. 1, pp. 34–42, Jan. 2009.
- [122] “Mathematical modelling of cancer cell invasion of tissue - ScienceDirect.” [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0895717707001823>. [Accessed: 08-Dec-2017].
- [123] G. Stamatakos *et al.*, “The technologically integrated oncosimulator: combining multiscale cancer modeling with information technology in the in silico oncology context,” *IEEE J. Biomed. Health Inform.*, vol. 18, no. 3, pp. 840–854, May 2014.
- [124] G. S. Stamatakos *et al.*, “In silico oncology: exploiting clinical studies to clinically adapt and validate multiscale oncosimulators,” *Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.*, vol. 2013, pp. 5545–5549, 2013.
- [125] G. S. Stamatakos, D. D. Dionysiou, E. Georgiadi, E. Kolokotroni, S. Giatili, and N. Graf, “In Silico Oncology: Multiscale Modelling of Clinical Tumour Response to Treatment Based on Discrete Entity - Discrete Event Simulation. The Oncosimulator concept. 1st Virtual Physiological Human Conference, Organized by the Virtual Physiological Human Network of Excellence www.vph-noe.eu, FP7-ICT-2007-2, Project. 2010,” Brussels, Belgium.
- [126] Marias *et al.*, “Clinically driven design of multi-scale cancer models: the ContraCancrum project paradigm,” *Interface Focus*, vol. 1, no. 3, pp. 450–461, 2011.
- [127] E. C. Georgiadi, D. D. Dionysiou, N. Graf, and G. S. Stamatakos, “Towards in silico oncology: adapting a four dimensional nephroblastoma treatment model to a clinical trial case based on multi-method sensitivity analysis,” *Comput. Biol. Med.*, vol. 42, no. 11, pp. 1064–1078, Nov. 2012.
- [128] G. S. Stamatakos, E. C. Georgiadi, N. Graf, E. A. Kolokotroni, and D. D. Dionysiou, “Exploiting Clinical Trial Data Drastically Narrows the Window of Possible Solutions to the Problem of Clinical Adaptation of a Multiscale Cancer Model,” *PLoS ONE*, vol. 6, no. 3, p. e17594, 2011.
- [129] E. A. Kolokotroni, D. D. Dionysiou, N. K. Uzunoglu, and G. S. Stamatakos, “Studying the growth kinetics of untreated clinical tumors by using an advanced discrete simulation model,” *Math. Comput. Model.*, vol. 54, no. 9–10, pp. 1989–2006, 2011.
- [130] E. Kolokotroni, E. Ouzounoglou, M. Stanulla, D. Dionysiou, and G. S. Stamatakos, “In Silico Oncology: Developing and Clinically Adapting the Acute Lymphoblastic Leukemia (ALL) Oncosimulator by Exploiting Pathway Based Gene Expression Analysis in the Context of the

- ALL-BFM 2000 Clinical Study. Virtual Physiological Human Conference 2014 (VPH 2014).,” in *Virtual Physiological Human Conference 2014 (VPH 2014)*, Trondheim, Norway.
- [131] S. J. Morrison and J. Kimble, “Asymmetric and symmetric stem-cell divisions in development and cancer,” *Nature*, vol. 441, no. 7097, pp. 1068–1074, Jun. 2006.
- [132] R. G. Dale, “Radiobiological Considerations in Gynaecological Radiotherapy,” in *Radiation Oncology of Gynecological Cancers*, Springer, Berlin, Heidelberg, 1997, pp. 13–23.
- [133] E. Κολοκοτρώνη, “Υπολογιστικές προσομοιώσεις διαγνωστικών και θεραπευτικών τεχνικών που αφορούν σε φυσιολογικά και παθολογικά κυτταρικά συστήματα,” Διδακτορική Διατριβή, Πανεπιστήμιο Πατρών. Σχολή Επιστημών Υγείας. Τμήμα Ιατρικής, 2013.
- [134] G. S. Stamatakos and D. D. Dionysiou, “Introduction of Hypermatrix and Operator Notation into a Discrete Mathematics Simulation Model of Malignant Tumour Response to Therapeutic Schemes *In Vivo*. Some Operator Properties,” *Cancer Inform.*, vol. 2009, no. 7, pp. 239–251, Oct. 2009.
- [135] C. G. Mullighan, “Molecular genetics of B-precursor acute lymphoblastic leukemia,” *J. Clin. Invest.*, vol. 122, no. 10, pp. 3407–3415, Oct. 2012.
- [136] J. L. Young and R. W. Miller, “Incidence of malignant tumors in U. S. children,” *J. Pediatr.*, vol. 86, no. 2, pp. 254–258, Feb. 1975.
