

ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ



Πτυχιακή Εργασία

Υπολογιστική Επαναστόχευση Φαρμάκων

Τσέτσιλας Θεόδωρος, Π2018060

Επιβλέπων: Έξαρχος Θεμιστοκλής

Σεπτέμβριος 2024

Επιβλέπων

Έξαρχος Θεμιστοκλής, *Αναπληρωτής Καθηγητής,*
Ιόνιο Πανεπιστήμιο, Τμήμα Πληροφορικής

Τριμελής Επιτροπή

Έξαρχος Θεμιστοκλής, *Αναπληρωτής Καθηγητής,*
Ιόνιο Πανεπιστήμιο, Τμήμα Πληροφορικής
Βραχάτης Αριστείδης, *Επίκουρος Καθηγητής,*
Ιόνιο Πανεπιστήμιο, Τμήμα Πληροφορικής
Κροκίδης Μάριος, *Επίκουρος Καθηγητής,*
Ιόνιο Πανεπιστήμιο, Τμήμα Πληροφορικής

Περίληψη

Η επαναστόχευση του φαρμάκου περιλαμβάνει τον εντοπισμό νέων θεραπευτικών χρήσεων για υπάρχοντα φάρμακα και την ανακάλυψη θεραπειών για ασθένειες που δεν θεραπεύονταν προηγουμένως. Είναι ζωτικής σημασίας για την ενίσχυση της διαδικασίας προκλινικής ανάπτυξης νέων φαρμάκων, καθώς μπορεί να μειώσει σημαντικά το χρόνο και το κόστος σε σύγκριση με τις παραδοσιακές μεθόδους ανακάλυψης φαρμάκων de novo. Η αύξηση των διαθέσιμων στο κοινό μεγάλης κλίμακας βιολογικών, βιοϊατρικών και ηλεκτρονικών δεδομένων υγείας, σε συνδυασμό με προηγμένες υπολογιστικές δυνατότητες υψηλής απόδοσης, έχει επιταχύνει την ανάπτυξη υπολογιστικών προσεγγίσεων επαναστόχευσης φαρμάκων. Ερευνητές από διάφορους κλάδους έχουν αναλάβει πολυάριθμες προσπάθειες, με διάφορους βαθμούς επιτυχίας, για να διερευνήσουν υπολογιστικά τις δυνατότητες επαναστόχευσης φαρμάκων για εναλλακτικές ενδείξεις.

Αυτή η μελέτη εξετάζει την πρόσφατη πρόοδο στην υπολογιστική επαναστόχευση φαρμάκων. Αναλύοντας 26 σχετικές μελέτες που προέρχονται από βάσεις δεδομένων όπως το IEEE Xplore, το PubMed και το Google Scholar, αυτή η ανασκόπηση υπογραμμίζει τις εξελίξεις, τις μεθοδολογίες και τα αποτελέσματα σε αυτόν τον καινοτόμο τομέα. Η κύρια εστίαση είναι στην ενσωμάτωση δεδομένων γονιδιωμικής, μεταγραφικής και χημικής δομής για τον εντοπισμό νέων θεραπευτικών χρήσεων για υπάρχοντα φάρμακα. Αρχικά εξετάζουμε διαφορετικές στρατηγικές επαναστόχευσης φαρμάκων και παρέχουμε μια επισκόπηση των πόρων που χρησιμοποιούνται συνήθως. Στη συνέχεια, συνοψίζουμε τις υπολογιστικές μεθόδους που εφαρμόζονται ευρέως στην έρευνα επαναστόχευσης φαρμάκων. Στη συνέχεια παρουσιάζουμε διάφορα υπολογιστικά και πειραματικά μοντέλα που χρησιμοποιούνται για την επικύρωση αυτών των μεθόδων. Συζητάμε επίσης πιθανές ευκαιρίες και στοχεύουμε περιοχές για μελλοντική εξερεύνηση. Τέλος, αντιμετωπίζουμε τις προκλήσεις και τους περιορισμούς που αντιμετωπίζει η υπολογιστική επαναστόχευση φαρμάκων και περιγράφουμε κατευθύνσεις για μελλοντική έρευνα.

Τα ευρήματα καταδεικνύουν ότι η υπολογιστική επαναστόχευση φαρμάκων όχι μόνο ενισχύει την κατανόησή μας για τους μηχανισμούς φαρμάκων, αλλά επίσης επιταχύνει την ανακάλυψη νέων συσχετίσεων φαρμάκων-ασθένειας. Αυτή η ανασκόπηση υπογραμμίζει τις σημαντικές δυνατότητες των υπολογιστικών τεχνικών για τη βελτίωση της αποτελεσματικότητας της ανάπτυξης φαρμάκων και την αντιμετώπιση ανεκπλήρωτων ιατρικών αναγκών μέσω της επαναστόχευσης των υπαρχόντων φαρμακολογικών παραγόντων.

Λέξεις-κλειδιά: Υπολογιστική επαναστόχευση φαρμάκων, Στρατηγικές επαναστόχευσης φαρμάκων, Εξόρυξη δεδομένων, Μηχανική μάθηση, Ανάλυση δικτύου

Περιεχόμενα

Κατάλογος Εικόνων	iv
Κατάλογος Πινάκων	v
A΄ Εισαγωγή	1
A΄.1 Στόχοι της μελέτης.....	4
A΄.2 Σημασία της μελέτης	5
B΄ Βιβλιογραφική ανασκόπηση	6
B΄.1 Επαναστόχευση φαρμάκων.....	6
B΄.2 Συμβατικά φαρμάκων εναντίον επαναστόχευσης φαρμάκων.....	6
B΄.3 Στρατηγικές επαναστόχευσης φαρμάκων	8
B΄.3.1 Επαναστόχευση με βάση τη γνώση	8
B΄.3.2 Επαναστόχευση φαρμάκων βάσει στόχου	8
B΄.3.3 Επαναστόχευση φαρμάκων βάσει μονοπατιών.....	8
B΄.3.4 Επαναστόχευση φαρμάκων με βάση το μηχανισμό στόχου	8
B΄.3.5 Επαναστόχευση με βάση την υπογραφή.....	8
B΄.3.6 Επαναστόχευση βάσει φαινοτύπων	9
B΄.3.7 Στρατηγική γονιδιώματος	9
B΄.3.8 Στρατηγική πληροφοριών χημικής δομής και μορίων	12
B΄.3.9 Στρατηγική φαινότυπου	14
B΄.4 Μέθοδοι Υπολογιστικής Επαναστόχευσης Φαρμάκων	15
B΄.4.1 Μηχανική Μάθηση	15
B΄.4.2 Μοντέλα Δικτύων	16
B΄.4.3 Εξόρυξη Κειμένων και Σηματολογικό Συμπέρασμα	17
B΄.5 Επικύρωση για Υπολογιστική Επαναστόχευση.....	17
B΄.6 Υπολογιστικά εργαλεία που χρησιμοποιούνται για την επαναστόχευση φαρμάκων	18
Γ΄ Μεθοδολογία.....	19
Δ΄ Μελέτες υπολογιστικής επαναστόχευσης φαρμάκων.....	21
Δ΄.1 Εξόρυξη δεδομένων	21
Δ΄.2. Εξόρυξη κειμένου	21
Δ΄.3 Σηματολογικές τεχνολογίες.....	29
Δ΄.5 Μηχανική μάθηση.....	33
Δ΄.6 Συνεργατική αλληλεπίδραση	35

Δ'.7 Βαθιά μάθηση	39
Δ'.8 Ομαδοποίηση	45
Δ'.9 Μέτρα κεντρικότητας του δικτύου.....	48
Αποτελέσματα-Συζήτηση	52
Συμπεράσματα.....	58
Βιβλιογραφία	60

Κατάλογος Εικόνων

Εικόνα 1: Εννοιολογικό διάγραμμα <i>de novo</i> ανάπτυξης φαρμάκων και επαναστόχευσης φαρμάκων.....	2
Εικόνα 2: Παραδοσιακή ανακάλυψη φαρμάκων έναντι επαναστόχευσης φαρμάκων	7
Εικόνα 3: Βάσεις δεδομένων που χρησιμοποιούνται σε μελέτες επαναστόχευσης.....	18
Εικόνα 4: Διάγραμμα ροής που απεικονίζει τη διαδικασία επιλογής άρθρων	20
Εικόνα 5: Το MeSHDD	25
Εικόνα 6: Η ροή εργασιών της DrugQuest.....	26
Εικόνα 7: Προσέγγιση <i>deepDR</i>	28
Εικόνα 8: Προσέγγιση <i>SLAP</i>	30
Εικόνα 9: Δομικά στοιχεία για τη συνολική αρχιτεκτονική του <i>PharmGKB</i>	32
Εικόνα 10: Ροή εργασιών πρόβλεψης <i>IC50</i>	34
Εικόνα 11: Διάγραμμα ροής της ανάλυσης των <i>Napolitano et al</i>	36
Εικόνα 12: Μια γραφική απεικόνιση της κύριας ιδέας του <i>DDR</i>	37
Εικόνα 13: Σχέδιο προσέγγισης <i>CI-PMF</i>	38
Εικόνα 14: Σχεδιασμός μελέτης	41
Εικόνα 15: Σχήμα Αρχιτεκτονικής Δικτύων για εκμάθηση μίας λήψης (<i>one-shot learning</i>) στην ανακάλυψη φαρμάκων.....	42
Εικόνα 16: Προσέγγιση μεθόδου επαναστόχευσης φαρμάκων, <i>VGAEDR</i>	45
Εικόνα 17: Ροή εργασιών για τον εντοπισμό υποψηφίων για επαναστόχευση φαρμάκων με χρήση παθήσεων που βασίζονται σε γονίδια και χαρακτηριστικά και ετερογενές δίκτυο φαρμάκων	46
Εικόνα 18: Η ροή εργασίας της μεθόδου των <i>Tan et al</i>	48
Εικόνα 19: Διάγραμμα ροής <i>HeTDR</i>	50

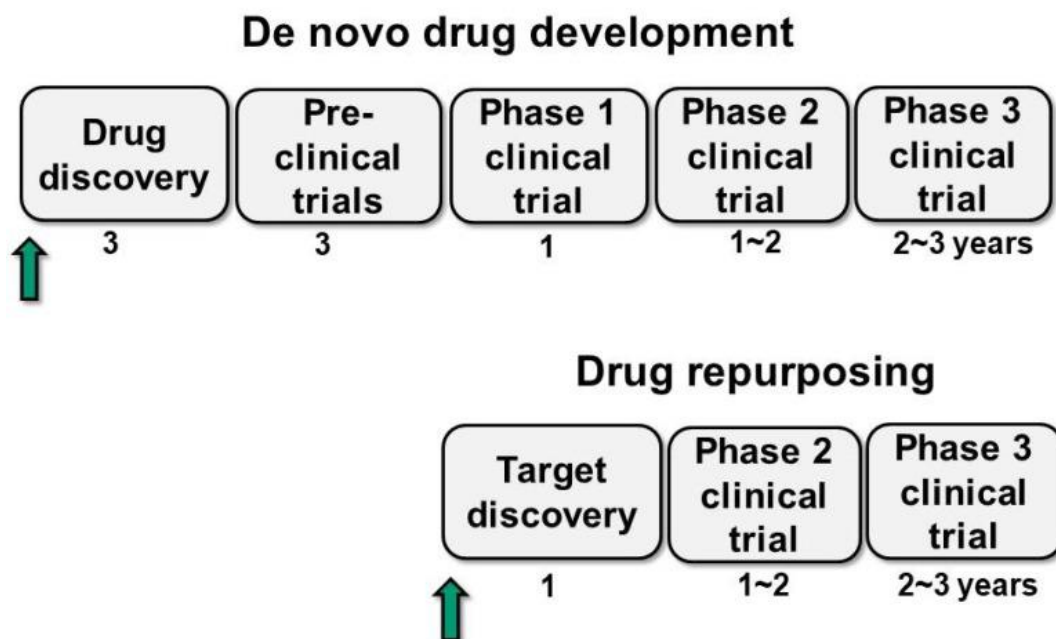
Κατάλογος Πινάκων

Πίνακας 1: Μερικά παραδείγματα επαναστοχευμένων φαρμάκων	2
Πίνακας 2: Αποτελέσματα βιβλιογραφικής ανασκόπησης της μελέτης.....	53

Α΄ Εισαγωγή

Η ανάπτυξη νέων φαρμάκων από την αρχή είναι μια δαπανηρή και χρονοβόρα προσπάθεια. Κατά μέσο όρο, χρειάζονται από 2 έως 3 δισεκατομμύρια δολάρια και τουλάχιστον 13-15 χρόνια για να κυκλοφορήσει ένα νέο φάρμακο. Επιπλέον, η διαδικασία έχει υψηλό ποσοστό αποτυχίας, με μόνο το 10% των φαρμάκων που εισέρχονται σε κλινικές δοκιμές φάσης I να λαμβάνουν τελικά έγκριση, κυρίως λόγω προβλημάτων τοξικότητας ή έλλειψης αποτελεσματικότητας ^{1,2}. Αυτή η υψηλή τριβή είναι συχνά αποτέλεσμα λανθασμένης αναγνώρισης του στόχου ή της απόκρισης του φαρμάκου. Παρά τις σημαντικές τεχνολογικές εξελίξεις και τις αυξημένες δαπάνες, ο αριθμός των φαρμάκων που εγκρίθηκαν πρόσφατα δεν έχει αυξηθεί. Στην ογκολογία, το ποσοστό έγκρισης είναι ακόμη χαμηλότερο, με μόνο το 5% των φαρμακευτικών ενώσεων σε κλινικές δοκιμές Φάσης I να λαμβάνει έγκριση. Για τα ορφανά φάρμακα, με περισσότερες από 8000 σχετικές ασθένειες, το τρέχον κόστος E&A καθιστά ανέφικτη την ανάπτυξη νέων φαρμάκων ^{3,4}.

