

Ιόνιο Πανεπιστήμιο
Τμήμα Πληροφορικής

Πτυχιακή Εργασία
Έρευνα Μεταφραστικών Δεδομένων (Translational Data Research)

«Τσαμκιράνης Ματθαίος»

Επιβλέπων Καθηγητής: «Βλάμος Παναγιώτης»

Κέρκυρα, Σεπτέμβριος 2018

Έρευνα Μεταφραστικών Δεδομένων (Translational Data Research)

Τσαμκιράνης Ματθαίος

Επιτροπή Επίβλεψης Πτυχιακής / Διπλωματικής Εργασίας

Επιβλέπων Καθηγητής:

Βλάμος Παναγιώτης

Κέρκυρα, Σεπτέμβριος 2018

Περίληψη

Ο στόχος της εργασίας αυτής είναι η ανάλυση της λειτουργίας και της σημαντικότητας των Μεταφραστικών Συστημάτων (Translational Research Systems), τα οποία μπορούν να χρησιμοποιηθούν για την συλλογή και αποθήκευση διαφόρων τύπων δεδομένων.

Η Έρευνα για τα Μεταφραστικά Δεδομένα (Translational Data Research) έχει εφαρμογή σε πολλούς τομείς όπως στην βιοϊατρική (που είναι και το πιο αξιοποιήσιμο λόγω του μεγάλου όγκου δεδομένων που συλλέγονται καθημερινά), στην γενετική ζώων, στον εντοπισμό και την πρόωρη πρόληψη διάφορων μορφών καρκίνου, ακόμα στον εντοπισμό των επιπέδων της εγκληματικότητας.

Το κύριο πρόβλημα στην έρευνα των μεταφραστικών δεδομένων είναι οι μεγάλες απαιτήσεις για διαχείριση και οργάνωση του μεγάλου όγκου των δεδομένων που συλλέγονται καθημερινά σε διάφορους τομείς.

Τα βασικά κριτήρια για την επιλογή και χρήση των δεδομένων αυτών είναι η ποιότητα τους, το πόσο διαθέσιμα είναι προς χρήση και το κατά πόσο είναι επαναχρησιμοποιήσιμα. Για τον λόγο αυτό οι μηχανισμοί που θα αναπτυχθούν για την υλοποίηση της μεταφραστικής έρευνας θα πρέπει να έχουν την δυνατότητα να επιλέγουν τα καλύτερα προς χρήση δεδομένα που να πληρούν τις παραπάνω απαιτήσεις.

Στην εργασία θα γίνει τόσο βιβλιογραφική αναφορά σε διάφορους τομείς στους οποίους μπορεί να γίνει υλοποίηση της έρευνας μεταφραστικών δεδομένων, όσο και αναφορά στα συστήματα που έχουν αναπτυχθεί ήδη για την υλοποίηση της έρευνας, τα οποία δέχονται πληροφορίες από εξουσιοδοτημένους χρήστες έτσι ώστε να γίνεται εισαγωγή, εξαγωγή και μετασχηματισμός των δεδομένων αυτών σε μια κεντρική βάση δεδομένων.

Λέξεις – Κλειδιά

Μεταφραστικά συστήματα, μεταφραστικά δεδομένα, βιοϊατρική, έρευνα για την αντιμετώπιση του καρκίνου

Abstract

This thesis aims at the analysis of the operation and significance of the translational research that can be used for the collection, storage and extraction for various types of data.

This research can find use at many sectors such as biomedical research, animal genetics, locating and early prevention of cancer and estimating true levels of crime.

The main problem of the translational research is the great requirements for managing and organizing of the big amount of data that are being collected on a daily basis. The main criteria for the selection and use of these data are their quality, availability and reusability. For that reason the mechanisms that are going to be created for the implementation of the translational research must have the ability to be able to choose the best data according to the above requirements.

At this thesis there will be a reference to some of the sectors that the translational research can be implemented and there will be named some mechanisms that have already been developed for the specific research. Those mechanisms collect data from some authorized users so that there will be import, extraction and modification of the data at one central database.

Keywords

Translational platforms, translational data, biomedicine, cancer research

Περιεχόμενα

Περίληψη.....	iv
Abstract	v
Περιεχόμενα	vi
Κατάλογος Εικόνων	ix
Κατάλογος Πινάκων	x
Συντομογραφίες & Ακρωνύμια.....	xi
1. Εισαγωγή.....	1
1.1. Περιγραφή Του Προβλήματος	1
1.2. Ορισμός Μεταφραστικής Έρευνας	1
1.3. Μοντέλο Μεταφραστικού Συνεχές	3
1.4. Τομείς Εφαρμογής	6
1.4.1. Βιοϊατρική	6
1.4.2. Γενετική Ζώων	7
1.4.3. Εντοπισμός Καρκίνου	8
1.4.4. Εγκληματικότητα	8
1.4.5. Εντοπισμός Γλωσσικών Δυσλειτουργιών	9
Σύνδρομο Down	10
Αυτισμός Και Συναφείς Αναπτυξιακές Διαταραχές	11
Σύνδρομο Williams.....	11
Συγκεκριμένες Γλωσσικές Δυσλειτουργίες	12
Δια- Πληθυσμιακές Μελέτες	12
Διαχρονικές Μελέτες	13
1.5. Στόχος της Εργασίας	13
1.6. Μεθοδολογία	14
1.7. Περιεχόμενα	14
2. Συστήματα Διαχείρισης Μεταφραστικών Δεδομένων.....	16
2.1. Γενικά Συστήματα.....	16
2.1.1. tranSMART.....	16
Δυνατότητες Πλατφόρμας	18
Εφαρμογές της πλατφόρμας	19
2.1.2. i2b2 (Informatics for Integrating Biology and the Bedside)	19
Αρχιτεκτονική του i2b2	20
2.1.3. iDASH (Integrating Data for Analysis, Anonymization and SHaring)	21
Τομείς Εφαρμογής	22
Διαφύλαξη Ιδιωτικότητας	23
Εργαλεία Ανάλυσης	24
2.1.4. BRISK (Biology-Related Information Storage Kit)	25
2.1.5. STRIDE (Stanford Translational Research Integrated Database Environment	27
Αρχιτεκτονική της στοίβα STRIDE	28
Εφαρμογές	30

Λειτουργίες	30
2.2. Συστήματα Εξειδικευμένα Στην Έρευνα Για Τον Καρκίνο.....	31
2.2.1. cBIO Cancer Genomics Portal	32
Υποστηριζόμενοι Τύποι Δεδομένων	32
Χαρακτηριστικά cBio.....	33
2.2.2. iCOD (Integrated Clinical Omics Database).....	35
Υποστηριζόμενοι Τύποι Δεδομένων	36
Εργαλεία iCOD.....	36
2.2.3. G-DOC (Georgetown Database of Cancer)	36
Υποστηριζόμενοι Τύποι Δεδομένων	37
Βασικά Χαρακτηριστικά του G-DOC	39
Βάση Δεδομένων	40
Στόχοι Της Πλατφόρμας.....	40
2.2.4. caTRIP (Cancer Translational Research Informatics Platform)	41
Εργαλεία και Εφαρμογές caBIG	41
2.2.5. Μοντέλα Ξενομοσχεύματων Όγκων Που Προέρχονται Από Τον Ασθενή (Patient-Derived Tumor Xenograft - PDX)	43
Μεθοδολογία Ανάπτυξης Των PDX Μοντέλων.....	44
Εφαρμογές των μοντέλων PDX στην έρευνα για τον καρκίνο	44
Περιορισμοί Των Μοντέλων PDX.....	48
3. Σύγκριση Μεταφραστικών Συστημάτων Διαχείρισης Δεδομένων	50
3.1. Περιεχόμενο Δεδομένων	50
3.1.1. Κλινικά Δεδομένα	50
3.1.2. Δεδομένα Omics	51
3.2. Υποστήριξη Διαλειτουργικότητας	51
3.3. Λειτουργίες Υποστήριξης Ανάλυσης.....	52
3.3.1. Στατιστικά και Αναλυτικά Εργαλεία	52
3.3.2. Εργαλεία Απεικόνισης	52
3.4. Εργαλεία Ασφάλειας Και Προστασίας Της Ιδιωτικής Ζωής	53
3.4.1. Ιδιωτικότητα Δεδομένων.....	53
3.5. Υποστήριξη Συστημάτων.....	54
3.5.1. Τεχνική Τεκμηρίωση.....	54
3.5.2. Εγκατάσταση Και Διαχείριση Της Πλατφόρμας	54
4. Απαιτήσεις των Μεταφραστικών Ερευνητικών Συστημάτων	56
4.1. Ιδιωτικότητα.....	56
4.2. Διαλειτουργικότητα Και Πρότυπα	56
4.3. Ετερογένεια Δεδομένων.....	57
4.4. Εγκατάσταση Και Συντήρηση	58
5. Προκλήσεις της Μεταφραστικής Έρευνας	59
5.1. Έλλειψη Σαφούς Μοντέλου Μετάφρασης.....	59
5.2. Πολιτιστικές Διαφορές.....	60
5.3. Έλλειψη Πόρων	60
5.4. Ρυθμιστικά Θέματα	60

5.5. Έλλειψη Προτύπων Ρόλων	61
5.6. Τα Μεταφραστικά Δεδομένα Έρευνας Είναι Μεγάλα Δεδομένα.....	61
5.7. Ηθική Της Έρευνας Για Την Μεταφραστική Επιστήμη	63
6. Συμπεράσματα	67
6.1. Συμπεράσματα	67
6.2. Προοπτικές Και Επεκτάσεις.....	67
6.2.1. Κατάρτιση Και Καθοδήγηση	68
6.2.2. Σταδιοδρομία.....	68
6.2.3. Εξορθολογισμός Των Ρυθμιστικών Θεμάτων.....	69
6.3. Μελλοντικές Ανάγκες	69
Βιβλιογραφία.....	70

Κατάλογος Εικόνων

Εικόνα 1: Μοντέλο μεταφραστικής έρευνας όπως έχει προταθεί από την Επιτροπής Αξιολόγησης της Ένωσης για την Εκπαίδευση στην Κλινική Έρευνα (Evaluation Committee of the Association for Clinical Research Training -ACRT) [3].	2
Εικόνα 2: Μεταφραστικό συνεχές στη βιοϊατρική έρευνα [1].	4
Εικόνα 3: Διάγραμμα του συστήματος tranSMART. Τα δημόσια και ιδιωτικά δεδομένα ευθυγραμμίζονται με τις τυπικές οντολογίες και επιτηρούνται και υποβάλλονται σε επεξεργασία ETL για αποθήκευση σε κεντρική αποθήκη δεδομένων [14].	17
Εικόνα 4: Περιγραφή σχήματος κεντρικής βάσης (i2b2 star schema) [29].	21
Εικόνα 5: Το iDASH, το νεότερο Εθνικό Κέντρο Βιοϊατρικής Πληροφορικής, είναι μια επιστημονική ομάδα που συγκεντρώνει μια πολυεθνική ομάδα επιστημόνων για την ανάπτυξη αλγορίθμων, εργαλείων και υπηρεσιών και βιοϊατρικής υποδομής στον κυβερνοχώρο για χρήση από βιοϊατρικούς και ερευνητές συμπεριφοράς. Στο σχήμα αυτό απεικονίζεται οι κυριότερες συνιστώσες σε κάθε περιοχή [30].	23
Εικόνα 6: Η δομή και λειτουργικότητα των τριών υπηρεσιών του BRISK [31].	26
Εικόνα 7: Αρχιτεκτονική της στοίβα STRIDE [33]	29
Εικόνα 9: Η πύλη γονιδιωματικής του καρκίνου του cBio. Η cBio Cancer Genomics Portal είναι μια ανοιχτή πλατφόρμα για την διαδραστική διερεύνηση πολυδιάστατων συνόλων δεδομένων γονιδιωματικής του καρκίνου στο πλαίσιο των κλινικών δεδομένων και των βιολογικών οδών [38].	32
Εικόνα 10: Βασικά στιγμιότυπα οθόνης του συστήματος iCOD [24].	35
Εικόνα 11: Επισκόπηση δεδομένων και χαρακτηριστικών ανάλυσης στο G-DOC. Τα δεδομένα (δημόσια ή ιδιωτικά) επεξεργάζονται ομοιόμορφα μέσω τυποποιημένων αγωγών βιοπληροφορικής και διατίθενται σε διάφορα εργαλεία ανάλυσης μέσω μιας Web διεπαφής φιλικής προς το χρήστη [44].	37
Εικόνα 12: Προβολή υψηλού επιπέδου της αρχιτεκτονικής caTRIP, συμπεριλαμβανομένων των Υπηρεσιών τομέα, του μηχανισμού κατανεμημένων ερωτημάτων και του γραφικού Διασύνδεση χρήστη. Domain Services, Distributed Query Engine, and Graphical User Interface [48].	42
Εικόνα 12 : Προκλινική διαλογή και μελέτη βιοδεικτών σε μοντέλα PDX [50].	46
Εικόνα 13: Συν-κλινική δοκιμή με μοντέλα PDX [50]	47
Εικόνα 13: Το πλήρες αρχείο υγείας - Περιλαμβάνει όλες τις πληροφορίες υγείας για έναν ασθενή, όχι μόνο αυτό που βρίσκεται στο ηλεκτρονικό ιατρικό αρχείο (EMR) [72].	62

Κατάλογος Πινάκων

Πίνακας 1: Συγκριτικός πίνακας συστημάτων μεταφραστικών δεδομένων [26].	55
---	----

Συντομογραφίες & Ακρωνύμια

ACRT	Association for Clinical Research Training
AllerGen	Allergy, Genes and Environment Network
BRISK	Biology-Related Information Storage Kit
caBIG	Cancer Biomedical Informatics Grid
CAE	Cancer Annotation Engine
caTIES	cancer Text Information Extraction System
caTRIP	cancer Translational Research Informatics Platform
CEnR	Community-Engaged Research
DCCC	Duke Comprehensive Cancer Center
EHR	Electronic Health Records
EMR	Electronic Medical Record
ETL	Extract, Transform, Load
FDA	Food and Drug Administration
G-DOC	Georgetown Database of Cancer
GWAS	Genome-Wide Association Studies
HIPAA	Health Insurance Portability and Accountability Act –
i2b2	informatics for Integrating Biology and the Bedside
ICGC	International Cancer Genome Consortium
iCOD	integrated Clinical Omics Database
iDASH	integrating Data for Analysis, Anonymization and Sharing
KD	Kawasaki Disease
KEGG	Kyoto Encyclopaedia of Genes and Genomes
MHRA	Medicines and Healthcare Regulatory Agency
MSKCC	Memorial Sloan-Kettering Cancer Centre
NCI	National Cancer Institute
NIH	National Institutes of Health
OCEANS	Observational Cohort Event Analysis and Notification System
SLIMS	Sample-Based Laboratory Information Management System
TCGA	The Cancer Genome Atlas
TR	Translational Research

1. Εισαγωγή

1.1. Περιγραφή Του Προβλήματος

Η διαδικασία της μετάφρασης βασικών επιστημονικών ανακαλύψεων σε κλινικές εφαρμογές και τελικά σε βελτιώσεις στη δημόσια υγεία έχει αναδειχθεί ως ένας σημαντικός αλλά και δύσκολος στόχος στη βιοϊατρική έρευνα. Η διαδικασία περιγράφεται καλύτερα ως "Μεταφραστικό Συνεχές " (Translation Continuum) [1], διότι διάφοροι πόροι και ενέργειες εμπλέκονται σε αυτή την εξέλιξη της γνώσης, η οποία προωθεί τις ανακαλύψεις από τον πάγκο του ερευνητή στο κρεβάτι του ασθενούς (bench to bedside).

Το τρέχον μοντέλο αυτού του συνεχούς επικεντρώνεται πρωτίστως στην μεταφραστική έρευνα, η οποία είναι απλώς ένα στοιχείο της συνολικής διαδικασίας μετάφρασης.

Η μεταφορά της επιστημονικής γνώσης από το «πάγκο στο κρεβάτι» [2] συμβαίνει μέσω μιας διαδικασίας που ονομάζεται μετάφραση. Αυτός ο όρος, μετάφραση, περιγράφει το μετασχηματισμό της γνώσης μέσω διαδοχικών πεδίων έρευνας από μια βασική επιστημονική ανακάλυψη στην επίδραση της δημόσιας υγείας - μια πολύπλοκη διαδικασία που απαιτεί τόσο έρευνα (π.χ. εργαστήρια και κλινικές δοκιμές) όσο και δραστηριότητες μη έρευνας (π.χ. υλοποίηση) .

Η σειρά αυτών των μεταφραστικών δραστηριοτήτων συμβαίνει κατά μήκος ενός «συνεχούς» (Εικόνα 2), το οποίο μπορεί να χρησιμοποιηθεί ως μοντέλο για την ακριβή περιγραφή ολόκληρης της διαδικασίας μετάφρασης.

1.2. Ορισμός Μεταφραστικής Έρευνας

Η Μεταφραστική Έρευνα (Translational Research) δεν είναι σαφώς καθορισμένη, επομένως οι συντονιστές ερευνητικών προγραμμάτων προσπαθούν να διατυπώσουν συγκεκριμένους στόχους του προγράμματος, να οριοθετήσουν τις γνώσεις και τις δεξιότητες (ικανότητες) που αναμένεται να αναπτύξουν οι εκπαιδευόμενοι, να δημιουργήσουν ένα κατάλληλο πρόγραμμα σπουδών και να παρακολουθήσουν τα αποτελέσματα για να αξιολογήσουν εάν οι στόχοι και οι απαιτήσεις έχουν επιτευχθεί.

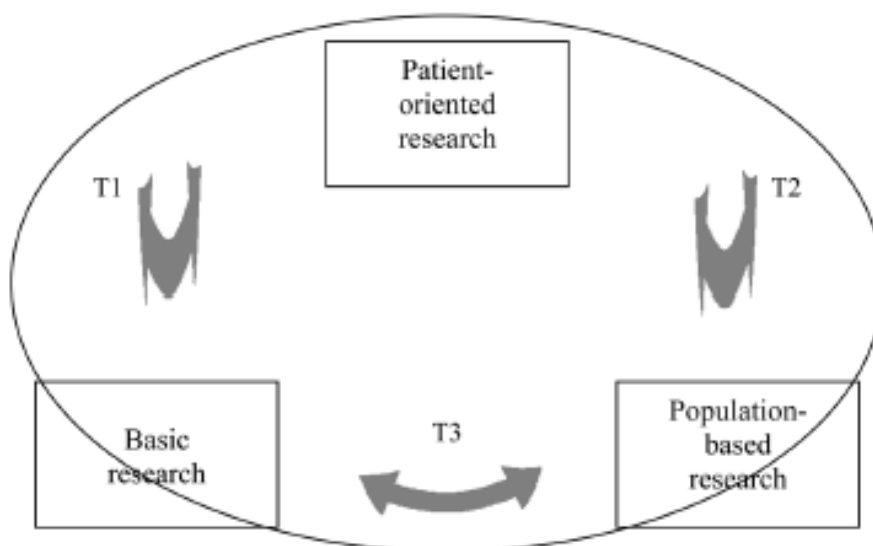
Τα μέλη της Επιτροπής Αξιολόγησης της Ένωσης για την Εκπαίδευση στην Κλινική Έρευνα (Association for Clinical Research Training - ACRT) εξέτασαν τους τρέχοντες ορισμούς της μεταφραστικής έρευνας και πρότειναν έναν επιχειρησιακό ορισμό που θα χρησιμοποιηθεί στο εκπαιδευτικό πλαίσιο.

Σύμφωνα με το [3], οι συγγραφείς υποστηρίζουν ότι η μεταφραστική έρευνα ενθαρρύνει την πολυδιάστατη και διεπιστημονική ενσωμάτωση της βασικής έρευνας, της έρευνας προσανατολισμένης στον ασθενή και της έρευνας με βάση τον πληθυσμό, με μακροπρόθεσμο στόχο τη βελτίωση της υγείας του κοινού.

Σε πρόσφατη ανακοίνωση, η NIH πρότεινε τον ακόλουθο ορισμό:

Η μεταφραστική έρευνα περιλαμβάνει δύο τομείς μετάφρασης. Η μία είναι η διαδικασία εφαρμογής ανακαλύψεων που προκύπτουν κατά τη διάρκεια της εργαστηριακής έρευνας και σε προ-κλινικών μελετών, στην ανάπτυξη δοκιμών και μελετών στον άνθρωπο. Ο δεύτερος τομέας της μετάφρασης αφορά την έρευνα με στόχο την ενίσχυση της υιοθέτησης βέλτιστων πρακτικών στην κοινότητα. Ο οικονομικά αποδοτικός τρόπος πρόληψης και θεραπείας αποτελεί επίσης σημαντικό μέρος της μεταφραστικής επιστήμης.

Σύμφωνα με αυτόν τον ορισμό, η μεταφραστική έρευνα αποτελεί μέρος μιας μονοκατευθυντικής συνέχειας στην οποία τα ερευνητικά ευρήματα μετακινούνται από τον ερευνητή στον ασθενή και την κοινότητα. Το πρώτο στάδιο της μεταφραστικής έρευνας (T1) μεταφέρει δεδομένα από τη βασική έρευνα στην κλινική έρευνα, ενώ το δεύτερο στάδιο (T2) μεταφέρει ευρήματα από κλινικές μελέτες ή κλινικές δοκιμές σε πρακτικό περιβάλλον και κοινότητες, όπου τα ευρήματα βελτιώνουν την υγεία (Εικόνα 1).



Εικόνα 1: Μοντέλο μεταφραστικής έρευνας όπως έχει προταθεί από την Επιτροπή Αξιολόγησης της Ένωσης για την Εκπαίδευση στην Κλινική Έρευνα (Evaluation Committee of the Association for Clinical Research Training -ACRT) [3].

Η Ομάδα Εργασίας Μεταφραστικής Έρευνας του Εθνικού Ινστιτούτου Καρκίνου (Translational Research Working Group of the National Cancer Institute -NCI) περιέλαβε βασική και κλινική έρευνα στο τμήμα T1 του συνεχούς:

"Η μεταγραφική έρευνα μετατρέπει τις επιστημονικές ανακαλύψεις που προκύπτουν από εργαστηριακές, κλινικές ή πληθυσμιακές μελέτες σε κλινικές εφαρμογές για τη μείωση της επίπτωσης, της νοσηρότητας και της θνησιμότητας από τον καρκίνο."

Ωστόσο, δεδομένου ότι η βασική και η κλινική έρευνα περιλαμβάνουν εγγενώς διαφορετικά σύνολα γνώσεων και μεθοδολογίες, το ίδιο τμήμα του συνεχούς (δηλαδή, στην T1) αποκρύπτει το γεγονός ότι διεπιστημονική μεταφραστική έρευνα μπορεί επίσης να συμβεί στη βασική και της κλινική επιστήμη.

Συμπερασματικά θα λέγαμε ότι η μεταφραστική έρευνα είναι η διαδικασία εφαρμογής ιδεών, γνώσεων και ανακαλύψεων που προκύπτουν μέσω βασικής επιστημονικής έρευνας για τη θεραπεία και την πρόληψη της ανθρώπινης νόσου - την κρίσιμη γέφυρα μεταξύ βασικής και κλινικής έρευνας.

Τα βήματα της μεταφραστικής έρευνας έχουν σχεδιαστεί για να εξασφαλίσουν ότι τα ευρήματα που εφαρμόζονται σε ανθρώπινες δοκιμές έχουν την υψηλότερη πιθανότητα επιτυχίας όσον αφορά την ασφάλεια και την ιδιωτικότητα.

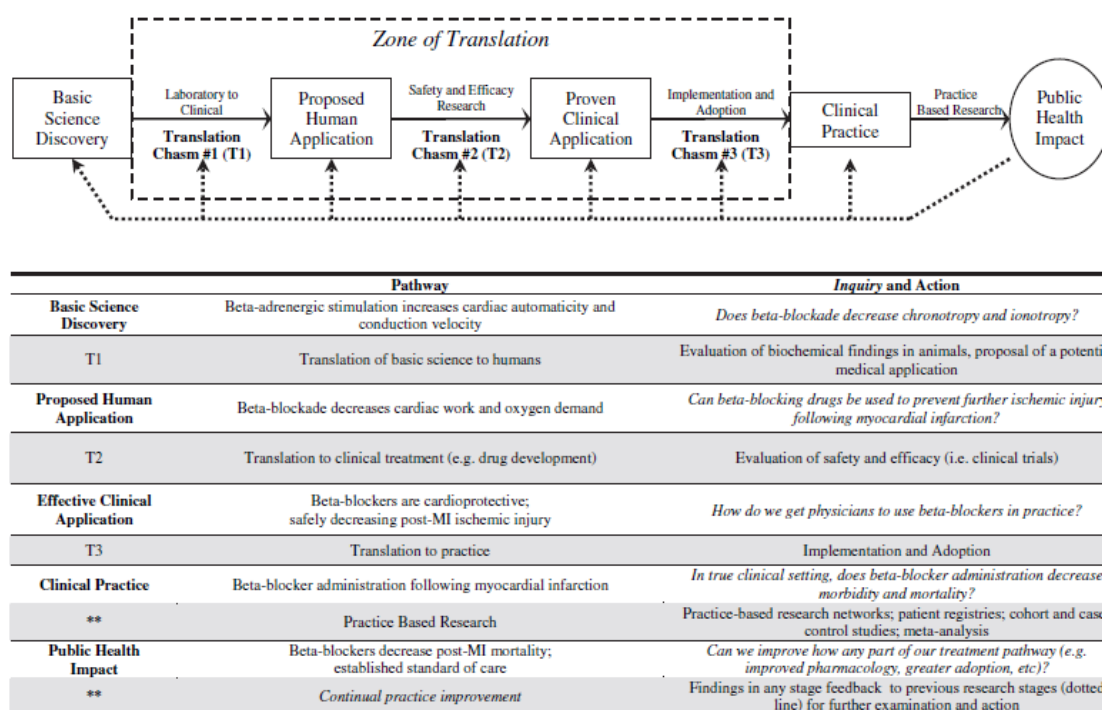
1.3. Μοντέλο Μεταφραστικού Συνεχές

Το μοντέλο του μεταφραστικού συνεχές (Εικόνα 2) αναπαριστά ολόκληρη τη διαδικασία μετάφρασης, από την βασική επιστημονική γνώση στη βελτίωση της δημόσιας υγείας. Αυτό το μοντέλο έχει 4 πρακτικά ορόσημα στη διαδικασία της γνώσης, τα οποία χωρίζονται από τα κενά που αντιστοιχούν στην προηγουμένως περιγραφείσα «T-ορολογία» (T1-T3) [4].

Τα κενά αυτά αναφέρονται ως «μεταφραστικά χάσματα», με σκοπό να αντιπροσωπεύσουν τις περιόδους κατά τις οποίες οι μεταφραστικές δραστηριότητες θα γεφυρώσουν τα 4 στάδια της ερευνητικής διαδικασίας.

Για παράδειγμα, για τη γεφύρωση του πρώτου μεταφραστικού χάσματος (T1), η μετάφραση πρέπει να λάβει χώρα στα δεδομένα της βασικής επιστημονικής έρευνας στο πλαίσιο μιας ενδεχόμενης ιατρικής εφαρμογής. Αν η επιστημονική ανακάλυψη είναι γενετική ή βιοχημική, τότε συνεχίζεται η εργαστηριακή διερεύνηση *in vitro* και μπορεί να

είναι απαραίτητα ζωικά μοντέλα για να γεφυρωθεί η T1 και να δημιουργηθεί μια πιθανή κλινική χρήση.



Εικόνα 2: Μεταφραστικό συνεχές στη βιοϊατρική έρευνα [1].

Επομένως, δραστηριότητες μετάφρασης T1 θα έχουν παρόμοιους στόχους (τη μετάφραση της βασικής γνώσης σε εφαρμογή στον άνθρωπο) αλλά θα είναι μοναδική για κάθε ανακάλυψη που εισάγεται στο μεταφραστικό συνεχές.

Το ίδιο ισχύει για κάθε επόμενο μεταφραστικό χάσμα και τη γεφύρωσή του. Αφού έχει προταθεί κάποια εφαρμογή στον άνθρωπο, μοντέλα ζώων και ανθρώπινες έρευνες (clinical trials) είναι απαραίτητες για την αξιολόγηση της ασφάλειας και της αποτελεσματικότητας των παρεμβάσεων. Αυτές οι δοκιμές γεφυρώνουν το χάσμα T2 σε μια αποδεκτή ανθρώπινη εφαρμογή.

Τέλος, η T3 γεφυρώνεται όταν εφαρμόζονται κλινικές πρακτικές και κατευθυντήριες γραμμές και υιοθετούνται τελικά για να επιφέρουν αλλαγές στη δημόσια υγεία.

Για την κατανόηση της μετάφρασης είναι απαραίτητη η κατανόηση ενός νέου λεξικού που αναπτύχθηκε στην βιβλιογραφία και ίσως είναι δύσκολο να μεταφραστεί χωρίς το κατάλληλο υπόβαθρο.

Η IOM Clinical Research Roundtable (IOM) περιγράφει για πρώτη φορά την τρέχουσα ορολογία και το μοντέλο της μεταφραστικής έρευνας (Translational Research) το 2003. Περιέγραψαν μια διαδικασία 2 φάσεων έρευνας που εξελίσσεται από

- (1) τη βασική επιστήμη στην κλινική επιστήμη και
- (2) από την κλινική επιστήμη στην επίδραση της δημόσιας υγείας.

Στο πλαίσιο τους, εντόπισαν «μεταφραστικά μπλοκ» μεταξύ αυτών των βημάτων που αποτελούν εμπόδια στην πρόοδο της έρευνας και αντιπροσωπεύουν σημαντικούς τομείς πρόκλησης για την επίτευξη βελτιώσεων της υγείας από τις βασικές επιστήμες.

Η δεύτερη φάση της μετάφρασης αργότερα υποδιαιρέθηκε δημιουργώντας ένα μοντέλο των ακόλουθων μεταφραστικών φάσεων:

- (T1) βασική επιστήμη στην κλινική επιστήμη
- στην κλινική πρακτική (T2) και
- στις βελτιώσεις υγείας (T3)

Οι συγγραφείς αυτού του μοντέλου χρησιμοποιούν τους όρους T1, T2 και T3 για να περιγράψουν τις 3 περιόδους μετάφρασης σε αντίστοιχη σειρά αλλά παραδόξως δεν αναλύουν αυτούς τους όρους στο χειρόγραφο τους [5]. Είτε εκ προθέσεως είτε όχι, αυτή η «ορολογία T» έχει γίνει πρότυπο, αν και δεν υπάρχει ρητή περιγραφή μέχρι σήμερα. Έτσι, δεν προκαλεί έκπληξη το γεγονός ότι συναντούμε τη διαφορετική χρήση και έννοια αυτών των «T» στη βιβλιογραφία.