- [137] G. M. Cooper, *The Cell*, 2nd ed. Sinauer Associates, 2000.
- [138] E. A. Nigg, “Cyclin-dependent protein kinases: key regulators of the eukaryotic cell cycle,” *BioEssays News Rev. Mol. Cell. Dev. Biol.*, vol. 17, no. 6, pp. 471–480, Jun. 1995.
- [139] T. Haberichter *et al.*, “A systems biology dynamical model of mammalian G1 cell cycle progression,” *Mol. Syst. Biol.*, vol. 3, p. 84, 2007.
- [140] A. Zetterberg, O. Larsson, and K. G. Wiman, “What is the restriction point?,” *Curr. Opin. Cell Biol.*, vol. 7, no. 6, pp. 835–842, Dec. 1995.
- [141] S. A. Ezhevsky, A. Ho, M. Becker-Hapak, P. K. Davis, and S. F. Dowdy, “Differential regulation of retinoblastoma tumor suppressor protein by G(1) cyclin-dependent kinase complexes in vivo,” *Mol. Cell. Biol.*, vol. 21, no. 14, pp. 4773–4784, Jul. 2001.
- [142] K. Leibundgut, N. Schmitz, A. Tobler, A. R. Lüthy, and A. Hirt, “In childhood acute lymphoblastic leukemia the hypophosphorylated retinoblastoma protein, p110RB, is diminished, as compared with normal CD34+ peripheral blood progenitor cells,” *Pediatr. Res.*, vol. 45, no. 5 Pt 1, pp. 692–696, May 1999.
- [143] S. M. Rubin, “Deciphering the retinoblastoma protein phosphorylation code,” *Trends Biochem. Sci.*, vol. 38, no. 1, pp. 12–19, 2013.
- [144] A. Hirt, A. M. Schmid, F. Julmy, N. M. R. Schmitz, and K. Leibundgut, “Expression of cyclin A in childhood acute lymphoblastic leukemia cells reveals undisturbed G1-S phase transition and passage through the S phase,” *Leuk. Off. J. Leuk. Soc. Am. Leuk. Res. Fund UK*, vol. 23, no. 2, pp. 414–417, Feb. 2009.
- [145] L. V. Nguyen, R. Vanner, P. Dirks, and C. J. Eaves, “Cancer stem cells: an evolving concept,” *Nat. Rev. Cancer*, vol. 12, no. 2, pp. 133–143, 2012.
- [146] M. L. O’Connor *et al.*, “Cancer stem cells: A contentious hypothesis now moving forward,” *Cancer Lett.*, vol. 344, no. 2, pp. 180–187, Mar. 2014.
- [147] A. Kreso and J. E. Dick, “Evolution of the cancer stem cell model,” *Cell Stem Cell*, vol. 14, no. 3, pp. 275–291, Mar. 2014.
- [148] K. Rehe *et al.*, “Acute B lymphoblastic leukaemia-propagating cells are present at high frequency in diverse lymphoblast populations,” *EMBO Mol. Med.*, vol. 5, no. 1, pp. 38–51, Jan. 2013.
- [149] J. S. McClellan and R. Majeti, “The cancer stem cell model: B cell acute lymphoblastic leukaemia breaks the mould,” *EMBO Mol. Med.*, vol. 5, no. 1, pp. 7–9, Jan. 2013.

- [150] A. Hirt, A.-M. Schmid, R. A. Ammann, and K. Leibundgut, "In pediatric lymphoblastic leukemia of B-cell origin, a small population of primitive blast cells is noncycling, suggesting them to be leukemia stem cell candidates," *Pediatr. Res.*, vol. 69, no. 3, pp. 194–199, Mar. 2011.
- [151] A. Hirt, E. M. Werren, A. R. Luethy, J. Gerdes, and H. P. Wagner, "Cell cycle analysis in lymphoid neoplasia of childhood: differences among immunologic subtypes and similarities in the proliferation of normal and leukaemic precursor B cells," *Br. J. Haematol.*, vol. 80, no. 2, pp. 189–193, Feb. 1992.
- [152] N. M. R. Schmitz, K. Leibundgut, and A. Hirt, "CDK2 catalytic activity and loss of nuclear tethering of retinoblastoma protein in childhood acute lymphoblastic leukemia," *Leuk. Off. J. Leuk. Soc. Am. Leuk. Res. Fund UK*, vol. 19, no. 10, pp. 1783–1787, Oct. 2005.
- [153] N. M. R. Schmitz, A. Hirt, M. Aebi, and K. Leibundgut, "Limited redundancy in phosphorylation of retinoblastoma tumor suppressor protein by cyclin-dependent kinases in acute lymphoblastic leukemia," *Am. J. Pathol.*, vol. 169, no. 3, pp. 1074–1079, Sep. 2006.