Υπό το φως αυτών των προκλήσεων, η επαναστόχευση των φαρμάκων έχει αναδειχθεί ως μια βιώσιμη εναλλακτική λύση. Συζητήθηκε για πρώτη φορά από τους Ashburn και Thor το 2004 και η προσέγγιση αυτή περιλαμβάνει την εύρεση νέων χρήσεων για φάρμακα που ήδη κυκλοφορούν στο εμπόριο. Το κύριο πλεονέκτημα της επαναστόχευσης φαρμάκων είναι ότι τα υπάρχοντα φάρμακα συνοδεύονται από προκλινικές και κλινικές πληροφορίες, συμπεριλαμβανομένων των προφίλ φαρμακοκινητικής, φαρμακοδυναμικής και τοξικότητας, γεγονός που μειώνει τους κινδύνους ανάπτυξης. Κατά συνέπεια, αυτά τα φάρμακα μπορούν γρήγορα να περάσουν σε κλινικές δοκιμές τελευταίου σταδίου, μειώνοντας τόσο το κόστος όσο και το χρόνο. Δεν αποτελεί έκπληξη το γεγονός ότι περίπου το 30% των πρόσφατα εγκεκριμένων φαρμάκων στις ΗΠΑ είναι πλέον φάρμακα που έχουν επανατοποθετηθεί ^{2,5}. Ο Πίνακας 1 παραθέτει μερικά παραδείγματα επαναστοχευμένων φαρμάκων και η Εικόνα 1 απεικονίζει τις διαδικασίες de novo ανάπτυξης φαρμάκων και επαναστόχευσης φαρμάκων.



Εικόνα 1: Εννοιολογικό διάγραμμα *de novo* ανάπτυξης φαρμάκων και επαναστόχευσης φαρμάκων

Το βέλος υποδηλώνει τον χρόνο έναρξης σε κάθε σενάριο και ο αριθμός υποδηλώνει μια χρονική περίοδο σε χρόνια. Σε αυτό το σχήμα, ο στόχος της επαναστόχευσης φαρμάκων θεωρήθηκε ότι ήταν ο εντοπισμός ή η ανακάλυψη νέων στόχων για ένα φάρμακο που διατίθεται στην αγορά. Να σημειωθεί ότι η επαναστόχευση του φαρμάκου ξεκινά με την ανακάλυψη στόχου για ένα υπάρχον φάρμακο, ακολουθούμενη απευθείας από κλινικές δοκιμές φάσης 2 και 3, ενώ κλινικές μελέτες σε ζώα και φάση 1 δεν πραγματοποιήθηκαν, καθώς τα αποτελέσματα για αυτές τις μελέτες είναι ήδη διαθέσιμα για ένα υπάρχον φάρμακο.

Πίνακας 1: Μερικά παραδείγματα επαναστοχευμένων φαρμάκων

Φάρμακο	Αρχική ένδειξη	Νέα ένδειξη
Αλλοπουρινόλη	Καρκίνος	ουρική αρθρίτιδα
Ασπιρίνη	Φλεγμογή, Πόνος	Αντιαιμοπεταλιακό
Βρωμοκρυπτίνη	Νόσος Πάρκινσον	Σακχαρώδης διαβήτης
βουπροπιόνη	Κατάθλιψη	Διακοπή καπνίσματος
ντουλοξετίνη	Κατάθλιψη	Ακράτεια ούρων από στρες
Finasteride	Καλοήθης υπερπλασία του προστάτη	Τριχόπτωση
γκαμπαπεντίνη	Επιληψία	Νευροπαθητικός πόνος
Γεμισιταβίνη	Αντιικό	Καρκίνο
Μεθοτρεξάτη	Ρευματοειδής αρθρίτιδα	Καρκίνος

Προπρανολόλη	Υπέρταση	Ημικρανία
Ραλοξιφαίνη	Οστεοπόρωση	Καρκίνος του μαστού
Sildenafil	Στηθάγχη	Στυτική δυσλειτουργία
θαλιδομίδη	Πρωινή ναυτία	Λέπρα, Πολλαπλό μυέλωμα
Ζιδοβουδίνη	Καρκίνος	AIDS

Η επαναστόχευση των φαρμάκων μπορεί να διεξαχθεί είτε πειραματικά είτε υπολογιστικά, με την τελευταία να είναι γνωστή ως "in silico drug repurposing". Αυτή η μέθοδος εμπίπτει στην υπολογιστική φαρμακολογία και χωρίζεται σε δύο κύριες στρατηγικές: την ανακάλυψη νέων χρήσεων για υπάρχοντα φάρμακα (φαρμακοκεντρικά) και τον εντοπισμό αποτελεσματικών φαρμάκων για συγκεκριμένες ασθένειες (ασθένειο-κεντρικά). Και οι δύο στρατηγικές περιλαμβάνουν την αξιολόγηση των ομοιοτήτων μεταξύ φαρμάκων ή/και ασθενειών. Διάφορες υπολογιστικές προσεγγίσεις για την επαναστόχευση έχουν αναθεωρηθεί από τους Jin & Wong (2014)⁶.

Η πρόοδος και η ευρεία υιοθέτηση του in silico repurposing φαρμάκων καθοδηγούνται από δύο τεχνολογικές τάσεις. Πρώτον, η δημιουργία και η συσσώρευση δεδομένων υψηλής απόδοσης από τη γονιδιωματική, την πρωτεϊνική, τη χημειο-πρωτεϊνωματική και τη φαινομενική έχει καταστήσει διαθέσιμους ολοκληρωμένους χάρτες πορείας και λεπτομερή προφίλ ασθενειών και φαρμάκων. Δεύτερον, οι πρόοδοι στις επιστήμες υπολογιστών και δεδομένων έχουν διευκολύνει την ανάπτυξη αλγορίθμων επαναστόχευσης και τη συντήρηση πειραματικών βάσεων δεδομένων⁷.

Στο εξελισσόμενο τοπίο της σύγχρονης ιατρικής, η επαναστόχευση φαρμάκων διαδραματίζει κρίσιμο ρόλο στην ενίσχυση των θεραπευτικών επιλογών. Με την αυξανόμενη συχνότητα εμφάνισης χρόνιων ασθενειών, αναδυόμενων μολυσματικών ασθενειών και αντοχής στα αντιβιοτικά, η ζήτηση για γρήγορες και αποτελεσματικές λύσεις θεραπείας είναι πρωταρχικής σημασίας. Η επαναστόχευση φαρμάκων προσφέρει μια βιώσιμη οδό για την κάλυψη αυτών των αναγκών διευκολύνοντας την ταχύτερη πρόσβαση σε θεραπείες, συχνά μέσω της επαναστόχευσης φαρμάκων που έχουν ήδη εγκριθεί για άλλες ενδείξεις.

Η παραδοσιακή διαδικασία ανακάλυψης φαρμάκων είναι γεμάτη προκλήσεις, όπως το υψηλό κόστος, τα μεγάλα χρονοδιαγράμματα ανάπτυξης και τα υψηλά ποσοστά αποτυχίας στις κλινικές δοκιμές. Η ανάπτυξη ενός νέου φαρμάκου μπορεί να διαρκέσει πάνω από μια δεκαετία και κοστίζει δισεκατομμύρια δολάρια, με πολλούς υποψήφιους να αποτυγχάνουν να κυκλοφορήσουν στην αγορά λόγω προβλημάτων ασφάλειας ή αποτελεσματικότητας. Σε αυτό το πλαίσιο, η φαρμακευτική βιομηχανία αντιμετωπίζει μια πιεστική ανάγκη για πιο αποτελεσματικές στρατηγικές που μπορούν να εξορθολογίσουν τον αγωγό ανάπτυξης φαρμάκων. Η επαναστόχευση φαρμάκων αναδεικνύεται ως μια πολλά υποσχόμενη λύση, μειώνοντας σημαντικά τόσο τον χρόνο όσο και το κόστος, ενώ παράλληλα μετριάσει ορισμένους από τους κινδύνους που συνδέονται με την ανάπτυξη νέων φαρμάκων^{1,3}.

Η επαναστόχευση των υπαρχόντων φαρμάκων μπορεί να εξοικονομήσει σημαντικά ποσά στο κόστος έρευνας και ανάπτυξης, καθώς πολλές μελέτες σχετικά με την ασφάλεια και την αποτελεσματικότητα έχουν ήδη διεξαχθεί.

Ταχύτερος χρόνος για την αγορά: Με καθιερωμένα προφίλ ασφάλειας, τα επαναστοχευμένα φάρμακα μπορούν συχνά να κινηθούν μέσω της ρυθμιστικής διαδικασίας πιο γρήγορα από τις νέες ενώσεις, επιτρέποντας ταχύτερη διαθεσιμότητα θεραπειών σε ασθενείς που έχουν ανάγκη.

Γνωστή Φαρμακοκινητική: Τα υπάρχοντα φάρμακα συνοδεύονται από πληθώρα φαρμακοκινητικών και φαρμακοδυναμικών δεδομένων, διευκολύνοντας την καλύτερη κατανόηση και διαχείριση πιθανών παρενεργειών και αλληλεπιδράσεων.

Δυνατότητα για τη θεραπεία σύνθετων ασθενειών: Η επαναστόχευση φαρμάκων επιτρέπει την εξερεύνηση θεραπειών για σύνθετες και πολυπαραγοντικές ασθένειες, όπως ο καρκίνος, οι νευροεκφυλιστικές διαταραχές και τα μεταβολικά σύνδρομα, όπου μπορεί να είναι δύσκολο να εντοπιστούν νέοι στόχοι.

Καινοτόμες λύσεις για επείγουσες ανάγκες υγείας: Σε περιόδους κρίσεων δημόσιας υγείας, όπως οι πανδημίες, η επαναστόχευση φαρμάκων μπορεί να προσφέρει γρήγορες θεραπευτικές λύσεις χρησιμοποιώντας υπάρχοντα φάρμακα για την αποτελεσματική αντιμετώπιση των αναδυόμενων απειλών.

Συνοπτικά, η επαναστόχευση φαρμάκων είναι μια κεντρική στρατηγική στη σύγχρονη ιατρική που όχι μόνο αντιμετωπίζει την επείγουσα ανάγκη για αποτελεσματική ανακάλυψη φαρμάκων, αλλά ενισχύει επίσης το θεραπευτικό τοπίο, καθιστώντας την πολύτιμη οδό για μελλοντική έρευνα και κλινικές εφαρμογές.

Α'.1 Στόχοι της μελέτης

Οι στόχοι αυτής της μελέτης για την Υπολογιστική Επαναστόχευση Φαρμάκων είναι:

- Να γίνει αναθεώρηση και ανάλυση των τρεχουσών υπολογιστικών μεθοδολογιών, ακολουθώντας μία ολοκληρωμένη ανασκόπηση των υπαρχουσών υπολογιστικών προσεγγίσεων που χρησιμοποιούνται στην επαναστόχευση φαρμάκων, συμπεριλαμβανομένης της μηχανικής εκμάθησης, των μεθόδων που βασίζονται σε δίκτυο και των τεχνικών εξόρυξης κειμένου.
- Επίσης θα εντοπιστούν οι κύριες προκλήσεις που αντιμετωπίζει η υπολογιστική επαναστόχευση φαρμάκων, όπως η ποιότητα των δεδομένων, οι αλγοριθμικοί περιορισμοί και τα ρυθμιστικά ζητήματα και προτείνετε πιθανές λύσεις για τον μετριασμό αυτών των προκλήσεων.
- Η εργασία θα επιδείξει και πρακτικές εφαρμογές της υπολογιστικής επαναστόχευσης φαρμάκων παρουσιάζοντας λεπτομερείς μελέτες, επισημαίνοντας τις μεθοδολογίες και τα αποτελέσματα των φαρμάκων που επαναπροορίζονται.