Το πιο σύγχρονο μοντέλο μετάφρασης στη βιβλιογραφία περιγράφει τις 3 μεταφραστικές περιόδους :

- Βασική επιστήμη που μεταφράζεται σε κλινική αποτελεσματικότητα (T1).
- Η αποτελεσματικότητα μεταφράζεται στην κλινική αποτελεσματικότητα (T2).
- Τελικά η αποτελεσματικότητα μεταφράζεται στην παροχή υγειονομικής περίθαλψης (T3) [4].

Αν και αυτό το μοντέλο βελτιώθηκε σε σχέση με τα προηγούμενα με μεγαλύτερη λεπτομέρεια, η ορολογία παραμένει ασαφής τόσο για τους ερευνητές όσο και για τους γιατρούς [6], [7].

1.4. Τομείς Εφαρμογής

Στις μέρες μας συλλέγεται ένας μεγάλος όγκος δεδομένων σε διάφορους τομείς ο οποίος μένει ανεκμετάλλευτος. Τα δεδομένα αυτά μπορούν να χρησιμοποιηθούν για να βοηθήσουν στο να βελτιώσουμε την καθημερινή μας ζωή και να την κάνουμε πιο εύκολη.

Για τον λόγο αυτό έχει ξεκινήσει μια καινούργια έρευνα η οποία ονομάζεται έρευνα μεταφραστικών δεδομένων (Translational Data Research) και σκοπός της είναι η αξιοποίηση του μεγάλου αυτού όγκου δεδομένων σε πολλούς τομείς όπως η βιοϊατρική, η γενετική ζώων, ο εντοπισμός και την πρόωρη πρόληψη διάφορων μορφών καρκίνου, η χρήση για τον εντοπισμό των επιπέδων της εγκληματικότητας και για πολλά άλλα.

1.4.1. Βιοϊατρική

Ο κυριότερος και πιο βασικός τομέας που γίνεται ανάπτυξη της μεταφραστικής έρευνας στις μέρες μας είναι αυτός της βιοϊατρικής. Η αύξηση της ιατρικής ακρίβειας και η διαθεσιμότητα πολλών δεδομένων σε καθημερινή βάση στο πλαίσιο της κλινικής περίθαλψης, έχουν αυξήσει την αναγκαιότητα για ανάπτυξη μεταφραστικών εργαλείων, έτσι ώστε να γίνει σωστή διαχείριση των δεδομένων αυτών.

Επίσης, ένας ακόμη λόγος που είναι απαραίτητη η μεταφραστική έρευνα στον τομέα της βιοϊατρικής είναι για μπορέσει να γίνει περαιτέρω διερεύνηση στις στους διαφορετικού τύπους ασθενειών που αντιμετωπίζουμε και στους τρόπους περίθαλψης των ασθενειών αυτών.

Η μεταφραστική έρευνα στον τομέα της βιοϊατρικής μας βοηθάει στο να εντοπίσουμε την ασθένεια που πρέπει να αντιμετωπιστεί πιο γρήγορα και με μεγαλύτερο ποσοστό ακρίβειας, με την χρήση κλινικών δεδομένων παλαιότερων περιπτώσεων της ίδιας ασθένειας. Μας δίνεται η δυνατότητα επίσης να ανακαλύψουμε και να δημιουργήσουμε καινούργιες μεθόδους διάγνωσης, πρόληψης και αντιμετώπισης σε συγκεκριμένες περιπτώσεις ασθενειών.

Με τον τρόπο αυτό μπορεί να γίνει γρηγορότερη φροντίδα για τον σωστό ασθενή, κάτι που βοηθάει στην εξοικονόμηση χρόνου στις κλινικές έτσι ώστε να γίνει γρηγορότερα η περίθαλψη του επόμενου ασθενή.

Μέσω της γρήγορης συλλογής, αποθήκευσης και ταξινόμησης όλων των ενεργών δεδομένων μπορεί επίσης να ενεργοποιηθεί η δυνατότητα να χτυπάει συναγερμός στις κλινικές, κάθε φορά που θα υπάρχει επιδημία ενός τύπου ασθένειας. Με τον τρόπο αυτό

μπορούν να εντοπίζονται με ταχύτερο ρυθμό οι επιδημίες και να αντιμετωπίζονται πιο γρήγορα.

1.4.2. Γενετική Ζώων

Η μεταφραστική επιστήμη έχει γίνει όλο και πιο δημοφιλής τα τελευταία χρόνια. Στον τομέα της ιατρικής έρευνας, τα ζωικά μοντέλα θεωρούνται πολύτιμα εργαλεία έρευνας που παρέχουν πληροφορίες για τον πολύπλοκο κόσμο των ανθρώπινων νόσων. Τα συγγενή ζώα εξαλείφουν αρκετές από τις προκλήσεις των ανθρώπινων μελετών και επιτρέπουν στους ερευνητές να διεξάγουν έρευνες για ασθένειες, με αναμφισβήτητο το πλέον ελεγχόμενο μοντέλο συστήματος.

Στον τομέα της γενετικής ζώων, η χρήση ζωικών μοντέλων στη μεταφραστική επιστήμη διευκολύνει την βαθύτερη κατανόηση της γονιδιακής λειτουργίας μέσω φυλογενικών διερευνήσεων μεταξύ των ειδών, που φωτίζουν τη φυσική εξέλιξη των ειδών και τη δομή και λειτουργία των ιστών και των οργάνων.

Τις τελευταίες δεκαετίες, το zebrafish (*Danio rerio*) υπήρξε ένας άξονας εργασίας για την αναπτυξιακή βιολογία και τη γενετική. Ταυτόχρονα, το zebrafish αποδείχθηκε ιδιαίτερα προσπελάσιμο και αποτελεσματικό για την μεταφραστική επιστήμη παρέχοντας ένα μοντέλο σπονδυλωτού ζώου χρήσιμο για την ανακάλυψη γονιδίων, τη μοντελοποίηση ασθενειών, τη χημική γενετική εξέταση και άλλες ιατρικές μελέτες.

Οι βασικοί πόροι, όπως μια σχολιασμένη και ολοκληρωμένη ακολουθία γονιδιώματος, και διάφορα εργαλεία για γενετικό χειρισμό, συνεχίζουν να ωθούν την ευρεία χρήση του zebrafish. Ως εκ τούτου, τα μοντέλα όπως το zebrafish έχουν σημαντικό ρόλο στην ενημέρωση, καθοδήγηση και βελτίωση των προβλέψεων για την κλινική παράλληλα με τα παραδοσιακά μοντέλα.

Το ειδικό τεύχος της Translational Research [8] υπογραμμίζει τη χρήση της zebrafish ως μεταφραστικού μοντέλου με έμφαση στην ανθρώπινη νόσο. Με πολλές άλλες πιθανές προσεγγίσεις και έναν ταχέως αυξανόμενο αριθμό ιατρικά σχετικών αναφορών, οι προσεγγίσεις του zebrafish ενισχύουν σημαντικά τα διαθέσιμα εργαλεία για τη μεταφραστική επιστήμη και έρευνα και βελτιώνουν ενεργά την κατανόηση της ανθρώπινης νόσου.

Ωστόσο, η ευκολία με την οποία τα ζωικά μοντέλα μας επιτρέπουν να προχωρήσουμε στην έρευνα συχνά μας οδηγεί στο να αγνοήσουμε το ενδεχόμενο ότι οι μελέτες αυτές να παρέχουν μόνο περιορισμένες μεταφραστικές δυνατότητες στους ανθρώπους [9].

1.4.3. Εντοπισμός Καρκίνου

Έχει επίσης ξεκινήσει έρευνα για τις διάφορες μορφές καρκίνου, όπως ο καρκίνος του μαστού και ο καρκίνος του προστάτη, έτσι ώστε να μπορεί να γίνει ευκολότερος εντοπισμός τους και να είναι πιο βέλτιστη η πρόληψη και η θεραπεία τους.

Στη συνέχεια περιγράφονται εκτενώς οι μελέτες της μεταφραστικής έρευνας για την αντιμετώπιση του καρκίνου.

1.4.4. Εγκληματικότητα

Ιστορικά, η πολιτική ποινικής δικαιοσύνης έχει επηρεαστεί σε μεγάλο βαθμό από την ιδεολογία, την κοινή γνώμη και την κάλυψη από τα μέσα μαζικής ενημέρωσης των απομονωμένων και προκλητικών περιστατικών.

Επιπλέον, η κατανόησή μας για τον τρόπο με τον οποίο χρησιμοποιούνται τα στοιχεία της έρευνας για τη διαμόρφωση των εκτιμήσεων των στρατηγικών για τα κοινωνικά προβλήματα και τις πιθανές λύσεις είναι περιορισμένη.

Η ιδέα της μεταφραστικής εγκληματολογίας (Translational Criminology) είναι απλή, αλλά ισχυρή: Αν θέλουμε να προλάβουμε, να περιορίσουμε και να διαχειριστούμε το έγκλημα, πρέπει να είμαστε σε θέση να μεταφράσουμε τις επιστημονικές ανακαλύψεις σε στρατηγικές και πράξεις.

Η μεταφραστική έρευνα βοηθάει στον προσδιορισμό του ποσοστού των εγκλημάτων που δεν αναφέρεται στην αστυνομία. Επίσης, βοηθά στην εκτίμηση των πραγματικών επιπέδων της εγκληματικότητας για ένα συγκεκριμένο τόπο εγκλήματος, βάση των προηγούμενων δεδομένων.

Ο στόχος της μεταφραστικής εγκληματολογίας είναι να καταργηθούν τα εμπόδια μεταξύ βασικής και εφαρμοσμένης έρευνας, δημιουργώντας μια δυναμική διεπαφή μεταξύ έρευνας και πρακτικής. Πρόκειται για δρόμο δύο κατευθύνσεων:

- Από τη μία, οι επαγγελματίες του χώρου περιγράφουν καθημερινά τις προκλήσεις που αντιμετωπίζουν στις δουλειές τους και
- Από την άλλη, οι επιστήμονες ανακαλύπτουν νέα εργαλεία και ιδέες για να ξεπεράσουν αυτές τις προκλήσεις και να αξιολογήσουν τον αντίκτυπό τους.

Συνοπτικά, στην έρευνα αυτή το ενδιαφέρον στρέφεται στην ανάλυση και οργάνωση μεγάλων συνόλων δεδομένων τα οποία έχουν συλλεχθεί από διάφορες ετερογενείς πηγές,

έτσι ώστε να γίνεται καλύτερη στατιστική ανάλυση και αναζήτηση, μέσω των μηχανισμών συλλογής και αποθήκευσης των δεδομένων.

Ωστόσο, η μεταφραστική εγκληματολογία υπερβαίνει τη συμβατική ιδέα "από την έρευνα στην πράξη". Αυτό επιτυγχάνεται μέσω συστηματικής μελέτης της διαδικασίας διάδοσης της γνώσης, αναγνωρίζοντας ότι η επιτυχής διάδοση των πορισμάτων της έρευνας μπορεί να απαιτεί πολλαπλές στρατηγικές.

Η επιτυχής διάδοση απαιτεί επίσης την ορθή εφαρμογή των αποδεικτικών στοιχείων. Με άλλα λόγια, δεν είναι μόνο η εύρεση αποδείξεων ότι κάτι λειτουργεί. Είναι το γιατί λειτουργεί και πώς μπορούν να εφαρμοστούν τα στοιχεία σε πραγματικά περιβάλλοντα.

Επιπλέον, αυτή η πτυχή της μεταφραστικής εγκληματολογίας θέτει μια προτεραιότητα στην εφαρμοσιμότητα - δηλαδή στην έρευνα με τις δυνατότητες πραγματικής εφαρμογής του, κάτι που είναι ιδιαίτερα ελκυστικό σε μια εποχή περιορισμένων πόρων [10], [11].

1.4.5. Εντοπισμός Γλωσσικών Δυσλειτουργιών

Υπάρχει επιτακτική ανάγκη για την έγκαιρη και ακριβή ταυτοποίηση γλωσσικών και άλλων αναπτυξιακών διαταραχών σε μικρά παιδιά και την παροχή έγκαιρης παρέμβασης, ανάλογα με την ηλικία, όπως επιτάσσει το μέρος Γ της πράξης της εκπαίδευσης των ατόμων με αναπηρίες (Part C of the Individuals with Disabilities Education Act).

Η Έρευνα έχει αποδείξει ότι η έγκαιρη παρέμβαση είναι αποτελεσματική για πολλά παιδιά με γλωσσικά προβλήματα, που προέρχονται από διαφορετικές αιτιολογικές ομάδες, και μπορεί να μειώσει τη επίδραση των διαταραχών στα παιδιά και τις οικογένειές τους.

Παρόλο αυτά, η ακριβής ταυτοποίηση σε βρέφη και νήπια που ανήκουν σε ομάδες κινδύνου για εκδήλωση γλωσσικών ανωμαλιών παραμένει ελλιπής, ειδικά για τα παιδιά με χρονική καθυστέρηση στην ομιλία χωρίς προφανή γενετικές ή νευρολογικές συνθήκες.

Επομένως, η ανάγκη για μεταγραφική έρευνα με ακριβείς και λεπτομερείς πληροφορίες σχετικά με πολλές πτυχές της έγκαιρη τυπικής και άτυπης λειτουργίας της γλώσσας και τους ποικίλους γνωστικούς, συμπεριφορικούς, και συγκυριακούς παράγοντες που τους επηρεάζουν, γίνεται επιτακτική.

Σε απάντηση, το NIH [12] διεξήγαγε συνεργατική μεταγραφική έρευνα («bench to bedside») πάνω στις βασικές αναπτυξιακές διεργασίες τόσο σε τυπικούς και άτυπους

πληθυσμούς με σκοπό να ενημερώσει καλύτερα και να καθοδηγήσει την κλινική πρακτική σε παιδιά με αναπτυξιακές διαταραχές, συμπεριλαμβανομένης της γλωσσικής δυσκολίας, και να ενθαρρύνει τους επιστήμονες μελέτης της συμπεριφοράς (behavioural scientists) να αναζητούν μια περαιτέρω κατανόηση του πώς οι συμπεριφορικές διεργασίες (π.χ. γλώσσα) τροποποιούνται από αναπτυξιακές διαταραχές.

Οι περισσότερες έρευνες για πρώιμη κατάκτηση της γλώσσας σε παιδιά με γλωσσικές διαταραχές έχει επικεντρωθεί σε ένα στενό φάσμα γλωσσικών συμπεριφορών (π.χ., λεξιλόγιο ή και μορφοσυντακτική ικανότητα) σε μία συγκεκριμένη κλινική ομάδα (π.χ. σύνδρομο Down, αυτισμό ή συγκεκριμένη γλωσσική δυσλειτουργία), περιγράφοντας την έλλειψη σε σχέση με διάφορες ομάδες ελέγχου ή αξιολογώντας τα άτομα της ίδιας ομάδας με συγκεκριμένες γλωσσικές και γνωστικές δεξιότητες [13], [14].

Παρά αυτή την περιορισμένη εστίαση, η έρευνα αυτή έχει παράσχει μια πληθώρα πολύτιμων πληροφοριών σχετικά με τη φύση και τη διαδικασία των συγκεκριμένων πτυχών γλωσσικής ανάπτυξης σε συγκεκριμένους πληθυσμούς.

Τα αποτελέσματα των αναλύσεων προφίλ σε διαφορετικούς πληθυσμούς έχουν καταστεί ιδιαίτερα χρήσιμα στην απομυθοποίηση στερεότυπων σχετικά με τις διαφορετικές αιτιολογικές ομάδες και τόνισαν μεθοδολογικά τους περιοριστικούς όρους για την μελλοντική έρευνα, συμπεριλαμβανομένης της ανάγκης για την διαχρονική μελέτη και τη συνεχιζόμενη ανάγκη για αξιολόγηση των πολλαπλών διαστάσεων της γλώσσας (π.χ., λεξιλόγιο/σημασιολογία, γραμματική/σύνταξη και πραγματολογία) σε διαφορετικές ομάδες.

Παρακάτω ακολουθεί μια σύντομη περίληψη ορισμένων βασικών πορισμάτων της έρευνας για τέσσερις κλινικές ομάδες παιδιών με καθυστερημένη ή διαταραγμένη γλωσσική ανάπτυξη : σύνδρομο Down, διαταραχές του φάσματος του αυτισμού, σύνδρομο Ουίλιαμς, και συγκεκριμένες γλωσσικές δυσλειτουργίες.

Σύνδρομο Down

Με ποσοστό περίπου 1 στα 700 έως 1.000, το σύνδρομο Down, είναι η πιο κοινή γενετική αιτία νοητικής υστέρησης. Παιδιά με σύνδρομο Down παρουσιάζουν ιδιαίτερο ενδιαφέρον για τους ερευνητές γλωσσικής ανάπτυξης επειδή το σύνδρομο Down διακρίνεται αιτιολογικά και μπορεί να εντοπιστεί και να μελετηθεί πολύ νωρίς στη ζωή.

Όπως ισχύει και για τα παιδιά με άλλες γενετικές διαταραχές, τα παιδιά με σύνδρομο Down εκδηλώνουν ένα ξεχωριστό προφίλ γλωσσικής, γνωστικής, συναισθηματικής και κοινωνικής ικανότητας στην πρώιμη παιδική ηλικία.

Παρόλο που καθυστερούν εξελικτικά, τα παιδιά με σύνδρομο Down διαφέρουν ως προς το επίπεδο της γενικής γνωστικής λειτουργίας από πολύ καθυστερημένη σε σχεδόν φυσιολογική [15]. Η καθυστερημένη, αλλά παρόλο αυτά μεταβλητή ανάπτυξη τους, επιτρέπει μια πιο ακριβή εξέταση των ακολουθιών των διαφόρων πτυχών της γλωσσικής και λεκτικής ανάπτυξης τους από ό, τι σε τυπικά παιδιά με πιο ταχεία ανάπτυξη [16].

Αυτισμός Και Συναφείς Αναπτυξιακές Διαταραχές

Παρά τη σχετικά χαμηλή συχνότητα εμφάνισης του αυτισμού (περίπου 4-10 περιπτώσεις σε κάθε 10.000 γεννήσεις), η πρώιμη γλωσσική και γνωστική ανάπτυξη των παιδιών με αυτισμό και συναφείς διαταραχές έχει μελετηθεί εντατικά κατά τη διάρκεια των τελευταίων 2 δεκαετιών [17].

Τα παιδιά με αυτισμό εμφανίζουν μια τριάδα από σοβαρές αναπηρίες, συμπεριλαμβανομένων της γλωσσικής και επικοινωνιακής έλλειψης, κοινωνικές γνωστικές (θεωρία του νου) δυσκολίες, και διαφορετικότητα κοινωνικής συμπεριφοράς, καθώς και μεγάλη διαφορά στην ατομικές τους λειτουργίες [18].

Σύνδρομο Williams

Το σύνδρομο Williams είναι μια σπάνια γενετική διαταραχή (συχνότητα 1/10.000 έως 1/20.000 γεννήσεις) που προκαλείται από μια μικροδιαγραφή περίπου 20 γονιδίων στο χρωμόσωμα 7q11.23 και που συνδέεται με ένα εξαιρετικά ασυνήθιστο νευροψυχολογικό προφίλ [19].

Παρά την μέτρια νοητική υστέρηση (μέσος όρος IQ κυμαίνεται από 50 έως 70), τα άτομα με αυτή τη διαταραχή έχουν εκπληκτικά σύνθετες γλωσσικές δεξιότητες και τείνουν να εμφανίζουν αρκετά επαρκής εκφραστική και κοινωνική συμπεριφορά [20]. Την ίδια στιγμή, τα άτομα αυτά παρουσιάζουν σοβαρή έλλειψη της λεπτής κινητικότητας και οπτικό-χωρικών γνωστικών λειτουργιών, που ξεπερνά το επίπεδο της γενικής γνωστικής δυσλειτουργίας τους.

Συγκεκριμένες Γλωσσικές Δυσλειτουργίες

Εξ ορισμού, τα παιδιά με συγκεκριμένες γλωσσικές δυσλειτουργίες δεν έχουν καμία απώλεια ακοής ή γνωστή γενετική ή νευρολογική διαταραχή. Αποδίδουν κάτω από τα αναμενόμενα γλωσσικά επίπεδα για την ηλικία τους αλλά εντός κανονικών ορίων σχετικά με τα μέτρα της λεκτικής νοημοσύνης. Έτσι, σε αντίθεση με άλλους πληθυσμούς που περιγράφεται παραπάνω, η γλωσσική δυσλειτουργία είναι το καθοριστικό χαρακτηριστικό αυτής της κατάστασης [21].

Η επικράτηση των παιδιών προσχολικής ηλικίας με συγκεκριμένη γλωσσική δυσλειτουργία κυμαίνεται από 2 έως 8% (διάμεσος 5,95%). Υψηλότερο ποσοστό εμφανίζεται :

- στα αγόρια
- στα παιδιά με οικογενειακό ιστορικό γλωσσικής, λεκτικής ή αναγνωστικής διαταραχής
- στα παιδιά που γεννιούνται πρόωρα και
- στα παιδιά από οικογένειες χαμηλής κοινωνικοοικονομικής κατάστασης

Περίπου το 25% με 50% των παιδιών προσχολικής ηλικίας με γλωσσική καθυστέρηση και κανονική λεκτική νοημοσύνη, θα συνεχίσουν να εκδηλώνουν προβλήματα γλωσσικής ανάπτυξης και στην μετέπειτα παιδική τους ηλικία [22].

Δια- Πληθυσμιακές Μελέτες

Δεδομένου του αυξημένου ενδιαφέροντος στον τομέα της γενετικής και της βιοσυμπεριφοράς (biobehavioral) σε συγκεκριμένες γλωσσικές διαταραχές κατά την τελευταία δεκαετία, ο Rice [14] και οι συνεργάτες του έχουν επισημάνει την ανάγκη για περαιτέρω συνεργατική δια-πληθυσμιακή έρευνα, με σκοπό την τεκμηρίωση των τρόπων που οι γλωσσικές διαταραχές εκδηλώνονται σε κλινικές συνθήκες.

Η εν λόγω έρευνα μπορεί να οδηγήσει σε μια βαθύτερη κατανόηση των ομοιοτήτων και των διαφορών των παθήσεων, των προφίλ συγκεκριμένων συμπτωμάτων που σχετίζονται με κάθε κατάσταση, και γενικά πώς μια διανοητική βλάβη με συγκεκριμένη γενετική βάση μπορεί να επηρεάσει τη διαδικασία της γλωσσικής ανάπτυξης σε διάφορους τομείς.

Οι δια-πληθυσμιακές συγκρίσεις επίσης θα διευκολύνουν την αναγνώριση των υποκατηγοριών των διαταραχών σε διάφορες κλινικές συνθήκες και θα ενισχύσουν την

ικανότητά έγκαιρης διάγνωσης σε αρχικά στάδια των παιδιών που είναι πιθανό να αναπτύξουν σημαντικά προβλήματα γλωσσικών διαταραχών [23].

Η τεκμηρίωση των υποκατηγοριών και των συγκεκριμένων συναφών παραγόντων που σχετίζονται με επιμέρους παθήσεις, θα αυξήσει την κατανόηση των διαφορετικών επιπέδων κινδύνου και προδιάθεσης, που μπορεί να εκδηλώνουν τα πολύ μικρά παιδιά διαφορετικών ομάδων, για επίμονες γλωσσικές δυσλειτουργίες. Γενικότερα, η έρευνα αυτή θα συμβάλει στην κατανόηση της φύσης και της μεταβλητότητας της ανθρώπινης γλωσσικής ανάπτυξης και θα ενισχύσει την κατανόηση των βασικών αρχών του εγκεφάλου, που σχετίζονται με τη γλώσσα και την γλωσσική ανάπτυξη [13].

Διαχρονικές Μελέτες

Απαιτούνται επιπλέον διαχρονικές μελέτες των πολλαπλών πτυχών της γλωσσικής ανάπτυξης, σε διαφορετικούς κλινικούς πληθυσμούς παιδιών με γλωσσικές διαταραχές, ιδιαίτερα κατά την περίοδο της Οντογένεσης, όταν αποκτιέται η γλώσσα [24]. Αν και δαπανηρή και χρονοβόρα διαδικασία, οι μελέτες αυτές θα μας επιτρέψουν την αξιολόγηση του τρόπου με τον οποίο προκύπτουν συγκεκριμένες γλωσσικές συμπεριφορές σε διαφορετικούς τομείς και μεταβάλλονται στον χρόνο.

Οι διαχρονικές έρευνες θα επιτρέψουν επίσης την αξιολόγηση των ατομικών διαφορές στα προφίλ της γλωσσικών, γνωστικών, και συμπεριφορικών δεξιοτήτων, στα άτομα την ίδιας ομάδας, με την πάροδο του χρόνου και την σαφέστερη και ακριβέστερη περιγραφή των γλωσσικών φαινοτύπων των υποομάδων των παιδιών [25].

Τέτοιες έρευνες είναι απαραίτητες για την υποστήριξη και ενημέρωση των μελετών που αξιολογούν τις γενετικές και νευροβιολογικές αιτιολογίες των διαφορετικών αναπτυξιακών γλωσσικών διαταραχών.

1.5. Στόχος της Εργασίας

Ο στόχος της εργασίας αυτής είναι η μελέτη της χρησιμότητας και οι δυνατότητες ανάπτυξης της μεταφραστικής έρευνας στους διάφορους τομείς που μπορεί να χρησιμοποιηθεί, έτσι ώστε να γίνει αξιοποιήσιμη χρήση και αποθήκευση του μεγάλου όγκου δεδομένων που μένουν ανεκμετάλλευτα καθημερινά. Επίσης, θα γίνει βιβλιογραφική αναφορά και μελέτη σε διάφορους μηχανισμούς που έχουν αναπτυχθεί ήδη και θα σχολιαστεί η απόδοσή τους.

1.6. Μεθοδολογία

Τα βήματα της μεθοδολογίας ήταν αρχικά η μελέτη της βιβλιογραφίας για άρθρα σχετικά με τα μεταφραστική επιστήμη που αντλήθηκαν από το PubMed, ResearchGate και GoogleScholar. Στη συνέχεια εστίασαμε στα συστήματα διαχείρισης βιοϊατρικών δεδομένων γενικά και ειδικότερα για αυτά που αναπτύχθηκαν για τη μελέτη του καρκίνου. Αναζητήθηκαν μελέτες για τα σημαντικότερα από αυτά.

Στο τέλος της παρουσίασης των συστημάτων ακολουθεί ένας συγκριτικός πίνακας αυτών με σχολιασμό των υπέρ και των κατά καθενός. Τέλος, αναφέρονται και σχολιάζονται οι προκλήσεις που αντιμετωπίζει η μεταφραστική έρευνα και τα συστήματα διαχείρισης μεταφραστικών δεδομένων και προτείνονται λύσεις με βάση τη βιβλιογραφία.

1.7. Περιεχόμενα

Στο κεφάλαιο 1 γίνεται μια εισαγωγή στην μεταφραστική έρευνα. Αρχικά δίνονται ενδεικτικά κάποιοι ορισμοί από διεθνείς οργανισμούς και περιγράφεται αναλυτικά το μοντέλο 3 φάσεων τη μεταφραστικής έρευνας. Αναφέρεται η χρήση της σε διαφόρους τομείς και στη συνέχεια η εργασία εστιάζει στα βιοϊατρικά δεδομένα.

Στο κεφάλαιο 2 αναλύονται τα συστήματα διαχείρισης μεταφραστικών δεδομένων. Παρουσιάζονται συστήματα για γενικά δεδομένα και έπειτα συστήματα προσανατολισμένα στη διαχείριση βιοϊατρικών δεδομένων για το καρκίνο. Για καθένα από αυτά περιγράφονται τα χαρακτηριστικά, η αρχιτεκτονική, οι υποστηριζόμενοι τύποι δεδομένων, οι δυνατότητες και τα εργαλεία και οι εφαρμογές που υποστηρίζουν.

Στο κεφάλαιο 3, ακολουθεί μια συγκριτική περιγραφή των μεταφραστικών συστημάτων διαχείρισης δεδομένων ως προς 5 άξονες: (1) περιεχόμενο δεδομένων, (2) υποστήριξη διαλειτουργικότητας, (3) υποστήριξη ανάλυσης, (4) υποστήριξη ασφάλειας και ιδιωτικότητας δεδομένων και (5) τεχνική υποστήριξη των συστημάτων.

αναλύονται οι απαιτήσεις που πρέπει να πληρούν τα μεταφραστικά ερευνητικά συστήματα.

Στο κεφάλαιο 4 περιγράφονται οι απαιτήσεις που πρέπει να πληρούν τα μεταφραστικά ερευνητικά συστήματα για την ομαλή τους λειτουργία. Στο κεφάλαιο 5 αναλύονται οι προκλήσεις που αντιμετωπίζει η μεταφραστική έρευνα και εμποδίζουν την εξασφάλιση της εξέλιξης της επιστημονικής γνώσης προς όφελος των ασθενών.

Και τέλος στο κεφάλαιο 5 εξάγονται συμπεράσματα για την μεταφραστική έρευνα και τα συστήματα διαχείρισης μεταφραστικών δεδομένων. Επίσης, υπογραμμίζονται οι προοπτικές εξέλιξης και οι επεκτάσεις, που θα μπορούσαν να βοηθήσουν στην καλύτερη αξιοποίηση των μεταφραστικών δεδομένων και συστημάτων.

2. Συστήματα Διαχείρισης Μεταφραστικών Δεδομένων

Όσον αφορά την χρήση στην βιοϊατρική ο όγκος των κλινικών και εργαστηριακών δεδομένων αυξάνεται όλο και περισσότερο, και εξαιτίας αυτού έχουν αυξηθεί και οι απαιτήσεις διαχείρισης και οργάνωσης των δεδομένων αυτών. Η ανάπτυξη μηχανισμών που μπορούν να διαχειρίζονται, να αποθηκεύουν και να ομαδοποιούν τα δεδομένα αυτά, μπορεί να οδηγήσει στην δημιουργία νέας γνώσης και στην βελτίωση της ήδη υπάρχουσας.

Κατά την διάρκεια των χρόνων έχουν αναπτυχθεί αρκετά συστήματα υποστήριξης της μεταφραστικής έρευνας στον τομέα της βιοϊατρικής. Τα συστήματα αυτά είναι ικανά να ενσωματώσουν μεγάλα σύνολα κλινικών δεδομένων με δεδομένα omics. Επομένως, θα πρέπει να παρέχουν:

- Αποθήκευση και ενσωμάτωση τόσο κλινικών όσο και δεδομένων omics.
- Ένα πλαίσιο ανάλυσης, που επιτρέπει στους επιστήμονες να διερευνήσουν τα δεδομένα τους και να δημιουργήσουν υποθέσεις.
- Πρόσθετες πληροφορίες που διασταυρώνονται από εξωτερικές βάσεις δεδομένων (π.χ. να συνδέονται με μια συγκεκριμένη περιγραφή γονιδίου σε δημοσιευμένη βιβλιογραφία ή δημόσιες βάσεις δεδομένων) [26].