- [154] E. Hulleman *et al.*, "Inhibition of glycolysis modulates prednisolone resistance in acute lymphoblastic leukemia cells," *Blood*, vol. 113, no. 9, pp. 2014–2021, Feb. 2009.
- [155] A. Hirt *et al.*, "Acute lymphoblastic leukaemia in childhood: cell proliferation without rest," *Br. J. Haematol.*, vol. 96, no. 2, pp. 366–368, Feb. 1997.
- [156] J. M. Boag *et al.*, "Altered glucose metabolism in childhood pre-B acute lymphoblastic leukaemia," *Leukemia*, vol. 20, no. 10, pp. 1731–1737, 2006.
- [157] A. B. Pardee, "G1 events and regulation of cell proliferation," *Science*, vol. 246, no. 4930, pp. 603–608, Nov. 1989.
- [158] B. Novák and J. J. Tyson, "A model for restriction point control of the mammalian cell cycle," *J. Theor. Biol.*, vol. 230, no. 4, pp. 563–579, Oct. 2004.
- [159] S. A. Ezhevsky, H. Nagahara, A. M. Vocero-Akbani, D. R. Gius, M. C. Wei, and S. F. Dowdy, "Hypo-phosphorylation of the retinoblastoma protein (pRb) by cyclin D:Cdk4/6 complexes results in active pRb," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 94, no. 20, pp. 10699–10704, Sep. 1997.
- [160] A. S. Lundberg and R. A. Weinberg, "Functional inactivation of the retinoblastoma protein requires sequential modification by at least two distinct cyclin-cdk complexes," *Mol. Cell. Biol.*, vol. 18, no. 2, pp. 753–761, Feb. 1998.
- [161] J. W. Harbour, R. X. Luo, A. Dei Santi, A. A. Postigo, and D. C. Dean, "Cdk phosphorylation triggers sequential intramolecular interactions that progressively block Rb functions as cells move through G1," *Cell*, vol. 98, no. 6, pp. 859–869, Sep. 1999.
- [162] A. Hirt, K. Leibundgut, A. R. Lüthy, N. von der Weid, and H. P. Wagner, "Cell birth and death in childhood acute lymphoblastic leukaemia: how fast does the neoplastic cell clone expand?," *Br. J. Haematol.*, vol. 98, no. 4, pp. 999–1001, Sep. 1997.
- [163] T. Zarkowska and S. Mittnacht, "Differential phosphorylation of the retinoblastoma protein by G1/S cyclin-dependent kinases," *J. Biol. Chem.*, vol. 272, no. 19, pp. 12738–12746, May 1997.
- [164] E. S. Knudsen and J. Y. Wang, "Dual mechanisms for the inhibition of E2F binding to RB by cyclin-dependent kinase-mediated RB phosphorylation," *Mol. Cell. Biol.*, vol. 17, no. 10, pp. 5771–5783, Oct. 1997.
- [165] E. S. Knudsen and J. Y. Wang, "Differential regulation of retinoblastoma protein function by specific Cdk phosphorylation sites," *J. Biol. Chem.*, vol. 271, no. 14, pp. 8313–8320, Apr. 1996.
- [166] K. Leibundgut, N. M. R. Schmitz, and A. Hirt, "Catalytic activities of G1 cyclin-dependent kinases and phosphorylation of retinoblastoma protein in mobilized peripheral blood CD34+

- hematopoietic progenitor cells,” *Stem Cells Dayt. Ohio*, vol. 23, no. 7, pp. 1002–1011, Aug. 2005.
- [167] F. A. Dick and S. M. Rubin, “Molecular mechanisms underlying RB protein function,” *Nat. Rev. Mol. Cell Biol.*, vol. 14, no. 5, pp. 297–306, May 2013.
- [168] J. I. Macdonald and F. A. Dick, “Posttranslational modifications of the retinoblastoma tumor suppressor protein as determinants of function,” *Genes Cancer*, vol. 3, no. 11–12, pp. 619–633, Nov. 2012.
- [169] F. A. Dick, “Structure-function analysis of the retinoblastoma tumor suppressor protein - is the whole a sum of its parts?,” *Cell Div.*, vol. 2, p. 26, 2007.
- [170] Y. Inoue, M. Kitagawa, and Y. Taya, “Phosphorylation of pRB at Ser612 by Chk1/2 leads to a complex between pRB and E2F-1 after DNA damage,” *EMBO J.*, vol. 26, no. 8, pp. 2083–2093, Apr. 2007.