Α΄.2 Σημασία της μελέτης

Αυτή η μελέτη μπορεί να μειώσει σημαντικά τον χρόνο και το κόστος που σχετίζεται με την ανάπτυξη παραδοσιακών φαρμάκων, εντοπίζοντας νέες χρήσεις για υπάρχοντα φάρμακα, επιταχύνοντας έτσι τη διαδικασία ανακάλυψης φαρμάκων. Με την επαναστόχευση των φαρμάκων, η μελέτη μπορεί να συμβάλει σε βελτιωμένες επιλογές θεραπείας για ασθενείς, ιδιαίτερα για ασθένειες που επί του παρόντος έχουν περιορισμένες θεραπευτικές επιλογές.

Η έρευνα προωθεί την ενσωμάτωση υπολογιστικών τεχνικών αιχμής, όπως η μηχανική μάθηση και η ανάλυση δικτύου, ενισχύοντας την καινοτομία στον φαρμακευτικό και βιοϊατρικό τομέα. Τα ευρήματα μπορούν να βοηθήσουν στην αντιμετώπιση επειγουσών προκλήσεων για τη δημόσια υγεία, συμπεριλαμβανομένης της ταχείας ανάπτυξης θεραπειών κατά τη διάρκεια πανδημιών ή της ανάγκης για θεραπείες για σπάνιες ασθένειες.

Η μελέτη μπορεί να προσφέρει πολύτιμες γνώσεις για τους μηχανισμούς δράσης των φαρμάκων, οδηγώντας σε καλύτερη κατανόηση της βιολογίας της νόσου και των αλληλεπιδράσεων των φαρμάκων. Εντοπίζοντας τις τρέχουσες προκλήσεις και τις μελλοντικές ευκαιρίες, η έρευνα μπορεί να ενημερώσει τους υπεύθυνους χάραξης πολιτικής και τους ερευνητές για τομείς που χρειάζονται περαιτέρω έρευνα και επένδυση. Η έμφαση στην ολοκλήρωση και την ανάλυση δεδομένων ενθαρρύνει μια πιο συστηματική και βασισμένη σε στοιχεία προσέγγιση για την επαναστόχευση των φαρμάκων, προς όφελος τόσο του ακαδημαϊκού κόσμου όσο και της βιομηχανίας. Τέλος, η μελέτη υπογραμμίζει τη σημασία της διεπιστημονικής συνεργασίας μεταξύ επιστημόνων υπολογιστών, βιολόγων και κλινικών γιατρών, ενθαρρύνοντας μια πιο ολιστική προσέγγιση στην ανακάλυψη φαρμάκων.

B' Βιβλιογραφική ανασκόπηση

B'.1 Επαναστόχευση φαρμάκων

Η επαναστόχευση φαρμάκων αναφέρεται στη διαδικασία χρήσης καθιερωμένων φαρμάκων που είναι ήδη διαθέσιμα στην αγορά για νέες ενδείξεις ασθένειας. Τα επαναστοχευμένα φάρμακα μπορούν να παρακάμψουν προηγούμενες φάσεις, προχωρώντας απευθείας στις κλινικές δοκιμές Φάσεων II και III και στη διαδικασία έγκρισης του FDA, μειώνοντας έτσι σημαντικά τόσο τον χρόνο όσο και το κόστος. Για παράδειγμα, μια συντηρητική εκτίμηση υποδηλώνει ότι η επαναστόχευση ενός φαρμάκου μπορεί να κοστίσει περίπου 300 εκατομμύρια δολάρια και να διαρκέσει περίπου έξι χρόνια ⁸. Ο κίνδυνος αποτυχίας είναι μικρότερος για τα επαναστοχευμένα φάρμακα, καθώς οι υποψήφιοι σε τελικό στάδιο ανάπτυξης έχουν ήδη επιδείξει ασφάλεια μέσω προκλινικών μοντέλων και μελετών σε ανθρώπους ⁹. Κατά συνέπεια, η επαναστόχευση φαρμάκων επιταχύνει το χρονοδιάγραμμα και μειώνει το κόστος που σχετίζεται με την εύρεση θεραπειών για διαφορετικές ασθένειες, επηρεάζοντας θετικά τα αποτελέσματα υγείας του πληθυσμού σε επίπεδο ασθενών.

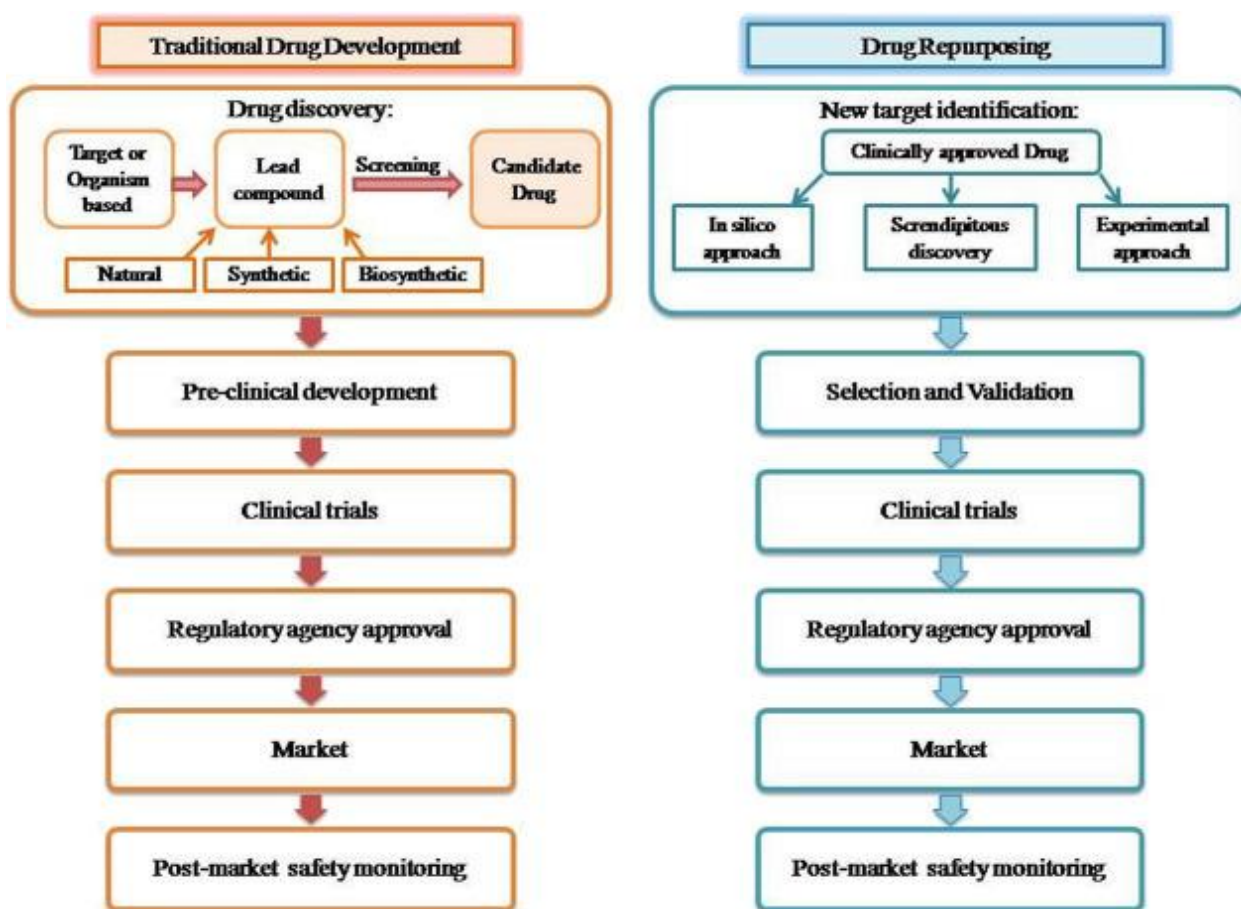
Δεδομένης της ασυνήθιστης φύσης των πρώιμων ανακαλύψεων, δίνεται μια αυξανόμενη έμφαση στην επαναστόχευση φαρμάκων βάσει δεδομένων, η οποία επιτρέπει τη δημιουργία πιο συνεπών υποθέσεων. Αυτή η προσέγγιση ανταποκρίνεται στην αυξανόμενη διαθεσιμότητα συνόλων βιοϊατρικών δεδομένων μεγάλης κλίμακας, όπως οι πολυμορφισμοί μεμονωμένων νουκλεοτιδίων κινδύνου (SNPs) που προσδιορίζονται σε μελέτες συσχέτισης σε όλο το γονιδίωμα (GWAS) και βάσεις δεδομένων αλληλεπίδρασης πρωτεϊνών, καθώς και κλινικά σύνολα δεδομένων όπως ηλεκτρονικά ιατρικά αρχεία (EMRs). Η υπολογιστική επαναστόχευση φαρμάκων χρησιμοποιεί συστηματική ανάλυση δεδομένων μέσω υπολογιστικών μεθόδων για τη δημιουργία υποθέσεων για επαναστόχευση φαρμάκων ^{8,9}. Ο αυστηρός αγωγός επαναστόχευσης φαρμάκων περιλαμβάνει πρωτίστως τη δημιουργία συνδέσεων μεταξύ υπαρχόντων φαρμάκων και ασθενειών που απαιτούν θεραπεία. Αυτή η σύνδεση διαμορφώνεται με βάση χαρακτηριστικά που συγκεντρώθηκαν μέσω βιολογικών ή κλινικών πειραμάτων, τα οποία μπορούν να περιγράψουν αυτά τα στοιχεία μέσω υπολογιστικών εργαλείων, ιδιαίτερα όταν πρόκειται για μεγάλα και υψηλών διαστάσεων σύνολα δεδομένων ^{10,11}.

B'.2 Συμβατικά φαρμάκων εναντίον επαναστόχευσης φαρμάκων

Το 2015, οι φαρμακευτικές εταιρείες επένδυσαν περίπου 2,6 δισεκατομμύρια δολάρια για την ανάπτυξη μιας νέας χημικής ένωσης που εγκρίθηκε από τον Οργανισμό Τροφίμων και Φαρμάκων των ΗΠΑ (FDA). Το υψηλό κόστος παραγωγής φαρμάκων οφείλεται στον μεγάλο αριθμό ενώσεων που αξιολογούνται σε προκλινικά στάδια και στον σημαντικό αριθμό τυχαιοποιημένων ελεγχόμενων δοκιμών που αποτυγχάνουν να επιδείξουν θεραπευτική αποτελεσματικότητα ή αντιμετωπίζουν προβλήματα τοξικότητας. Δεδομένων των αυξανόμενων ποσοστών αποτυχίας, του κλιμακούμενου κόστους και των μακρών χρονοδιαγραμμάτων ανάπτυξης νέων φαρμάκων, η βελτιστοποίηση της χρήσης των υπαρχόντων φαρμάκων και η ελαχιστοποίηση των παρενεργειών στις δοκιμές αποτελεί μια

βιώσιμη εναλλακτική λύση. Ο βραβευμένος με Νόμπελ Sir James Black δήλωσε περίφημα: «Η πιο γόνιμη βάση για την ανακάλυψη ενός νέου φαρμάκου είναι να ξεκινήσεις με ένα παλιό φάρμακο» ¹

Η επαναστόχευση των φαρμάκων περιλαμβάνει την εύρεση νέων χρήσεων για εγκεκριμένα ή υπό έρευνα φάρμακα. Τα επαναστοχευμένα φάρμακα μπορούν να παρέχουν θεραπεία πιο γρήγορα και με χαμηλότερο κόστος από την ανάπτυξη νέων φαρμάκων, καθώς η αποτελεσματικότητά τους έχει ήδη τεκμηριωθεί μέσω κλινικών δοκιμών για άλλες καταστάσεις ¹². Πολλά ερευνητικά και ακαδημαϊκά ιδρύματα υποστηρίζουν την εξερεύνηση των υφιστάμενων αποθετηρίων φαρμάκων, προτείνοντας ότι τα φάρμακα που έχουν σχεδιαστεί για μια ασθένεια μπορεί επίσης να είναι αποτελεσματικά για μια άλλη. Ένα αξιοσημείωτο πρόσφατο παράδειγμα είναι η επαναστόχευση της ρεμδεσιβίρης για τη θεραπεία του COVID-19 ¹².