Παρακάτω αναλύονται τα σημαντικότερα συστήματα διαχείρισης μεταφραστικών δεδομένων. Τα συστήματα αυτά χωρίστηκαν σε συστήματα διαχείρισης γενικών ιατρικών δεδομένων και σε συστήματα διαχείρισης δεδομένων εστιασμένα για τον εντοπισμό και τη πρόληψη του καρκίνου.

2.1. Γενικά Συστήματα

2.1.1. tranSMART

Το tranSMART [27] είναι μια ανοιχτού κώδικα πλατφόρμα διαχείρισης της γνώσης για τη μεταφραστική ιατρική με γνώμονα την κοινότητα. Πρόκειται για ένα έργο το οποίο αναπτύσσεται σε συνεργασία από περισσότερους από 100 επιστήμονες υπολογιστών και ιατρούς από περισσότερους από 20 οργανισμούς από όλο τον κόσμο.

Η πλατφόρμα αναπτύχθηκε αρχικά από επιστήμονες της Johnson & Johnson and Recombinant Data Corporation το 2009 για να υποστηρίξει την εσωτερική έρευνα στο τμήμα Centocor R & D της Johnson & Johnson και αργότερα επεκτάθηκε για να

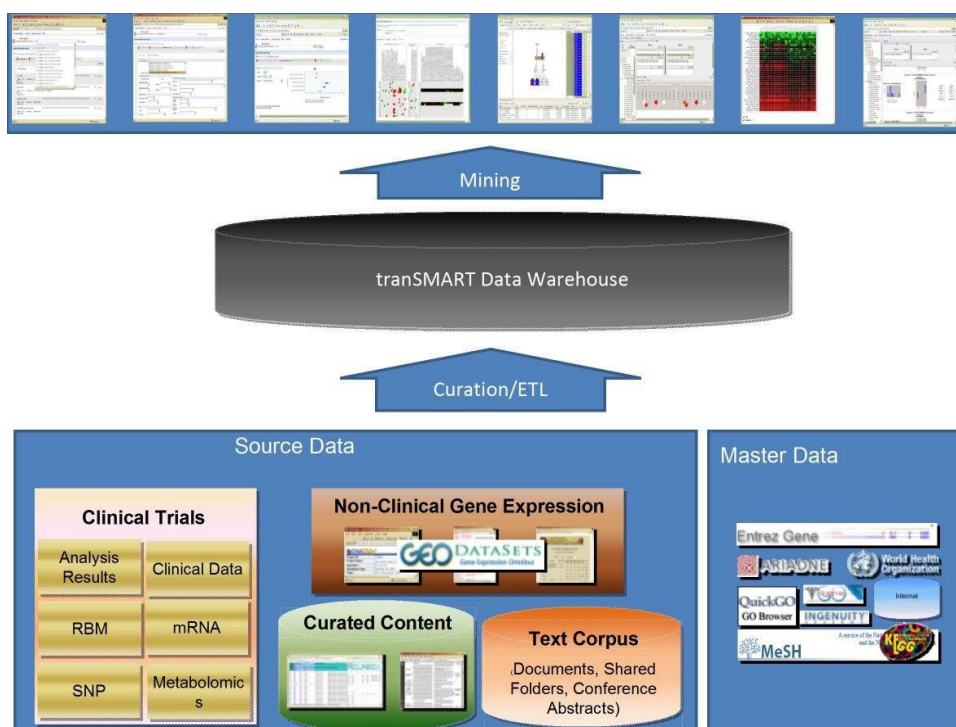
υποστηρίξει εξωτερικές συνεργασίες. Τελικά προσφέρεται ως πλατφόρμα λογισμικού ανοικτού κώδικα από το 2012.

Η πλατφόρμα χρησιμοποιείται από περισσότερους από 150 ακαδημαϊκούς, βιομηχανικούς, κυβερνητικούς και ιδιωτικούς ερευνητικούς οργανισμούς, καθώς και από βασικούς προμηθευτές στο οικοσύστημα tranSMART.

Το tranSMART διαχειρίζεται από το Ίδρυμα tranSMART, έναν μη κερδοσκοπικό οργανισμό, που επιτρέπει την τεχνολογική και επιστημονική συνεργασία για την προώθηση της μεταφραστικής βιοϊατρικής έρευνας.

Το Ίδρυμα tranSMART δημιουργήθηκε το 2013 για να δημιουργήσει και να διατηρήσει μια κοινότητα ανοικτού κώδικα γύρω από την πλατφόρμα και κυκλοφόρησε δύο εκδόσεις:

- Η έκδοση 1.1 (τέλη 2013) παρείχε ολοκληρωμένη υποστήριξη της πλατφόρμας βάσης δεδομένων ανοικτού κώδικα PostgreSQL της και
- Η έκδοση 1.2 (στα μέσα του 2014) παρείχε πρόσθετη υποστήριξη για επιπλέον τύπους δεδομένων υψηλής ευκρίνειας και νέα χαρακτηριστικά που προέρχονται από 12 διαφορετικούς κλάδους του βασικού κώδικα.



Εικόνα 3: Διάγραμμα του συστήματος tranSMART. Τα δημόσια και ιδιωτικά δεδομένα ευθυγραμμίζονται με τις τυπικές οντολογίες και επιτηρούνται και υποβάλλονται σε επεξεργασία ETL για αποθήκευση σε κεντρική αποθήκη δεδομένων [14].

Το tranSMART αποτελεί μια ενιαία πλατφόρμα ανάλυσης και ανταλλαγής δεδομένων για κλινική και μεταφραστική έρευνα που είναι διαθέσιμη σε όλους τους επιστήμονες ανά τον κόσμο. Αυτό επιτρέπει την ανταλλαγή προ-ανταγωνιστικών και ιδιωτικών δεδομένων και προωθεί τη συνεργασία ανάμεσα στα μέλη ενός οργανισμού ή με συνεργάτες σε όλο τον κόσμο.

Δυνατότητες Πλατφόρμας

Η πλατφόρμα tranSMART παρέχει εργαλεία για τη φόρτωση ποικίλων τύπων δεδομένων υψηλού περιεχομένου, χημικών, βιολογικών, omics και κλινικών δεδομένων, από διάφορες πηγές.

Το tranSMART μπορεί να διαχειριστεί

- Δομημένα δεδομένα από κλινικές δοκιμές (δημογραφικά στοιχεία, εργαστηριακά αποτελέσματα και κλινικούς φαινοτύπους) και
- Ευθυγραμμισμένα δεδομένα βιοδεικτών υψηλής περιεκτικότητας, όπως προφίλ γονιδιακής έκφρασης, γονότυπους, μεταβολικοποιητικά και πρωτεϊνωματικά δεδομένα.

Η πλατφόρμα επιτρέπει στους επιστήμονες να ανακαλύπτουν και να εξειδικεύουν υποθέσεις διερευνώντας τις σχέσεις μεταξύ γενετικών και φαινοτυπικών δεδομένων για ομάδες ασθενών και να αξιολογούν τα αναλυτικά αποτελέσματα τους στο πλαίσιο της δημοσιευμένης βιβλιογραφίας καθώς και της δικής τους εργασίας.

Σχεδιάστηκε για να ανταποκρίνεται στις ανάγκες των βασικών επιστημόνων, κλινικών επιστημόνων, φαρμακολόγων, τοξικολόγων και άλλων που εμπλέκονται στην ανακάλυψη θεραπευτικών και διαγνωστικών μεθόδων και στην ανάπτυξη νέων θεραπειών για τους ασθενείς.

Η πλατφόρμα tranSMART υλοποιεί μια σειρά από μοντέλα δεδομένων, κοινά σύνολα δεδομένων, βοηθητικά προγράμματα μετασχηματισμού δεδομένων και αναλυτικές εφαρμογές ιστού που εστιάζουν στον ασθενή, επιταχύνοντας τις ανακαλύψεις δημιουργώντας μια τυποποιημένη και σημασιολογικά ενσωματωμένη βάση δεδομένων των ερευνητικών αποτελεσμάτων που συνδέονται με επαναχρησιμοποιούμενες και κλιμακούμενες αναλύσεις αυτοεξυπηρέτησης .

Ένα API έχει προστεθεί στην πλατφόρμα tranSMART που επιτρέπει τη σύνδεση με εξωτερικά πακέτα για:

- Εκτεταμένη προβολή δεδομένων (Genome Browser κ.λπ.).
- Στατιστική ανάλυση (R)
- Οπτικοποίηση και ανάλυση (Spotfire).

Αυτά μπορεί να είναι μέρος της διανεμημένης πλατφόρμας ανοιχτού κώδικα ή να διανέμονται ιδιωτικά από προμηθευτές που χρησιμοποιούν το API.

Εφαρμογές της πλατφόρμας

Η πλατφόρμα tranSMART οργανώνει κλινικά και ερευνητικά δεδομένα ανά ασθενή. Αυτό επιτρέπει στους επιστήμονες να εκτελούν τους ακόλουθους τύπους ερευνητικών δραστηριοτήτων, με στόχο τη βελτίωση της κατανόησής τους σε υγιείς και νοσηρές διεργασίες στην ανακάλυψη νέων θεραπευτικών και διαγνωστικών εργαλείων και προϊόντων για τους ασθενείς.

- Σύγκριση δεδομένων από την πρωτεϊνωματική, τη μεταβολική και άλλες μελέτες "omics".
- Αντισταθμιστικά πρότυπα γονιδιακής έκφρασης σε υγιή και άρρωστα άτομα και σε δείγματα ανθρώπινου ιστού.
- Διερεύνηση συσχετισμών μεταξύ γονότυπου και φαινοτύπου σε δεδομένα κλινικών δοκιμών.
- Προ-κλινικά δεδομένα για τη γνώση της βιολογίας της νόσου του ανθρώπου.
- Μελέτη γενετικών και περιβαλλοντικών παραγόντων που εμπλέκονται στην ανθρώπινη ασθένεια.
- Οπτικοποίηση δεδομένων χρησιμοποιώντας μια γραφική διεπαφή.
- Συνεργασία μεταξύ ακαδημαϊκών, κυβερνητικών και εταιρικών ερευνητικών τομέων.

Η πλατφόρμα tranSMART βασίζεται στο σύστημα ανοιχτού κώδικα i2b2 CDW [28], που περιγράφεται στη συνέχεια.

2.1.2. i2b2 (Informatics for Integrating Biology and the Bedside)

Το i2b2 διαχειρίζεται και οργανώνει τα κλινικά, εργαστηριακά και διαγνωστικά δεδομένα από διάφορες περιπτώσεις ασθενών. Τα δεδομένα αυτά αποθηκεύονται ανώνυμα σε ένα κεντρικό σύστημα, έτσι ώστε να μπορούν να χρησιμοποιηθούν οποιαδήποτε στιγμή αυτό είναι απαραίτητο.

Δίνει την δυνατότητα στους χρήστες να αναζητούν για πληροφορίες στην κεντρική βάση μέσω ερωτήσεων που θέτουν, έτσι ώστε να συλλέγουν στοιχεία για χρήση σε μελέτες. Τα αποτελέσματα στις αναζητήσεις που κάνουν οι χρήστες είναι άκρως ανώνυμα.

Με την χρήση του i2b2 μπορεί να γίνει γρήγορη και εύκολη αναζήτηση στη βάση δεδομένων, όπως επίσης και διαχείριση και αλλαγή των προνομίων στα άτομα που έχουν πρόσβαση στη βάση δεδομένων. Είναι ένα ολοκληρωμένο σύστημα που παρέχει την δυνατότητα για ενοποίηση κλινικών και ερευνητικών δεδομένων, τα οποία είναι αποθηκευμένα σε ηλεκτρονικούς φακέλους.

Αρχιτεκτονική του i2b2

Hive

Στο σύστημα αυτό υπάρχει μια εφαρμογή με την ονομασία hive, μέσα στην οποία συμπεριλαμβάνεται και οργανώνεται ένα σύνολο υπηρεσιών, που ονομάζονται cells. Τα cells βοηθούν στην διαχείριση και την αποθήκευση των δεδομένων, καθώς προσφέρουν και μεθόδους για την ανάλυση και την επεξεργασία των δεδομένων.

Project Management Cell

Η υπηρεσία χρησιμοποιείται για τη διαχείριση των εργασιών στο i2b2. Τα δεδομένα που χειρίζεται βρίσκονται αποθηκευμένα στο σχήμα i2b2pm της βάσης δεδομένων. Συγκεκριμένα αποθηκεύονται οι πληροφορίες για τις εργασίες που έχουν οριστεί στο κεντρικό σύστημα και τους χρήστες / ρόλους του συστήματος περιγράφοντας τα προνόμια πρόσβασής τους στο σύστημα και σε κάθε επιμέρους εργασία. Τέλος, στην υπηρεσία αυτή περιγράφονται οι υπόλοιπες υπηρεσίες του i2b2.

Identity Management Cell

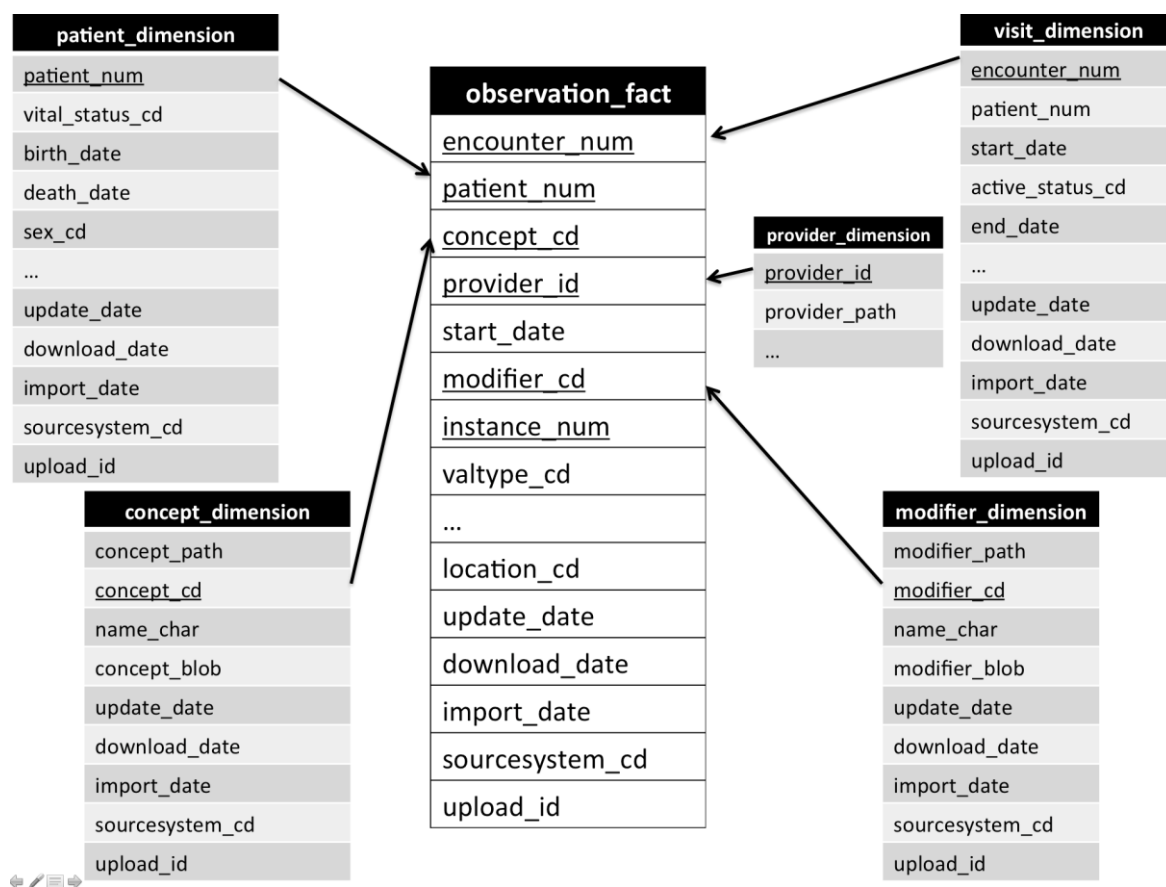
Η υπηρεσία χρησιμοποιείται για την διαχείριση της πληροφορίας για τους χρήστες του συστήματος, τα ιατρικά – ερευνητικά κέντρα που συμμετέχουν σε κάθε εργασία (im_project_sites), καθώς και τους ασθενείς που συμμετέχουν σε μια εργασία (im_project_patients). Τα δεδομένα που χειρίζεται βρίσκονται αποθηκευμένα στο σχήμα i2b2imdata της βάσης δεδομένων

Με αυτόν τον τρόπο και με την ανωνυμοποίηση των δεδομένων των ασθενών, ελέγχεται η πρόσβαση των χρηστών σε συγκεκριμένο υποσύνολο ασθενών οι οποίοι θα συμμετέχουν σε μια συγκεκριμένη εργασία. Ελέγχεται έτσι τόσο η διαχείριση όσο και η ασφάλεια των δεδομένων των ασθενών.

Επειδή το i2b2 χρησιμοποιείται για να περιγράφει βιοϊατρικά δεδομένα εντός ενός σημασιολογικού πλαισίου, όλα τα δεδομένα που χρειάζονται για την περιγραφή μίας πληροφορίας ταξινομούνται σε μια ιεραρχική δομή, έτσι ώστε να γίνεται ευκολότερη αναζήτηση μέσα στο σύστημα.

Η βάση δεδομένων που χρησιμοποιείται για να διαχειρίζονται και να αποθηκεύονται τα δεδομένα είναι οργανωμένη σε πίνακες (Εικόνα 4). Ξεκινάει από έναν κεντρικό πίνακα με το όνομα `observation_fact`, ο οποίος χωρίζεται σε μικρότερους πίνακες.

Κάποιοι από αυτούς τους πίνακες είναι αυτοί οι οποίοι αναφέρονται στους ασθενείς (`patient_dimension`), στους όρους που χρησιμοποιούνται για την δημιουργία των παρατηρήσεων (`concept_dimension`), στις επισκέψεις των ασθενών (`visit_dimension`) και πίνακες που αφορούν τους διαχειριστές των δεδομένων (`provider_dimension`).



Εικόνα 4: Περιγραφή σχήματος κεντρικής βάσης (i2b2 star schema) [29]

2.1.3. iDASH (Integrating Data for Analysis, Anonymization and SHaring)

Το iDASH [30] είναι το νεότερο Εθνικό Κέντρο Βιοϊατρικής Πληροφορικής (National for Biomedical Computing) και χρηματοδοτείται από το NIH (National Institutes of Health). Η ανάπτυξη του ξεκίνησε το 2011 με σκοπό να επιτρέπει την ανταλλαγή δεδομένων και

στοιχείων με τέτοιο τρόπο ώστε να διαφυλάσσεται η ιδιωτικότητα και ζωή του κάθε ατόμου.

Η τεχνολογική αυτή έρευνα, που βασίζεται στην διαφύλαξη του απορρήτου, χρησιμοποιεί έναν μηχανισμό ανάπτυξης συνεργατικών εργαλείων σε ένα ιδιωτικό Cloud για νόμους περί φορητότητας και λογοδοσίας για την ασφάλιση υγείας (Health Insurance Portability and Accountability Act - HIPAA).

Οι προσπάθειες για εκμάθηση και διάδοση βοηθούν στην κατάρτιση των κατόχων και των ατόμων που χειρίζονται τα δεδομένα, ως προς τον τρόπο που μπορούν να μοιραστούν και να χρησιμοποιήσουν τα κλινικά και βιολογικά αυτά δεδομένα.

Με την χρήση όλων αυτών των μηχανισμών το iDASH μπορεί και παρέχει στους ερευνητές πρόσβαση σε ένα υψηλής απόδοσης περιβάλλον πληροφορικής, δεδομένα και λογισμικό, έτσι ώστε να μπορούν να αναπτύσσουν και να δοκιμάζουν νέες υποθέσεις.

Τομείς Εφαρμογής

Η πλατφόρμα του iDASH χρησιμοποιείται από ερευνητές σε διάφορους τομείς.

➤ **Στον Μοριακό Φαινοτυπικό Προσδιορισμό Της Νόσου Του Kawasaki (Molecular Phenotyping of Kawasaki Disease (KD))**

Βοηθάει στην έρευνα για τους μοριακούς μηχανισμούς της αγγειίτιδας και στον σχηματισμό του ανευρύσματος, έτσι ώστε να βρεθούν νέοι θεραπευτικοί στόχοι και νέοι τρόποι θεραπείας και αντιμετώπισης. Λόγω του ότι η ασθένεια αυτή είναι πολύ σπάνια, είναι πολύ χρήσιμο να συγκεντρώνονται δεδομένα από όσο το δυνατόν περισσότερες πηγές.

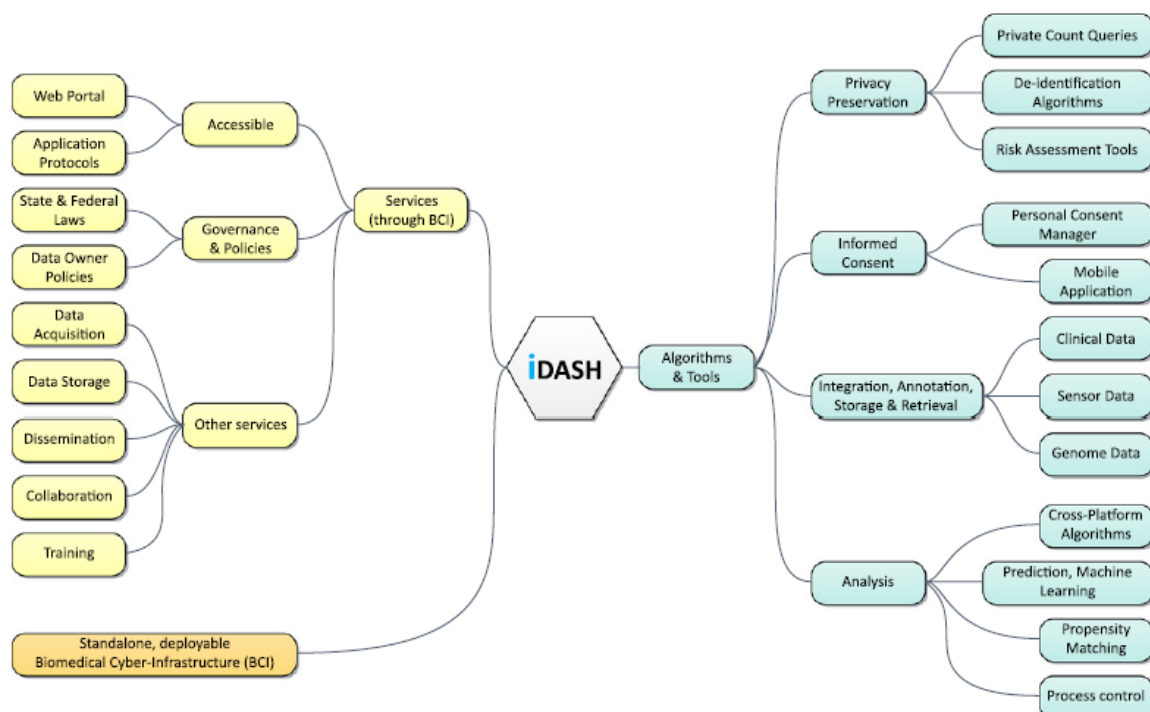
➤ **Στην Εξατομικευμένη Παρέμβαση Για Τη Βελτίωση Της Σωματικής Δραστηριότητας (Individualized Invention to Enhance Physical Activity)**

Η ομάδα ερευνητών μελετά τα προσωπικά πρότυπα κίνησης κατά την διάρκεια της ημέρας και το πώς αυτά συνδέονται με την νοσηρότητα, έτσι ώστε να αναπτύξουν μια συσκευή η οποία μπορεί να αναγνωρίσει πρότυπα συμπεριφοράς σε πραγματικό χρόνο χρησιμοποιώντας αλγορίθμους μηχανικής μάθησης.

➤ **Πολυεθνική Παρακολούθηση των Φαρμάκων (Multi - institutional Surveillance)**

Ερευνάται το πόσο ασφαλή είναι τα φάρμακα σε ένα καταναμεμημένο περιβάλλον. Όταν εγκρίνονται και κυκλοφορούν νέα φάρμακα η παρακολούθηση της ασφάλειας τους είναι ένα δύσκολο έργο, μιας και πρέπει να υπάρχει εποπτεία σε

πληθυσμούς ασθενών που δεν είχαν μελετηθεί προηγουμένως. Οι ερευνητές στον τομέα αυτό χρησιμοποιούν το iDASH για την παρακολούθηση των αποτελεσμάτων στις μελέτες αυτές.



Εικόνα 5: :Το iDASH, το νεότερο Εθνικό Κέντρο Βιοϊατρικής Πληροφορικής, είναι μια επιστημονική ομάδα που συγκεντρώνει μια πολυεθνική ομάδα επιστημόνων για την ανάπτυξη αλγορίθμων, εργαλείων και υπηρεσιών και βιοϊατρικής υποδομής στον κυβερνοχώρο για χρήση από βιοϊατρικούς και ερευνητές συμπεριφοράς. Στο σχήμα αυτό απεικονίζεται οι κυριότερες συνιστώσες σε κάθε περιοχή [30].

Διαφύλαξη Ιδιωτικότητας

Όσον αφορά την διαφύλαξη της ιδιωτικής ζωής των ατόμων που συμμετέχουν, οι ερευνητές του iDASH έχουν αναπτύξει αλγόριθμους και εργαλεία τα οποία να αντικατοπτρίζουν όλες τις ανησυχίες και τους φόβους που έχει το κοινό που θα βοηθήσει στην έρευνα.

Τα εργαλεία αυτά έχουν ως στόχο να διαφυλάσσουν την προσωπική ταυτότητα του κάθε ατόμου και να προσφέρουν προστασία και ενημέρωση ως προς το ποιοι ερευνητές και ιδρύματα θα μπορέσουν να χρησιμοποιήσουν τα δεδομένα αυτά στο μέλλον.

Η αρχιτεκτονική του iDASH έχει αναπτυχθεί σε ένα περιβάλλον το οποίο ακολουθεί τους κανόνες του πρότυπου HIPPA, και στεγάζεται σε ένα ασφαλές κέντρο δεδομένων στο San Diego Supercomputive Center. Αυτό είναι ένα υψηλής απόδοσης ίδρυμα το οποίο προστατεύεται με παρακολούθηση με βίντεο, ελεγχόμενη πρόσβαση και βιομετρικούς

ελέγχους. Οι πόροι που χρησιμοποιεί το iDASH προστατεύονται και από φυσικά αλλά και από εικονικά τοίχοι προστασίας και λειτουργούν μέσω ενός μεγάλου δικτύου διαχείρισης.

Τα εργαλεία και οι αλγόριθμοι του iDASH αναπτύσσονται για την εξυπηρέτηση των ατόμων που ανήκουν στην βιοϊατρική έρευνα. Επίσης παρέχει ένα περιβάλλον για τους μηχανικούς και επιστήμονες υπολογιστών, για να μπορούν να δοκιμάζουν και να αξιολογούν νέους αλγόριθμους.

Στις μέρες μας τα δεδομένα που παράγονται είναι πάρα πολλά και αυτά περιορίζονται σε μικρές ομάδες ερευνητών για μελέτη. Στόχος του iDASH είναι να προσφέρει μια μεγαλύτερη κοινότητα ερευνητών έτσι ώστε να μπορέσει να ξεκινήσει ο αγώνας για καινούργιες ανακαλύψεις και θεραπείες.

Εργαλεία Ανάλυσης

Για τις ανάγκες του iDASH έχουν αναπτυχθεί διαφόρων ειδών εργαλεία. Έχουν αναπτυχθεί αλγόριθμοι και εργαλεία για την ενσωμάτωση των δεδομένων για ανάλυση, καθώς και για την διαφύλαξη της ιδιωτικής ζωής.

Στην κατηγορία των εργαλείων για την ενσωμάτωση δεδομένων για ανάλυση οι ερευνητές του iDASH έχουν αναπτύξει τα εξής εργαλεία :

SlimGene

Για τις απαιτήσεις του έργου KD (Kawasaki Disease) αναπτύχθηκε μια εφαρμογή για την συμπίεση συνόλων δεδομένων υψηλής ανάλυσης, το SlimGene. Για ομάδες δεδομένων που αποτελούνται μόνο από πληροφορίες ανάγνωσης, το SlimGene μπορεί να πετύχει μέχρι 400 φορές συμπίεση. Όσον αφορά τις τιμές ποιότητας η συμπίεση τους είναι πιο δύσκολη από αυτή των πληροφοριών ανάγνωσης, ωστόσο το ανεπτυγμένο εργαλείο μπορεί να πετύχει 4-5 φορές συμπίεση.

Ένας άλλος στόχος του εργαλείου είναι να επιτρέπει την αναζήτηση γενετικής ποικιλίας. Για τον στόχο αυτό έγινε αφαίρεση των στρώσεων λογισμικού, κάτι που δίνει στα εργαλεία ανάλυσης γονιδιώματος την δυνατότητα να αλληλοεπιδρούν με καταναεμημένα και συμπεισμένα σύνολα δεδομένων.

Επίσης, αναπτύχθηκε μια γλώσσα ερωτημάτων γονιδιώματος έτσι ώστε οι ερευνητές να μπορούν να επεξεργαστούν τα γονιδιωματικά δεδομένα με έναν αποτελεσματικό τρόπο.

AnyExpress

Για το φιλτράρισμα δεδομένων που αφορούν τον τρόπο έκφρασης γονιδίων μεταξύ πλατφόρμων αναπτύχθηκε το εργαλείο AnyExpress. Το AnyExpress διαχειρίζεται πολλές πλατφόρμες χρησιμοποιώντας λογισμικό στο οποίο μπορούν να γίνουν διάφορες αλλαγές, όπως διάφορες στατιστικές μέθοδοι για επεξεργασία και νέες εκδόσεις πλατφόρμων.

Επίσης επιτρέπει στους διάφορους χρήστες που χρησιμοποιούν το AnyExpress να μπορούν να επιλέξουν πηγές αναφοράς σύμφωνα με την κρίση και τις προτιμήσεις τους.