- [171] T. Hattori, C. Uchida, H. Takahashi, N. Yamamoto, M. Naito, and Y. Taya, “Distinct and Site-Specific Phosphorylation of the Retinoblastoma Protein at Serine 612 in Differentiated Cells,” *PLoS ONE*, vol. 9, no. 1, Jan. 2014.
- [172] J. J. Tyson and B. Novak, “Regulation of the eukaryotic cell cycle: molecular antagonism, hysteresis, and irreversible transitions,” *J. Theor. Biol.*, vol. 210, no. 2, pp. 249–263, May 2001.
- [173] K. Iwamoto, Y. Tashima, H. Hamada, Y. Eguchi, and M. Okamoto, “Mathematical modeling and sensitivity analysis of G1/S phase in the cell cycle including the DNA-damage signal transduction pathway,” *Biosystems*, vol. 94, no. 1–2, pp. 109–117, Nov. 2008.
- [174] K. Iwamoto, H. Hamada, Y. Eguchi, and M. Okamoto, “Mathematical modeling of cell cycle regulation in response to DNA damage: exploring mechanisms of cell-fate determination,” *Biosystems*, vol. 103, no. 3, pp. 384–391, Mar. 2011.
- [175] B. Pfeuty, “Strategic cell-cycle regulatory features that provide mammalian cells with tunable G1 length and reversible G1 arrest,” *PloS One*, vol. 7, no. 4, p. e35291, 2012.
- [176] M. Swat, A. Kel, and H. Herzog, “Bifurcation analysis of the regulatory modules of the mammalian G1/S transition,” *Bioinforma. Oxf. Engl.*, vol. 20, no. 10, pp. 1506–1511, Jul. 2004.
- [177] G. Yao, T. J. Lee, S. Mori, J. R. Nevins, and L. You, “A bistable Rb-E2F switch underlies the restriction point,” *Nat. Cell Biol.*, vol. 10, no. 4, pp. 476–482, Apr. 2008.
- [178] G. Yao, T. J. Lee, S. Mori, J. R. Nevins, and L. You, “A bistable Rb-E2F switch underlies the restriction point,” *Nat Cell Biol*, vol. 10, no. 4, pp. 476–482, 2008.
- [179] R. Conradie *et al.*, “Restriction point control of the mammalian cell cycle via the cyclin E/Cdk2:p27 complex,” *FEBS J.*, vol. 277, no. 2, pp. 357–367, Jan. 2010.
- [180] B. D. Aguda and Y. Tang, “The kinetic origins of the restriction point in the mammalian cell cycle,” *Cell Prolif.*, vol. 32, no. 5, pp. 321–335, Oct. 1999.
- [181] J. Robert C Bast, D. W. Kufe, R. E. Pollock, R. R. Weichselbaum, J. F. Holland, and I. I. I. Emil Frei, “Holland-Frei Cancer Medicine,” 2000. [Online]. Available: <http://www.ncbi.nlm.nih.gov/books/NBK20777/>.
- [182] S. X. Guo *et al.*, “Hypermethylation of p16 and p15 genes and RB protein expression in acute leukemia,” *Leuk. Res.*, vol. 24, no. 1, pp. 39–46, Jan. 2000.
- [183] V. Chelliah, C. Laibe, and N. Novère, “BioModels Database: A Repository of Mathematical Models of Biological Processes,” in *In Silico Systems Biology*, vol. 1021, M. V. Schneider, Ed. Totowa, NJ: Humana Press, 2013, pp. 189–199.
- [184] S. Hoops *et al.*, “COPASI--a COMplex PATHway SIMulator,” *Bioinforma. Oxf. Engl.*, vol. 22, no. 24, pp. 3067–3074, Dec. 2006.

- [185] N. L. Novère *et al.*, “The Systems Biology Graphical Notation,” *Nat. Biotechnol.*, vol. 27, no. 8, pp. 735–741, 2009.
- [186] M. Hucka *et al.*, “The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models,” *Bioinformatics*, vol. 19, no. 4, pp. 524–531, Mar. 2003.
- [187] J. Kennedy and R. Eberhart, “Particle swarm optimization,” in *IEEE International Conference on Neural Networks, 1995. Proceedings, 1995*, vol. 4, pp. 1942–1948 vol.4.
- [188] J. W. Ludlow, C. L. Glendening, D. M. Livingston, and J. A. DeCarpio, “Specific enzymatic dephosphorylation of the retinoblastoma protein,” *Mol. Cell. Biol.*, vol. 13, no. 1, pp. 367–372, Jan. 1993.
- [189] E. Rubin, S. Mittnacht, E. Villa-Moruzzi, and J. W. Ludlow, “Site-specific and temporally-regulated retinoblastoma protein dephosphorylation by protein phosphatase type 1,” *Oncogene*, vol. 20, no. 29, pp. 3776–3785, Jun. 2001.