Εικόνα 2: Παραδοσιακή ανακάλυψη φαρμάκων έναντι επαναστόχευσης φαρμάκων

Σε σύγκριση με τις παραδοσιακές μεθόδους ανακάλυψης φαρμάκων, η επαναστόχευση φαρμάκων προσφέρει πολλά πλεονεκτήματα. Υπάρχει σημαντική μείωση του χρόνου που απαιτείται για έρευνα και ανάπτυξη. Ενώ η ανάπτυξη συμβατικών φαρμάκων συνήθως διαρκεί περίπου 10 έως 16 χρόνια, η επαναστόχευση του φαρμάκου μπορεί να ολοκληρωθεί σε περίπου 3 έως 12 χρόνια ¹³. Τα επαναστοχευμένα φάρμακα μπορούν να παρακάμψουν τις δοκιμές πρώιμου σταδίου, μειώνοντας έτσι τον χρόνο, το κόστος και τον κίνδυνο. Οι μελέτες δείχνουν ότι τα επαναστοχευμένα φάρμακα

χρειάζονται συνήθως 3 έως 12 χρόνια για να λάβουν την έγκριση του FDA, με κόστος περίπου το μισό από αυτό που απαιτείται για εγκρίσεις νέων φαρμάκων ^{14,15}. Η Εικόνα 2 απεικονίζει τη ροή εργασίας για την ανάπτυξη συμβατικών φαρμάκων έναντι της επαναστόχευσης φαρμάκων. Η επαναστόχευση φαρμάκων εστιάζει κυρίως στην ανάπτυξη θεραπειών για αναδυόμενες μολυσματικές ασθένειες και παραμελημένες καταστάσεις, ενώ η παραδοσιακή ανακάλυψη φαρμάκων στοχεύει κυρίως χρόνια ζητήματα υγείας ¹⁶.

B'3 Στρατηγικές επαναστόχευσης φαρμάκων

B'3.1 Επαναστόχευση με βάση τη γνώση

Αυτή η στρατηγική περιλαμβάνει τη δημιουργία μοντέλων χρησιμοποιώντας πληροφορίες που σχετίζονται με φάρμακα, όπως στόχους φαρμάκων, χημικές δομές, μονοπάτια και δυσμενείς επιπτώσεις, για την πρόβλεψη άγνωστων στόχων, βιοδεικτών ή μηχανισμών για ασθένειες. Περιλαμβάνει επαναστόχευση φαρμάκων βάσει στόχου, οδού και μηχανισμού στόχου ¹⁷.

B'3.2 Επαναστόχευση φαρμάκων βάσει στόχου

Αυτή η μέθοδος περιλαμβάνει διαλογή υψηλής απόδοσης και/ή υψηλού περιεχομένου (HTS/HCS) ενώσεων φαρμάκων, ακολουθούμενη από έλεγχο σε πυρίτιο από βιβλιοθήκες φαρμάκων χρησιμοποιώντας διαλογή ή σύνδεση με βάση συνδέτη. Σε αντίθεση με τις τυφλές αναζητήσεις, ο επαναπροσανατολισμός βάσει στόχου συνδέει τους στόχους απευθείας με μηχανισμούς ασθενειών, αυξάνοντας την πιθανότητα ανακάλυψης αποτελεσματικών φαρμάκων. Ωστόσο, δεν μπορεί να αναγνωρίσει άγνωστους μηχανισμούς πέρα από τους γνωστούς στόχους ¹⁸.

B'3.3 Επαναστόχευση φαρμάκων βάσει μονοπατιών

Αυτή η προσέγγιση χρησιμοποιεί πληροφορίες για μεταβολικές οδούς, μονοπάτια σηματοδότησης και δίκτυα αλληλεπίδρασης πρωτεϊνών για να προβλέψει τις συνδέσεις μεταξύ ασθενειών και φαρμάκων. Οι ειδικές για τη νόσο μονοπάτια που ανακατασκευάζονται από δεδομένα omics μπορούν να χρησιμεύσουν ως νέοι στόχοι για επαναστοχευμένα φάρμακα ¹⁹.

B'3.4 Επαναστόχευση φαρμάκων με βάση το μηχανισμό στόχου

Αυτή η μέθοδος ενσωματώνει πληροφορίες μονοπατιού σηματοδότησης, δεδομένα θεραπείας omics και δίκτυα αλληλεπίδρασης πρωτεϊνών για να ανακαλύψει νέους μηχανισμούς δράσης για τα φάρμακα. Στοχεύει στην εύρεση μηχανισμών που σχετίζονται με ασθένειες, φάρμακα και συγκεκριμένες θεραπείες, υποστηρίζοντας την αυξανόμενη ανάγκη για ιατρική ακριβείας.

B'3.5 Επαναστόχευση με βάση την υπογραφή

Αυτή η στρατηγική χρησιμοποιεί πληροφορίες γονιδιακής υπογραφής από δεδομένα ασθενειών για την ανακάλυψη νέων εκτός στόχων ή μηχανισμών ασθένειας. Συγκρίνει τα προφίλ γονιδιακής έκφρασης φαρμάκων και ασθενειών για να εντοπίσει αντίστροφες σχέσεις. Για παράδειγμα, ο DURDAĞI (2020)²⁰ διερεύνησε πιθανά ζεύγη φαρμάκων-νόσων για τη φλεγμονώδη νόσο του εντέρου

(IBD) συγκρίνοντας τα προφίλ γονιδιακής έκφρασης από τη βάση δεδομένων omnibus γονιδιακής έκφρασης με αυτά από τον χάρτη συνδεσιμότητας, οδηγώντας στην ανακάλυψη νέων ζευγών φαρμάκων-νόσων επικυρωθεί σε προκλινικά μοντέλα. Αυτές οι μέθοδοι εμβαθύνουν σε μοριακούς και γενετικούς μηχανισμούς πιο βαθιά από τις μεθόδους που βασίζονται στη γνώση.

B'.3.6 Επαναστόχευση βάσει φαινοτύπων

Αυτή η στρατηγική αξιοποιεί φαινοτυπικές πληροφορίες, που χρησιμοποιούνται όλο και περισσότερο από προσεγγίσεις συστημάτων για τον εντοπισμό γενετικών χαρακτηριστικών που σχετίζονται με ασθένειες. Η επεξεργασία φυσικής γλώσσας που εφαρμόζεται στα ηλεκτρονικά αρχεία υγείας (EHRs) μπορεί να αποκαλύψει πρόσθετες ανεπιθύμητες ενέργειες φαρμάκων που δεν παρατηρήθηκαν κατά την ανάπτυξη φαρμάκων. Για παράδειγμα, τα EHR εξόρυξης αναγνώρισαν τη μεταφορμίνη ως πιθανή θεραπεία για τον καρκίνο ¹⁶.

B'.3.7 Στρατηγική γονιδιώματος

Το γονιδίωμα είναι ένας όρος που χρησιμοποιείται για να περιγράψει όλα τα γονίδια που αφορούν έναν συγκεκριμένο οργανισμό. Με άλλα λόγια, τα βιολογικά δεδομένα που αποθηκεύονται σε ένα γονιδίωμα αντιπροσωπεύονται από το DNA του και χωρίζονται σε ξεχωριστές μονάδες που ονομάζονται γονίδια. Η εισαγωγή του έργου αλληλουχίας του ανθρώπινου γονιδιώματος σηματοδότησε ένα σημείο καμπής στην απόκτηση γνώσης σε μοριακό επίπεδο σχετικά με τον τρόπο λειτουργίας των ζωντανών οργανισμών και έφερε επανάσταση στις μελέτες επαναστόχευσης φαρμάκων. Πιο συγκεκριμένα, με την εύρεση των γονιδίων ή των πρωτεϊνών που διαδραματίζουν σημαντικό ρόλο στις μοριακές δράσεις των φαρμάκων και των ασθενειών, η πρωτοβουλία του έργου αλληλουχίας ανθρώπινου γονιδιώματος επέτρεψε την καλύτερη κατανόηση του τρόπου δράσης των φαρμάκων και των ασθενειών. Αυτά τα γονίδια και οι πρωτεΐνες έχουν γίνει δελεαστικοί στόχοι για τις κυβερνήσεις και τη φαρμακευτική βιομηχανία, γεγονός που οδήγησε στο να έχει αυτό το πεδίο της επιστήμης ως έναν από τους πιο εντατικά μελετημένους ερευνητικούς τομείς σε ερευνητικά εργαστήρια σε όλο τον κόσμο ²¹.

Οι τεράστιοι όγκοι δημοσίως διαθέσιμων γονιδιωματικών και μεταγραφικών δεδομένων που παράγονται για δείγματα ασθενειών, καθώς και κλινικές βάσεις δεδομένων, παρέχουν μια μοναδική ευκαιρία για την κατανόηση των μηχανισμών δράσης της νόσου και των φαρμάκων και την ανακάλυψη νέων χρήσεων για υπάρχοντα φάρμακα. Ωστόσο, λόγω της τεράστιας πολυπλοκότητας των βιολογικών συστημάτων, η ολοκληρωμένη κατανόηση τέτοιων συστημάτων εξακολουθεί να είναι ελλιπής. Ως αποτέλεσμα, η έρευνα για μια μοριακή εξήγηση βιολογικών συστημάτων συνεχίζεται ακόμη εκτενώς. Αξίζει να σημειωθεί ότι το προφίλ έκφρασης γονιδίου μικροσυστοιχίας είναι το πιο ευρέως χρησιμοποιούμενο μεταγραφικό προφίλ μεταξύ των μεθόδων γενετικού προφίλ που έχουν διερευνηθεί για την επαναστόχευση του φαρμάκου. Σε αντίθεση με τα περισσότερα παραδοσιακά εργαλεία μοριακής βιολογίας που επιτρέπουν τη μελέτη ενός μόνο γονιδίου ή ενός μικρού συνόλου γονιδίων, το προφίλ έκφρασης γονιδίων μικροσυστοιχιών καταγράφει τις δυναμικές ιδιότητες ενός κυττάρου και

μετρά όλη τη μεταγραφική δραστηριότητα χιλιάδων γονιδίων ταυτόχρονα, οδηγώντας σε επανάσταση στον ερευνητικό τομέα της μοριακής βιολογίας. Η εφαρμογή του προφίλ έκφρασης γονιδίων σε μικροσυστοιχίες έχει, ως εκ τούτου, λάβει μεγάλη προσοχή για τον ζωτικό ρόλο της στην κατανόηση του τρόπου με τον οποίο τα γονίδια δρουν την ίδια στιγμή και υπό τις ίδιες συνθήκες²².

Οι υπολογιστικές μελέτες επαναστόχευσης φαρμάκων που χρησιμοποιούν γονιδιακά ρυθμιστικά δεδομένα υποθέτουν ότι τα φάρμακα στοχεύουν τις ίδιες πρωτεΐνες με συγκρίσιμα προφίλ γονιδιακής έκφρασης. Αυτή η κατανόηση οδήγησε στην ανακάλυψη ενός τεράστιου αριθμού νέων και απροσδόκητων λειτουργικών γονιδιακών αλληλεπιδράσεων, στην ανίχνευση νέων υποτύπων ασθενειών και στον εντοπισμό υποκείμενων μηχανισμών απόκρισης ασθένειας ή φαρμάκων²³

Το έργο Χάρτης Συνδεσιμότητας (CMap) και η εκτεταμένη Βιβλιοθήκη Ολοκληρωμένων Κυψελών Υπογραφών που Βασίζονται σε Δίκτυα (LINCS) θεωρούνται βασική ιδέα πίσω από διάφορες καλά αναγνωρισμένες μελέτες επαναστόχευσης φαρμάκων. Ο χάρτης συνδεσιμότητας μπορεί να οριστεί ως ένας συνδυασμός δεδομένων μεταγραφικής έκφρασης σε όλο το γονιδίωμα που βοηθά στην αποκάλυψη λειτουργικών συνδέσεων μεταξύ φαρμάκων, γονιδίων και ασθενειών. Το εκτεταμένο έργο του CMap παρήγαγε προφίλ γονιδιακής έκφρασης μεγάλης κλίμακας από ανθρώπινα καρκινικά κύτταρα που στοχεύονταν από διάφορες φαρμακευτικές ενώσεις σε διαφορετικά περιβάλλοντα. Οι Lamb et al. (2006)²⁴ χρησιμοποίησαν δεδομένα έκφρασης γονιδίων μικροσυστοιχιών για να δημιουργήσει έναν χάρτη συνδεσιμότητας που χρησιμοποιείται για την ανακάλυψη σχέσεων μεταξύ της λίστας γονιδίων που σχετίζονται με μια συγκεκριμένη ασθένεια ή φάρμακο, που ονομάζεται υπογραφή ερωτήματος, και ένα σύνολο προφίλ γονιδιακής έκφρασης που ονομάζεται βάση δεδομένων αναφοράς. Τα προφίλ έκφρασης που συσχετίζονται πολύ θετικά με την υπογραφή ερωτήματος θεωρείται ότι έχουν πολύ παρόμοιο τρόπο δράσης με την υπογραφή του ερωτήματος. Τα προφίλ έκφρασης που συσχετίζονται σε μεγάλο βαθμό αρνητικά με την υπογραφή του ερωτήματος λαμβάνονται υπόψη για περαιτέρω διερεύνηση θεραπείας.