OCEANS

Το OCEANS (Observational Cohort Event Analysis And Notification System), είναι μια σειρά από εύελικτες στατιστικές εφαρμογές, που έχουν σκοπό να εντοπίζουν τον αριθμό συμβάντων που έχουν σχέση με ένα συγκεκριμένο κρούσμα πάνω σε ένα σώμα που παρακολουθείται.

Προσφέρει στατιστικές μεθόδους, όπως στατιστικές διαδικασίες οι οποίες μπορούν να προσαρμοστούν ανάλογα με τον κίνδυνο, και διαδοχικά τεστ αναλογίας πιθανότητας επίσης προσαρμοσμένα ανάλογα με τον κίνδυνο.

Η προσαρμογή κινδύνου που προσφέρουν οι αλγόριθμοι αυτοί βασίζεται στην χρήση μοντέλων πρόβλεψης κινδύνου, καθώς και μηχανισμούς για την προσαρμογή για επαναλαμβανόμενη ρύθμιση σε περίπτωση σφάλματος.

2.1.4. BRISK (Biology-Related Information Storage Kit)

Το Biology-Related Information Storage Kit (BRISK) είναι ένα πακέτο τριών ανοιχτού κώδικα εφαρμογών ιστού, που παρέχουν μια συνεκτική πλατφόρμα ολοκλήρωσης και διαχείρισης δεδομένων. Αρχικά αναπτύχθηκε για ανταλλαγή δεδομένων για τους ερευνητές στο Consortium AllerGen (The Allergy, Genes and Environment Network) (<http://www.allergen-nce.ca>) [31].

Το BRISK είναι ένα πακέτο λογισμικού ανοικτού κώδικα που παρέχει την αναγκαία υποδομή για την αποτελεσματική επικοινωνία μεταξύ συνεργατών σχετικά με την αποθήκευση δεδομένων και δειγμάτων, τις μορφές και την ανάκτηση δεδομένων. Έχει την ικανότητα να χειρίζεται και να οργανώνει μεγάλα σύνολα δεδομένων, συμπεριλαμβανομένων των δεδομένων που παράγονται από διαχρονικές μελέτες. Αυτή η

λύση περιλαμβάνει επίσης μια εργαστηριακή προσανατολισμένη εφαρμογή που διαχειρίζεται το φυσικό δείγμα, το υποκείμενο και τα δεδομένα των δοχείων.

Το BRISK μπορεί να διαχειριστεί δεδομένα περιγραφής κλινικών φαινοτύπων και σωματικής μετάλλαξης (μονοπυρηνικοί πολυμορφισμοί). Παρέχει στους ερευνητές δυνατότητες ανάλυσης Μελετών Συσχετισμού Γονιδιώματος (Genome-Wide Association Study - GWAS).

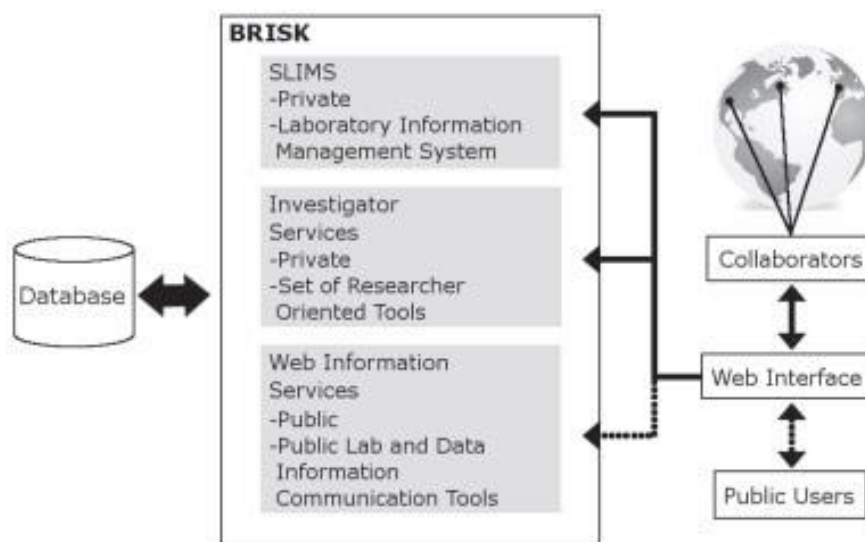
Το BRISK χρησιμοποιήθηκε από τους ερευνητές για την ανάκτηση πληροφοριών από μεγάλες μελέτες GWAS με εκατοντάδες χιλιάδες markers (>500.000), σε μικρότερες κλίμακες Illumina Golden Gate με 1536 δείκτες.

Υπηρεσίες BRISK

Το πακέτο λογισμικού BRISK έχει τρεις κύριες υπηρεσίες (Εικόνα 6):

- Σύστημα Διαχείρισης Εργαστηριακών Πληροφοριών βασισμένο στο Δείγμα (Sample-Based Laboratory Information Management System)
- Υπηρεσίες Ερευνών (Investigator Services)
- Υπηρεσίες Πληροφοριών Ιστού (Web Information Services)

Οι υπηρεσίες αναλύονται παρακάτω.



Εικόνα 6: Η δομή και λειτουργικότητα των τριών υπηρεσιών του BRISK [31].

Σύστημα Διαχείρισης Εργαστηριακών Πληροφοριών Που Βασίζεται Σε Δείγματα (Sample-Based Laboratory Information Management System - SLIMS)

Το SLIMS είναι ένα εμπλουτισμένο με χαρακτηριστικά εργαστηριακό σύστημα διαχείρισης πληροφοριών [32]. Προορίζεται να χρησιμοποιηθεί εσωτερικά από το

εργαστηριακό προσωπικό για τη διαχείριση δειγμάτων, αντικειμένων και δεδομένων δοχείων, υποστηρίζοντας τέσσερα νέα βασικά χαρακτηριστικά

- Υποστήριξη barcode
- Απομακρυσμένη είσοδος δείγματος
- Επισύναψη αρχείων
- Εισαγωγέας illumina

Υπηρεσίες Έρευνας (Investigator Services)

Προοριζόμενο να χρησιμοποιείται κυρίως από τους επικεφαλείς έρευνας, στατιστικούς και αναλυτές δεδομένων. Οι Υπηρεσίες Έρευνας λειτουργούν ως κεντρική βάση για την ανάκτηση δεδομένων. Επιτρέπει σε κάθε χρήστη να είναι αυτοδύναμος, χωρίς να χρειάζεται να βασίζεται σε άλλο μέλος του προσωπικού για τη δημιουργία αρχείων ανάλυσης για αυτά.

Υπηρεσίες Πληροφοριών Ιστού (Web Information Services)

Οι Υπηρεσίες Πληροφοριών Ιστού είναι ο συνδετικός κρίκος των υπηρεσιών του BRISK με το κοινό και αναπτύχθηκε ως αυτόματο σύστημα ανάκτησης δεδομένων για τα τεράστια ποσά των δεδομένων που παράγονται στο GWAS. Έχει δύο κύριες λειτουργίες:

- Παρέχει στο εργαστήριο ένα μέρος για να περιγράψει την έρευνά του και να παρουσιάσει τα ευρήματά του.
- Επιτρέπει στο κοινό να διερευνήσει και να αλληλεπιδράσει με τα δεδομένα.

2.1.5. STRIDE (Stanford Translational Research Integrated Database Environment

Το STRIDE [33] (Stanford Translational Research Integrated Database Environment) είναι ένα ερευνητικό πρόγραμμα του Πανεπιστημίου του Στάνφορντ για την ανάπτυξη μιας πλατφόρμας βασισμένης σε πρότυπα για την υποστήριξη κλινικής και μεταφραστικής έρευνας.

Το 2003, το SCCI (Stanford Center for Clinical Informatics) ξεκίνησε μια πολυετή έρευνα για την δημιουργία μιας ολοκληρωμένης, βασισμένης σε πρότυπα πληροφορικής πλατφόρμας με σκοπό την υπερπήδηση τριών μεγάλων εμποδίων για αποτελεσματικότερη κλινική και μεταφραστική έρευνα :

1. αποτελεσματική πρόσβαση σε κλινικά δεδομένα για ερευνητικούς σκοπούς
2. παράδοση στιβαρών λύσεων διαχείρισης ερευνητικών δεδομένων και
3. διαθεσιμότητα του συστήματος για τη διαχείριση και την ανακάλυψη δεδομένων βιοστοιχείων (biospecimen)

Το σύστημα χρησιμοποιείται καθημερινά στο Πανεπιστήμιο του Στάνφορντ και είναι ένα σημαντικό μέρος του Προγράμματος πληροφορική CTSA (Clinical and Translational Science Award) του πανεπιστημίου [34] .

Το STRIDE αποτελείται από τρία ολοκληρωμένα στοιχεία :

1. μια αποθήκη κλινικών δεδομένων, βασισμένη στο HL7 Reference Information Model (RIM)
2. ένα πλαίσιο ανάπτυξης εφαρμογών για την δημιουργία εφαρμογών διαχείρισης ερευνητικών δεδομένων πάνω στην πλατφόρμα STRIDE και
3. ένα σύστημα διαχείρισης δεδομένων βιοστοιχείων (biospecimen)

Μέχρι σήμερα, έχουν αναπτυχθεί διάφορες εφαρμογές διαχείριση ερευνητικών δεδομένων και βάσεις δεδομένων βασισμένες στην πλατφόρμα του STRIDE, που κυμαίνονται από ένα μεγάλο πολυμέσικο σύστημα ικανό να υποστηρίζει ένα ολόκληρο ακαδημαϊκό τμήμα, διάφορα διαχρονικά μητρώα ασθενών και αρκετά λιγότερο σύνθετα συστήματα παροχής στήριξης σε μικρά και μεσαίου μεγέθους ερευνητικά έργα.

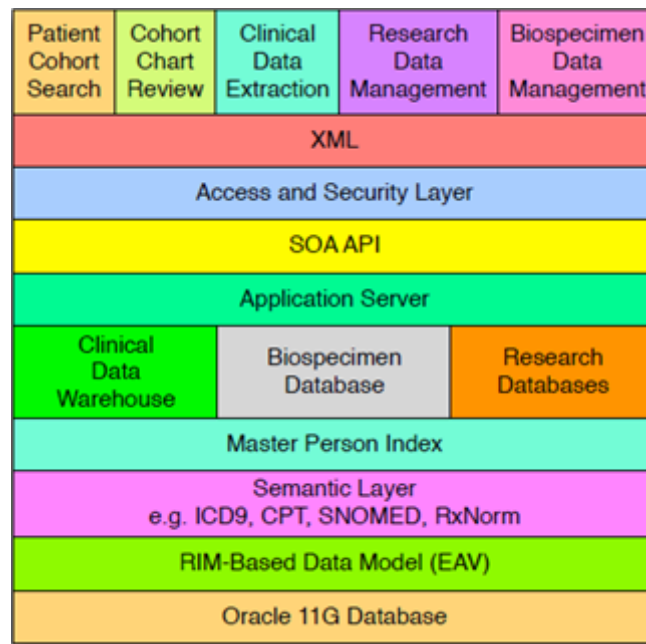
Αρχιτεκτονική της στοίβα STRIDE

Το STRIDE (Εικόνα 7:) είναι χτισμένο πάνω στην πλατφόρμα σχεσιακής βάσης δεδομένων Oracle 11g και χρησιμοποιεί μια αρχιτεκτονική n-επιπέδων. Τα δεδομένα αποθηκεύονται χρησιμοποιώντας ένα μοντέλο Entity-Attribute-Value (EAV) [35] και αντικειμενοστραφείς δομές δεδομένων (πρόσωπα, ρόλους και πράξεις) που προέρχεται από το HL7 Reference Information Model (RIM) [36]. Το σύστημα περιλαμβάνει ένα Master Person Index (MPI) που συμπληρώνεται δυναμικά από κλινικά, ερευνητικά και δεδομένα βιοστοιχείων.

Physical Database

Το επίπεδο της φυσικής βάσης (physical database) δεδομένων του STRIDE είναι οργανωμένη σε τρία τμήματα:

1. μια αποθήκη κλινικών δεδομένων Clinical Data Warehouse (CDW), βασισμένη στο HL7 Reference Information Model (RIM)
2. διαχείριση ερευνητικών βάσεων δεδομένων (Research Database Management) που υποστηρίζει πολλαπλές ξεχωριστές λογικές ερευνητικές βάσεις δεδομένων και
3. το αποθετήριο δεδομένων βιοστοιχείων (Biospecimen Data Repository), που υποστηρίζει πολλές ξεχωριστές βάσεις δεδομένων βιοστοιχείων



Εικόνα 7: Αρχιτεκτονική της στοίβα STRIDE [33]

Τα τρία αυτά τμήματα βασίζονται στην ίδια αρχιτεκτονική και υπηρεσίες. Τα δεδομένα μπορούν να διαμοιράζονται μεταξύ όλων των τμημάτων. Για παράδειγμα, τα ερευνητικά ή/και δεδομένα βιοστοιχείων μπορούν να συγχωνευθούν με κλινικά δεδομένα από την αποθήκη κλινικών δεδομένων CDW.

Semantic Layer

Το σημασιολογικό επίπεδο του STRIDE (Semantic Layer) αποτελείται από ένα πλαίσιο που υποστηρίζει πολλές ορολογίες, συμπεριλαμβανομένων ICD9 CM, ICDO, CPT, RxNorm και SNOMED. Το μοντέλο αυτό μικτών ορολογιών υποστηρίζει καταχώρηση δεδομένων που βασίζονται σε πρότυπα, ενσωμάτωση δεδομένων, ιεραρχική ανάκτηση και διαλειτουργικότητα δεδομένων.

Για παράδειγμα, χρησιμοποιώντας την ορολογία RxNorm για την αναπαράσταση δεδομένων φαρμάκων, το σύστημα μπορεί να συγχωνεύσει πληροφορίες για φάρμακα από

δύο μοντέλα φαρμάκων διαφορετικού προμηθευτή, με δυναμική ολοκλήρωση παραγγελία παιδιατρικών και φαρμάκων ενήλικων εντός του CDW.

Εφαρμογές

Οι εφαρμογές του STRIDE χρησιμοποιούν τη σουίτα Java Swing και επικοινωνούν με ένα σύνολο υπηρεσιών για τη διαχείριση των δημογραφικών, κλινικών, ερευνητικών και δεδομένων αναφοράς. Αυτές οι υπηρεσίες λογισμικού οργανώνονται σε μία Service Oriented αρχιτεκτονική (Service Oriented Architecture - SOA), με τις υπηρεσίες γραμμής να εξυπηρετούν όχι μόνο εφαρμογές προγράμματος-πελάτη αλλά επίσης κλινικών και ερευνητικών υπηρεσιών δεδομένων της στοίβας του πυρήνα του συστήματος. Οι εφαρμογές πελάτη STRIDE, είναι γραμμένες σε Java και αναπτύσσονται χρησιμοποιώντας επαναχρησιμοποιήσιμα κατασκευαστικά στοιχεία εξασφαλίζοντας τη γρήγορη ανάπτυξη εφαρμογών.

Η χρήση Application Programming Interface (API) επιτρέπει σε εφαρμογές προγράμματος-πελάτη να καθορίζουν δυναμικά το περιεχόμενο των ερωτημάτων ανάκτησης δεδομένων. Στη συνέχεια αυτά μεταφράζονται στη πλευρά του σέρβερ σε ερωτήματα SQL από τους κανόνες ορατότητας του επιπέδου πρόσβασης και ασφάλειας του STRIDE (Access and Security Layer) με αποτέλεσμα την δημιουργία δυναμικά αναπροσδιορισμών σύνολα δεδομένων Virtual Private Databases (VPD) [37].

Από τον Ιούλιο του 2009 το STRIDE CDW περιέχει πλήρως προσδιορίζονται συγχωνευμένες δημογραφικές και κλινικές πληροφορίες για πάνω από 1,3 εκατομμύρια παιδιατρικών και ενηλίκων ασθενών που έχουν νοσηλευτεί στο Stanford University Medical Centre (SUMC).

Λειτουργίες

Αναζήτηση Δεδομένων (Patient Cohort Search)

Το STRIDE παρέχει τη δυνατότητα αναζήτησης του CDW για εντοπισμό πιθανών ερευνητικών κλινικών ασθενών, που θα έχουν συγκεκριμένα χαρακτηριστικά και θα είναι κατάλληλοι για συγκεκριμένη έρευνα. Αυτό γίνεται με τη χρήση ενός εργαλείου οπτικών ερωτημάτων, που ονομάζεται Anonymous Patient Cohort Discovery Tool.

Αναθεώρηση Δεδομένων (Cohort Chart Review)

Το STRIDE παρέχει επίσης ένα εργαλείο αναθεώρησης κλινικών δεδομένων σχεδιασμένο για την εξέταση κλινικών δεδομένων για να προσδιοριστεί εάν οι ασθενείς είναι κατάλληλοι για συμπερίληψη σε ερευνητικές μελέτες.

Εξαγωγή Δεδομένων (Clinical Data Extraction)

Οι κλινικές μελέτες των ασθενών μπορούν να αποθηκευτούν μέσα στο κεντρικό σύστημα του STRIDE και να χρησιμοποιηθούν μελλοντικά για αναθεώρηση ή εξαγωγή του σύνολο δεδομένων .

Διαχείριση Ερευνητικών Βάσεων Δεδομένων (Research Data Management)

Η πλατφόρμα STRIDE παρέχει τη δυνατότητα ανάπτυξης εφαρμογών διαχείρισης ερευνητικών δεδομένων, με ενσωματωμένη απεικόνιση δεδομένων. Τα εργαλεία ανάπτυξης εφαρμογών του STRIDE προσφέρουν μεγάλη ποικιλία στοιχείων, από απλές λίστες τιμών, αριθμητικά ή πεδία εισαγωγής κειμένου και widgets με αναζήτηση-καθώς-πληκτρολογείς, που επιτρέπουν στους χρήστες που εισάγουν ερευνητικά δεδομένα την ταχεία αναζήτηση δεδομένων αναφοράς για κατάλληλη τιμή.

Αποθετήριο Δεδομένων Βιοστοιχείων (Biospecimen Data Repository)

Το αποθετήριο δεδομένων βιοστοιχείων του STRIDE παρέχει λειτουργίες καταγραφής, παρακολούθησης και διαχείρισης για πολλαπλές ξεχωριστές τράπεζες βιοστοιχείων. Μεμονωμένες τράπεζες μπορεί να αποθηκεύσουν και να διαχειριστούν τα δεδομένα βιοστοιχείων ιδιωτικά στο STRIDE, ενώ καθολικά δεδομένα βιοστοιχείων μπορούν να αναζητηθούν σε όλες τις τράπεζες. Αυτό επιτρέπει τη δημιουργία ενός εικονικού αποθετηρίου δεδομένων βιοστοιχείων, που μπορούν να προστεθούν βαθμιαία δεδομένα από επιπλέον τράπεζες.

2.2. Συστήματα Εξειδικευμένα Στην Έρευνα Για Τον Καρκίνο

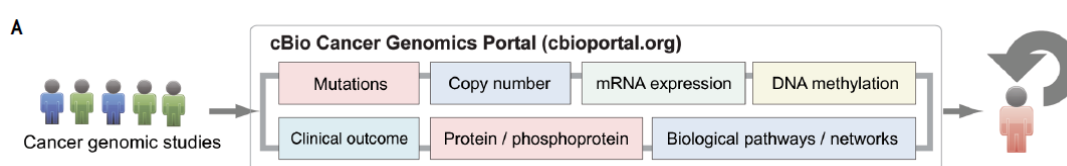
Τα μεταφραστικά δεδομένα μπορούν να προσφέρουν μεγάλη βοήθεια στον πρόωρο εντοπισμό, την πρόληψη και την αντιμετώπιση διάφορων μορφών καρκίνου. Η λειτουργία αυτή είναι εξαιρετικά χρήσιμη μιας και ο καρκίνος είναι μια μεγάλη μάστιγα στις μέρες μας και όσο μεγαλύτερη είναι η συνεισφορά στην καταπολέμηση του, τόσο πιο ραγδαίο θα είναι το ποσοστό των ατόμων που θα προλαβαίνει να το αντιμετωπίσει.

Παρακάτω αναλύονται τα συστήματα μεταφραστικών δεδομένων εξειδικευμένα στην έρευνα για τον καρκίνο.

2.2.1. cBIO Cancer Genomics Portal

Το cBio (<http://cbioportal.org>) είναι μια πύλη γονιδιώματος του καρκίνου η οποία προσφέρει πρόσβαση σε δεδομένα τα οποία έχουν συλλεχθεί από πάνω από 20 μελέτες για τον καρκίνο και πάνω από 5.000 δείγματα όγκων από τις μελέτες αυτές.

Η χρήση του cBio προσφέρει στους ερευνητές, που χειρίζονται και χρησιμοποιούν τα δεδομένα αυτά για έρευνες, την δυνατότητα για γρηγορότερη και ακριβέστερη αναζήτηση και πρόσβαση πάνω σε μοριακά προφίλ καθώς και σε κλινικά δεδομένα, έτσι ώστε να τα χρησιμοποιήσουν για βιολογική επίγνωση όπως επίσης και σε κλινικές εφαρμογές.



Εικόνα 8: Η πύλη γονιδιωματικής του καρκίνου του cBio. Η cBio Cancer Genomics Portal είναι μια ανοιχτή πλατφόρμα για την διαδραστική διερεύνηση πολυδιάστατων συνόλων δεδομένων γονιδιωματικής του καρκίνου στο πλαίσιο των κλινικών δεδομένων και των βιολογικών οδών [38].

Οι έρευνες στο πεδίο του γονιδιώματος του καρκίνου αναπτύσσονται και εξελίσσονται με ραγδαίο ρυθμό, μέσω των μεγάλων προσπάθειών του The Cancer Genome Atlas (TCGA) και του International Cancer Genome Consortium (ICGC) [39]. Τα δεδομένα που παραγόntonτουσαν από τις έρευνες αυτές όμως δεν ήταν εύκολα διαθέσιμα για τους υπόλοιπους επιστήμονες της ερευνητικής κοινότητας του καρκίνου και για τον λόγο αυτό ήταν δύσκολη η μετάφραση των δεδομένων αυτών και η μετατροπή τους σε φάρμακα, νέες βιολογικές γνώσεις και κλινικές δοκιμές.

Για τον λόγο αυτό αναπτύχθηκε στο Memorial Sloan-Kettering Cancer Center (MSKCC) η πύλη γονιδιώματος του καρκίνου του cBio, έτσι ώστε να μπορεί να γίνεται ευκολότερη η συλλογή και ταξινόμηση του μεγάλου όγκου δεδομένων από έρευνες πάνω στο γονιδίωμα του καρκίνου, καθώς και στο να είναι ευκολότερη η πρόσβαση των δεδομένων αυτών σε όλη την ερευνητική κοινότητα του τομέα αυτού (Εικόνα 8).

Υποστηριζόμενοι Τύποι Δεδομένων

Η πλατφόρμα cBio περιέχει 5 δημοσιευμένα σύνολα δεδομένων και 15 προσωρινά σύνολα δεδομένων TCGA. Τα προσωρινά σύνολα δεδομένων TCGA ενημερώνονται μηνιαίως, με βάση τις τελευταίες εκδόσεις παραγωγής TCGA και η πλατφόρμα θα ενημερώνεται συνεχώς καθώς προστίθενται νέοι τύποι καρκίνου TCGA.

Τα δημοσιευμένα σύνολα δεδομένων περιλαμβάνουν δεδομένα μετάλλαξης, αλλά δεν υπάρχουν προσωρινά σύνολα δεδομένων. Καθώς ο κάθε τύπος καρκίνου εντός του TCGA οριστικοποιείται και οι σωματικές μεταλλάξεις επικυρώνονται, τα δεδομένα μετάλλαξης θα δημοσιευθούν και θα προστεθούν στην πύλη.

Εκτός από τα δεδομένα μετάλλαξης, η πύλη περιλαμβάνει:

- Αλλοιώσεις αριθμών αντιγράφων.
- Μεταβολές έκφρασης mRNA που βασίζονται σε μικροσυστοιχίες και RNA με ανάλυση αλληλουχίας.
- Τιμές μεθυλίωσης DNA.
- Επίπεδα πρωτεϊνών και φωσφοπρωτεϊνών.

Για τον κάθε τύπο δεδομένων υπάρχει ένα επίπεδο γονιδίου (gene level) στο οποίο αποθηκεύονται, και έπειτα συνδυάζονται με άλλα διαθέσιμα κλινικά δεδομένα τα οποία είναι μη αναγνωρισμένα, όπως συνολικό διάστημα επιβίωσης του ασθενή και συνολικό διάστημα που ο ασθενής επιβίωσε χωρίς καθόλου συμπτώματα της ασθένειας. Στην συνέχεια τα δεδομένα που προκύπτουν οργανώνονται ως μια συνάρτηση ασθενή και γονιδίου.

Χαρακτηριστικά cBio

Τροποποιημένα Γονίδια

Το βασικό χαρακτηριστικό της πύλης είναι η έννοια των τροποποιημένων γονιδίων. Συγκεκριμένα ένα γονίδιο ταξινομείται ως τροποποιημένο πάνω σε κάποιον ασθενή σε περίπτωση που έχει μεταλλαχθεί, αφαιρεθεί, ενισχυθεί ή η σχετική έκφρασή του mRNA είναι μικρότερη ή μεγαλύτερη από ένα καθορισμένο από το χρήστη όριο.

Η έννοια των τροποποιημένων γονιδίων επιτρέπει στους χρήστες να αναλύουν πολύπλοκα σύνολα δεδομένων και να αναπτύσσουν βιολογικές υποθέσεις σχετικά με επαναλαμβανόμενα τροποποιημένα σύνολα γονιδίων και να χαράσσουν νέα βιολογικά μονοπάτια γνώσης.

Ευκολία Χρήσης

Το cBio χαρακτηρίζεται από την ευκολία χρήσης του. Όλα τα εργαλεία της πύλης του cBio είναι διαθέσιμα μέσω μιας διεπαφής ιστού (web interface) 4 βημάτων, στο οποίο οι χρήστες καλούνται να επιλέξουν:

1. Την μελέτη καρκίνου για την οποία ενδιαφέρονται (π.χ. GBM - TCGA Glioblastoma Multiforme).
2. Ένα ή περισσότερα γονιδιακά προφίλ (π.χ. μεταλλάξεις και αλλοιώσεις).
3. Μια περίπτωση ασθενούς (π.χ. ασθενείς με μετάλλαξη GBM, τον αριθμό αντιγράφου και τα δεδομένα mRNA του ασθενή).
4. Ένα γονιδιακό σύνολο το οποίο ενδιαφέρει τον χρήστη.

Οι χρήστες έχουν επίσης τη δυνατότητα να υπολογίζουν αυτόματα την αμοιβαία αποκλειστικότητα και την συνύπαρξη μεταξύ όλων των ζευγών γονιδίων. Επίσης, δίνεται στους χρήστες η δυνατότητα να κάνουν αναζήτηση με ένα πιο απλό ερώτημα 2 βημάτων, όπου οι χρήστες καλούνται να επιλέξουν μόνο "όλες της μελέτες για τον καρκίνο" και να διαλέξουν το σύνολο γονιδίων για το οποίο ενδιαφέρονται

Συμπερασματικά, η πλατφόρμα cBio διευκολύνει την πρόσβαση σε σύνολα γονιδιωματικών δεδομένων για τον καρκίνο για ολόκληρη τη βιοϊατρική κοινότητα. Παρέχει μια απλή αλλά ευέλικτη διασύνδεση με ολοκληρωμένα σύνολα δεδομένων, έξυπνες επιλογές απεικόνισης και μια προγραμματική διαδικτυακή διασύνδεση, τα οποία μπορούν να βοηθήσουν τους ερευνητές στη μετάφραση των γονιδιωματικών δεδομένων του καρκίνου σε βιολογικές γνώσεις και πιθανές κλινικές εφαρμογές.

Με την ενσωμάτωση πολλαπλών τύπων γονιδιωματικών δεδομένων και την ευκολότερη πρόσβαση στα δεδομένα, η πλατφόρμα επιτρέπει στους ερευνητές να ανακτήσουν ευκολότερα γονιδιωματικά δεδομένα, να δοκιμάσουν υποθέσεις σχετικά με τις γενετικές αλλοιώσεις στον καρκίνο και να προσθέσουν γονιδιωματικά δεδομένα στο πλαίσιο προηγούμενων βιολογικών ερευνών.

Η πλατφόρμα cBio συμπληρώνει τα υπάρχοντα εργαλεία, συμπεριλαμβανομένων των πυλών δεδομένων TCGA και ICGC [39], του IGV [40], του UCSC Cancer Genomics Browser [41] και του IntOGen [42] εστιάζοντας στην ανάλυση διακριτών γονιδιακών δεδομένων, ευκολία χρήσης, υποστήριξη διερευνητικής ανάλυσης δεδομένων και ανάλυση διαδραστικού δικτύου [38].

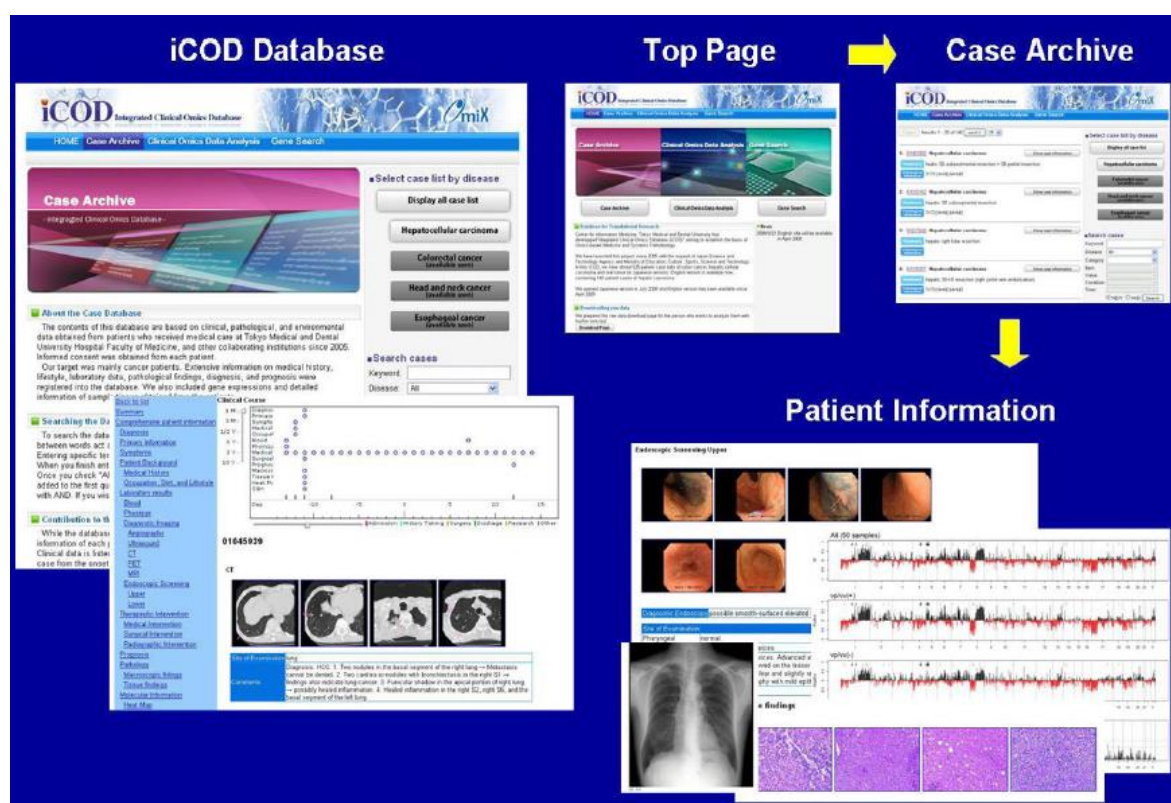
2.2.2. iCOD (Integrated Clinical Omics Database)

Η Ολοκληρωμένη Βάση Δεδομένων Κλινικής Omics (Integrated Clinical Omics Database - iCOD) [43] αναπτύχθηκε για να συνδυάζει ολοκληρωμένες κλινικοπαθολογικές και μοριακές πληροφορίες ασθενών για να παρέχει μια ολιστική κατανόηση των ασθενειών (

Εικόνα 9: Βασικά στιγμιότυπα οθόνης του συστήματος iCOD [24].