- [190] K. Renner, M. J. Ausserlechner, and R. Kofler, “A conceptual view on glucocorticoid-induced apoptosis, cell cycle arrest and glucocorticoid resistance in lymphoblastic leukemia,” *Curr. Mol. Med.*, vol. 3, no. 8, pp. 707–717, Dec. 2003.
- [191] K. Rhee, W. Bresnahan, A. Hirai, M. Hirai, and E. A. Thompson, “c-Myc and cyclin D3 (CcnD3) genes are independent targets for glucocorticoid inhibition of lymphoid cell proliferation,” *Cancer Res.*, vol. 55, no. 18, pp. 4188–4195, Sep. 1995.
- [192] M. J. Ausserlechner, P. Obexer, G. Böck, S. Geley, and R. Kofler, “Cyclin D3 and c-MYC control glucocorticoid-induced cell cycle arrest but not apoptosis in lymphoblastic leukemia cells,” *Cell Death Differ.*, vol. 11, no. 2, pp. 165–174, Feb. 2004.
- [193] D. Fernandes *et al.*, “Glucocorticoids inhibit proliferation, cyclin D1 expression, and retinoblastoma protein phosphorylation, but not activity of the extracellular-regulated kinases in human cultured airway smooth muscle,” *Am. J. Respir. Cell Mol. Biol.*, vol. 21, no. 1, pp. 77–88, Jul. 1999.
- [194] R. Addeo *et al.*, “Glucocorticoids induce G1 arrest of lymphoblastic cells through retinoblastoma protein Rb1 dephosphorylation in childhood acute lymphoblastic leukemia in vivo,” *Cancer Biol. Ther.*, vol. 3, no. 5, pp. 470–476, May 2004.
- [195] G. I. Lambrou *et al.*, “Prednisolone exerts late mitogenic and biphasic effects on resistant acute lymphoblastic leukemia cells: Relation to early gene expression,” *Leuk. Res.*, vol. 33, no. 12, pp. 1684–1695, Dec. 2009.
- [196] K. B. Petersen, W. J. Jusko, M. Rasmussen, and K. Schmiegelow, “Population pharmacokinetics of prednisolone in children with acute lymphoblastic leukemia,” *Cancer Chemother. Pharmacol.*, vol. 51, no. 6, pp. 465–473, Jun. 2003.
- [197] M. H. Magee, R. A. Blum, C. D. Lates, and W. J. Jusko, “Pharmacokinetic/pharmacodynamic model for prednisolone inhibition of whole blood lymphocyte proliferation,” *Br. J. Clin. Pharmacol.*, vol. 53, no. 5, pp. 474–484, May 2002.
- [198] J. Xu, J. Winkler, S. N. Sabarinath, and H. Derendorf, “Assessment of the impact of dosing time on the pharmacokinetics/pharmacodynamics of prednisolone,” *AAPS J.*, vol. 10, no. 2, pp. 331–341, Jun. 2008.
- [199] J. Xu, J. Winkler, and H. Derendorf, “A pharmacokinetic/pharmacodynamic approach to predict total prednisolone concentrations in human plasma,” *J. Pharmacokinet. Pharmacodyn.*, vol. 34, no. 3, pp. 355–372, Jun. 2007.
- [200] E. A. Gehan and S. L. George, “Estimation of human body surface area from height and weight,” *Cancer Chemother. Rep.*, vol. 54, no. 4, pp. 225–235, Aug. 1970.
- [201] G. Gatti, E. Perucca, G. M. Frigo, L. D. Notarangelo, L. Barberis, and A. Martini, “Pharmacokinetics of prednisone and its metabolite prednisolone in children with nephrotic

- syndrome during the active phase and in remission,” *Br. J. Clin. Pharmacol.*, vol. 17, no. 4, pp. 423–431, Apr. 1984.
- [202] E. A. Gehan and S. L. George, “Estimation of human body surface area from height and weight,” *Cancer Chemother. Rep.*, vol. 54, no. 4, pp. 225–235, Aug. 1970.
- [203] G. S. Stamatakos, N. Graf, and R. Radhakrishnan, “Multiscale Cancer Modeling and In Silico Oncology: Emerging Computational Frontiers in Basic and Translational Cancer Research,” *J Bioeng. Biomed. Sci* 3 E114 Doi, 2013.
- [204] G. S. Stamatakos, E. A. Kolokotroni, D. D. Dionysiou, E. C. Georgiadi, and C. Desmedt, “An advanced discrete state-discrete event multiscale simulation model of the response of a solid tumor to chemotherapy: Mimicking a clinical study,” *J. Theor. Biol.*, vol. 266, no. 1, pp. 124–139, Sep. 2010.