Οι Iorio et al., (2010)²⁵ ανέπτυξαν μια αυτόματη προσέγγιση που εκμεταλλεύεται την ομοιότητα στα προφίλ γονιδιακής έκφρασης προκειμένου να ανακαλύψει φάρμακα που έχουν κοινό αποτέλεσμα και τρόπο δράσης. Αρχικά, οι συγγραφείς δημιούργησαν ένα δίκτυο φαρμάκων όπου οι κόμβοι αντιπροσωπεύουν φάρμακα και οι άκρες δείχνουν ομοιότητες μεταξύ ενός ζεύγους φαρμάκων. Στη συνέχεια, χρησιμοποίησαν τεχνικές γραφημάτων για τον εντοπισμό κοινοτήτων ναρκωτικών. Τα ναρκωτικά σε καθεμία από αυτές τις κοινότητες έχουν παρόμοιο τρόπο δράσης. Οι Hu & Agarwal, (2009)²⁶ διεξήγαγαν μια εκτενή ανάλυση της διαταραχής του ανθρώπινου φαρμάκου και της έκφρασης του γονιδίου της νόσου με βάση μια αρνητική συσχέτιση για την κατασκευή ενός δικτύου ασθένειας-φαρμάκων για την πρόβλεψη νέων εφαρμογών για ήδη εγκεκριμένα φάρμακα. Οι Altae-Tran et al., (2017)²⁷ πραγματοποίησαν μια ολοκληρωμένη συστηματική ανάλυση των προφίλ γονιδιακής έκφρασης για διαφορετικές ασθένειες και φάρμακα που οδήγησε στην ανακάλυψη νέων υποψηφίων για επαναστόχευση φαρμάκων.

Οι Jiang et al., (2012)²⁸ χρησιμοποίησαν πηγές δεδομένων CMap για να προσδιορίσει τις σχέσεις μεταξύ μικρών μορίων και miRNA σε ανθρώπινους καρκίνους, προκειμένου να βρει θεραπευτικές δυνατότητες και νέες ενδείξεις για υπάρχοντα φάρμακα. Οι Jahchan et al., (2013)⁹ χρησιμοποίησαν προφίλ γονιδιακής έκφρασης για να αναγνωρίσει μόρια φαρμάκων για τη θεραπεία του μικροκυτταρικού καρκίνου του πνεύμονα, ο οποίος δεν είχε αποτελεσματικές θεραπείες. Οι Huang et al., (2015)³¹ εισήγαγαν έναν νέο χάρτη συνδεσιμότητας που ονομάζεται (DMAP) που ξεπερνά τον περιορισμό δεδομένων CMap προτείνοντας έναν χάρτη συνδεσιμότητας φαρμάκου-πρωτεΐνης. Το DMAP αποτελείται από κατευθυνόμενες επιδράσεις φαρμάκου σε πρωτεΐνη και τις βαθμολογίες τους. Όλες οι προηγουμένως παρατηρηθείσες σχέσεις μεταξύ του συσχετιζόμενου φαρμάκου και πρωτεΐνης σε διάφορες πηγές δεδομένων χρησιμοποιήθηκαν για τον υπολογισμό των βαθμολογιών επίδρασης από όλες τις καταχωρήσεις της βάσης δεδομένων μεταξύ του φαρμάκου και της πρωτεΐνης καθώς και του επιπέδου εμπιστοσύνης της ποιότητας αυτών των υπολογισμένων βαθμολογιών επίδρασης.

Οι τεράστιες ποσότητες δημοσίως διαθέσιμων συνόλων δεδομένων προφίλ γονιδιακής έκφρασης έχουν ενθαρρύνει τους ερευνητές να εξετάσουν την έννοια της ενοχής για τη διερεύνηση των συσχετίσεων φαρμάκων-φαρμάκων και φαρμάκων-ασθένειας για τον εντοπισμό θεραπευτικών ενδείξεων για υπάρχοντα φάρμακα. Οι Iorio et al. (2010)²⁵ υιοθέτησε την έννοια της ενοχής κατά συσχέτιση για τη σύγκριση διαφορετικών φαρμάκων προκειμένου να εντοπιστεί οποιαδήποτε ομοιότητα μεταγραφικών αποκρίσεων, υποθέτοντας ότι αυτά τα φάρμακα θα έχουν παρόμοιο τρόπο δράσης (MoA).

Πρόσφατα, τα microRNA (miRNAs) έχουν λάβει μεγάλη προσοχή σε βιολογικές και βιοϊατρικές μελέτες για τους ρόλους τους στη ρύθμιση διαφορετικών τύπων κυτταρικών δραστηριοτήτων. Ως εκ τούτου, τα miRNAs έχουν γίνει βασικοί παίκτες στον εντοπισμό θεραπευτικών στόχων επαναστόχευσης φαρμάκων, καθώς τα miRNA είναι ζωτικής σημασίας για την ομοιοστάση των κυττάρων και είναι ενεργά σε πολλά στάδια της νόσου.

Οι Jiang et al. (2012)²⁸ χρησιμοποίησαν επίσης miRNA μαζί με μικρά μόρια, ως πιθανά φάρμακα, για να δημιουργήσουν δίκτυα για διαφορετικούς τύπους καρκίνου, προκειμένου να εντοπίσουν συσχετίσεις μικρών μορίων-miRNA για επαναστόχευση φαρμάκων με βάση τους κανονισμούς miRNA και τις μεταγραφικές αποκρίσεις. Έχουν γίνει αρκετές προσπάθειες για τη δημιουργία δημόσιων αποθετηρίων με στόχο την αύξηση της ανάπτυξης θεραπευτικών miRNA βασισμένων σε μικρά μόρια.

Οι Liu et al. (2013)³² επιμελήθηκαν χειροκίνητα την επιστημονική βιβλιογραφία αναζητώντας πώς τα μικρά μόρια επηρέασαν τις εκφράσεις του miRNA και ανέπτυξαν μια βάση δεδομένων (SM2miR) προκειμένου να συλλάβει τις υπάρχουσες συσχετίσεις μικρού μορίου-miRNA. Οι Li et al. (2014)³³ ανέκτησαν με μη αυτόματο τρόπο υποστηριζόμενους πειραματικά συσχετισμούς ασθενειών miRNA από επιστημονικά άρθρα και δημιούργησε τη βάση δεδομένων ανθρώπινων MicroRNA Disease (HMDD v2.0) για να διευκολύνει την εξερεύνηση δεδομένων. Οι Huang et al., (2019)³¹ εισήγαγαν το (HMDD v3.0) προσθέτοντας έναν σημαντικό αριθμό συσχετισμών miRNA-ασθένειας

στο (HMDD v2.0) και βελτιώνοντας την ακρίβεια αυτών των συσχετίσεων με βάση στοιχεία που βασίζονται στη βιβλιογραφία. Οι Rukon et al. (2014)³⁴ δημιούργησαν μια βάση δεδομένων (PharmacomiR) για τον εντοπισμό συσχετίσεων τριπλού συνόλου miRNA-γονιδίου-φαρμάκου συνδυάζοντας δεδομένα σχετικά με τη στόχευση miRNA και τις αλληλεπιδράσεις πρωτεΐνης-φαρμάκου.

Οι Okada et al. (2014)³⁵ ανέπτυξαν μια νέα προσέγγιση in silico διενεργώντας μια ανάλυση πολλαπλών σταδίων GWAS ασθενών με στοχευμένη νόσο για να αποκαλύψει ένα σύνολο άγνωστων τόπων κινδύνου που σχετίζονται με τη στοχευόμενη ασθένεια και να προσδιορίσει περαιτέρω ένα σύνολο βιολογικών υποψηφίων γονιδίων που στοχεύουν ήδη εγκεκριμένα φάρμακα. Επίσης, μια συλλογή υπαρχόντων φαρμάκων που έχουν εγκριθεί για άλλες ενδείξεις εντοπίστηκε και συνδέθηκε με τη μελετώμενη νόσο για πιθανές πιθανότητες επαναστόχευσης του φαρμάκου.

Οι Garnett et al. (2012)³⁶ διεξήγαγαν μια μεγάλης κλίμακας πολυπαραγοντική ανάλυση της γενετικής καρκινικών κυτταρικών γραμμών και φαρμάκων σε έργα φαρμακευτικών αγωγών για να αποκαλύψει νέους βιοδείκτες ευαισθησίας και αντοχής στα θεραπευτικά του καρκίνου. Σε κάποιο βαθμό, τα μεταλλαγμένα γονίδια επιδεικνύουν τη μοριακή δραστηριότητα των φαρμάκων και μπορούν να θεωρηθούν ως βιοδείκτες φαρμάκων κατά τη διαδικασία επαναστόχευσης του φαρμάκου. Μερικά μεταλλαγμένα γονίδια καρκινικής κυτταρικής σειράς βρέθηκαν να σχετίζονται με την ευαισθησία στο φάρμακο, η οποία μπορεί να χρησιμεύσει ως πιθανοί βιοδείκτες για την επαναστόχευση του φαρμάκου.

B'3.8 Στρατηγική πληροφοριών χημικής δομής και μορίων

Δεδομένου ότι η στρατηγική που βασίζεται σε φάρμακα στο γονιδίωμα υποθέτει ότι τα φάρμακα μοιράζονται κοινές ενδείξεις λόγω των παρόμοιων προφίλ, η χημική δομή και οι πληροφορίες μορίων των φαρμάκων θεωρούνται επίσης αξιόλογες πηγές για τον εντοπισμό οποιασδήποτε ομοιότητας μεταγραφικών αποκρίσεων μεταξύ φαρμάκων για ευκαιρίες επαναστόχευσης, καθώς αυτά τα φάρμακα συνήθως επηρεάζουν τα γονίδια, πρωτεΐνες και άλλες βιολογικές οντότητες σε παρόμοιες μορφές. Η ομοιότητα της χημικής δομής μπορεί να μετρηθεί με διάφορους τρόπους, όπως τα δισδιάστατα (2-D) τοπολογικά δακτυλικά αποτυπώματα και τα τρισδιάστατα (3-D) διαμορφωτικά δακτυλικά αποτυπώματα³⁷.

Οι Keiser et al. (2009)³⁸ πρότειναν μια συστηματική προσέγγιση ομοιότητας χημικής δομής για τη διαλογή ενώσεων υπαρχόντων και υπό επεξεργασία φαρμάκων έναντι εκατοντάδων προσδεμάτων που δεσμεύουν πρωτεϊνικούς στόχους. Η ομοιότητα της χημικής δομής μεταξύ των ενώσεων φαρμάκων και των στόχων συνδέτη αποκάλυψε χιλιάδες απρόβλεπτες συσχετίσεις, μερικές από τις οποίες δοκιμάστηκαν και επιβεβαιώθηκαν πειραματικά. Η προτεινόμενη προσέγγιση μπορεί να εξηγήσει ορισμένες από τις παρενέργειες των υπαρχόντων φαρμάκων και μπορεί επίσης να συμβάλει στον εντοπισμό νέων εφαρμογών επαναστόχευσης για υπάρχοντα φάρμακα.

Ο Swamidass (2011) πρότειναν τη χρήση χημικών δομών για να προσδιοριστεί ποιοι στόχοι φαρμάκων θα ρυθμίζουν τον σχετικό με την ασθένεια φαινότυπο. Μια τέτοια τακτική θα έδινε ενδείξεις

για το πώς άλλα φάρμακα, με παρόμοιες χημικές δομές, ρυθμίζουν τον σχετικό με την ασθένεια φαινότυπο και ως εκ τούτου αντιμετωπίζουν την ασθένεια.