).

Το iCOD είναι η πρώτη βάση δεδομένων που παρέχει τις ολοκληρωμένες κλινικές, παθολογικές πληροφορίες και πληροφορίες για τον τρόπο ζωής εκτός από τις πληροφορίες μοριακής βιολογίας καθώς και την εκτιμώμενη αλληλεξάρτησή τους.



Εικόνα 9: Βασικά στιγμιότυπα οθόνης του συστήματος iCOD [24].

Η βάση δεδομένων iCOD είναι χρήσιμη τόσο για τους κλινικούς ερευνητές που προτίθενται να αποκτήσουν γνώσεις σχετικά με τη μοριακή βάση της νόσου που θα μπορούσαν να χρησιμοποιηθούν για τη διάγνωση, τη θεραπεία και την πρόγνωση των ασθενειών και για τους μοριακούς βιολόγους που σκοπεύουν να γνωρίσουν τη λειτουργία και τον φαινότυπο των μοριακών οδών τη σχέση τους μέσω της γνώσης σε περιπτώσεις δυσλειτουργίας τους.

Υποστηριζόμενοι Τύποι Δεδομένων

Τα περιεχόμενα του iCOD βασίζονται σε κλινικά, παθολογικά και περιβαλλοντικά δεδομένα που λαμβάνονται από ασθενείς που έλαβαν ιατρική περίθαλψη στο Tokyo Medical and Dental University Hospital και σε άλλα συνεργαζόμενα ιδρύματα από το 2005.

Το iCOD μπορεί να χειριστεί

- Δεδομένα omics όπως προφίλ γονιδιακής έκφρασης.
- Ετερογενείς κλινικές πληροφορίες, όπως λεπτομερείς φαινότυποι, εικόνες ακτινολογίας ή αποτελέσματα εργαστηριακών εξετάσεων.

Εργαλεία iCOD

Το iCOD παρέχει στους χρήστες μια βολική μηχανή αναζήτησης για την αναζήτηση λέξεων-κλειδιών που σχετίζονται με παθολογικά / κλινικά ευρήματα και αναγνωριστικό ασθενούς που είναι αποθηκευμένα στη βάση δεδομένων.

Οι τοπικοί χάρτες ολοκληρωμένης απεικόνισης των ασθενειών παρέχονται για να παρατηρηθεί η συσχέτιση μεταξύ κλινικοπαθολογικού φαινοτύπου και γονιδιακής έκφρασης χρησιμοποιώντας πολυπαραγοντική στατιστική ανάλυση που εφαρμόζονται στις μοριακές και κλινικοπαθολογικές πληροφορίες των ασθενών και αντιπροσωπεύουν πιθανές οδούς ασθένειας.

Συμπερασματικά, το iCOD είναι η πρώτη βάση δεδομένων που παρέχει ολοκληρωμένες κλινικές, παθολογικές πληροφορίες και πληροφορίες για τον τρόπο ζωής επιπρόσθετα με τις πληροφορίες μοριακής βιολογίας, καθώς και την εκτιμώμενη αλληλεξάρτησή τους.

Η βάση δεδομένων iCOD είναι χρήσιμη τόσο για τους κλινικούς ερευνητές, που ερευνούν τη μοριακή βάση της νόσου και θα μπορούσαν να χρησιμοποιηθούν για τη διάγνωση, τη θεραπεία και την πρόγνωση των ασθενειών, αλλά και για τους μοριακούς βιολόγους, που μελετούν τη λειτουργία και τον φαινότυπο των μοριακών μονοπατιών και τη σχέση τους σε περιπτώσεις δυσλειτουργίας τους.

Η ενσωμάτωση δεδομένων omics στη βάση δεδομένων, μπορεί να καλύψει τις πολλαπλές ιεραρχικές (από μοριακό έως κλινικό επίπεδο) πληροφορίες σχετικά με τις ασθένειες.

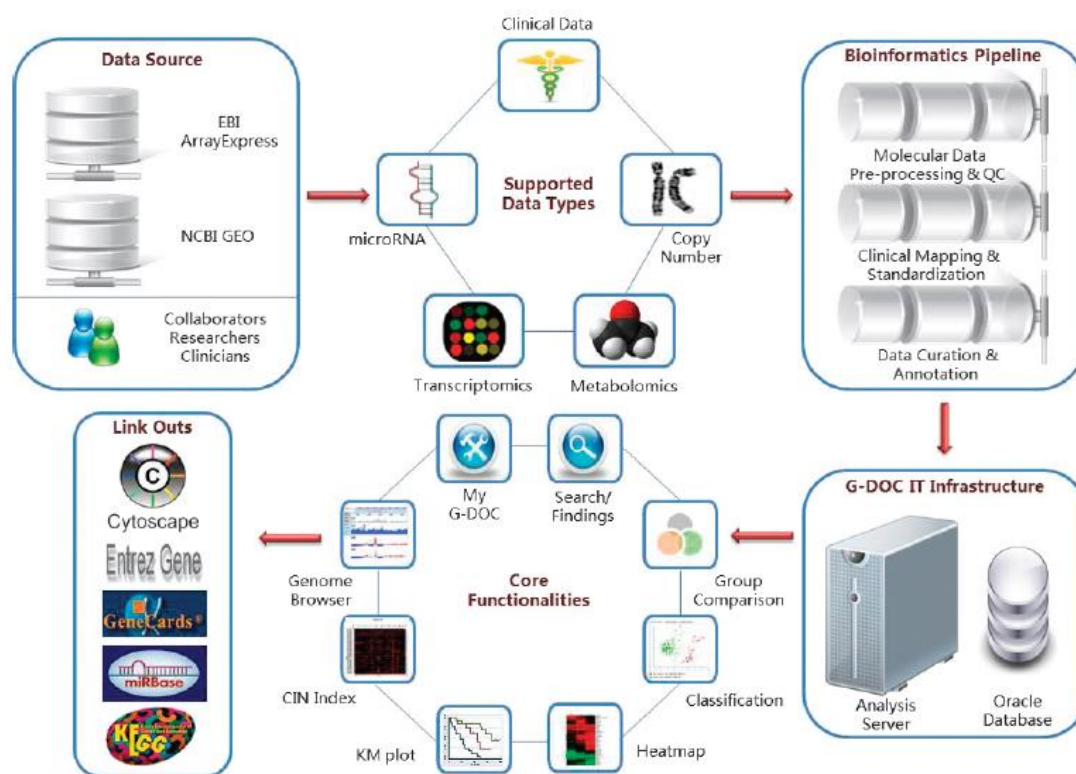
2.2.3. G-DOC (Georgetown Database of Cancer)

Το Georgetown Database of Cancer (G-DOC), που αναπτύχθηκε στο πλήρες κέντρο καρκίνου του Lombardi στο Πανεπιστήμιο του Georgetown [44], είναι μια ενοποιημένη

πλατφόρμα ιστού που επιτρέπει τη βασική και κλινική έρευνα με την ενσωμάτωση των χαρακτηριστικών του ασθενούς και των δεδομένων κλινικών αποτελεσμάτων σε συνδυασμό με μια ποικιλία ερευνητικών δεδομένων υψηλής απόδοσης.

Η πρώτη δημόσια έκδοση της δικτυακής πύλης G-DOC εγκαινιάστηκε τον Απρίλιο του 2011 από την ερευνητική κοινότητα του Πανεπιστημίου Georgetown και τους συνεργάτες του, παρέχοντας στους ερευνητές του καρκίνου ένα ευρύ φάσμα εργαλείων για την απεικόνιση και την ανάλυση δεδομένων, omics από προηγούμενως δημοσιευμένες κλινικές μελέτες για τον καρκίνο και ένα μικρότερο σύνολο συνόλων ιδιωτικών δεδομένων.

Επί του παρόντος, το G-DOC περιλαμβάνει δεδομένα που συλλέχθηκαν από περισσότερα από 2500 δείγματα καρκίνου του μαστού και σχεδόν 800 δείγματα καρκίνων του γαστρεντερικού (ήπατος, παχέος εντέρου, στομάχου και παγκρέατος).



Εικόνα 10: Επισκόπηση δεδομένων και χαρακτηριστικών ανάλυσης στο G-DOC. Τα δεδομένα (δημόσια ή ιδιωτικά) επεξεργάζονται ομοιόμορφα μέσω τυποποιημένων αγωγών βιοπληροφορικής και διατίθενται σε διάφορα εργαλεία ανάλυσης μέσω μιας Web διεπαφής φιλικής προς το χρήστη [44].

Υποστηριζόμενοι Τύποι Δεδομένων

Το G-DOC περιλαμβάνει μια ευρεία συλλογή βιοπληροφορικών συστημάτων και βιολογίας για ανάλυση και απεικόνιση τεσσάρων μεγάλων τύπων δεδομένων omics (Εικόνα 10).

Δεδομένα Έκφρασης mRNA.

Μεγάλο μέρος των δεδομένων που βρίσκονται επί του παρόντος στο G-DOC είναι δεδομένα έκφρασης mRNA που παράγονται από πειράματα υβριδισμού συστοιχίας. Περιλαμβάνουν τόσο δεδομένα αναλογίας δύο καναλιών όσο και δεδομένα εντάσεως ενός καναλιού.

Δεδομένα Έκφρασης miRNA

Το MicroRNA αποτελεί αντικείμενο αυξανόμενου ενδιαφέροντος για τις κλινικές, μεταφραστικές και βασικές επιστημονικές κοινότητες του καρκίνου και το G-DOC υποστηρίζει αυτόν τον τύπο δεδομένων. Έχει αναπτυχθεί ένας αγωγός προεπεξεργασίας δεδομένων για δεδομένα έκφρασης miRNA που υποστηρίζει τις μεγάλες μορφές πλατφόρμας υψηλής απόδοσης.

Δεδομένα Μεταβολισμού

Η μεταβολική (metabolics) είναι μία από τις νεότερες επιστήμες omics και στοχεύει στη μελέτη των παγκόσμιων προφίλ μεταβολιτών μικρών μορίων εντός ενός βιολογικού συστήματος κάτω από ένα δεδομένο σύνολο συνθηκών. Τυπικά, αυτά τα πειράματα εκτελούνται σε βιολογικά υγρά όπως ούρα, σάλιο ή πλάσμα αίματος, αλλά μπορούν επίσης να χρησιμοποιηθούν απομονωμένα κύτταρα και ιστοί.

Δεδομένα Αριθμού Αντιγράφων DNA

Έχει δοθεί σημαντική προσοχή στην κατανόηση των γενετικών συμβάντων χρωμοσωματικής τροποποίησης που είναι κοινά σε πολλά είδη καρκίνου. Σημαντικό ενδιαφέρον δίνεται στην αναγνώριση των αποκλίσεων που εμφανίζονται στην εξέλιξη και στην πρόοδο των νεοπλασιών.

Για να εξασφαλιστεί ότι η G-DOC μπορεί να επιτρέψει στους ερευνητές να χρησιμοποιήσουν καλύτερα αυτόν τον τύπο δεδομένων, αναπτύχθηκε ένας αγωγός επεξεργασίας δεδομένων με χρήση του R (<http://www.r-project.org/>) για ανάλυση και απεικόνιση των δεδομένων αριθμού αντιγράφων DNA που λαμβάνονται από μια ποικιλία πλατφόρμων.

Όλα συνδέονται με κλινικές πληροφορίες ασθενούς που δεν έχουν αναγνωριστεί, αυξάνοντας σημαντικά την αξία τους. Τα αποτελέσματα της ανάλυσης δεδομένων omics χαρτογραφούνται σε έναν ενσωματωμένο φυλλομετρητή ανθρώπινου γονιδιώματος στο επίπεδο είτε των μεμονωμένων ασθενών ή / και ομάδων ασθενών.

Το G-DOC επιτρέπει στους ερευνητές να συνδυάσουν το συνηθισμένο κλινικό προσωπικό ιστορικό, τη φυσική εξέταση, τις εργαστηριακές μελέτες, τις μελέτες ακτινολογίας, το οικογενειακό ιστορικό και άλλα σχετικά δεδομένα - με λεπτομερή ανάλυση omics του καρκίνου του ασθενούς, προκειμένου να διευκολυνθεί η εξερεύνηση των κλινικών και μοριακών παραγόντων που καθορίζουν την έκβαση της νόσου.

Βασικά Χαρακτηριστικά του G-DOC

Αναζήτηση

Το G-DOC διαθέτει εύχρηστες δυνατότητες αναζήτησης για κλινικά δεδομένα, μελέτες, biospecimen, δεδομένα omics, μικρά μόρια και βασικά δημοσιευμένα ευρήματα. Το σχετικό πλαίσιο, η Georgetown Clinical and Omics Development Engine (G-CODE) [45] περιέχει ένα ευρύ φάσμα εργαλείων βιοπληροφορικής και συστημάτων βιολογίας για την ανάλυση και απεικόνιση δεδομένων σε χρήστες με ποικίλα υπόβαθρα και επίπεδα δεξιοτήτων.

Συστήματα Βιολογικών Αναλύσεων και Οπτικοποίηση

Το G-DOC υποστηρίζει εργασίες υψηλής έντασης υπολογιστών, όπως σύγκριση τάξεων, ιεραρχική συσσώματωση, ανάλυση κύριων συστατικών και ανάλυση δικτύου για μεταγραφικά, γονιδιωματικά και μεταβολικά δεδομένα.

Εξατομικευμένο G-DOC (My G-DOC)

Η εφαρμογή G-DOC επιτρέπει τον καθορισμό του επιπέδου για τις μελέτες και τα δεδομένα. Έτσι αυτά μπορούν:

- να δημοσιοποιηθούν (προσβάσιμα σε όλους τους χρήστες)
- να περιοριστούν σε μία ή περισσότερες ομάδες συνεργασίας
- να είναι διαθέσιμα μόνο στον κάτοχο δεδομένων.

Οι ομάδες συνεργασίας είναι περιορισμένες ομάδες χρηστών, οι οποίες καθορίζονται από τον διαχειριστή ομάδας, και μπορούν να μοιράζονται πληροφορίες μεταξύ των μελών μόνο. Ο διαχειριστής ομάδας συνεργασίας έχει πλήρη έλεγχο της πρόσβασης σε περιορισμένα δεδομένα και είτε να εγκρίνει είτε να απορρίπτει αιτήσεις πρόσβασης.

Οι ομάδες συνεργασίας επιτρέπουν στους χρήστες να μοιράζονται τους καταλόγους των ασθενών, των γονιδίων, των αναφορών και των αποτελεσμάτων της ανάλυσης σε ένα ασφαλές και συνεργατικό περιβάλλον που προάγει την επικοινωνία και την συνεργασία των επιστημονικών ομάδων.

Το σύνολο αυτών των χαρακτηριστικών παρέχει ένα ασφαλές περιβάλλον εργασίας, όπου τα μέλη των ομάδων μπορούν να μοιράζονται αποτελέσματα ανάλυσης, εργαλεία και λίστες βιοδεικτών μεταξύ τους.

Βάση Δεδομένων

Επί του παρόντος, η βάση δεδομένων του G-DOC απαρτίζεται από ερευνητικά και κλινικά δεδομένα από το Πανεπιστήμιο του Georgetown και ως εκ τούτου περιέχει δεδομένα από ασθενείς με καρκίνο του μαστού και του παχέος εντέρου. Ωστόσο, με την πάροδο του χρόνου, η βάση θα εμπλουτιστεί σημαντικά με την ανατροφοδότηση που λαμβάνει από την κοινότητα των χρηστών.

Η βάση δεδομένων της G-DOC αποτελείται από δύο αποκλίνοντα σύνολα:

- **Δημόσια Δεδομένα**

Είναι διαθέσιμα σε όλους. Τα σύνολα δημόσιων δεδομένων λαμβάνονται συνήθως από αποθετήρια όπως το National Center for Biotechnology Information Gene Expression Omnibus [46] και το EBI Array Express [47].

- **Ιδιωτικά Δεδομένα**

Παρέχονται από και είναι προσιτά σε μεμονωμένους ερευνητές και τις ομάδες συνεργασίας τους. Τα ιδιωτικά σύνολα δεδομένων μεταφορτώνονται σε ένα ασφαλή διακομιστή πρωτοκόλλου μεταφοράς αρχείων SSH.

Ένα σύνολο τυπικών διαδικασιών λειτουργίας ακολουθείται πριν από την αποδοχή των δεδομένων από τους συνεργάτες ώστε να διασφαλιστεί ότι όλα τα δεδομένα είναι αναγνωρισμένα σύμφωνα με τους κανονισμούς περί Φορητότητας Και Λογοδοσίας Ασφαλιστικών Φορέων Υγείας (Health Insurance Portability and Accountability Act - HIPAA). Το G-DOC δεν αποθηκεύει ούτε διανέμει πληροφορίες ασθενούς

Μόλις ανακτηθούν, είτε από δημόσια αποθετήρια είτε μέσω ιδιωτικής μεταφοράς, τα δεδομένα omics προεπεξεργάζονται και οι κλινικές πληροφορίες αντιστοιχίζονται στις υπάρχουσες δομές δεδομένων ως προεργασία για τη φόρτωση εντός της βάσης δεδομένων G-DOC.

Στόχοι Της Πλατφόρμας

Οι κύριοι στόχοι από την χρήση της πλατφόρμας G-DOC είναι:

- Να παρέχει μια ισχυρή αλλά εύχρηστη διασύνδεση για τη χρήση εργαλείων βιοϊατρικής πληροφορικής από βασικούς, κλινικούς και μεταφραστικούς ερευνητές.
- Να παρέχει μια μεγάλη ποικιλία αναλυτικών εργαλείων και δυνατοτήτων, ενσωματωμένων προβολών για γονιδιωματικά χαρακτηριστικά και τρισδιάστατες δομές φαρμάκων.
- Να υποστηρίζει αποτελεσματικές συγκρίσεις ομάδων ασθενών, παρέχοντας ευέλικτα κλινικά κριτήρια περιήγησης για να καταστήσει δυνατή την επιλογή συγκεκριμένων ομάδων ασθενών και να διευκολύνει τη δημιουργία αναλυτικών αναφορών και ενημερωτικών εκθέσεων.
- Να εντοπιστούν νέα υποψήφια θεραπευτικά μόρια, χρησιμοποιώντας την ενσωματωμένη δομή και τις λεπτομερείς λειτουργίες αναζήτησης μοριακών ιδιοτήτων στις βιβλιοθήκες εσωτερικών χημικών ενώσεων.
- Να μοιράζονται οι ερευνητές με ασφάλεια τη γνώση με άλλους μέσω μιας ισχυρής σουίτας δυνατοτήτων συνεργασίας που λειτουργούν μέσα στο ασφαλές τους περιβάλλον.

Συμπερασματικά, πιστεύεται ότι η G-DOC θα βοηθήσει στη διευκόλυνση της ιατρικής, παρέχοντας τον προσδιορισμό των τάσεων και των μοντέλων σε ολοκληρωμένα σύνολα δεδομένων και, ως εκ τούτου, θα διευκολύνει τη χρήση πιο στοχευμένων θεραπειών για τον καρκίνο.

2.2.4. caTRIP (Cancer Translational Research Informatics Platform)

Το κέντρο Duke Comprehensive Cancer Center (DCCC), σε συνεργασία με την SemanticBits LLC, έχει αναπτύξει την Πλατφόρμα Πληροφορικής Ερευνητικής Μετάφρασης του Καρκίνου (Cancer Translational Research Informatics Platform - caTRIP) [48], μια μεταγραφική έρευνα για τη μελέτη για τη Cancer Biomedical Informatics Grid (caBIG) [49].

Εργαλεία και Εφαρμογές caBIG

Το caTRIP βασίζεται στη μεγάλη ποικιλία εφαρμογών caBIG ανοιχτού κώδικα και συμπεριλαμβάνει τα παρακάτω εργαλεία και εφαρμογές.

Tumor Registry

Είναι ένα κλινικό σύστημα που χρησιμοποιείται για τη συλλογή δεδομένων τελικού σημείου.

cancer Text Information Extraction System (caTIES)

<https://cabig.nci.nih.gov/tools/caties>

Πρόκειται για έναν εντοπιστή πόρων ιστών που λειτουργεί μέσω της εξαγωγής κλινικών πληροφοριών από εκθέσεις χειρουργικής παθολογίας ελεύθερου κειμένου, ενώ χρησιμοποιεί ελεγχόμενες ορολογίες για τη δημοσίευση δομών δεδομένων συμβατών με το caBIG.

cTissue CORE

<https://cabig.nci.nih.gov/tools/catissuecore>

Είναι ένα εργαλείο αποθεματοποίησης ιστών για απογραφή βιοστοιχείων, παρακολούθηση και βασικό σχολιασμό.

Cancer Annotation Engine (CAE)

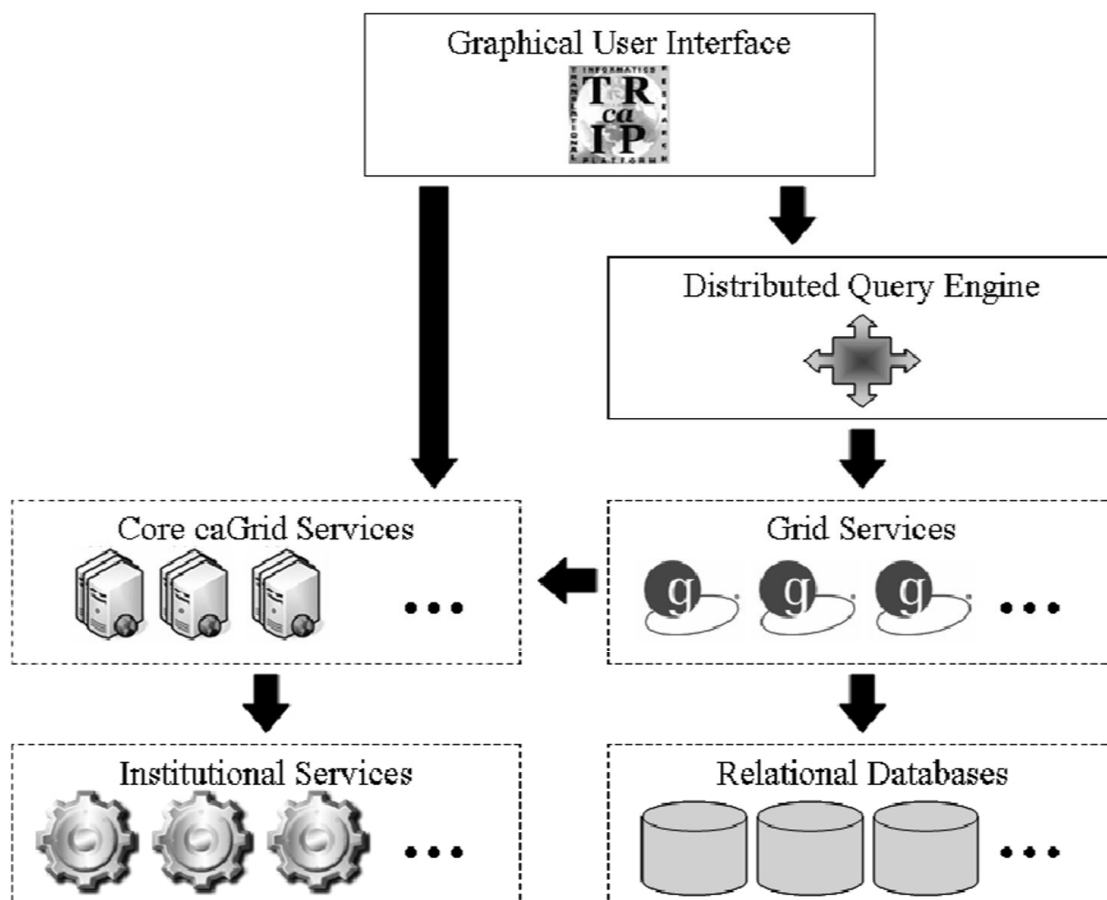
<https://cabig.nci.nih.gov/tools/cae>

Είναι ένα σύστημα αποθήκευσης και αναζήτησης σχολιασμών παθολογίας.

caIntegrator

<http://gforge.nci.nih.gov/projects/caintegrator>

Είναι ένα εργαλείο για την αποθήκευση ερωτημάτων και την ανάλυση δεδομένων μετάφρασης, συμπεριλαμβανομένων των δεδομένων SNP.



Εικόνα 11: Προβολή υψηλού επιπέδου της αρχιτεκτονικής caTRIP, συμπεριλαμβανομένων των Υπηρεσιών τομέα, του μηχανισμού κατανεμημένων ερωτημάτων και του γραφικού Διασύνδεση χρήστη. Domain Services, Distributed Query Engine, and Graphical User Interface [48].

Το caTRIP έχει σχεδιαστεί ως μια αρχιτεκτονική N-tier, με τρεις πρωτογενείς βαθμίδες (Εικόνα 11):

- τις Υπηρεσίες Τομέα (Domain Services)
- τον Κατανεμημένο Μηχανισμό Αναζήτησης (Distributed Query Engine) και
- το Γραφικό Περιβάλλον Χρήστη (Graphical User Interface)

Χρησιμοποιεί κυρίως την υποδομή caGrid για να εξασφαλίσει τη συμβατότητα με άλλα εργαλεία που αναπτύσσονται επί του παρόντος από το caBIG.

Η διεπαφή εφαρμογών σχεδιάστηκε έτσι ώστε οι χρήστες να μπορούν να δημιουργούν ερωτήματα χρησιμοποιώντας είτε το απλό περιβάλλον εργασίας μέσω αναπτυσσόμενων μενού είτε το Advanced Interface για πιο εξελιγμένες στρατηγικές αναζήτησης με τη χρήση drag-and-drop. Επιπλέον, η εφαρμογή αντιμετωπίζει τα θέματα ασφαλείας σχετικά με την πιστοποίηση ταυτότητας, και την εξουσιοδότηση.

2.2.5. Μοντέλα Ξενομοσχεύματων Όγκων Που Προέρχονται Από Τον Ασθενή (Patient-Derived Tumor Xenograft - PDX)

Τα μοντέλα ξενομοσχεύματων όγκων που προέρχονται από τον ασθενή [50] (Patient-Derived Tumor Xenograft - PDX) δεν είναι κάτι νέο. Μελέτες που διεξήχθησαν ήδη από τη δεκαετία του 1980, έχουν δείξει υψηλό βαθμό συσχετισμού μεταξύ κλινικής ανταπόκρισης σε κυτταροτοξικούς παράγοντες σε ενήλικες ασθενείς με καρκίνο του πνεύμονα και την ανταπόκριση στον ίδιο παράγοντα σε μοντέλα PDX, που έχουν αναπτυχθεί από αυτούς τους ασθενείς [51].

Τα τελευταία χρόνια, έχει υπάρξει ένα ανανεωμένο ενδιαφέρον για την ανάπτυξη μοντέλων PDX από τύπους διαφορετικών όγκων. Πράγματι, αυτά τα πρότυπα γίνονται το εργαλείο που προτιμούν προκλινικά τόσο η βιομηχανία όσο και πανεπιστημιακές ομάδες σε μια προσπάθεια βελτίωσης της διαδικασίας ανάπτυξης φαρμάκων.

Τα PDX μοντέλα ως επί το πλείστον διατηρούν τα κύρια ιστολογικά και γενετικά χαρακτηριστικά του όγκου του δότη και παραμένουν σταθερά κατά την επεξεργασία. Τα μοντέλα αυτά έχουν αποδειχθεί να λειτουργούν προγνωστικά των κλινικών αποτελεσμάτων και χρησιμοποιούνται για προκλινική αξιολόγηση φαρμάκων, ταυτοποίηση βιοδεικτών, βιολογικές μελέτες και εξατομικευμένη ιατρική.

Επί του παρόντος, υπάρχουν πολλές εκτενείς συλλογές PDX μοντέλων για διαφορετικές εφαρμογές μεταγραφικής έρευνας. Οι συλλογές αυτές αντιπροσωπεύουν ευρέως την

πολύπλοκη ετερογένεια των κλινικών όγκων και τη μοριακή ποικιλότητα των ανθρώπινων καρκίνων.

Καθώς οι μελέτες αποδεικνύουν την τεράστια ετερογένεια των ανθρώπινων καρκίνων και καθώς μεγάλες συλλογές PDX μοντέλων δεν είναι προσιτές στις επιμέρους ομάδες, η δημιουργία δικτύων συνεργασίας, είναι το κλειδί για την αντιμετώπιση της πρόκληση της ανάπτυξης φαρμάκων ακριβείας.

Μεθοδολογία Ανάπτυξης Των PDX Μοντέλων

Αν και κάθε επιμέρους ομάδα μελέτης έχει αναπτύξει συγκεκριμένη προσέγγιση, η μεθοδολογία και οι βασικές αρχές ανάπτυξης PDX μοντέλων, είναι κοινές. Εν συντομία, κομμάτια από μεταστατικούς στερεούς όγκους, που διατηρούνται ως δομές ιστών συλλέγονται μέσω χειρουργικής επέμβασης ή βιοψίας.