- [205] E. Kolokotroni *et al.*, “In Silico Oncology: Quantification of the In Vivo Antitumor Efficacy of Cisplatin-Based Doublet Therapy in Non-Small Cell Lung Cancer (NSCLC) through a Multiscale Mechanistic Model,” *PLoS Comput. Biol.*, vol. 12, no. 9, Sep. 2016.
- [206] G. Cario *et al.*, “Distinct gene expression profiles determine molecular treatment response in childhood acute lymphoblastic leukemia,” *Blood*, vol. 105, no. 2, pp. 821–826, Jan. 2005.
- [207] M. Stanulla and M. Schrappe, “Treatment of Childhood Acute Lymphoblastic Leukemia,” *Semin. Hematol.*, vol. 46, no. 1, pp. 52–63, Jan. 2009.
- [208] J. Cooperman, R. Neely, D. T. Teachey, S. Grupp, and J. K. Choi, “Cell division rates of primary human precursor B cells in culture reflect in vivo rates,” *Stem Cells Dayt. Ohio*, vol. 22, no. 6, pp. 1111–1120, 2004.
- [209] M. Tsurusawa, M. Ito, Z. Zha, S. Kawai, Y. Takasaki, and T. Fujimoto, “Cell-cycle-associated expressions of proliferating cell nuclear antigen and Ki-67 reactive antigen of bone marrow blast cells in childhood acute leukemia,” *Leukemia*, vol. 6, no. 7, pp. 669–674, Jul. 1992.
- [210] M. Tsurusawa, M. Aoyama, K. Saeki, and T. Fujimoto, “Cell cycle kinetics in childhood acute leukemia studied with in vitro bromodeoxyuridine labeling, Ki67-reactivity, and flow cytometry,” *Leuk. Off. J. Leuk. Soc. Am. Leuk. Res. Fund UK*, vol. 9, no. 11, pp. 1921–1925, Nov. 1995.
- [211] T. Ginsberg, “Modellierung und Simulation der Proliferationsregulation und Strahlentherapie normaler und maligner Gewebe,” *Fortschr.-Berichte VDI Reihe 17 Biotech.*, vol. 140, pp. 103–107, 1996.
- [212] M. Tsurusawa, M. Niwa, N. Katano, and T. Fujimoto, “Flow cytometric analysis by bromodeoxyuridine/DNA assay of cell cycle perturbation of methotrexate-treated mouse L1210 leukemia cells,” *Cancer Res.*, vol. 48, no. 15, pp. 4288–4293, Aug. 1988.
- [213] T. Ociepa *et al.*, “Simultaneous assessment of p53 and MDM2 expression in leukemic cells in response to initial prednisone therapy in children with acute lymphoblastic leukemia,” *Pol. J. Pathol. Off. J. Pol. Soc. Pathol.*, vol. 61, no. 4, pp. 199–205, 2010.
- [214] R Core Team, *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2016.
- [215] H. E. Skipper and S. Perry, “Kinetics of Normal and Leukemic Leukocyte Populations and Relevance to Chemotherapy,” *Cancer Res.*, vol. 30, no. 6, pp. 1883–1897, Jun. 1970.
- [216] K. Price, R. M. Storn, and J. A. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization (Natural Computing Series)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2005.
- [217] K. M. Mullen, D. Ardia, D. L. Gil, D. Windover, and J. Cline, “DEoptim: An R package for global optimization by differential evolution,” *R J.*, vol. 3, no. 1, pp. 27–34, 2010.

- [218] “Building Predictive Models in R Using the caret Package | Kuhn | Journal of Statistical Software.” [Online]. Available: <https://www.jstatsoft.org/article/view/v028i05>. [Accessed: 15-Dec-2016].
- [219] E. Ouzounoglou, E. Kolokotroni, M. Stanulla, and G. S. Stamatakos, “In Silico Oncology: Evaluating the Predictability of Acute Lymphoblastic Leukemia Patients’ Response to Treatment Utilizing a Multiscale Oncosimulator Model in Conjunction with Machine Learning Methods,” in *A.G. Hoekstra (Editor), VPH2016, book of abstracts*, Amsterdam, The Netherlands, 2016.
- [220] L. Breiman, “Random Forests,” *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001.
- [221] K. Hechenbichler and K. Schliep, “Weighted k-Nearest-Neighbor Techniques and Ordinal Classification,” 2004. [Online]. Available: <https://epub.ub.uni-muenchen.de/1769/>. [Accessed: 21-Dec-2016].
- [222] A. Liaw and M. Wiener, “Classification and Regression by randomForest,” *R News*, vol. 2, no. 3, pp. 18–22, 2002.