Τις περισσότερες φορές, η ομοιότητα χημικής δομής ενσωματώνεται με τη μοριακή δραστηριότητα και άλλες βιολογικές πληροφορίες για τον εντοπισμό νέων συσχετίσεων και πιθανών επιπτώσεων εκτός στόχου για εγκεκριμένα και υπό έρευνα φάρμακα. Οι Yamanishi et al. (2008)⁴⁰ ανέπτυξαν ένα μοντέλο εποπτευόμενης μάθησης για ένα διμερές γράφημα για τον εντοπισμό πιθανών αλληλεπιδράσεων φαρμάκου-στόχου. Οι συγγραφείς ενσωμάτωσαν πληροφορίες χημικής δομής φαρμάκου, δίκτυο αλληλεπίδρασης πρωτεΐνης-πρωτεΐνης και δίκτυο αλληλεπίδρασης φαρμάκου-στόχου για να προβλέψουν τις θεραπευτικές δυνατότητες και να αποκαλύψουν εφαρμογές επαναστόχευσης φαρμάκων. Οι Kinnings et al. (2009)⁴¹ χρησιμοποίησαν χημικές αλληλεπιδράσεις φαρμάκων κάτω από διαφορετικές μεταβλητές περιβάλλοντος για να δημιουργήσει ένα δίκτυο ομοιότητας φαρμάκων όπου τα φάρμακα ορίζονται ως κόμβοι και δημιουργείται ένα πλεονέκτημα μεταξύ δύο φαρμάκων όταν έχουν υψηλή βαθμολογία ομοιότητας. Στη συνέχεια, οι συγγραφείς ανέλυσαν το δίκτυο ναρκωτικών για να ανιχνεύσουν διαφορετικές κοινότητες ναρκωτικών και ερεύνησαν φάρμακα σε κάθε κοινότητα για πιθανή επαναστόχευση φαρμάκων. Οι Bleakley & Yamanishi (2009)⁴² πρότειναν μια στατιστική μέθοδο για την πρόβλεψη των αλληλεπιδράσεων φαρμάκου-στόχου χρησιμοποιώντας πληροφορίες χημικής δομής και πληροφορίες γονιδιωματικής αλληλουχίας. Οι συγγραφείς δημιούργησαν ένα διμερές μοντέλο γραφήματος εποπτευόμενης μάθησης που βασίζεται σε ανεξάρτητα τοπικά εποπτευόμενα μαθησιακά προβλήματα για την πρόβλεψη των πρωτεϊνών-στόχων ενός δεδομένου φαρμάκου και στη συνέχεια για την πρόβλεψη φαρμάκων που στοχεύουν μια δεδομένη πρωτεΐνη.

Ερευνητές συνδύασαν πληροφορίες για τη χημική δομή του φαρμάκου με πληροφορίες στόχους φαρμάκων και αλληλεπιδράσεις για να αναπτύξουν ένα νέο μοντέλο διμερούς γραφήματος για τον υπολογισμό της ομοιότητας ανά ζεύγη φαρμάκων. Τα αποτελέσματα εμπλούτισαν σημαντικά τόσο τη βιοϊατρική βιβλιογραφία όσο και τις κλινικές δοκιμές σε σύγκριση με μια ομάδα ελέγχου χρήσεων φαρμάκων. Η αναπτυγμένη προσέγγιση ξεπέρασε τις άλλες προσεγγίσεις που χρησιμοποιούν μόνο προφίλ στόχων φαρμάκων και συγκέντρωσε τις σιωπηρές πληροφορίες μεταξύ των στόχων ναρκωτικών.

Οι Novick et al., (2013)⁴³ βρήκαν μια νέα μορφή «προφίλ έκφρασης» που βασίζεται σε τρισδιάστατες πληροφορίες χημικής δομής φαρμάκου, πληροφορίες σημασιολογικής ομοιότητας γονιδίων και δίκτυα αλληλεπίδρασης φαρμάκου-στόχου. Οι συγγραφείς έδωσαν βαθμολογίες απόκρισης συναίνεσης (CRS) μεταξύ κάθε φαρμάκου και πρωτεΐνης και χρησιμοποίησαν την απόλυτη τιμή του συντελεστή συσχέτισης μεταξύ κάθε δύο φαρμάκων ως ομοιότητας βαθμού για να δημιουργήσουν ένα δίκτυο ομοιότητας φαρμάκων (DNS), το οποίο οδήγησε στον εντοπισμό νέων ενδείξεων φαρμάκων. Η προτεινόμενη προσέγγιση έλαβε υπόψη τις τρισδιάστατες πληροφορίες χημικής δομής του φαρμάκου για να ξεπεραστεί η αστάθεια των προφίλ γονιδιακής έκφρασης που

αποκτήθηκαν από διαφορετικά πειράματα λόγω πειραματικών συνθηκών όπως το περιβάλλον και η ηλικία του ασθενούς.

B'3.9 Στρατηγική φαινότυπου

Ο φαινότυπος περιγράφεται ως το συνολικό σύνολο πληροφοριών φαινοτυπικών χαρακτηριστικών και έχει προκύψει ως μια νέα στρατηγική για τη σύνδεση των φαρμάκων με κλινικές επιδράσεις για την επαναστόχευση φαρμάκων λόγω του επιχειρήματος ότι αντιπροσωπεύει τα ακούσια αποτελέσματα ενός φαρμάκου και καθορίζει τις φυσιολογικές συνέπειες του βιολογικές δραστηριότητες. Επιπλέον, η φαινοτυπική έκφραση της παρενέργειας ενός φαρμάκου μπορεί να σχετίζεται στενά με τη φαινοτυπική έκφραση μιας νόσου, γεγονός που υποδηλώνει ότι τόσο το φάρμακο όσο και η ασθένεια μπορεί να μοιράζονται παρόμοιες υποκείμενες οδούς.

Οι κλινικές παρενέργειες και οι απροσδόκητες δραστηριότητες που προέρχονται από εκτός στόχων έχει αποδειχθεί ότι έχουν την ικανότητα να χαρακτηρίζουν ανθρώπινα φαινοτυπικά χαρακτηριστικά που σχετίζονται με φάρμακα και μπορεί τελικά να βοηθήσουν στην αποκάλυψη πιθανών θεραπευτικών χρήσεων για αυτά τα φάρμακα. Οι Brouwers et al., (2011)⁴⁴ πρότειναν ένα μέτρο ομοιότητας παρενεργειών που βασίζεται στην ισχυρή συσχέτιση μεταξύ των προφίλ δέσμευσης στοχευμένων μερών και της ομοιότητας παρενεργειών και επαλήθευσε πειραματικά ότι η ομοιότητα παρενεργειών υποδηλώνει νέες θεραπευτικές χρήσεις για υπάρχοντα φάρμακα. Οι Yang & Agarwal (2011)⁴⁵ απέδειξαν ότι οι κλινικές παρενέργειες θα μπορούσαν να χρησιμοποιηθούν για τη δημιουργία ενός φαινοτυπικού προφίλ φαρμάκων και τον εντοπισμό πιθανών νέων ενδείξεων ασθένειας. Ένα σύνολο δεδομένων σχέσης παρενεργειών-φαρμάκων ενσωματώθηκε με ένα σύνολο δεδομένων σχέσης φαρμάκου-ασθένειας για να προκύψουν σχέσεις παρενεργειών-ασθένειας. Στη συνέχεια, οι παρενέργειες χρησιμοποιήθηκαν ως χαρακτηριστικά για την κατασκευή ενός μοντέλου πρόβλεψης για ενδείξεις ασθένειας.

Οι Ye et al. (2014)⁴⁶ κατασκεύασαν ένα δίκτυο ομοιότητας φαρμάκου-φαρμάκου βασισμένο σε κλινικές παρενέργειες υποθέτοντας ότι φάρμακα με παρόμοιες παρενέργειες μπορεί να έχουν παρόμοιες θεραπευτικές ενδείξεις. Εντοπίστηκαν νέες ενδείξεις φαρμάκων επιπλέον των ήδη γνωστών ενδείξεων. Οι Bisgin et al. (2014)⁴⁷ χρησιμοποίησαν πληροφορίες για παρενέργειες για να δημιουργήσει ένα μοντέλο πρόβλεψης νέων θεραπευτικών ενδείξεων για υπάρχοντα φάρμακα. Αξίζει να αναφερθεί ότι απαιτείται βαθύ υπόβαθρο στους μοριακούς μηχανισμούς για τη χρήση πληροφοριών φαινοτυπικών χαρακτηριστικών στην πρόβλεψη νέων ενδείξεων φαρμάκων. Ενώ το μεγαλύτερο μέρος της φαινοτυπικής έρευνας αξιοποιεί δεδομένα από κλινικές μελέτες και ετικέτες φαρμάκων, οι Plachouras et al. (2016)⁴⁸ χρησιμοποίησαν δεδομένα παρενεργειών που εξάχθηκαν από τα μέσα κοινωνικής δικτύωσης για τον εντοπισμό νέων θεραπευτικών ενδείξεων επιπλέον των ενδείξεων που είχαν προσδιοριστεί προηγουμένως.

Τελικά, οι πληροφορίες φαινοτυπικών χαρακτηριστικών μπορούν να ενσωματωθούν με άλλες πηγές δεδομένων, όπως το γονιδίωμα για θεραπευτικές δυνατότητες και νέες ενδείξεις φαρμάκων. Οι

Hoehndorf et al. (2012)⁴⁹ χρησιμοποίησαν φαινοτυπική ομοιότητα για να προσδιορίσει συσχετίσεις γονότυπου-ασθένειας που αργότερα συνδυάστηκαν με δεδομένα συσχέτισης γονότυπου-ασθένειας για να προβλέψουν νέες συσχετίσεις φαρμάκων-ασθένειας. Ένα τέτοιο μοντέλο μπορεί να θεωρηθεί ως εισαγωγή σε ένα ολοκληρωμένο σύστημα για τον εντοπισμό συσχετίσεων φαρμάκων-ασθένειας για ασθένειες με άγνωστη μοριακή βάση. Οι Gottlieb et al. (2012)⁵⁰ ανέπτυξαν ένα μοντέλο χρησιμοποιώντας διάφορα μέτρα ομοιότητας φαρμάκου-φαρμάκου, συμπεριλαμβανομένης της ομοιότητας με βάση το φαινότυπο, για να προβλέψει τις νέες αλληλεπιδράσεις φαρμάκου-φαρμάκου και το επίπεδο σοβαρότητας που σχετίζεται με καθεμία από αυτές τις αλληλεπιδράσεις. Οι Sridhar et al. (2016)⁵¹ ενσωμάτωσαν διαφορετικά μέτρα ομοιότητας φαρμάκου-φαρμάκου, συμπεριλαμβανομένης της φαινοτυπικής ομοιότητας με ήδη γνωστές αλληλεπιδράσεις φαρμάκου-φαρμάκου για να αποκαλύψει τις αλληλεπιδράσεις φαρμάκου-φαρμάκου, συμπεριλαμβανομένων αρκετών νέων αλληλεπιδράσεων.

Το πλεονέκτημα αυτών των προσεγγίσεων έγκειται στην ικανότητά τους να αποκαλύπτουν νέους μηχανισμούς δράσης για τα φάρμακα και να περιλαμβάνουν πιο λεπτομερείς γνώσεις σε μοριακό και γενετικό επίπεδο από τις παραδοσιακές μεθόδους.

B'4 Μέθοδοι Υπολογιστικής Επαναστόχευσης Φαρμάκων

B'4.1 Μηχανική Μάθηση

Οι τεχνικές μηχανικής μάθησης (ML) που εφαρμόζονται στην επαναστόχευση φαρμάκων περιλαμβάνουν λογιστική παλινδρόμηση, μηχανή διανυσμάτων υποστήριξης (SVM), τυχαίο δάσος, νευρωνικό δίκτυο (NN) και βαθιά μάθηση (DL).

Για την λογιστική παλινδρόμηση, το PREDICT είναι ένα πλαίσιο ML που βασίζεται στην ομοιότητα όπου η ομοιότητα φαρμάκου-φαρμάκου ενσωματώνεται με την ομοιότητα ασθένειας-ασθένειας και αυτές οι ενσωματωμένες τιμές ομοιότητας χρησιμοποιούνται ως χαρακτηριστικά στην πρόβλεψη παρόμοιων φαρμάκων για παρόμοιες ασθένειες χρησιμοποιώντας λογιστική παλινδρόμηση. Το SPACE, μια άλλη μέθοδος ML που βασίζεται στην ομοιότητα, χρησιμοποιεί επίσης λογιστική παλινδρόμηση για να προβλέψει τη θεραπευτική χημική κατηγορία ενός φαρμάκου ενσωματώνοντας πολλαπλές πηγές δεδομένων.