Μερικές μελέτες επίσης έχουν χρησιμοποιήσει υγρό στραγγισμένο από κακοήγη ασκίτη ή πλευριτική συλλογή. Στη συνέχεια, οι όγκοι εμφυτεύονται ως κομμάτια ή μονοκύτταροι αναστολείς, είτε μόνοι τους είτε σε ορισμένες μελέτες, επιχρισμένα με Matrigel ή αναμεμειγμένα με ανθρώπινους ινοβλάστους ή μεσεγχυματικά βλαστικά κύτταρα.

Η πιο κοινή τοποθεσία εμφύτευσης είναι στην ραχιαία περιοχή των ποντικών (υποδόρια εμφύτευση), αν και εμφύτευση στο ίδιο όργανο όπως του αρχικού όγκου μπορεί να είναι μια επιλογή (εμφύτευση ορθοτροπική, δηλαδή, παγκρέατος, στοματικής κοιλότητας, ωοθηκών, εγκεφάλου, κ.λπ.). Επιπλέον, ανεξάρτητα από την προέλευση του όγκου, διάφορες προσεγγίσεις έχουν εμφυτεύσει πρωτοπαθείς όγκους στην νεφρική κάψα, σε μια προσπάθεια να αυξήσουν τα ποσοστά επιτυχίας της μεταμόσχευσης.

Εφαρμογές των μοντέλων PDX στην έρευνα για τον καρκίνο

Έλεγχος Φαρμάκων Και Ανάπτυξη Βιοδεικτών

Είναι γνωστό ότι ένα από τα μείζονα θέματα στην ανάπτυξη φαρμάκων Ογκολογίας είναι το χαμηλό ποσοστό επιτυχίας νέων παραγόντων. Αυτό οφείλεται κατά ένα μέρος των στο ότι τα συμβατικά προκλινικά μοντέλα που χρησιμοποιείται για την παρακολούθηση των νέων παραγόντων για κλινική ανάπτυξη έχουν πολύ μικρή προγνωστική ικανότητα [52].

Επιπλέον, τα νέα φάρμακα, γενικότερα, δοκιμάζονται χωρίς τους κατάλληλους βιοδείκτες για την επιλογή των ασθενών και την παρακολούθηση της απόκρισης. Ως εκ τούτου, απαιτούνται στρατηγικές για να μειωθούν αυτά τα υψηλά ποσοστά λάθους.

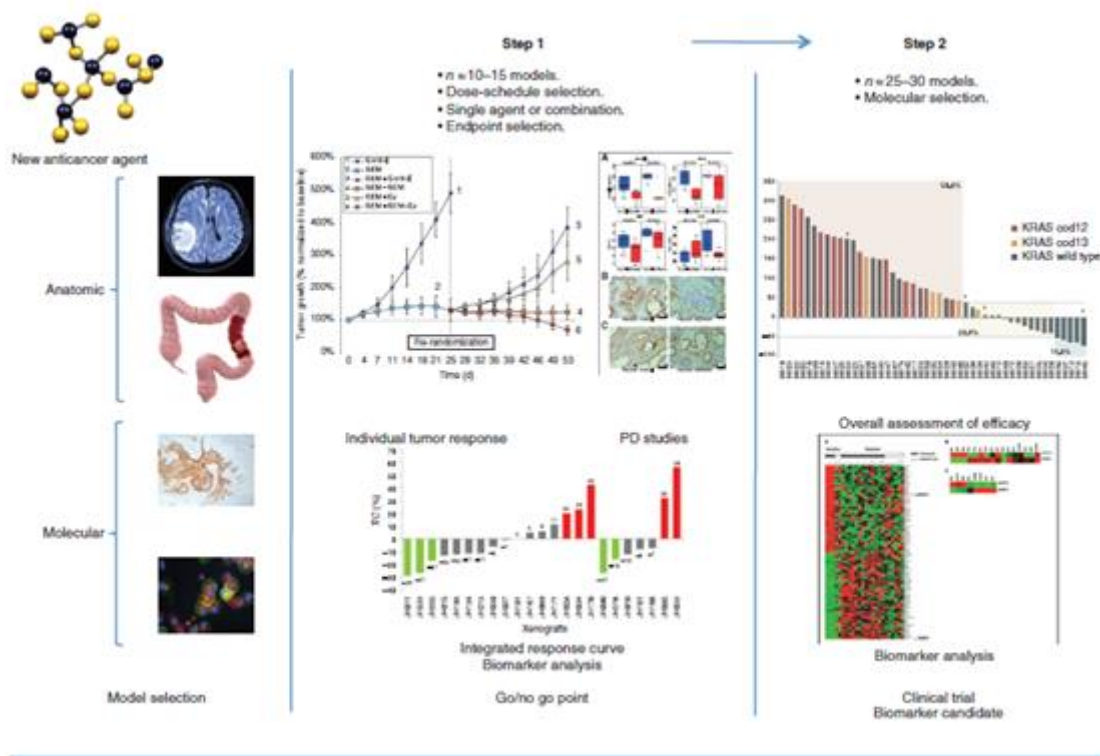
Επομένως, η διαθεσιμότητα προκλινικών μοντέλων με υψηλή προγνωστική αξία παρουσιάζει μεγάλο ενδιαφέρον, καθώς θα επιτρέψει τη διεξαγωγή προκλινικών μελετών φάσης II για την επιλογή πιθανών ενδείξεις για μεταγενέστερες κλινικές μελέτες.

Πλεονεκτήματα PDX μοντέλων

Το σκεπτικό για την υλοποίηση PDX μοντέλα για την επίτευξη αυτού του στόχου στηρίζεται στο γεγονός ότι :

- τα PDX μοντέλα παρέχουν τη δυνατότητα πρόβλεψη της κλινικής έκβασης
- τα PDX μοντέλα όχι μόνο βοηθούν στην ιεράρχηση πιθανών κλινικών ενδείξεων, αλλά μπορούν επίσης να διευκολύνουν τον εντοπισμό των πιθανών αποτελεσματικότερων βιοδεικτών φαρμάκων
- βοηθούν στην ανακάλυψη της αντίστασης βιοδεικτών που μπορεί να βοηθήσουν στο σχεδιασμό συνδυασμού κλινικών δοκιμών.
- Προκλινικές δοκιμές σε PDX μοντέλα μπορούν επίσης να διευκολύνουν την βελτιστοποίηση του σχεδιασμού των κλινικών δοκιμών.

Με βάση αυτά τα δεδομένα, τα μοντέλα PDX μπορούν να παίξουν σημαντικό ρόλο στις μελέτες ανταπόκρισης φαρμάκων, καθώς βοηθούν στην επιλογή πληθυσμών ασθενών που είναι πιο πιθανό να είναι ευαίσθητα σε ένα νέο παράγοντα, καθώς και να δώσουν προτεραιότητα στην ανάπτυξη νέων βιοδεικτών.



Εικόνα 12 : Προκλινική διαλογή και μελέτη βιοδεικτών σε μοντέλα PDX [50].

Η Εικόνα 12 απεικονίζει μια προτεινόμενη διαδρομή για την ένταξη των μοντέλων PDX στην ανάπτυξη νέων φαρμάκων και μερικά από τα βασικά στοιχεία της προκλινική μελέτη σε μοντέλα PDX.

Στο βήμα 1, ένας περιορισμένος αριθμός μοντέλων μπορεί να ελεγχθεί σε δόσεις του παράγοντα που είναι γνωστό ότι είναι αποτελεσματικές και φαρμακολογικά ενεργές σε προηγούμενες προκλινικές μελέτες.

Δεδομένα από το βήμα 1 μπορούν να χρησιμοποιηθούν, καθώς προχωρούμε στο βήμα 2, για τον επαναπροσδιορισμό και την επιλογή μοντέλου με βάση την μοριακή κατανόηση των μοντέλων απόκρισης. Στο βήμα 2, ένα μεγαλύτερο ρεπερτόριο των μοντέλων μπορεί να χρησιμοποιηθεί. Μετά το πέρας της μελέτης, μια απόφαση χρειάζεται να παρθεί για να προχωρήσει στην κλινική ανάπτυξη και να ιεράρχηση των βιοδεικτές που πρέπει να διερευνηθούν στην κλινική φάση.

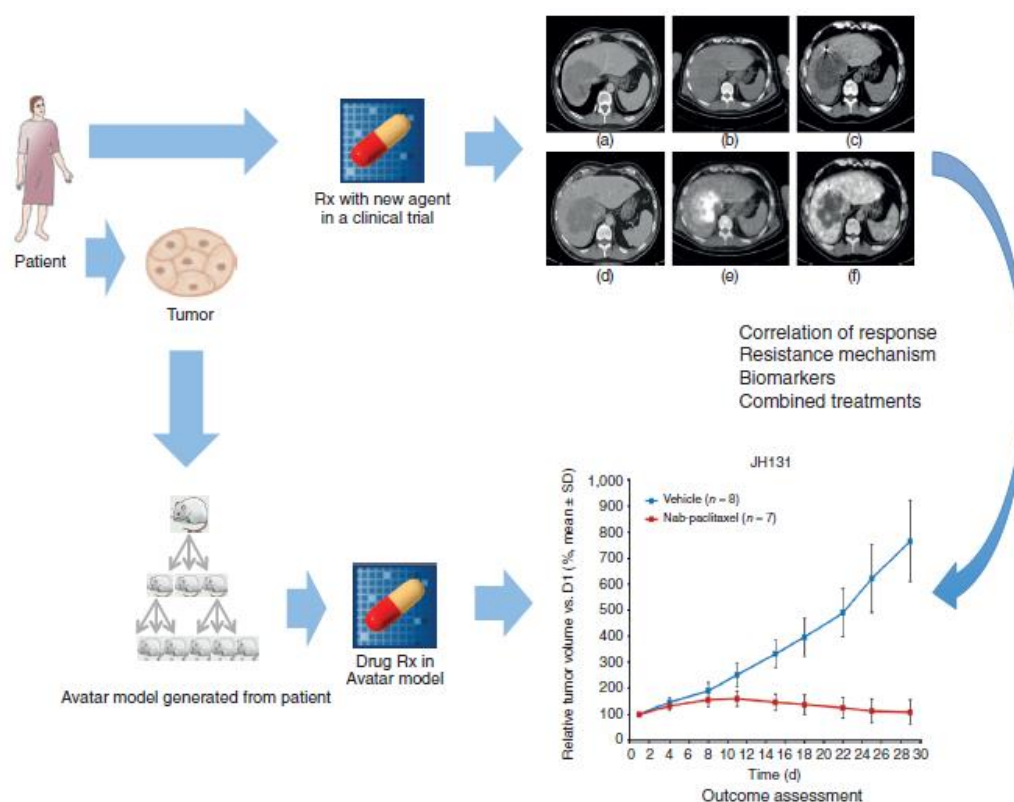
Συν-Κλινικές Δοκιμές

Όταν ένα φάρμακο εισέλθει σε κλινικές δοκιμές, υπάρχουν περιορισμένες ευκαιρίες να αναλύσει και να ενσωματώσει σε πραγματικό χρόνο πληροφορίες που μπορεί να είναι χρήσιμες για την ανάπτυξη αυτού του παράγοντα. Αυτό οφείλεται εν μέρει στην εγγενή φύση των κλινικών δοκιμών, στις οποίες οι ασθενείς υποβάλλονται σε θεραπεία με ένα φάρμακο ή σκεύασμα κάθε φορά, και ακολουθούνται με πολύ συγκεκριμένα κριτήρια αλλά και λόγω της έλλειψης επαρκών και εύκολα προσβάσιμων βιολογικών υλικών για πιο εμπεριστατωμένες μελέτες κλινικών παρατηρήσεων.

Για την επίλυση ορισμένων από αυτά τα ζητήματα, έχει προταθεί η έννοια των συν-κλινικών δοκιμών. Τα μοντέλα PDX έχουν χρησιμοποιηθεί σε αυτό το πλαίσιο σε παράλληλες μελέτες σε μοντέλα τρωκτικών και ασθενών και έχουν χρησιμεύσει στην αναγνώριση πιθανών βιοδεικτών [53].

Επιπλέον, τα μοντέλα PDX μπορούν να χρησιμοποιηθούν σε άλλη εφαρμογή της ιδέας της συν-κλινικής δοκιμής (Εικόνα 13). Σε αυτή την προσέγγιση, ένα εξατομικευμένο μοντέλο PDX, που αποκαλείται "Avatar", αναπτύσσεται από έναν ασθενή που παίρνει μέρος σε μια κλινική δοκιμή και αντιμετωπίζεται με τους ίδιους πειραματικούς παράγοντες για να μιμηθεί κλινική ανταπόκριση. Αυτή η στρατηγική επιτρέπει την ταυτόχρονη εκτίμηση της ανταπόκρισης του φαρμάκου στο μοντέλο του ασθενούς και του

ποντικού, παρέχοντας μια ενδιαφέρουσα πλατφόρμα για τη διερεύνηση βιοδεικτών ευαισθησίας και αντοχής.



Εικόνα 13: Συν-κλινική δοκιμή με μοντέλα PDX [50]

Εξατομικευμένη Ιατρική

Ο τομέας της ογκολογίας εξελίσσεται με ταχείς ρυθμούς από «μία θεραπεία για όλους» σε μια εποχή στην οποία οι όγκοι του ασθενούς εξετάζονται με μεγαλύτερη λεπτομέρεια για να επιλέγει η πιο κατάλληλη θεραπεία. Αν και η προσαρμοσμένη αυτή στρατηγική αντιπροσωπεύει μια σημαντική εξέλιξη στην μεταφραστική έρευνα του καρκίνου, περαιτέρω πρόοδοι απαιτούνται.

Μία τέτοια εξαιρετική πρόοδος συμπεριλαμβάνει την εξέταση των ασθενών για τους οποίους, παρά τις εκτεταμένες δοκιμές, δεν ανιχνεύονται βιοδείκτες αποτελεσματικότητας των φαρμάκων. Οι ασθενείς αυτοί δεν μπορούν να έχουν εξατομικευμένη θεραπεία. Η αντίθετη κατάσταση ισχύει επίσης: καθώς η σκιαγράφιση του καρκίνου εξελίσσεται και γίνεται πιο ολοκληρωμένη, πολλοί δυνητικοί όγκοι εντοπίζονται σε ορισμένους ασθενείς, προκαλώντας σύγχυση στην επιλογή της πιο κατάλληλης θεραπείας.

Είναι πιθανό ότι η συμβολή των μοντέλων PDX για εξατομικευμένη ιατρική κατά του καρκίνου θα αυξηθεί με την ένταξή τους σε προσεγγίσεις παγκόσμιας εξατομικευμένης

ιατρικής και όχι ως μια αυτόνομη πλατφόρμας. Η σημαντική επανάσταση στη γενετική του καρκίνου επιτρέπει, για πρώτη φορά, τη συγκέντρωση των τεράστια ποσοτήτων γονιδιωματικών πληροφοριών, περιλαμβανομένης της πλήρης εκτίμησης του γονιδιώματος του καρκίνου, την ενίσχυση της λήψη κλινικών αποφάσεων [54]. Επιπλέον, η ενσωμάτωση των παρατηρημένων αποκρίσεων σε ποντίκια με τις γενετικές πληροφορίες όγκου, θα οδηγήσει τελικά στην ανακάλυψη νέων δεικτών της αποτελεσματικότητας των φαρμάκων.

Για τους ασθενείς, των οποίων οι όγκοι δεν ενσωματώνονται σε ποντίκια ή εκείνους που απαιτεί πολύ χρόνο για να συσταθεί και να χαρακτηριστεί, μια εναλλακτική λύση από τη στρατηγική του Avatar θα μπορούσε να είναι η επιλογή θεραπείας που βασίζεται στην ανταπόκριση των φάρμακα από ένα παρόμοιο μοντέλο PDX. Οι βιοψίες σε πρωτοπαθείς όγκους ή μεταστάσεις θα χαρακτηρίζονται μοριακά και θα συγκρίνονται με τις διαθέσιμες συλλογές PDX με την ίδια παθολογία, για τις οποίες αποκρίσεις σε χημειοθεραπείες και στοχευμένους παράγοντες θα έχουν προηγουμένως προσδιοριστεί.

Περιορισμοί Των Μοντέλων PDX

Παρόλο που η ενσωμάτωση των μοντέλων PDX στην έρευνα για τον καρκίνο επιφέρει κάποιες βελτιώσεις, όπως έχει αναλυθεί παραπάνω, είναι σαφές ότι εξακολουθούν να έχουν σημαντικούς περιορισμούς που πρέπει να αντιμετωπιστούν για να βελτιώσουν τη χρήση τους στην μεταγραφική έρευνα για τον καρκίνο. Ορισμένοι από αυτούς τους περιορισμούς είναι τεχνικού χαρακτήρα και περιλαμβάνουν πολλά θέματα, όπως :

- (i) η εξέταση του πλέον κατάλληλου ιστού από τον οποίο δημιουργείται ένα μοντέλο PDX και η επεξεργασία αυτού του ιστού. Οι περισσότερες από τις δημοσιευμένες μελέτες βασίστηκαν σε χειρουργικά δείγματα που φυσικά παρέχουν μεγάλες ποσότητες ιστών. Αν και αυτή η προσέγγιση είναι χρήσιμη για τη δημιουργία συλλογών μοντέλων PDX, μικρότερα δείγματα, όπως βιοψίες όγκου, είναι καλύτερα προσαρμοσμένα στις εξατομικευμένες εφαρμογές φαρμάκων
- (ii) Είναι σημαντικό να καθοριστεί η καλύτερη στρατηγική για την ενσωμάτωση σε ποντίκια (υποδόρια έναντι ορθοτοπικής εμφύτευσης) για διαφορετικούς τύπους όγκων
- (iii) Η καθυστέρηση μεταξύ του χρόνου εμφύτευσης σε ποντίκια και των προγραμμάτων θεραπείας για ασθενείς είναι επίσης ένας περιοριστικός παράγοντας για εφαρμογές εξατομικευμένων ιατρικών εφαρμογών σε πραγματικό χρόνο. Διάρκει συνήθως 4 έως 8 μήνες για να αναπτυχθεί ένα μοντέλο PDX έτοιμο για μια προκλινική μελέτη, ένα χρονικό πλαίσιο που πολλοί ασθενείς δεν έχουν

(iv) Ένα άλλο πρόβλημα είναι η αποτυχία εμφύτευσης που είναι ακόμα υψηλή για ορισμένους τύπους όγκων με συγκεκριμένους φαινότυπους, όπως ο καρκίνος του ανθρώπινου μαστού με θετικό υποδοχέα ορμονών. Για στρατηγικές εξατομικευμένης ιατρικής, είναι υποχρεωτικό να βελτιωθούν τα ποσοστά λήψης όγκων σε αποδεκτό 60% έως 70%, γεγονός που αποτελεί μία από τις κύριες πτυχές που απαιτούν βελτίωση. Αυτό δεν είναι μόνο ένα πρόβλημα στην εξατομικευμένη ιατρική, καθώς οι περισσότεροι ασθενείς δεν έχουν συνδεδεμένο μοντέλο PDX, αλλά και σε μελέτες απεικόνισης φαρμάκων, καθώς οι σημερινές συλλογές μοντέλων PDX είναι επικλινές προς ορισμένους υποτύπους καρκίνου και δεν αντιπροσωπεύουν ευρέως την ετερογένεια της νόσου [50].

3. Σύγκριση Μεταφραστικών Συστημάτων Διαχείρισης Δεδομένων

Ακολουθεί μια περιγραφή και σύγκριση των δυνατοτήτων και της αρχιτεκτονικής των μεταφραστικών συστημάτων διαχείρισης δεδομένων (Πίνακας 1).

3.1. Περιεχόμενο Δεδομένων

3.1.1. Κλινικά Δεδομένα

Ο όρος κλινικά δεδομένα περιλαμβάνει ένα ευρύ φάσμα δεδομένων:

- Δημογραφικά στοιχεία (π.χ. Ηλικία, φύλο και εθνικότητα).
- Φυσικές εξετάσεις, αποτελέσματα εργαστηριακών εξετάσεων (από απλή εξέταση αίματος έως προχωρημένο βιομοριακό προσδιορισμό) και εικόνες ακτινολογίας.
- Ιστορικό ασθενούς, ιατρικές διαγνώσεις (χρησιμοποιώντας τυποποιημένες ορολογίες, συμπεριλαμβανομένων των κωδικών icd10),.
- Θεραπείες, αναφορές παθολογίας σε ελεύθερο κείμενο και κλινικές εκβάσεις (π.χ. Ποσοστά επιβίωσης)

Η συλλογή και διαχείριση τέτοιων εξαιρετικά πολύπλοκων δεδομένων, για κάθε ασθενή, είναι το ίδιο προκλητικό για τους βιοπληροφορικούς και τους ερευνητές. Επιπλέον, η προέλευση των πληροφοριών είναι διαφορετική. Τα κλινικά δεδομένα "περίθαλψης" αποθηκεύονται συχνά σε ηλεκτρονικά αρχεία υγείας (Electronic Health Records -EHR) ή CDW, ενώ τα κλινικά δεδομένα «έρευνας» συλλέγονται σε ηλεκτρονικά έντυπα αναφοράς περιπτώσεων ή σε συστήματα διαχείρισης κλινικών δεδομένων.

Συνεπώς, η διαχείριση των δεδομένων πρέπει να προσαρμοστεί (συμπεριλαμβανομένων των πτυχών μοντελοποίησης, της μορφοποίησης των δεδομένων και των διαδικασιών εκχύλισης, μετασχηματισμού, φόρτωσης (Extract, Transform, Load -ETL)).

Το BRISK και το cBio Cancer Genomics Portal επικεντρώνονται κυρίως στην εξερεύνηση των δεδομένων omics. Σε αυτές τις πλατφόρμες συλλέγονται και αποθηκεύονται κλινικά δεδομένα για να καταστεί δυνατή η κατηγοριοποίηση δειγμάτων και να πραγματοποιηθεί ειδική ανάλυση (π.χ. τύπος παθολογίας για ανάλυση GWAS σε BRISK και διαστήματα χωρίς ασθένεια για ανάλυση επιβίωσης στην cBio Cancer Genomics Portal).

Τα caTRIP, G-DOC, iCOD, iDASH και tranSMART επικεντρώνονται επίσης στην εξερεύνηση κλινικών δεδομένων. Το iDASH παρέχει πολλά εργαλεία ανάλυσης εικόνας

και διαχειρίζεται τα έγγραφα χρησιμοποιώντας τη λύση MIDAS (<http://midasplatform.org/>), ένα λογισμικό ανοιχτού κώδικα.

Στο tranSMART, τα φαινοτυπικά δεδομένα αποθηκεύονται χρησιμοποιώντας το μοντέλο δεδομένων i2b2 που αποτελείται από ένα σχηματισμό αστεριών που προέρχεται από την τιμή της οντότητας [55]. Το G-DOC και το iCOD χρησιμοποιούν τη δική τους μορφή βάσης δεδομένων.

3.1.2. Δεδομένα Omics

Όσον αφορά τα δεδομένα omics, κάθε πλατφόρμα υποστηρίζει ένα συγκεκριμένο σύνολο δεδομένων, ανάλογα με τους αρχικούς στόχους της πλατφόρμας και τις ανάγκες της ερευνητικής ομάδας που διευθύνει το έργο.

Το G-DOC υποστηρίζει τέσσερις τύπους δεδομένων omics: mRNA, microRNA, παραλλαγή αριθμού αντιγράφων και φασματομετρία μάζας μεταβολίτη.

Ως μεταφραστική ερευνητική πλατφόρμα που στοχεύει αρχικά στον τομέα ανάπτυξης φαρμάκων, το tranSMART υποστηρίζει πολλαπλά σύνολα omics χρήσιμα για φαρμακευτικές εταιρείες: προφίλ γονιδιακής έκφρασης, γονότυπους, πίνακες πρωτεϊνών ορού, μεταβολομικής και πρωτεϊνωματικά δεδομένα.

Η πλατφόρμα BRISK επικεντρώνεται στη μελέτη συσχέτισης GWAS: οι πολυμορφισμοί ενός νουκλεοτιδίου είναι τα μόνα υποστηριζόμενα δεδομένα omics.

Η cBio Cancer Genomics Portal είναι σε θέση να υποστηρίζει ένα ευρύ φάσμα δεδομένων omics που παράγεται από μελέτες μεγάλης κλίμακας: δεδομένα μετάλλαξης, αλλοιώσεις αριθμών αντιγράφων, αλλαγές σε έκφραση mRNA με βάση μικροσυστοιχίες και RNA αλληλούχισης, τιμές μεθυλίωσης DNA και επίπεδα πρωτεϊνών και φωσφοπρωτεϊνών .

Το iCOD περιλαμβάνει δεδομένα μοριακών omics όπως συγκριτικό γονιδιωματικό υβριδισμό και προφίλ γονιδιακής έκφρασης.

3.2. Υποστήριξη Διαλειτουργικότητας

Οι περισσότερες πλατφόρμες δεν παρέχουν υποστήριξη για τυποποιημένες ορολογίες και οντολογίες. Μόνο το iDASH και το caTRIP δημιουργήθηκαν για να υποστηρίξουν με τη σειρά τους μια περιορισμένη σειρά ορολογιών. Το TranSMART χειρίζεται σήμερα τη χρήση ορολογιών (π.χ. ICD10 ή LOINC). Οι απλές χαρτογραφήσεις μπορούν να διαχειριστούν μέσω των λειτουργιών i2b2 [56], [57] .

Ένα συνεργατικό και ασφαλές περιβάλλον παρέχεται επίσης από κάθε πλατφόρμα, εκτός από το iCOD (οι πληροφορίες δεν είναι διαθέσιμες). Αυτό επιτρέπει στους ερευνητές να μοιράζονται με ασφάλεια και να εργάζονται ταυτόχρονα σε αποθηκευμένα σύνολα δεδομένων, ενδεχομένως επιταχύνοντας την ερευνητική διαδικασία.

Παραδόξως, καμία από τις πλατφόρμες δεν μπορεί να ενσωματωθεί πλήρως σε ένα παγκόσμιο πλαίσιο: τα τυποποιημένα μορφότυπα όπως το CDISC ODM [58] ή το HL7 CDA [59] δεν αντιμετωπίζονται ως μορφή εισόδου και τα εξαγόμενα δεν είναι πάντοτε συμβατά με τους υπάρχοντες αγωγούς βιοπληροφορικής.

3.3. Λειτουργίες Υποστήριξης Ανάλυσης

3.3.1. Στατιστικά και Αναλυτικά Εργαλεία

Τα αναλυτικά χαρακτηριστικά που παρέχονται από το cBio Cancer Genomics Portal, το G-DOC, το iCOD, το iDASH και το tranSMART στηρίζονται κυρίως σε ένα εργαλείο τρίτου μέρους, όπως το στατιστικό λογισμικό R, ενσωματωμένο απευθείας στις πλατφόρμες. Παρέχουν έτοιμες αναλυτικές δέσμες ενεργειών που υλοποιούν τις κύριες δοκιμές και τα αναλυτικά εργαλεία που χρησιμοποιούνται από τους ερευνητές (συμπεριλαμβανομένων, ενδεικτικά, του t-test και της κύριας ανάλυσης συστατικών).

Τα σενάρια ανάλυσης είναι διαθέσιμα μέσω φιλικών προς το χρήστη γραφικών διεπαφών. Ως εκ τούτου, ο τελικός χρήστης δεν χρειάζεται προηγμένες γνώσεις υπολογιστών για να είναι σε θέση να αξιοποιήσει τα χαρακτηριστικά ανάλυσης. Για πολλές πτυχές του πλαισίου ανάλυσης omics, το tranSMART χρησιμοποιεί το Bioconductor [60] και το GenePattern [61] (ένα σύστημα που παρέχεται από το Broad Institutes), ενώ τα G-DOC και iCOD χρησιμοποιούν κυρίως εργαλεία που αναπτύσσονται εσωτερικά.

Τα εργαλεία ανάλυσης που χρησιμοποιούνται από το BRISK δεν αναφέρονται. Δεν συνεχίσουμε σε λεπτομερή περιγραφή των διαθέσιμων τύπων ανάλυσης, καθώς εξαρτώνται σε μεγάλο βαθμό από τα κλινικά και omics δεδομένα που διαχειρίζονται τα συστήματα και επομένως μπορεί να εξελίσσονται με γρήγορο ρυθμό.

3.3.2. Εργαλεία Απεικόνισης

Επιπρόσθετα από το πλαίσιο ανάλυσης, παρέχονται πολλαπλά εργαλεία απεικόνισης. Αυτός ο τύπος εργαλείων περιλαμβάνει λογισμικό τρίτου μέρους (π.χ. Integrative Genome Viewer) και εσωτερικά συστατικά στοιχεία, όπως το «OncoPrint» για την πλατφόρμα του cBio Cancer Genomics Portal και την «Web Information Service» στο BRISK.

Το G-DOC αξιοποιεί το εργαλείο ανοιχτού κώδικα Cytoscape [62] για την εμφάνιση του δικτύου αλληλεπίδρασης, του Java TreeView για χάρτες θερμότητας καθώς και διάφορα εσωτερικά στοιχεία. Το MetacoreTM (<http://thomsonreuters.com/metacore/>) από την Thomson Reuters διατίθεται στο tranSMART.

Αυτά τα είδη εργαλείων απεικόνισης είναι σημαντικά στοιχεία για μια μεταφραστική πλατφόρμα έρευνας, καθώς διευκολύνουν την ερμηνεία σύνθετων αποτελεσμάτων ανάλυσης. Εκτός από τα αναλυτικά εργαλεία, τα περισσότερα συστήματα εφαρμόζουν λειτουργίες εξαγωγής συμβατές με το λογισμικό SAS, R ή MS Excel, επιτρέποντας την προηγμένη ανάλυση από ειδικούς στατιστικούς.

Για να διευκολυνθεί περαιτέρω η ερμηνεία των αποτελεσμάτων, οι πλατφόρμες ενσωμάτωσαν πρόσθετες πληροφορίες πλαισίου που συγκεντρώθηκαν από δημόσιες πηγές. Για παράδειγμα, εργαλεία χαρτογράφησης ενσωματώθηκαν στο BRISK για πρόσβαση σε πληροφορίες συμφραζομένων από εξωτερικές βάσεις δεδομένων, όπως η Kyoto Encyclopedia of Genes and Genomes (KEGG) [63] . Παρόμοια εργαλεία υλοποιήθηκαν σε όλες τις πλατφόρμες.

3.4. Εργαλεία Ασφάλειας Και Προστασίας Της Ιδιωτικής Ζωής

3.4.1. Ιδιωτικότητα Δεδομένων

Η διαχείριση των δεδομένων κλινικής έρευνας σε επίπεδο ασθενούς είναι ένα πολύ ευαίσθητο ζήτημα, όσον αφορά θέματα δεοντολογίας και προστασίας της ιδιωτικής ζωής. Επομένως, είναι αναμενόμενο ότι όλες οι πλατφόρμες περιλαμβάνουν βασικές πολιτικές ασφαλείας.