- [223] K. Schliep and S. Hechenbichler, *kknn: Weighted k-Nearest Neighbors. R package version 1.3.1*. 2016.
- [224] N. S. Altman, “An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression,” *Am. Stat.*, vol. 46, no. 3, pp. 175–185, Aug. 1992.
- [225] R. J. Samworth, “Optimal weighted nearest neighbour classifiers,” *Ann. Stat.*, vol. 40, no. 5, pp. 2733–2763, Oct. 2012.
- [226] “Decision tree learning - Wikipedia.” [Online]. Available: https://en.wikipedia.org/wiki/Decision_tree_learning. [Accessed: 04-Feb-2018].
- [227] “Ensemble learning,” *Wikipedia*. 30-Dec-2017.
- [228] J. A. Miller *et al.*, “Strategies for aggregating gene expression data: The collapseRows R function,” *BMC Bioinformatics*, vol. 12, no. 1, p. 322, Aug. 2011.
- [229] P. Langfelder and S. Horvath, “WGCNA: an R package for weighted correlation network analysis,” *BMC Bioinformatics*, vol. 9, no. 1, p. 559, Dec. 2008.
- [230] T. Aittokallio, “Dealing with missing values in large-scale studies: microarray data imputation and beyond,” *Brief. Bioinform.*, vol. 11, no. 2, pp. 253–264, Mar. 2010.
- [231] G. N. Brock, J. R. Shaffer, R. E. Blakesley, M. J. Lotz, and G. C. Tseng, “Which missing value imputation method to use in expression profiles: a comparative study and two selection schemes,” *BMC Bioinformatics*, vol. 9, no. 1, p. 12, Jan. 2008.
- [232] R. Jörnsten, H.-Y. Wang, W. J. Welsh, and M. Ouyang, “DNA microarray data imputation and significance analysis of differential expression,” *Bioinformatics*, vol. 21, no. 22, pp. 4155–4161, Nov. 2005.
- [233] A. W.-C. Liew, N.-F. Law, and H. Yan, “Missing value imputation for gene expression data: computational techniques to recover missing data from available information,” *Brief. Bioinform.*, vol. 12, no. 5, pp. 498–513, Sep. 2011.
- [234] J. Luengo, S. García, and F. Herrera, “On the choice of the best imputation methods for missing values considering three groups of classification methods,” *Knowl. Inf. Syst.*, vol. 32, no. 1, pp. 77–108, Jul. 2012.
- [235] O. Troyanskaya *et al.*, “Missing value estimation methods for DNA microarrays,” *Bioinformatics*, vol. 17, no. 6, pp. 520–525, Jun. 2001.
- [236] S. Hwang, “Comparison and evaluation of pathway-level aggregation methods of gene expression data,” *BMC Genomics*, vol. 13, no. Suppl 7, p. S26, Dec. 2012.
- [237] P. Khatri, M. Sirota, and A. J. Butte, “Ten Years of Pathway Analysis: Current Approaches and Outstanding Challenges,” *PLoS Comput Biol*, vol. 8, no. 2, p. e1002375, 2012.

- [238] V. Varadan, P. Mittal, C. J. Vaske, and S. C. Benz, “The Integration of Biological Pathway Knowledge in Cancer Genomics: A review of existing computational approaches,” *IEEE Signal Process. Mag.*, vol. 29, no. 1, pp. 35–50, 2012.
- [239] S. Hänzelmann, R. Castelo, and J. Guinney, “GSVA: gene set variation analysis for microarray and RNA-seq data,” *BMC Bioinformatics*, vol. 14, p. 7, 2013.
- [240] M. Morgan, S. Falcon, and R. Gentleman, *GSEABase: Gene set enrichment data structures and methods*. 2016.
- [241] M. L. Delignette-Muller and C. Dutang, “fitdistrplus: An R Package for Fitting Distributions,” *J. Stat. Softw.*, vol. 064, no. i04, 2015.
- [242] J. J. M. van Dongen, V. H. J. van der Velden, M. Brüggemann, and A. Orfao, “Minimal residual disease diagnostics in acute lymphoblastic leukemia: need for sensitive, fast, and standardized technologies,” *Blood*, vol. 125, no. 26, pp. 3996–4009, Jun. 2015.
- [243] T. Flohr *et al.*, “Minimal residual disease-directed risk stratification using real-time quantitative PCR analysis of immunoglobulin and T-cell receptor gene rearrangements in the international multicenter trial AIEOP-BFM ALL 2000 for childhood acute lymphoblastic leukemia,” *Leukemia*, vol. 22, no. 4, pp. 771–782, Apr. 2008.