Για το SVM, οι Fabris et al. (2016)⁵² προέβλεψαν τη θεραπευτική κατηγορία φαρμάκων χρησιμοποιώντας μια προσέγγιση SVM που βασίζεται στον μοριακό στόχο, τη χημική δομή του φαρμάκου και την ομοιότητα γονιδιακής έκφρασης. Αυτά τα χαρακτηριστικά συγχωνεύτηκαν σε μια ενιαία μήτρα ομοιότητας φαρμάκου για να χρησιμοποιηθεί ως πυρήνας για την ταξινόμηση SVM. Ομοίως, ερευνητές πρότειναν ένα μοντέλο SVM που ενσωματώνει μοριακή δραστηριότητα, χημική δομή φαρμάκου και παρενέργειες, ενσωματώνοντας αυτούς τους τύπους δεδομένων για την κατασκευή μιας συνάρτησης πυρήνα για τον ταξινομητή SVM, με αποτέλεσμα υψηλότερη απόδοση από άλλες μεθόδους.

Για το NN, οι Menden et al., (2013)⁵³ ανέπτυξαν ένα μοντέλο πρόβλεψης βασισμένο σε NN για την ανταπόκριση της καρκινικής κυτταρικής σειράς στη φαρμακευτική θεραπεία, παραμετροποιημένο με IC50. Το μοντέλο τους ανέλυσε γονιδιωματικά χαρακτηριστικά (π.χ. κατάσταση μικροδορυφόρου και κατάσταση μετάλλαξης 77 ογκογονιδίων) και χημικά χαρακτηριστικά (π.χ. δομικό δακτυλικό αποτύπωμα) καρκινικών κυτταρικών γραμμών για να δημιουργήσει ένα perceptron NN και τυχαία παλινδρόμηση δασών.

Το DL, σε σύγκριση με τη ρηχή μάθηση, μπορεί να ανακαλύψει λανθάνουσες και πολύπλοκες δομές σε μεγάλα σύνολα δεδομένων. Χρησιμοποιώντας αλγόριθμους backpropagation, προσαρμόζει τα βάρη σύνδεσης, επιτρέποντας τον υπολογισμό της αναπαράστασης κάθε στρώματος με βάση το προηγούμενο επίπεδο. Οι Aliper et al. (2016)⁵⁴ χρησιμοποίησαν μια προσέγγιση DL για να αναλύσουν τα δεδομένα προφίλ γονιδιακής έκφρασης για να προβλέψουν θεραπευτικές κατηγορίες φαρμάκων, διαπιστώνοντας ότι τα βαθιά νευρωνικά δίκτυα (DNN) ξεπέρασαν το SVM, προτείνοντας το DL ως χρήσιμο εργαλείο στην ανάπτυξη φαρμάκων. Πρόσθετες αναφορές πρότειναν ότι οι προσεγγίσεις που βασίζονται σε DL που χρησιμοποιούν μάθηση πολλαπλών εργασιών ξεπερνούν τους παραδοσιακούς αλγόριθμους ML στην πρόβλεψη της τοξικότητας.

B'.4.2 Μοντέλα Δικτύων

Στα μοντέλα δικτύου, οι κόμβοι αντιπροσωπεύουν φάρμακα, ασθένειες ή γονιδιακά προϊόντα, ενώ οι ακμές αντιπροσωπεύουν αλληλεπιδράσεις ή σχέσεις μεταξύ κόμβων. Αυτά τα δίκτυα, τα οποία είτε βασίζονται στη γνώση είτε συνάγονται υπολογιστικά από πολλαπλές πηγές δεδομένων, αντιπροσωπεύουν διάφορες αλληλεπιδράσεις, όπως φάρμακο-φάρμακο, φάρμακο-στόχος, φάρμακο-νόσος, ασθένεια-νόσος, ασθένεια-γονίδιο, αλληλεπιδράσεις πρωτεΐνης-πρωτεΐνης και μεταγραφικές και δίκτυα σηματοδότησης. Χρησιμοποιώντας την αρχή «ενοχή-από-σύνδεση», η ενσωμάτωση ετερογενών δεδομένων βοηθά στην ανακάλυψη άγνωστων ή κρυφών σχέσεων φαρμάκου-ασθένειας. Σύμφωνα με αυτή την αρχή, φάρμακα που προκαλούν παρόμοιες μεταγραφικές αποκρίσεις θα μπορούσαν να έχουν παρόμοιο τρόπο δράσης.

Χρησιμοποιώντας ένα διμερές δίκτυο, οι Cheng et al., (2012)²⁸ συνέκριναν τις ομοιότητες μεταξύ των αλληλεπιδράσεων που βασίζονται σε φάρμακα, του στόχου και του δικτύου για να προβλέψουν την αλληλεπίδραση φαρμάκου-στόχου, βρίσκοντας ότι το συμπέρασμα βάσει δικτύου έδειξε την καλύτερη απόδοση σε σύγκριση με το περιοχή κάτω από την καμπύλη λειτουργίας του δέκτη (AUROC). Επιπλέον, χρησιμοποιώντας ένα μοντέλο ετερογενούς δικτύου φαρμάκων-ασθένειας, οι Wu et al. (2013)⁵⁶ εντόπισαν στενά συνδεδεμένες ενότητες φαρμάκων και ασθενειών για την εξαγωγή πιθανών υποψηφίων ζευγών φαρμάκων-νόσων για επαναστόχευση φαρμάκων. Οι Jin et al. (2012)¹¹ πρότειναν μια μέθοδο επαναστόχευσης για τα αντικαρκινικά φάρμακα αξιοποιώντας πιθανές επιδράσεις εκτός στόχου στα μονοπάτια σηματοδότησης των καρκινικών κυττάρων.

B'4.3 Εξόρυξη Κειμένων και Σηματολογικό Συμπέρασμα

Η βιοϊατρική και φαρμακευτική βιβλιογραφία περιέχει τεράστιο όγκο πληροφοριών για φάρμακα και ασθένειες. Πιθανές ενδείξεις για υπάρχοντα φάρμακα μπορούν να ανιχνευθούν μέσω σχημάτων εξόρυξης κειμένου. Μια βάση για αυτό το σχήμα είναι η βιολογική οντολογία, η οποία επιτρέπει τη σύγκριση και την ανάλυση βιολογικών πληροφοριών από διάφορες πηγές. Διάφορες προσεγγίσεις εξόρυξης κειμένου για την επαναστόχευση φαρμάκων συνοψίζονται σε μια πρόσφατη μελέτη. Για παράδειγμα, εάν μια μελέτη διαπίστωσε ότι η ανεπάρκεια διατροφής B προκαλούσε τη νόσο A, ενώ μια άλλη μελέτη βρήκε ότι το φάρμακο C ενεργοποιεί τη διατροφή B, η εξόρυξη κειμένου θα συνιστούσε την επαναστόχευση του φαρμάκου C για τη νόσο A.

Το σηματολογικό συμπέρασμα περιλαμβάνει τεχνολογίες όπως η μοντελοποίηση θεμάτων για την ανακάλυψη ενδείξεων φαρμάκων με την ενσωμάτωση διαφόρων πηγών δεδομένων. Για παράδειγμα, οι Guney et al. (2016)⁵⁷ πρότειναν την επαναστόχευση φαρμάκου με βάση την λανθάνουσα κατανομή Dirichlet, ενσωματώνοντας ένα θεματικό μοντέλο για την επεξεργασία πληροφοριών φαινομένων για παρενέργειες φαρμάκου. Με τη μοντελοποίηση των σχέσεων μεταξύ των φαρμάκων για τον καρκίνο του μαστού που έχουν εγκριθεί από τον FDA και των σχετικών γονιδίων, οδών, SNPs και ασθενειών, ερευνητές ανέπτυξαν ένα εργαλείο γνώσης που βασίζεται στην οντολογία για την πρόβλεψη πιθανών ζευγαριών ασθένειας-φαρμάκου στον καρκίνο του μαστού. Οι Cheng et al. (2012)⁵⁵ πρότειναν μια σηματολογική συνδεδεμένη προσέγγιση σε δίκτυο για την αξιολόγηση των συσχετίσεων φαρμάκων-στόχων, η οποία περιελάμβανε φάρμακα, πρωτεϊνικούς στόχους, χημικές ενώσεις, ασθένειες, μονοπάτια, παρενέργειες και τις σχέσεις τους. Στο μοντέλο τους, η τοπολογία και η σηματολογία αντιπροσωπεύονταν από το υπογράφημα μεταξύ ενός φαρμάκου και ενός στόχου, όπου τα ζεύγη φαρμάκου-στόχου που βρίσκονται σε διαφορετικές περιοχές ασθένειας, αλλά παρουσιάζουν ομοιότητα, συσχετίστηκαν με μια πιθανή ευκαιρία επαναστόχευσης.

B'5 Επικύρωση για Υπολογιστική Επαναστόχευση

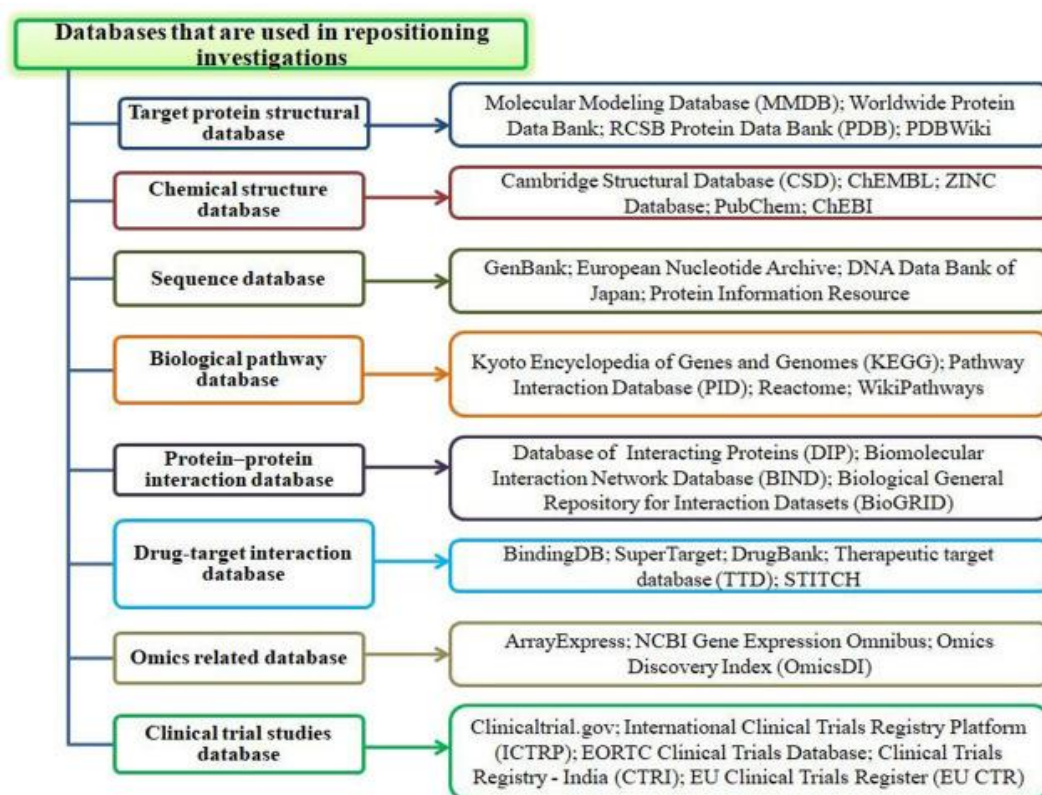
Η επικύρωση των αποτελεσμάτων επαναστόχευσης μπορεί να γίνει υπολογιστικά και πειραματικά. Για την υπολογιστική επικύρωση, ένας απλός τρόπος είναι να αξιολογηθούν οι τιμές AUROC, η ειδικότητα, η ευαισθησία και η θετική προγνωστική αξία (PPV). Η ανάκληση και η ακρίβεια μπορούν να υπολογιστούν για να ληφθεί η καμπύλη ακριβείας-ανάκλησης (PRC), ακολουθούμενη από την περιοχή κάτω από αυτήν την καμπύλη (AUPRC). Επιπλέον, η εγκυρότητα μπορεί να αξιολογηθεί συγκρίνοντας προβλεπόμενους στόχους σε PubMed, ClinicalTrials ή EHR. Για παράδειγμα, οι EHR χρησιμοποιήθηκαν για την επικύρωση των επιδράσεων της μετοφορμίνης που σχετίζονται με τη θνησιμότητα από καρκίνο.