Η μη εξουσιοδοτημένη πρόσβαση σε σύνολα δεδομένων εμποδίζεται από μια διαδικασία επαλήθευσης ταυτότητας χρήστη (προσωπικά διαπιστευτήρια για κάθε ερευνητή), σε συνδυασμό με συγκεκριμένα δικαιώματα πρόσβασης.

Το iDASH έχει σχεδιαστεί ειδικά για να αντιμετωπίσει την πρόκληση παραβίασης της ιδιωτικής ζωής. Το σύστημα προτείνει εργαλεία NLP για αποχαρακτηρισμό, καθώς και πολλά στατιστικά εργαλεία προσαρμοσμένα για τη διαφύλαξη της ιδιωτικότητας των ασθενών.

Στα άλλα συστήματα, τα βήματα αποχαρακτηρισμού πρέπει να εκτελούνται είτε πριν από την φόρτωση των δεδομένων στο σύστημα είτε, όταν χρειάζεται, κατά τη διαδικασία εξαγωγής.

3.5. Υποστήριξη Συστημάτων

3.5.1. Τεχνική Τεκμηρίωση

Οι περισσότερες από τις πλατφόρμες που περιγράφονται βρίσκονται ακόμη σε πρώιμα στάδια της ανάπτυξής τους και δεν διαθέτουν επαρκή τεκμηρίωση. Ωστόσο, οι ενεργές κοινότητες χρηστών παρέχουν πολύτιμη τεχνική βοήθεια τόσο για την εγκατάσταση όσο και για τη χρήση των συστημάτων. Οι πλατφόρμες έτοιμες προς χρήση (π.χ. G-DOC ή cBio Cancer Genomics Portal) παρέχουν μαθήματα ή εκπαίδευση για τους χρήστες τους.

3.5.2. Εγκατάσταση Και Διαχείριση Της Πλατφόρμας

Η επιλογή μιας πλατφόρμας εξαρτάται σε μεγάλο βαθμό από διάφορους παράγοντες: τους στόχους, τους πόρους και επίσης από τις πρακτικές πτυχές που καθοδηγούνται από δεοντολογικές και νομικές απαιτήσεις.

Ορισμένες πλατφόρμες παρέχουν αποθήκευση δεδομένων και ανάλυση 'ως υπηρεσία' για μεταφραστικούς ερευνητές. Για παράδειγμα, το G-DOC δεν είναι λογισμικό ανοιχτού κώδικα και ο κωδικός του δεν είναι διαθέσιμος στο κοινό. Ωστόσο, οι ερευνητικές ομάδες μπορούν να χρησιμοποιήσουν το σύστημα υπογράφοντας συμφωνία. Σε αυτή την περίπτωση, τα δεδομένα των ερευνητών πρέπει να μοιραστούν με το Πανεπιστήμιο του Georgetown και θα αποθηκευτούν στους διακομιστές της G-DOC.

Η πλατφόρμα είναι πλήρως λειτουργική και μπορεί να χρησιμοποιηθεί απευθείας από τον χρήστη μετά τη φόρτωση δεδομένων. Τα θέματα ανάπτυξης μειώνονται στο απολύτως ελάχιστο, καθώς η εγκατάσταση και η διαχείριση εκτελούνται από την ομάδα του G-DOC. Η ηλεκτρονική έκδοση της πύλης cBio Cancer Genomics Portal βασίζεται στην ίδια αρχή.

Το BRISK, το iDASH, η τοπική έκδοση του cBio Cancer Genomics Portal και το tranSMART αναπτύσσονται «on-site» και συνεπώς απαιτούν σημαντικές υποδομές (π.χ. εξυπηρετητές ιστού και Java, βάσεις δεδομένων) οι οποίες πιθανώς είναι ανεπαρκείς για τον μέσο μεταφραστικό κλινικό ή ερευνητή.

Σημαντικές προσπάθειες έγιναν για να διευκολυνθεί η εγκατάσταση και η χρήση τέτοιων περίπλοκων συστημάτων. Για παράδειγμα, το tranSMART παρέχει μια έκδοση έτοιμη προς χρήση μέσω της εικονικής συσκευής tranSMART [64].

Platform	tranSMART	cBioPortal	BRISK	iDASH	iCOD	G-DOC	caTRIP
Community							
Institution/project initiators	Johnson & Johnson, USA	Memorial Sloan-Kettering Cancer Center, NY, USA	University of British Columbia, Canada	iDASH team, USA	Tokyo Medical Dental University, Japan	Georgetown University, DC, USA	caBIG consortium, USA
Funding	Initially Johnson & Johnson funded; now public-private consortia	NIH and industry grants	Public and private consortium	NIH	Public	HSS	NIH
Reference	Szalma <i>et al.</i> 2010	Cerami <i>et al.</i> 2012	Tan <i>et al.</i> 2011	Ohno-Machado <i>et al.</i> 2012	Shimokawa <i>et al.</i> 2010	Madhavan <i>et al.</i> 2011	McConnell <i>et al.</i> 2008
PMID	20642836	22588877	21712248	22081224	21143802	21969811	19108734
Software availability	Open source	Free for academic use, commercial licenses available	Open source	Open source	Not distributed	Not distributed	Not distributed
Licensing	GPL v3	Free for academic use, commercial licenses available	GPL v2	BSD-like	Unknown	Proprietary	Open source
User mailing list or support	Yes	Yes	No	Yes	Unknown	Yes	No
URL	http://transmartfoundation.org	http://www.cbioportal.org/	http://genapha.icapture.ubc.ca/brisk/index.do	http://idash.ucsd.edu	http://omics.tmd.ac.jp/icod/	https://gdoc.georgetown.edu/gdoc/	https://cabig-stage.nci.nih.gov/community/tools/
Information content							
Clinical data							
Demographics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Outcomes	Yes	Yes	No	Yes	Yes	Yes	Yes
Biological results	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Images	No	No	No	Yes	Yes	No	Unknown
Structured clinical research data	Yes	Yes	Yes	Yes	Yes	Yes	Yes
'omics' data							
mRNA expression	Yes	Yes	No	Yes	Yes	Yes	Unknown
miRNA expression	No	Yes	No	Yes	No	Yes	Unknown
SNPs	Yes	Yes	Yes	Yes	No	Yes	Unknown
Copy number variations	Yes	Yes	Yes	Unknown	No	Yes	Unknown
DNA methylation	No	Yes	Yes	Unknown	No	Yes	Unknown
Protein/phosphoprotein expression	Yes	Yes	No	Unknown	No	Yes	Unknown
Structural rearrangements	No	Yes	No	Unknown	No	No	Unknown
Privacy management environment							
Anonymization/de-anonymization	No	No	No	Yes	No	No	Yes
Analysis supports							
Statistical framework	R	R, Matlab	<i>In-house components</i>	R	R	R	<i>In-house components</i>
Analytical features	GenePattern, Bioconductor Plink	<i>In-house components</i>	<i>In-house components</i>	<i>Privacy protecting analytical features</i>	<i>In-house components</i>	Bioconductor	<i>In-house components</i>
Visualization tools	Haploview, IGV and output of	<i>In-house components</i>	<i>In-house components</i>	<i>In-house components</i>	<i>In-house components</i>	Java TreeView, Jbrowse,	<i>In-house components</i>
Interoperability support							
Ontologies/standard terminologies	i2b2 ontology features	No	No	Yes	No	No	Yes
Collaborative environment	Yes	Yes	Yes	Yes	Unknown	Unknown	Yes
Secure environment	Yes	Yes	Yes	Yes	No	Yes	Yes
Support of multisite requests	No	No	No	Yes	Unknown	No	Yes
APIs/web services interface	JBoss	Yes	Unknown	Yes	Unknown	Unknown	Yes
System requirements					Unknown	PostgreSQL	Unknown
Operating system	Linux	Linux	Linux	Linux		Linux	
Database management system	Oracle 11/PostgreSQL	MySQL	MySQL	Multiple		Oracle	Oracle, MySQL
Database type							
Software dependencies	Centralized SOLR, R, i2b2	Centralized None	Local & external None	Per application/Web service MIDAS, varies per application	Centralized R	Centralized Adobe Flex framework, Java TreeView, Cytoscape, Jbrowse, Lucene	Distributed caGRID tools
Main programming language							
Server side	Java/GRails Tomcat/JBoss	Java, Javascript Tomcat	Java Tomcat	Multiple Multiple	Java Tomcat	GRails Unknown	Java Unknown
Client-side interface	Web browser	Web browser	Web browser	Web browser	Unknown	Web browser	Java
Platform support							
Installation procedures	Yes	Yes	Yes	Per application/web service	Not available	N/A	Not available
Configuration documentation	Yes	Yes	Yes	Wiki per application	Not available	N/A	Not available
User documentation	Yes	Yes	Yes	Wiki per application	Not available	yes	Not available

Πίνακας 1: Συγκριτικός πίνακας συστημάτων μεταφραστικών δεδομένων [26].

4. Απαιτήσεις των Μεταφραστικών Ερευνητικών Συστημάτων

4.1. Ιδιωτικότητα

Πολλά από τα συστήματα που αναλύθηκαν στην παρούσα εργασία δεν ήταν διαθέσιμα τοπικά. Αντιθέτως, χρησιμοποιούν μια αρχιτεκτονική πελάτη / διακομιστή, για την οποία ο διακομιστής δεν ελέγχεται από τον τελικό χρήστη.

Αυτό θα μπορούσε να οδηγήσει σε πιθανά προβλήματα όσον αφορά τους κανονισμούς περί απορρήτου δεδομένων. Η αποθήκευση και η διανομή κλινικών και omics δεδομένων είναι πολύ ευαίσθητα θέματα στον τομέα της μεταφραστικής έρευνας, καθώς εγείρουν θέματα ηθικής και ιδιωτικότητας.

Οι ομάδες κλινικών ερευνών συχνά δεσμεύονται από αυστηρούς κανόνες προστασίας της ιδιωτικής ζωής (π.χ. νόμος για την ασφάλεια φορητότητας και λογοδοσίας στις Ηνωμένες Πολιτείες - HIPAA). Η χρήση απομακρυσμένων λύσεων, συμπεριλαμβανομένων των λύσεων σε νέφος (cloud) εξακολουθεί να είναι υπό συζήτηση.

Παρόλο που προσφέρει τεράστιες δυνατότητες, ειδικά όσον αφορά την αποδοτικότητα, την υπολογιστική ισχύ και την ευελιξία, μια πλατφόρμα που βασίζεται σε νέφος (cloud) δεν θα πρέπει να χρησιμοποιείται μέχρι να αξιολογηθούν προσεκτικά τα ζητήματα απορρήτου και ανταλλαγής δεδομένων.

Οι ίδιοι κανόνες θα πρέπει να ισχύουν για ιδιωτικές ή δημόσιες απομακρυσμένες πλατφόρμες. Αναμφισβήτητα, οι λύσεις ενσωμάτωσης δεν χρειάζεται μόνο να επιτρέπουν την αποθήκευση και την εξερεύνηση των δεδομένων, αλλά και να διαθέτουν αυτές τις λειτουργίες σε ελεγχόμενο περιβάλλον συμβατό με κυβερνητικούς κανονισμούς και ορθές πρακτικές.

Κατά συνέπεια, τα ζητήματα ιδιωτικού απορρήτου συχνά απαιτούν την ικανότητα εγκατάστασης μεταφραστικών ερευνητικών συστημάτων εντός των ορίων του θεσμικού οργάνου.

4.2. Διαλειτουργικότητα Και Πρότυπα

Μεταξύ των άλλων στόχων, οι μεταφραστικές πλατφόρμες ισχυρίζονται ότι:

- Καταστούν δυνατή την αποτελεσματική ανταλλαγή δεδομένων, για παράδειγμα, αυξάνουν τη ποσότητα των διαθέσιμων δεδομένων για τις σπάνιες ασθένειες και
- Διευκολύνουν την ενσωμάτωση δεδομένων.

Και οι δύο στόχοι απαιτούν διαλειτουργικότητα και συμβατότητα των δεδομένων.

Ωστόσο, καμία από τα συστήματα που περιεγράφηκαν σε αυτήν την εργασία δεν παρείχε τη δυνατότητα άμεσης αλληλεπίδρασης με ηλεκτρονικά ή προσωπικά αρχεία υγείας. Τα συστήματα δεν είναι σε θέση να εισάγουν δεδομένα σε διεθνή πρότυπα ανταλλαγής όπως το HL7 CDA [59] ή το CDISC ODM [58].

Η αποτελεσματική ενσωμάτωση δεδομένων απαιτεί επίσης ότι τα μεταφραστικά ερευνητικά συστήματα μπορούν να συνδυαστούν με υπάρχουσες διαδικασίες συλλογής δεδομένων εντός των ιδρυμάτων. Ωστόσο, οι ενότητες πλατφόρμας δεν σχεδιάζονται συχνά για να μοιράζονται εύκολα με άλλες πλατφόρμες.

Η αυξημένη ανάπτυξη προσαρμοσίμων και επαναχρησιμοποιήσιμων εργαλείων και βιβλιοθηκών θα αποτελέσει μεγάλη βοήθεια για τον τομέα. Ομοίως, η υιοθέτηση των API δεν έχει φτάσει ακόμα στο πλήρες δυναμικό της και θα επέτρεπε τέτοιες προσαρμοσμένες συνδέσεις, για παράδειγμα, επιτρέποντας απλή πρόσβαση στα δεδομένα ή εύκολη ρύθμιση αναλυτικών ροών εργασίας.

Επιπλέον, η χρήση τυποποιημένων ορολογιών και οντολογιών είναι ένα άλλο βασικό στοιχείο της διαλειτουργικότητας και της ανταλλαγής δεδομένων. Παραδόξως, οι εξεταζόμενες πλατφόρμες προσφέρουν περιορισμένη ικανότητα χειρισμού τέτοιων χαρακτηριστικών.

Επιπλέον, η χρήση διεθνών τυποποιημένων ορολογιών (π.χ. ICD-10, SNOMED CT) θα επέτρεπε τη χρήση των ιδιοτήτων υποσυγκέντρωσης και των σημασιολογικών ζευγών στις ορολογίες, επιτρέποντας έτσι τη συλλογιστική του υπολογιστή.

4.3. Ετερογένεια Δεδομένων

Η συλλογή δεδομένων κλινικών ερευνών περιορίζεται με ένα πρωτόκολλο κατάλληλης συγκρισιμότητας και τα δεδομένα omics αποθηκεύονται και παράγονται σε τυποποιημένη μορφή (π.χ. MIAME), ενώ τα δεδομένα κλινικής περίθαλψης συλλέγονται όποτε χρειάζεται φροντίδα των ασθενών. Τα περισσότερα από τα συστήματα που παρουσιάστηκαν χρησιμοποιούν μια αναπαράσταση βασισμένη σε μοντέλο δεδομένων κλινικής έρευνας.

Κατά συνέπεια, η ενσωμάτωση των δεδομένων φροντίδας στα συστήματα απαιτεί μετασχηματισμό όχι μόνο της μορφής αλλά και του μοντέλου. Τα δεδομένα της κλινικής

έρευνας και φροντίδας είναι δύσκολο να ενσωματωθούν, εν μέρει λόγω του ετερογενούς χαρακτήρα τους και λόγω των χρονικών προβλημάτων.

Πρέπει να σημειωθεί ότι καμία από τις πλατφόρμες που περιλαμβάνονται σε αυτή την εργασία δεν ήταν σε θέση να χειριστεί σύνθετα δεδομένα του χρόνου (π.χ. διαστήματα φαρμακευτικής αγωγής), αφήνοντας περιθώρια βελτίωσης.

4.4. Εγκατάσταση Και Συντήρηση

Παρόλο που τα περισσότερα από τα συστήματα που περιγράφονται είναι ώριμα έργα και ήδη παρέχουν μεταφραστικούς ερευνητές με προηγμένες δυνατότητες, πρέπει να καταστήσουμε σαφές ότι η εγκατάσταση και η συντήρηση αυτών των συστημάτων απαιτεί τη βοήθεια μιας ομάδας τεχνικών, υποδομής υπολογιστών και δικτύου.

Τα περισσότερα συστήματα ενσωματώνουν έτοιμα προς χρήση εργαλεία ανάλυσης και απεικόνισης. Η προσθήκη νέων χαρακτηριστικών, που διευκολύνεται από την αρθρωτή αρχιτεκτονική των συστημάτων, μπορεί να απαιτεί ανάπτυξη από στατιστικούς ή ειδικούς υπολογιστών.

Ωστόσο, για σημαντικά χαρακτηριστικά, τα συστήματα συχνά αξιοποιούν ευρέως υιοθετημένες λύσεις (π.χ. το στατιστικό λογισμικό R και το μοντέλο i2b2 CDW). Αυτό παρέχει το πλεονέκτημα μιας ενεργής κοινότητας προγραμματιστών και χρηστών στον τομέα της μεταφραστικής έρευνας καθώς και συνεισφορές από άλλους τομείς.

5. Προκλήσεις της Μεταφραστικής Έρευνας

Ο αναγκαίος ρόλος της μεταφραστικής έρευνας στην εξασφάλιση της ταχείας εξέλιξης της βασικής επιστημονικής γνώσης για τα οφέλη του ασθενούς έχει τονιστεί για περισσότερο από 3 δεκαετίες [2]. Κατά τη διάρκεια αυτής της περιόδου, εντοπίστηκαν επίσης τα εμπόδια στην μεταφραστική έρευνα. Τα κυριότερα θέματα που συζητήθηκαν επανειλημμένα αναλύονται στη συνέχεια.

5.1. Έλλειψη Σαφούς Μοντέλου Μετάφρασης

Αν και η βελτίωση της υγείας είναι ο κοινός στόχος της βασικής επιστήμης και της βιοϊατρικής έρευνας, ο στόχος αυτός δεν έχει επιτευχθεί όσο συχνά επιθυμείτε ή χρειάζεται. Το NIH χρηματοδότησε περίπου 15 δισεκατομμύρια δολάρια βασικής επιστημονικής έρευνας το 2009, αλλά το ποσοστό μετάφρασης αυτής στην κλινική πρακτική είναι χαμηλό και αργό.

Μια καλά σχεδιασμένη μελέτη [65] διαπίστωσε ότι λιγότερο από το 25% των πολύ υποσχόμενων βιοϊατρικών ανακαλύψεων οδήγησε σε δημοσιευμένη τυχαιοποιημένη κλινική δοκιμή και λιγότερο από 10% διαπιστώθηκε στην κλινική πρακτική εντός 20 ετών.

Επιπλέον, ακόμη και όταν αποδείχθηκε ότι έχουν θετικά οφέλη για την υγεία, οι θεραπείες δεν χορηγούνται καθολικά στην κλινική πρακτική. Πολύ γνωστά παραδείγματα περιλαμβάνουν τη χορήγηση β-αναστολέων μετά από έμφραγμα του μυοκαρδίου (MI) ή διαλογή λιπιδίων και παράγοντες μείωσης της χοληστερόλης σε ασθένεια στεφανιαίας αρτηρίας [66], [67].

Ο λόγος για όλα τα παραπάνω περιστρέφεται γύρω από δύσκολες ή άγνωστες διαδικασίες μετακίνησης της βασικής επιστημονικής έρευνας σε μεταφραστικό τρόπο (π.χ. μη βασική έρευνα που είναι εντελώς διαφορετική έρευνα).

Δεν υπάρχει σαφές μοντέλο μετάφρασης για την περιγραφή αυτών των διαδικασιών που μεταφέρουν τη βασική επιστήμη σε μετάφραση. Η έλλειψη επαρκούς πλαισίου εμποδίζει τον προσδιορισμό των βημάτων κατά τα οποία η έρευνα και η γνώση «χάνονται στη μετάφραση» και συνεπώς δεν επιτυγχάνουν οφέλη για τη δημόσια υγεία.

Η κατανόηση και ο καθορισμός της συνέχειας από το εργαστήριο στη δημόσια υγεία είναι ζωτικής σημασίας για τη βελτίωση της επιτυχημένης μετάφρασης της έρευνας. Χωρίς ένα

πλαίσιο για τον εντοπισμό βημάτων στη διαδικασία μετάφρασης όπου η γνώση χάνεται, είναι δύσκολο ή αδύνατο να γεφυρωθούν αυτά τα κενά.

Οι συνέπειες της κακής μετάφρασης και της τήρησης των κατευθυντήριων γραμμών είναι διπλάσιες. Πρώτον, δισεκατομμύρια δολάρια έρευνας «χάνονται στη μετάφραση» κατά μήκος του μονοπατιού από το πάγκο στο κομοδίνο όπως συζητήθηκε προηγουμένως. Το πιο σημαντικό είναι ότι ενδέχεται να χαθούν δυνητικά σημαντικά οφέλη για τη δημόσια υγεία και σε αυτή τη διαδικασία [1].

5.2. Πολιτιστικές Διαφορές

Οι πολιτισμικές διαφορές μεταξύ των δύο ομάδων ερευνητών οφείλονται σε μεγάλο βαθμό στην έλλειψη επικοινωνίας, στις διαφορές στην εκπαίδευση και την κατάρτιση και σε διαφορετικούς στόχους και μηχανισμούς ανταμοιβής.

5.3. Έλλειψη Πόρων

Τα προβλήματα πόρων, όσον αφορά το εργατικό δυναμικό και την υποδομή, περιλαμβάνουν την έλλειψη εκπαιδευμένου διεπιστημονικού προσωπικού για την υποστήριξη των ερευνών σε όλο τον κύκλο της μεταφραστικής έρευνας, την προστασία του χρόνου για την έρευνα, ιδίως για τους κλινικούς ιατρούς, καθώς και την πρόσβαση σε κοινούς πόρους.

5.4. Ρυθμιστικά Θέματα

Τα ρυθμιστικά θέματα αποτελούν πρόβλημα τόσο για τους βασικούς επιστήμονες όσο και τους κλινικούς γιατρούς και περιλαμβάνουν την ηθική δεοντολογία που εμπλέκεται στον τομέα της ανθρώπινης έρευνας, των κανονισμών για την αποθήκευση ιστών και τη μεταφορά υλικών, των δικαιωμάτων διανοητικής ιδιοκτησίας, της τοξικολογίας και των κανονισμών παρασκευής, εγκρίσεις της FDA (Food and Drug Administration) και της MHRA (Medicines and Healthcare Regulatory Agency), τη χορηγία σπουδών, καθώς και την παρακολούθηση των δοκιμών και των δεδομένων [68].

Αυτά γίνονται ακόμη πιο πολύπλοκα με την επέκταση στους τομείς των κυτταρικών και γονιδιακών θεραπειών και της μηχανικής ιστών.

Μαζί, αυτά τα θέματα συμβάλλουν σε διάφορα σημεία ελέγχου μεταξύ των φάσεων της μεταφραστικής έρευνας, συμπεριλαμβανομένης της "κοιλιάδας του θανάτου" που υπάρχει μεταξύ της προκλινικής έρευνας και των κλινικών δοκιμών [69].

5.5. Έλλειψη Προτύπων Ρόλων

Η έλλειψη προτύπων ρόλων και μεντόρων σε αυτόν τον τομέα έχει εντοπιστεί ως σημαντικό εμπόδιο από τους ερευνητές που προσπαθούν να συμμετάσχουν στην μεταφραστική έρευνα. Η καθοδήγηση αναγνωρίζεται όλο και περισσότερο ως ανεκτίμητη για την ανάπτυξη ερευνητών πρώιμης σταδιοδρομίας σε ανεξάρτητους ερευνητές και τα περισσότερα ιδρύματα διαθέτουν προγράμματα καθοδήγησης.

Στο Ηνωμένο Βασίλειο, η Ακαδημία Ιατρικών Επιστημών διαθέτει ένα αποτελεσματικό εθνικό πρόγραμμα καθοδήγησης και εξωστρέφειας, το οποίο υποστηρίζεται από το Υπουργείο Υγείας και το NIHR, το οποίο λειτουργεί από το 2002 [70].

5.6. Τα Μεταφραστικά Δεδομένα Έρευνας Είναι Μεγάλα Δεδομένα

Με τις σημερινές τεχνολογικές εξελίξεις, είναι τώρα εφικτό να δημιουργηθεί ένα μητρώο ασθενών που ενεργοποιείται από το πλήρες αρχείο υγείας του ασθενούς (Complete Health Record), συνδυάζοντας τα γονιδιωματικά δεδομένα, τον τεράστιο όγκο δεδομένων που διατίθενται μέσω των Συνδεδεμένων Ιατρικών Συσκευών (Connected Medical Devices), τον συνεχώς αυξανόμενο αριθμό προσωπικών συσκευών φυσικής κατάστασης μέσω του διαδικτύου των πραγμάτων – Internet of Things), μαζί με δεδομένα από το ηλεκτρονικό ιατρικό αρχείο (Electronic Medical Record - EMR) και άλλους τύπους διαθέσιμων δεδομένων ασθενών. Οι γνώσεις που αποκτήθηκαν από το πλήρες αρχείο υγείας στοχεύουν να επιλύσουν ένα ευρύ φάσμα προβλημάτων υγείας (Εικόνα 14).

Εδώ, το πρόβλημα της βελτίωσης της φροντίδας των ασθενών μέσω της μεταφραστικής επιστήμης διασταυρώνεται με το πρόβλημα των μεγάλων δεδομένων (Big Data). Το πρόβλημα έγκειται στη διαθεσιμότητα και ανάλυση, στην εύρεση και κατανόηση των σημάτων, των τάσεων, των αιτιωδών σχέσεων, και στη συνέχεια η μετατροπή αυτών σε χρήσιμες πληροφορίες.

Και ενώ έχουμε τώρα τη δυνατότητα να επεξεργαζόμαστε και να αποθηκεύουμε τεράστια ποσά δεδομένων, καθώς και τα πρότυπα και τις τεχνολογίες που επιτρέπουν την ενσωμάτωση αυτών των δεδομένων, εξακολουθούν να υπάρχουν προβλήματα στα μεγάλα δεδομένα γενικά και πιο συγκεκριμένα στην εφαρμογή τεχνικών αυτών στη μεταφραστική ιατρική, που εμποδίζουν το όραμα της εξεύρεσης αυτών των κρίσιμων στοιχείων.

Οι επικριτές των μεγάλων δεδομένων χρησιμοποιούν την συχνά αναφερόμενη φράση:

«Το πρόβλημα με τα Μεγάλα Δεδομένα (Big Data) αφορά τα σωστά δεδομένα που έχουν δοθεί στους σωστά άτομα την σωστή στιγμή» [71].

- **Τα Σωστά Δεδομένα**

Δεν αμφισβητεί μόνο την αξιοπιστία των δεδομένων, αλλά την εφαρμογή των δεδομένων και την προσβασιμότητα των δεδομένων.

- **Σωστά Άτομα**

Αφορά όχι μόνο τους ίδιους τους αναλυτές αλλά επίσης την τεχνολογία ανάλυσης που μπορεί να χρησιμοποιήσει ο αναλυτής για να δώσει πληροφορίες.

- **Σωστή Στιγμή**

Ένα από τα βασικά στοιχεία των μεγάλων δεδομένων είναι η παραδοχή ότι τα δεδομένα αλλάζουν ταχύτατα.



Εικόνα 14: Το πλήρες αρχείο υγείας - Περιλαμβάνει όλες τις πληροφορίες υγείας για έναν ασθενή, όχι μόνο αυτό που βρίσκεται στο ηλεκτρονικό ιατρικό αρχείο (EMR) [72].

Συμπερασματικά θα λέγαμε ότι, υπάρχουν τα πρότυπα για την ενσωμάτωση τόσο των δεδομένων υγείας που εστιάζονται στον πάροχο όσο και στον ασθενή. Η τεχνολογία υπάρχει και αναπτύσσεται πιο εξελιγμένη τεχνολογία, που επιτρέπει στους ερευνητές πιο πολύπλοκη επεξεργασία των δεδομένων. Υπάρχουν πλατφόρμες για τη συλλογή δεδομένων εστιασμένων στον ασθενή και την αποθήκευση αυτών παράλληλα με τα δεδομένα EMR.

Αυτό που χρειάζεται για την επίλυση των προβλημάτων με τα μεγάλα δεδομένα στη μεταφραστική ιατρική επιστήμη είναι ένας συνδυασμός προσεγγίσεων: εφαρμογές που εστιάζουν στην έρευνα σε συνδυασμό με την πρόσβαση σε πλήρη αρχεία υγείας ασθενών, επιτρέποντας σε μεγαλύτερο αριθμό ασθενών να συνεισφέρουν και επιτρέποντας στους ερευνητές να έχουν πρόσβαση σε πλήρη δεδομένα υγείας, τα οποία θα οδηγήσουν στην κατανόηση και την ικανότητα ανίχνευσης των μοτίβων, παρά στην απλή υποθετική εξέταση, οδηγώντας τελικά σε πιο έγκαιρες και στενευμένες θεραπείες [72].

5.7. Ηθική Της Έρευνας Για Την Μεταφραστική Επιστήμη

Η μεταφραστική έρευνα, που ορίζεται από το Εθνικό Ινστιτούτο Υγείας (NIH) ως μια *«διαδικασία εφαρμογής ιδεών και ανακαλύψεων που δημιουργούνται μέσω βασικής επιστημονικής έρευνας για τη θεραπεία ή την πρόληψη ανθρώπινων ασθενειών»*, θεωρείται ως ένας από τους τρόπους επιτυχημένης ενσωμάτωσης των αποτελεσμάτων της έρευνας στην ιατρική περίθαλψη

Σήμερα, η μεταφραστική έρευνα επικεντρώνεται όχι μόνο στην εφαρμογή βασικών επιστημονικών ανακαλύψεων στην συνήθη κλινική πρακτική (bed to bench), αλλά και στην εξασφάλιση της υιοθέτησης τεκμηριωμένων θεραπειών σε κοινοτικούς χώρους (bench to curbside translation)

Η μεταφορά νέων ανακαλύψεων τόσο στην κλινική πρακτική όσο και στην ευρύτερη κοινότητα απαιτεί την εξάρτηση από διεπιστημονικές μεταφραστικές ομάδες που περιλαμβάνουν ερευνητές με διαφορετικούς τομείς εμπειρογνωμοσύνης, εκπροσώπους συστημάτων υγειονομικής περίθαλψης και κοινοτικών οργανώσεων και ασθενών.

Ωστόσο, η συμμετοχή νέων ενδιαφερομένων στην έρευνα απαιτεί την επανεξέταση ή την επέκταση της έννοιας της ηθικής στη μεταφραστική επιστήμη. Για να γίνει αυτό, οι συγγραφείς του [73], διεξήγαγαν 2 διαδικτυακές, τροποποιημένες ομάδες Delphi με 63 εμπειρογνώμονες. Ο στόχος ήταν να αξιολογήσαν και να συζητήσουν 9 δεοντολογικές αρχές, που παραδοσιακά χρησιμοποιούνται συνήθως στην έρευνα που ασχολείται με την κοινότητα (Community-Engaged Research – CenR), όσον αφορά τη σημασία και τη σκοπιμότητά τους για χρήση στη μεταφραστική έρευνα.