- [244] D. Campana, “Minimal residual disease in acute lymphoblastic leukemia,” *Hematol. Am. Soc. Hematol. Educ. Program*, vol. 2010, pp. 7–12, 2010.
- [245] E. Ouzounoglou, D. Dionysiou, and G. S. Stamatakis, “Differentiation resistance through altered retinoblastoma protein function in acute lymphoblastic leukemia: in silico modeling of the deregulations in the G1/S restriction point pathway,” *BMC Syst. Biol.*, 2016.
- [246] J. C. Panetta, A. Sparreboom, C.-H. Pui, M. V. Relling, and W. E. Evans, “Modeling Mechanisms of In Vivo Variability in Methotrexate Accumulation and Folate Pathway Inhibition in Acute Lymphoblastic Leukemia Cells,” *PLoS Comput. Biol.*, vol. 6, no. 12, Dec. 2010.
- [247] G. Clapp and D. Levy, “A review of mathematical models for leukemia and lymphoma,” *Drug Discov. Today Dis. Models*, vol. 16, pp. 1–6, 2015.
- [248] A. Bucur *et al.*, “Workflow-driven clinical decision support for personalized oncology,” *BMC Med. Inform. Decis. Mak.*, vol. 16, no. Suppl 2, Jul. 2016.

Δημοσιεύσεις στα πλαίσια της διατριβής

Διεθνή Περιοδικά με Κριτές

- [1] E. Ouzounoglou, D. Dionysiou, and G. S. Stamatakos, “Differentiation resistance through altered retinoblastoma protein function in Acute Lymphoblastic Leukemia: In silico modeling of the deregulations in the G1/S restriction point pathway,” *BMC Syst. Biol.*
- [2] E. Ouzounoglou, E. Kolokotroni, M. Stanulla, and G. S. Stamatakos, “A study on the predictability of acute lymphoblastic leukaemia response to treatment using a hybrid oncosimulator,” *Interface Focus*, vol. 8, no. 1, p. 20160163, Feb. 2018.
- [3] V. Sakkalis et al., “Web-Based Workflow Planning Platform Supporting the Design and Execution of Complex Multiscale Cancer Models,” *IEEE J. Biomed. Health Inform.*, vol. 18, no. 3, pp. 824–831, 2014.

Διεθνή Συνέδρια με Κριτές

- [1] E. N. Ouzounoglou, D. D. Dionysiou, M. Stanulla, and G. S. Stamatakos, “Towards patient personalization of an Acute Lymphoblastic Leukemia Model during the oral administration of prednisone in children: Initiating the ALL Oncosimulator,” in *Advanced Research Workshop on In Silico Oncology and Cancer Investigation - The TUMOR Project Workshop (IARWISOCI), 2012 5th International*, 2012, pp. 1–4.
- [2] E. Kolokotroni, E. Ouzounoglou, M. Stanulla, D. Dionysiou, and G. S. Stamatakos, “In Silico Oncology: Developing and Clinically Adapting the Acute Lymphoblastic Leukemia (ALL) Oncosimulator by Exploiting Pathway Based Gene Expression Analysis in the Context of the ALL-BFM 2000 Clinical Study. Virtual Physiological Human Conference 2014 (VPH 2014).,” in *Virtual Physiological Human Conference 2014 (VPH 2014)*, Trondheim, Norway.
- [3] E. Ouzounoglou, E. Kolokotroni, M. Stanulla, and G. S. Stamatakos, “In Silico Oncology: Evaluating the Predictability of Acute Lymphoblastic Leukemia Patients’ Response to Treatment Utilizing a Multiscale Oncosimulator Model in Conjunction with Machine Learning Methods,” in A.G. Hoekstra (Editor), *VPH2016*, book of abstracts, Amsterdam, The Netherlands, 2016.

Παραδοτέα Ερευνητικών Προγραμμάτων (που σχετίζονται άμεσα με τη διατριβή)

- [1] D3.2 Finalized versions of the EU-US models for the specified clinical scenarios & integrated scenarios - Transatlantic TUMour MOdel Repositories (FP7--IST-247754)
- [2] D12.1 Architecture and information flow diagrams of the Oncosimulator and the biomechanism models – P-medicine (270089)
- [3] D12.2 Algorithms and initial versions of the Oncosimulator and the biomechanism models. – P-medicine (270089)
- [4] D12.3 Report on the development of the Oncosimulator and the utilization of the biomechanism models – P-medicine (270089)
- [5] D12.4 Report on the clinical adaptation and optimization of the Oncosimulator models. *In silico* studies – P-medicine (270089)
- [6] D12.5 The in silico experimentation system. Clinical validation and overall evaluation – P-medicine (270089)