Παρόλο που αυτές οι επαναλαμβανόμενες και διευρυμένες ηθικές αρχές είναι σημαντικές για την έρευνα CEnR, δεν είναι γνωστό πόσο σημαντικό είναι για την μεταφραστική επιστήμη γενικά. Ως προκαταρκτικό βήμα στην ανάλυση της μεταφραστικής ερευνητικής δεοντολογίας, οι συγγραφείς του [73] ,εξετάσανε τη γνώμη των εμπειρογνομόνων σχετικά

με την εφαρμογή των ηθικών αρχών, που ασκούνται συνήθως στην CEnR) στη μεταφραστική επιστήμη.

Λόγω του εξερευνητικού χαρακτήρα αυτού του έργου, προσέλαβαν ένα ευρύ φάσμα εμπειρογνομόνων, που ασχολούνται με διάφορα είδη μεταφραστικών ερευνητικών προγραμμάτων και ζήτησαν σκόπιμα από τους συμμετέχοντες να σκεφτούν έναν ευρύ ορισμό της μεταφραστικής επιστήμης.

Η μέθοδος καταλληλότητας RAND / UCLA χρησιμοποιήθηκε για την ανάλυση των δεδομένων και για τον καθορισμό της συμφωνίας και της διαφωνίας μεταξύ των συμμετεχόντων εμπειρογνομόνων.

Τα ευρήματά παρουσιάζουν νέες δεοντολογικές προκλήσεις που συνδέονται με την ενεργό εμπλοκή των ενδιαφερομένων στην έρευνα και προτείνουν ότι οι μεταφραστικοί ερευνητές μπορεί να έχουν ανησυχίες σχετικά με τη σκοπιμότητα της ενεργού συμμετοχής των ενδιαφερομένων. Ένας σχετικά υψηλός βαθμός συναίνεσης μεταξύ των συμμετεχόντων σχετικά με τη σημασία αυτών των νέων δεοντολογικών αρχών βρίσκεται σε αντίθεση με τις αποκλίνουσες απόψεις τους όσον αφορά την εφικτότητά τους.

Ωστόσο, είναι σημαντικό να σημειωθεί ότι οι διαφορές μεταξύ των ομάδων, σχετικά με τη σκοπιμότητα ορισμένων από αυτές τις αρχές και ένα σημαντικό αριθμό αρχών που θεωρούνται σταθερά αβέβαιες, δεν σημαίνει ότι η μετάφραση έρευνα θα πρέπει να τις αγνοεί.

Αντίθετα, ο υψηλός επιπολασμός των αβέβαιων αξιολογήσεων σκοπιμότητας μπορεί να υποδηλώνει ότι οι συμμετέχοντες δεν ήταν σίγουροι για τους καλύτερους τρόπους επίτευξης και μέτρησης αυτών των σκοπών, διότι αυτές οι αρχές είναι σχετικά νέες και σπάνια έχουν συζητηθεί στο πλαίσιο της μεταφραστικής ερευνητικής ηθικής στη βιβλιογραφία.

Επιπλέον, επειδή οι μεταφραστικές έρευνες είναι ένα διαφοροποιημένο πεδίο, ορισμένοι από τους συμμετέχοντες ήταν αβέβαιοι σχετικά με τη δυνατότητα εφαρμογής αυτών των νέων δεοντολογικών αρχών σε όλα τα μεταφραστικά έργα, ανεξάρτητα από το γεγονός ότι η μεταφραστική επιστήμη ως επιστημονική πειθαρχία έχει γίνει περισσότερο εμπλεκόμενη.

Η πιο ισχυρή διαπίστωση της μελέτης είναι η σημασία της εμπιστοσύνης ως ηθικής αρχής για τη μεταφραστική επιστήμη. Η διαφάνεια και η αξιοπιστία στις σχέσεις μεταξύ των ακαδημαϊκών και των κοινοτικών εταίρων είναι ένα εξέχον θέμα στην βιβλιογραφία της

CEnR και της μεταφραστικής επιστήμης, που προσπαθούν να εξηγήσουν τα χαμηλά επίπεδα δημόσιας εμπιστοσύνης στην επιστήμη, ιδιαίτερα μεταξύ των μειονοτικών και των υποσιτισμένων κοινοτήτων.

Η έλλειψη εμπιστοσύνης στην επιστήμη και η υποεκπροσώπηση των εθνοτικών μειονοτήτων στην έρευνα στον τομέα της υγείας έχουν αποδοθεί εν μέρει στην ηθική παράβαση στην έρευνα που αφορά συμμετέχοντες μειονοτήτων. Η συμμετοχή των κοινοτικών εταίρων σε όλες τις πτυχές της έρευνας θεωρείται ότι διευκολύνει την οικοδόμηση εμπιστοσύνης στην επιστήμη μεταξύ της κοινότητας .

Ως εκ τούτου, τα αποτελέσματα της μελέτης δείχνουν ότι η μεταφραστική έρευνα θα μπορούσε να γίνει ένας κρίσιμος χώρος για τη μείωση της δυσπιστίας στην επιστήμη και τους επιστήμονες. Αυτό το εύρημα είναι ιδιαίτερα ενδιαφέρον, επειδή οι ειδικοί κατέληξαν σε συμφωνία σχετικά με τη σημασία και τη σκοπιμότητα της χρήσης της εμπιστοσύνης ως ηθικής αρχής για την καθοδήγηση της μετάφρασης.

Τα ευρήματα δείχνουν ότι εκτός από την προστασία των ατόμων, τα σύγχρονα μοντέλα μεταφραστικών επιστημών πρέπει να λογοδοτούν για τα συμφέροντα και έχουν δεοντολογικές υποχρεώσεις προς τα μέλη της ερευνητικής ομάδας και της κοινότητας γενικότερα.

Ανεξάρτητα από αυτούς τους περιορισμούς, τα συμπεράσματά έχουν ορισμένες πιθανές επιπτώσεις στην έρευνα για την υγεία, στην επισκόπηση και την εποπτεία της έρευνας.

Οι μεταφραστικοί ερευνητές, οι χρηματοδότες και οι χορηγοί τους και τα όργανα αξιολόγησης των θεσμικών οργάνων θα πρέπει να γνωρίζουν το βαθμό στον οποίο οι δεοντολογικές αρχές, που ξεπερνούν τις παραδοσιακές αρχές Belmont και εκτείνονται πέρα από τα θέματα σπουδών, επαναπροσδιορίζουν τη δεοντολογία της μεταφραστικής έρευνας.

Το πεδίο της μεταφραστικής έρευνας μπορεί να ωφεληθεί από την ανάπτυξη στρατηγικών που μπορούν να αξιολογήσουν την εφαρμογή νέων ηθικών διαστάσεων, όπως η εμπιστοσύνη, η πολιτιστική καταλληλότητα και η ισονομία.

Οι μεταφραστικοί ερευνητές μπορούν να επωφεληθούν από περαιτέρω κατάρτιση και ευκαιρίες διαβούλευσης για να μάθουν περισσότερα για δεοντολογικές αρχές όπως η εμπιστοσύνη, η ισονομία, ο προσανατολισμός δράσης και οι κοινότητες που χρησιμοποιούνται ευρέως στην CEnR.

Συγκεκριμένα, οι στρατηγικές διαβούλευσης και η προσανατολισμένη προς τη διαδικασία εκπαίδευση θα μπορούσαν να είναι χρήσιμες για να υποστηρίξουν τους μεταφραστικούς ερευνητές στις προσπάθειές τους να επεκτείνουν την προστασία των μελών της ερευνητικής ομάδας και της κοινότητας γενικότερα.

6. Συμπεράσματα

6.1. Συμπεράσματα

Η απώλεια των επιστημονικών ανακαλύψεων κατά τη διάρκεια του μετασχηματισμού μεταφράσεων της βιοϊατρικής έρευνας αποτελεί σημαντικό πρόβλημα της δημόσιας υγείας.

Παρά τη χρηματοδότηση πολλών δισεκατομμυρίων δολαρίων έρευνας και τη σημαντική εργαστηριακή παραγωγικότητα, μόνο ένα κλάσμα των πολλά υποσχόμενων βασικών ανακαλύψεων επιστημών έχει ως αποτέλεσμα εφαρμοσμένες κλινικές πρακτικές και οφέλη για την υγεία [65].

Αν θέλουμε να εκμεταλλευτούμε στο έπακρο τα βασικά και κλινικά δεδομένα προς όφελος της δημόσιας υγείας, θα πρέπει να γεφυρωθεί το χάσμα κατά το μονοπάτι του μοντέλου του μεταφραστικού συνεχές. Θα πρέπει να κατανοηθούν οι προκλήσεις που αντιμετωπίζει η μεταφραστική έρευνα και να προταθούν λύσεις.

6.2. Προοπτικές Και Επεκτάσεις

Ένα σαφές πλαίσιο για τον ορισμό και την στερεοποίηση των εννοιών και της ορολογίας της μετάφρασης θα ήταν χρήσιμο, οι γιατροί και άλλοι ερευνητές δεν ακολουθούν τη συνεχώς μεταβαλλόμενη βιβλιογραφία για τη μετάφραση.

Η αυξημένη ανάπτυξη προσαρμόσιμων και επαναχρησιμοποιήσιμων εργαλείων και βιβλιοθηκών θα αποτελέσει μεγάλη βοήθεια για τον τομέα.

Ομοίως, η υιοθέτηση των API δεν έχει φτάσει ακόμα στο πλήρες δυναμικό της και θα επέτρεπε τέτοιες προσαρμόσιμες συνδέσεις, για παράδειγμα, επιτρέποντας απλή πρόσβαση στα δεδομένα ή εύκολη ρύθμιση αναλυτικών ροών εργασίας.

Επιπλέον, η χρήση τυποποιημένων ορολογιών και οντολογιών είναι ένα άλλο βασικό στοιχείο της διαλειτουργικότητας και της ανταλλαγής δεδομένων. Παραδόξως, οι εξεταζόμενες πλατφόρμες προσφέρουν περιορισμένη ικανότητα χειρισμού τέτοιων χαρακτηριστικών.

Επιπλέον, η χρήση διεθνών τυποποιημένων ορολογιών (π.χ. ICD-10, SNOMED CT) θα επέτρεπε τη χρήση των ιδιοτήτων υποσυγκέντρωσης και των σημασιολογικών ζευγών στις ορολογίες, επιτρέποντας έτσι τη συλλογιστική του υπολογιστή [26].

6.2.1. Κατάρτιση Και Καθοδήγηση

Ενώ τα ρυθμιστικά θέματα αποτελούν σημαντικό παράγοντα για την επιτυχία των μεταφραστικών προγραμμάτων, είναι επίσης επιτακτική η εκπαίδευση και ανάπτυξη ενός αγωγού βασικών και κλινικών επιστημόνων ώστε να είναι προετοιμασμένοι για την αντιμετώπιση των συνεχιζόμενων προκλήσεων της μεταφραστικής έρευνας.

Κατά τη διάρκεια της εκπαίδευσής τους, οι κλινικοί επιστήμονες θα επωφεληθούν από την εμπλοκή τους στην εργαστηριακή έρευνα ενώ οι βασικοί επιστήμονες θα πρέπει να έχουν έκθεση στο κλινικό περιβάλλον υπό την καθοδήγηση έμπειρων κλινικών ιατρών.

Και οι δύο ομάδες θα πρέπει επίσης να παρακολουθήσουν κατάρτιση σε μεθοδολογία έρευνας συμπεριλαμβανομένου του σχεδιασμού και της συμπεριφοράς των κλινικών δοκιμών και των ιατρικών στατιστικών.

Επιπλέον, στην κλινική ιατρική εξακολουθούν να υπάρχουν ορισμένοι μύθοι και παρανοήσεις σχετικά με το ρόλο της έρευνας σε προγράμματα κατάρτισης όπως:

- (i) Η έρευνα δεν έχει θέση στο πρόγραμμα εκπαίδευσης ενός κλινικού ιατρού.
- (ii) Δεν υπάρχει χώρος για βασική έρευνα στο πρόγραμμα κλινικής κατάρτισης.
- (iii) Κάθε χειρουργός πρέπει να κάνει κάποια βασική έρευνα.
- iv) Οι κλινικοί γιατροί θα πρέπει να συμμετέχουν μόνο στην κλινική έρευνα.

Αυτά πρέπει να αντιμετωπιστούν σε επίπεδο σχολών, θεσμών και κρατών.

6.2.2. Σταδιοδρομία

Μια άλλη πρόκληση είναι η εξέλιξη της σταδιοδρομίας και η προώθηση και η απόκτηση μόνιμης θέσης στην ακαδημαϊκή ιατρική ή σε πανεπιστημιακό επιστημονικό τμήμα. Τα τρέχοντα κριτήρια για την προώθηση εξακολουθούν να βασίζονται σε μεμονωμένα ερευνητικά αποτελέσματα, όπως δημοσιεύσεις με υψηλό αντίκτυπο, επιχορηγήσεις και προσκεκλημένες διαλέξεις.

Οι ερευνητές που εμπλέκονται στην μεταφραστική έρευνα ενδέχεται να μην είναι σε θέση να προσκομίσουν τα απαιτούμενα αποδεικτικά στοιχεία για τη συμβολή τους, όπως ένα επαρκές αρχείο δημοσίευσης, δεδομένου ότι τα μεταφραστικά έργα γενικά χρειάζονται περισσότερο χρόνο για να ολοκληρωθούν. Είναι επίσης πιθανό να είναι μέρος μιας αρκετά μεγάλης διεπιστημονικής ομάδας και ο ρόλος τους στις δημοσιεύσεις και στις επιχορηγήσεις μπορεί να μην είναι εμφανής

6.2.3. Εξορθολογισμός Των Ρυθμιστικών Θεμάτων

Εκτός από τις πολιτισμικές διαφορές, έχουν καταβληθεί προσπάθειες για τον εξορθολογισμό των εγκρίσεων της έρευνας στον τομέα της υγείας και τη μείωση του ρυθμιστικού φόρτου.

Το NIHR ανέπτυξε το 2008 το Ολοκληρωμένο Εφαρμοσμένο Ερευνητικό Σύστημα (Integrated Research Application System (IRAS)), ένα ολοκληρωμένο ηλεκτρονικό σύστημα που περιλαμβάνει τις πληροφορίες που απαιτούνται για την έγκριση από διάφορα όργανα ανασκόπησης.

Στις ΗΠΑ, το Εθνικό Κέντρο Προώθησης Μεταφραστικών Επιστημών (National Center for Advancing Translational Sciences NCATS, <http://www.ncats.nih.gov/>) ιδρύθηκε το 2012

με αποστολή

"να δράσει καταλυτικά στην παραγωγή καινοτόμων μεθόδων και τεχνολογιών που θα ενισχύσουν την ανάπτυξη, τη δοκιμή και την εφαρμογή των διαγνωστικών και θεραπευτικών μεθόδων σε ένα ευρύ φάσμα ανθρώπινων ασθενειών και καταστάσεων. "

Αυτό θα περιλάμβανε την αντιμετώπιση των σημείων συμφόρησης όπως ρυθμιστικά ζητήματα στη διαδικασία μεταφραστικής έρευνας για να καταστεί " πιο αποτελεσματική, λιγότερο δαπανηρή και λιγότερο επικίνδυνη " [74].

6.3. Μελλοντικές Ανάγκες

Η άνοδος της ιατρικής ακρίβειας και η διαθεσιμότητα μοριακών αναλύσεων υψηλής απόδοσης στο πλαίσιο της κλινικής περίθαλψης έχουν αυξήσει την ανάγκη των ερευνητών μεταφραστικών δεδομένων για επαρκή εργαλεία για τη διαχείριση και διερεύνηση αυτών των δεδομένων.

Βιβλιογραφία

- [1] B. Drolet και N. Lorenzi, «Translational research: understanding the continuum from bench to bedside».
- [2] S. Wolf, «Editorial: the real gap between bench and bedside,» *New Engl J Med* 1974;290:802–3..
- [3] D. Rubio, E. Schoenbaum, E. Lee και S. Linda, «Defining Translational Research: Implications for Training,» *Academic Medicine*, 85(3), pp. 470-475.
- [4] D. Dougherty και P. Conway, «The “3T’s” road map to transform US health care: the “how” of high-quality care.,» *JAMA* ;, τόμ. 2319, αρ. 21, p. 299, 2008.
- [5] J. Westfall, J. Mold και L. Fagnan, «Practice-based research—“blue highways” on the NIH roadmap.,» *JAMA* 2007;297:403–6..
- [6] S. Woolf, «The meaning of translational research and why it matters,» *JAMA* 2008;299:211–3.
- [7] K. Fiscella, N. Bennett και P. Szilagyi, «Nomenclature in translational research,» *JAMA* 2008;299. 2148-a–9.
- [8] P. Michael, E. W. Klee, W. Mequon και M. Rochester, «Zebrafish approaches enhance the translational research tackle box,» *Translational Research*, Φεβρουάριος 2014.
- [9] A. Burkhardt και A. Zlotnik, «Translating Translational Research: Mouse Models of Human Disease,» *Cellular and Molecular Immunology*, τόμ. 10, αρ. 5, p. 373–374., 2013.
- [10] «National Institute of Justice,» Office of Justice Programs, [Ηλεκτρονικό]. Available: <https://www.nij.gov/journals/268/Pages/criminology.aspx>. [Πρόσβαση Δεκέμβριος 2017].
- [11] «Center for criminology and public policy research,» Florida State University, [Ηλεκτρονικό]. Available: <http://criminology.fsu.edu/center-for-criminology-public-policy-research/current-projects/translational-criminology-research-and-public-policy/> . [Πρόσβαση Δεκέμβριος 2017].
- [12] M. Beeghly, « Translational research on early language development: Current challenges and future directions,» *Development and Psychopathology*, 18(3), 737-757., 2006.
- [13] «Rice, ML.; Warren, SF. Developmental language disorders: From phenotypes to etiologies. Mahwah, NJ: Erlbaum; 2005a.».
- [14] «Rice ML, Warren SF, Betz SK. Language symptoms of developmental language disorders: An overview of autism, Down syndrome, fragile X, specific language impairment, and Williams syndrome. *Applied Psycholinguistics*. 2005; 26:7–27».
- [15] «Chapman RS, Hesketh L. Language, cognition, and short-term memory in individuals with Down».
- [16] «Cicchetti, D.; Beeghly, M. An organizational approach to the study of Down syndrome: Contributions».
- [17] S. Baron-Cohen, « Mindblindness: An essay on autism and theory of mind. Cambridge, MA: MIT Press;».
- [18] «Dawson, G.; Castelloe, P. Autism. In: Walker, DE.; Roberts, MC., editors. Handbook

-] of child clinical».
- [19 «Karmiloff-Smith A, Thomas M. What can developmental disorders tell us about
] the».
- [20 «Reilly J, Klima ES, Bellugi U. Once more with feeling: Affect and language in atypical
] populations.».
- [21 «Leonard, L. Children with specific language impairment. Cambridge, MA: MIT Press;
] 1998.».
- [22 «Feldman HM. Evaluation and management of language and speech disorders in
] preschool children.».
- [23 «Reilly J, Losh M, Bellugi U, Wulfeck B. “Frog, where are you?” Narratives in children
] with specific».
- [24 E. Bates, «Explaining and interpreting deficits in language development across clinical
] groups: Where do we go from here?,» *Brain and Language*, τόμ. 88, αρ. PubMed: 14965545, p. 248–253, 2003.
- [25 «Tager-Flusberg H. Strategies for conducting research on language in autism. Journal
] of Autism and Developmental Disorders. 2004b; 34:75–80. [PubMed: 15098960]».
- [26 V. Canuel, B. Rance, P. Avillach, P. Degoulet και A. Burgun, «Translational research
] platforms integrating clinical and omics data: a review of publicly available solutions,» *Briefings in Bioinformatics*, 16(2), pp. 280-290, 2014.
- [27 The transSMART Foundation, «TransSMART Project,» [Ηλεκτρονικό]. Available:
] <https://wiki.transmartfoundation.org/display/transmartwiki/transSMART+Project+wiki+Home>. [Πρόσβαση Δεκέμβριος 2017].
- [28 S. Murphy, G. Weber και M. Mendis, «Serving the enterprise and beyond with
] informatics for integrating biology and the bedside (i2b2)». *J Am Med Inform Assoc* 2010;17(2): 124–30.
- [29 R. Bellazzi, M. Masseroli, S. Murphy, A. Shabo και R. P. , «Clinical Bioinformatics:
] challenges and opportunities,» *BMC bioinformatics*, Τόμ. 13 από 213(14), S1., 2012.
- [30 L. Ohno-Machado, V. Bafna, A. Boxwala και e. al, «iDASH: integrating data for
] analysis, anonymization and sharing,» 2012.
- [31 A. Tan, B. Tripp και D. Daley, «BRISK—research-oriented storage kit for biology-
] related data,» *Bioinformatics*, 27(17), pp.2422-2425., 2011.
- [32 T. Van Rossum και e. al, «SLIMS—a user-friendly sample operations and inventory,»
] 2010.
- [33 H. J. Lowe, T. A. Ferris, P. M. Hernandez και S. C. Weber, «STRIDE—An integrated
] standards-based translational research informatics platform,» *In AMIA Annual Symposium Proceedings (. American Medical Informatics Association*, τόμ. 2009, p. 391, 2009.
- [34 Z. EA., « US biomedical research: Basic, translational, and clinical sciences,» *JAMA*
] 2005, Sep 21;294(11):1352-8..
- [35 «Brandt CA, Morse R, Matthews K, Sun K, Deshpande AM, Gadagkar R, et al.
] Metadata-Driven creation of data marts from an eav-modeled clinical research database. *Int J Med Inform* 2002, Nov 12;65(3):225-41.».

- [36 «Lyman JA, Scully K, Tropello S, Boyd J, Dalton J, Pelletier S, Egyhazy C. Mapping from a clinical data warehouse to the HL7 reference information model. AMIA Annu Symp Proc 2003:920».
- [37 «<http://www.oracle-base.com/articles/8i/VirtualPrivateDatabases.php>,»
] [Ηλεκτρονικό].
- [38 E. Cerami, J. Gao, U. Dogrusoz, B. Gross, S. Sumer, B. Aksoy, A. Jacobsen, C. Byrne, M. Heuer, E. Larsson, Y. Antipin, B. Reva, A. Goldberg και C. a. S. N. Sander, «The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data:», *Cancer Discovery*, 2(5), pp. 401-404., 2012.
- [39 J. Zhang, J. Baran, A. Cros, J. Guberman, S. Haider, J. Hsu και e. a. , «International Cancer Genome Consortium Data Portal—a one-stop shop for cancer genomics data», *Database (Oxford)*, τόμ. bar026, 2011.
- [40 J. Robinson, H. Thorvaldsdottir, W. Winckler, M. Guttman, E. Lander, G. Getz και e. a. , «Integrative genomics viewer», *Nat Biotechnol*, τόμ. 29, p. 24–6, 2011.
- [41 J. Sanborn, S. Benz, B. Craft, C. Szeto, K. Kober, L. Meyer και e. a. , «The UCSC Cancer Genomics Browser: update 2011», *Nucleic Acids Res*, τόμ. 39, p. D951–9., 2011.
- [42 G. Gundem, C. Perez-Llamas, A. Jene-Sanz, A. Kedzierska, A. Islam, J. Deu-Pons και et al., «IntOGen: integration and data mining of multidimensional oncogenomic data», *Nat Methods*, τόμ. 7, pp. 92-3, 2010.
- [43 K. Shimokawa, K. Mogushi, S. Shoji, A. Hiraishi, K. Ido και H. a. T. H. Mizushima, «iCOD : an integrated clinical omics database based on the systems-pathology view of disease. BMC Genomics, 11(Suppl 4), p.S19.,» 2010.
- [44 S. Madhavan, Y. Gusev, M. Harris, D. Tanenbaum, R. Gauba, K. Bhuvaneshwar, A. Shinohara, K. Rosso, L. Carabet, L. Song, R. Riggins, S. W. Y. B. S. Dakshanamurthy και R. a. W. L. Clarke, «G-DOC: A Systems Medicine Platform fo», 2011.
- [45 S. Madhavan, Y. Gusev, M. Harris και e. al., «G-CODE: enabling systems medicine through innovative informatics.» *Genome Biol* 2011;12(Suppl 1):P38.
- [46 T. Barrett, T. Suzek, D. Troup, S. Wilhite, W.-C. Ngau και e. a. , «NCBI GEO: mining millions of expression profiles—database and tools», *Nucleic Acids* , τόμ. Res 33, p. D562–D566, 2005.
- [47 H. Parkinson, U. Sarkans, M. Shojatalab, N. Abeygunawardena και e. a. , «ArrayExpress—a public repository for microarray gene expression data at the EBI», *Nucleic Acids* , τόμ. 33, p. D553–D555, 2005.
- [48 P. McConnell, D. R. R. Chilukuri, R. Pietrobon, K. Johnson και R. a. C. A. Annechiarico, «The cancer translational research informatics platform. BMC Medical Informatics and Decision Making, 8(1).», 2008.
- [49 D. Fenstermacher, C. Street, T. McSherry, V. Nayak, C. Overby και M. Feldman, «The Cancer Biomedical Informatics Grid (caBIG™)», σε *IEEE Eng Med Biol Soc* 1:743-6., 2005.
- [50 «Hidalgo, M., Amant, F., Biankin, A. V., Budinská, E., Byrne, A. T., Caldas, C., ... & Roman-Roman, S. (2014). Patient-derived xenograft models: an emerging platform for translational cancer research. *Cancer discovery*, 4(9), 998-1013.».
- [51 N. H. H. H. K. H. K. D. A. H. Fiebig HH, «Development of three human small cell lung cancer models in nude mice. *Recent Results Cancer Res* 1985;97:77–86».

- [52 D. S. Z. D. R. L. V. J. S. S. e. a. Johnson JI, «Relationships between drug activity in NCI preclinical in vitro and in vivo models and early clinical trials. *Br J Cancer* 2001;84:1424–31.».
- [53 T. A. U. M. A. M. M. W. V. M. e. a. Garrido-Laguna I, «Integrated preclinical and clinical development of mTOR inhibitors in pancreatic cancer. *Br J Cancer* 2010;103:649–55.».
- [54 H. M. Valencia A, «Getting personalized cancer genome analysis into the clinic: the challenges in bioinformatics. *Genome Med* 2012;4:61.».
- [55 «Murphy SN, Weber G, Mendis M, et al. Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2). *J Am Med Inform Assoc* 2010;17(2): 124–30».
- [56 S. Wu, H. Liu, D. Li και e. a. , « Unified Medical Language System term occurrences in clinical notes: a large-scale corpus analysis,» *JAmMed Inform Assoc*, p. 149–156, 2012;19(e1).
- [57 D. Segagni, V. Tibollo, A. Dagliati και e. al., «The ONCO-I2b2 project: integrating biobank information and clinical data to support translational research in oncology,» *Stud Health Technol Inform*, τόμ. 169, p. 887–91, 2011.
- [58 ODM, «Certification & Archiving and Interchange of Metadata,» 11 2017. [Ηλεκτρονικό]. Available: <http://www.cdisc.org/odm>.
- [59 R. Dolin, L. Alschuler, C. Beebe και e. al, «The HL7 clinical document architecture,» *J Am Med Inform Assoc*, τόμ. 8(6), p. 552–69, 2001.
- [60 J. Zhang, V. Carey και R. Gentleman, « An extensible application for assembling annotation for genomic data,» *Bioinformatics*, τόμ. 19(1), pp. 155-6, 2003.
- [61 M. Reich, T. Liefeld, J. Gould και e. a. , «GenePattern 2.0. *Nat Genet*,» τόμ. 38(5), p. 500–1, 2006.
- [62 «Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;13(11):2498–504.».
- [63 «Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 2000;28(1):27–30.».
- [64 «TranSMART Wiki. tranSMART 1.1 Virtual Appliance,» [Ηλεκτρονικό]. Available: <https://wiki.transmartfoundation.org/display/TSMTGPL/>. [Πρόσβαση Δεκέμβριος 2017].
- [65 D. Contopoulos-Ioannidis, E. Ntzani και J. Ioannidis, «Translation of highly promising basic science research into clinical applications,» *Am J Med* , τόμ. 114, p. 477–84, 2003.
- [66 N. Freemantle, J. Cleland, P. Young, J. Mason και J. Harrison, «beta Blockade after myocardial infarction: systematic review and meta regression analysis,» *BMJ*, τόμ. 318, pp. 1730-7, 1999.
- [67 P. McBride, H. Schrott, M. Plane, G. Underbakke και R. Brown, «Primary care practice adherence to National Cholesterol Education Program guidelines for patients with coronary heart disease,» *Arch Intern Med*, τόμ. 158, pp. 1238-44, 1998.
- [68 W. N. Hait, «Translating research into clinical practice: deliberations from the American Association for Cancer Research,» *Clinical Cancer Research*, τόμ. 11, αρ. 12,

pp. 4275-4277, 2005.

- [69 D. Butler, «Translational research: crossing the valley of death,» *Nature*, τόμ. 453, αρ. 7197, pp. 840-842, 2008.]
- [70 H.-V. Shervanthi και J. Tsui, «The continuing challenges of translational research: clinician-scientists' perspective,» *Cardiology research and practice*, 2012.]
- [71 F. Wares, «Failure to Launch: From Big Data to Big Decisions,» 2014. [Ηλεκτρονικό]. Available: http://www.fortewares.com/Administrator/userfiles/Banner/forte-wares-pro-active-reporting_EN.pdf.]
- [72 L. Jordan, «The problem with Big Data in Translational Medicine. A review of where we've been and the possibilities ahead. *Applied & Translational Genomics*, 6, pp.3-6.,» 2015.]
- [73 D. Khodyakov, L. Mikesell, R. Schraiber και M. B. E. Booth, «On using ethical principles of community-engaged research in translational science. *Translational Research*, 171, pp.52-62.e1.,» *Translational Research*, τόμ. 17, αρ. 1, pp. 52 - 62, 2016.]
- [74 «National Centre for Advancing Translational Sciences,» National Institute of Health, [Ηλεκτρονικό]. Available: <http://www.ncats.nih.gov/>. [Πρόσβαση Δεκέμβριος 2017].]
- [75 «ODM Certification & Archiving and Interchange of Metadata.,» [Ηλεκτρονικό]. Available: <http://www.cdisc.org/odm>. [Πρόσβαση Δεκέμβριος 2017].]