Για πειραματική επικύρωση, πρέπει να πραγματοποιηθούν στοχευμένες αναλύσεις με βάση τα κύτταρα (in vitro και in vivo) και πειράματα σε ζώα. Για παράδειγμα, η ευεργετική επίδραση της ατορβαστατίνης στην επιβίωση του μοσχεύματος παρατηρήθηκε σε μια κοόρτη ενός κέντρου 2.515

ασθενών που έλαβαν μεταμόσχευση νεφρού με αναδρομική ανάλυση των EHR που ακολουθήθηκε για έως και 22 χρόνια, η οποία επικυρώθηκε σε μια μετα-ανάλυση χρησιμοποιώντας δημόσια σύνολα δεδομένων μικροσυστοιχιών όπου η ατορβαστατίνη ήταν βρέθηκε επίσης ότι είναι ευεργετικό για τη μεταμόσχευση οργάνων.

Β.6 Υπολογιστικά εργαλεία που χρησιμοποιούνται για την επαναστόχευση φαρμάκων

Δεδομένου ότι αυτή η προσέγγιση είναι συνήθως πιο οικονομική, η επαναστόχευση φαρμάκων θεωρείται η πιο πολλά υποσχόμενη τεχνική για την ανακάλυψη νέων κλινικών εφαρμογών για ήδη αδειοδοτημένα ή υπάρχοντα φάρμακα. Με τις τεχνολογικές εξελίξεις, συμπεριλαμβανομένης της μεταβολομικής, της πρωτεϊνικής, της γονιδιωματικής, της μεταγραφικής, υπάρχει μια πληθώρα προοπτικών για την ανακάλυψη φαρμάκων συνδυάζοντας όλες τις προαναφερθείσες προσεγγίσεις. Οι επιστήμονες διαθέτουν πλέον τα πιο ενημερωμένα εργαλεία και δεδομένα για τη διερεύνηση νέων ανεξερεύνητων τρόπων δράσης/μονοπατιών που βασίζονται σε πρωτεΐνες/γονίδια στόχους και/ή μοναδικούς βιοδείκτες που συνδέονται με την ανάπτυξη ασθενειών. Έχουν ήδη επινοηθεί πολυάριθμες υπολογιστικές προσεγγίσεις για την επιτάχυνση και την απλούστευση της διαδικασίας επαναστόχευσης. Η Εικόνα 3 παραθέτει μερικές από τις πιο εξέχουσες βάσεις δεδομένων που χρησιμοποιούνται σε αναλύσεις επαναστόχευσης φαρμάκων ¹⁶. Αυτές οι βάσεις δεδομένων παρέχουν σημαντικές πληροφορίες για τις τρισδιάστατες δομές των φαρμάκων και τις θέσεις δέσμευσής τους.

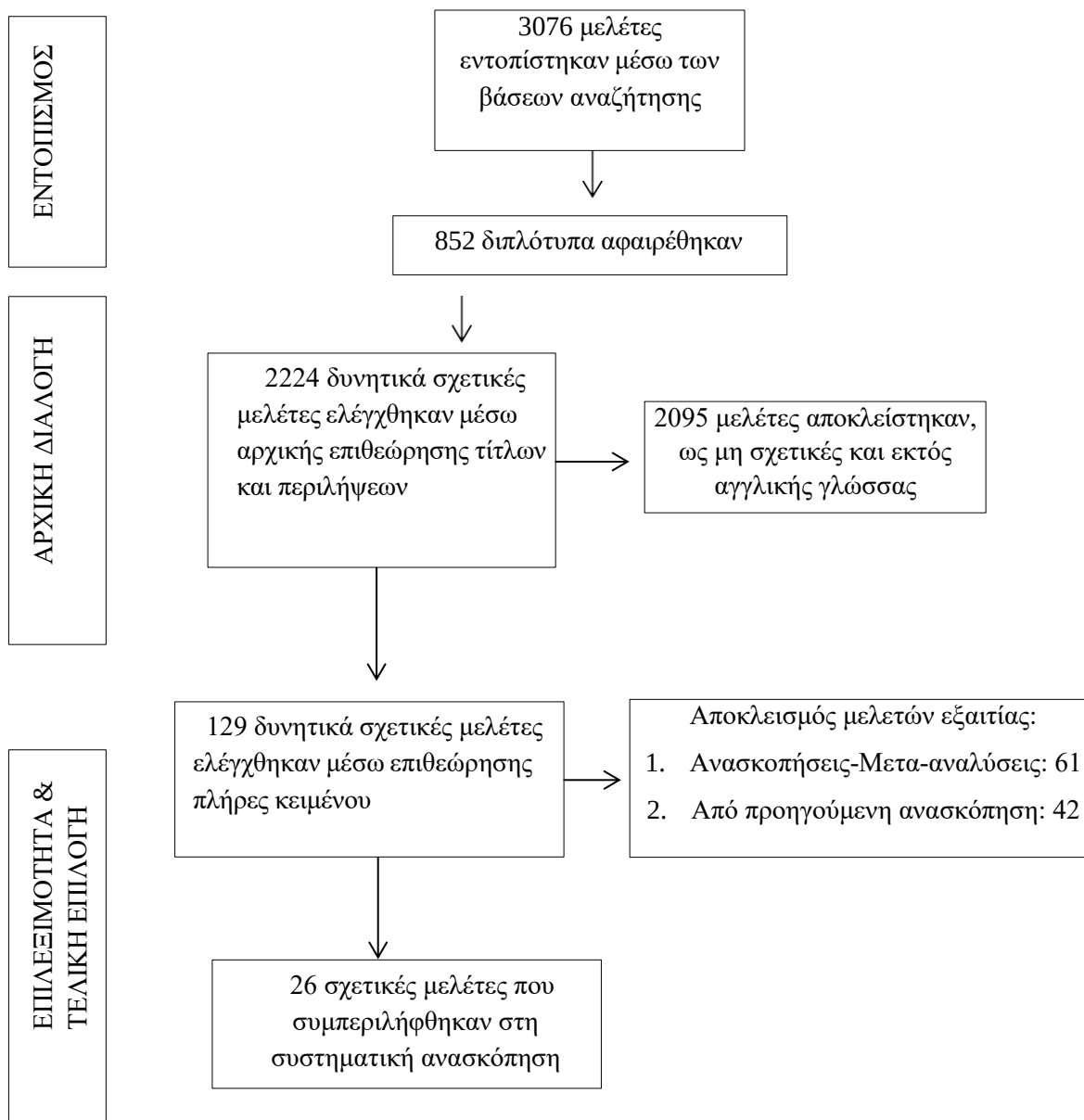


Εικόνα 3: Βάσεις δεδομένων που χρησιμοποιούνται σε μελέτες επαναστόχευσης

Γ' Μεθοδολογία

Η μεθοδολογία που παρουσιάζεται για τις βιβλιογραφικές ανασκοπήσεις ακολουθεί το διάγραμμα PRISMA, το οποίο και χρησιμοποιήθηκε για τη δημιουργία της στρατηγικής αναζήτησης. Πραγματοποιήθηκε μια ολοκληρωμένη αναζήτηση σε βάσεις δεδομένων, όπως IEEE xplore, pubmed και google scholar, για όλα τα σχετικά άρθρα που σχετίζονται με υπολογιστικές μεθόδους για τον επαναστόχευση φαρμάκων. Στην ανασκόπηση συμπεριλήφθηκαν άρθρα που είχαν δημοσιευθεί σε έγκριτα περιοδικά με κριτές. Οι λέξεις-κλειδιά που χρησιμοποιήθηκαν στην αναζήτηση ήταν: 'drug reposition' AND 'computational' OR 'computation' OR 'algorithm' OR 'network' OR 'machine learning' OR 'deep learning' OR 'prediction'. Η αναζήτηση κατέληξε σε 3076 δημοσιεύσεις. Το διάγραμμα ροής PRISMA παρουσιάζεται στην Εικόνα 4.

Μετά την αφαίρεση 852 διπλών άρθρων, παρέμειναν 2.224 μελέτες. Αυτές οι μελέτες εξετάστηκαν μέσω μιας αρχικής ανασκόπησης των τίτλων και των περιλήψεων, με αποτέλεσμα να αποκλειστούν 2.095 μελέτες που ήταν άσχετες ή αποτελούνταν από ανασκοπήσεις (review), μετα-αναλύσεις ή άλλους τύπους μη εμπειρικών εργασιών. Οι υπόλοιπες 129 μελέτες υποβλήθηκαν στη συνέχεια σε ανασκόπηση πλήρους κειμένου για να διαπιστωθεί εάν πληρούσαν τα κριτήρια συμπερίληψης. Αυτή η διαδικασία ανασκόπησης είχε ως αποτέλεσμα την τελική επιλογή 26 μελετών που συμπεριλήφθηκαν στη συστηματική ανασκόπηση. Επομένως ο τελικός αριθμός των μελετών που θα συμπεριληφθούν στην βιβλιογραφική ανασκόπηση αναγράφεται στο διάγραμμα ροής που ακολουθεί..



Εικόνα 4: Διάγραμμα ροής που απεικονίζει τη διαδικασία επιλογής άρθρων

Δ΄ Μελέτες υπολογιστικής επαναστόχευσης φαρμάκων

Δ΄.1 Εξόρυξη δεδομένων

Η τεράστια ποσότητα γονιδίων, φαρμάκων και πληροφοριών που σχετίζονται με ασθένειες που αποθηκεύονται σε βάσεις δεδομένων εκτός από τη ζωντανή βιβλιογραφία που αναπτύσσεται από την ταχεία αύξηση του αριθμού των βιολογικών, βιοϊατρικών και φαρμακευτικών μελετών έχουν οδηγήσει στην ανάγκη για εξόρυξη δεδομένων όπου οι ερευνητές μπορεί να ανακαλύψει έναν τεράστιο όγκο πληροφοριών που κρύβονται στη βιβλιογραφία. Η πλειοψηφία των μελετών που υιοθετούν την προσέγγιση της εξόρυξης δεδομένων χρησιμοποιούν τεχνικές εξόρυξης κειμένου.

Δ΄.2. Εξόρυξη κειμένου

Η εξόρυξη κειμένου όπως εφαρμόζεται στο πρόβλημα επαναστόχευσης φαρμάκων χρησιμοποιείται συνήθως για την εύρεση δεδομένων που σχετίζονται με ένα συγκεκριμένο γονίδιο, ασθένεια ή φάρμακο που καθορίζεται και στη συνέχεια ταξινομεί τις σχετικές οντότητες ή γνώσεις από τα ανακτημένα δεδομένα με βάση τη συνύπαρξη μεταξύ των σχετικών οντοτήτων ή χρησιμοποιώντας επεξεργασία φυσικής γλώσσας. Για παράδειγμα, εάν το φάρμακο R συνδέεται με το γονίδιο G και το γονίδιο G σχετίζεται με τη νόσο D, τότε το φάρμακο R μπορεί να έχει μια νέα σύνδεση με τη νόσο D. Γενικά, η εξόρυξη κειμένου περιλαμβάνει τέσσερα στάδια που είναι: (1) Ανάκτηση πληροφοριών (IR), (2) Αναγνώριση οντοτήτων (NER), (3) Εξαγωγή πληροφοριών (IE) και (4) Ανακάλυψη γνώσης (KD) [94].

Άρθρο 1 Systematic identification of pharmacogenomics information from clinical trials.

Οι Li και Lu⁵⁸ εισήγαγαν ένα μοντέλο για τον προσδιορισμό των σχέσεων κλινικής φαρμακογονιδιοματικής (PGx) γονιδίου-φαρμάκου-ασθένειας από δεδομένα κλινικών δοκιμών. Οι συγγραφείς προσδιόρισαν κείμενο ενδιαφέροντος σε αρχεία κλινικών δοκιμών που ανακτήθηκαν από το ClinicalTrials.gov και χρησιμοποίησαν ένα λεξικό για να προσδιορίσουν τις έννοιες PGx. Στη συνέχεια, εξέτασαν τη συνύπαρξη των εννοιών PGx σε κάθε κλινική δοκιμή για να καθορίσουν τις σχέσεις γονιδίου-φαρμάκου- ασθένειας. Τέλος, ευρετηρίασαν κάθε κλινική δοκιμή χρησιμοποιώντας τις ταυτοποιημένες σχέσεις γονιδίου-φαρμάκου-ασθένειας. Επομένως, δεδομένου ενός γονιδίου PGx, το εισαγόμενο μοντέλο μπορεί να αναγνωρίσει σχετικές ασθένειες και φάρμακα εντός των αντίστοιχων κλινικών δοκιμών. Ομοίως, δεδομένου ενός ζεύγους γονιδίου PGx-φαρμάκου ή γονιδίου-ασθένειας, το εισαγόμενο μοντέλο μπορεί να επιστρέψει κλινικές δοκιμές στις οποίες το ζεύγος PGx έχει μελετηθεί ή έχει μελετηθεί.

Στην αξιολόγηση, διαπιστώθηκε ότι οι ταυτοποιημένες αυτές σχέσεις επικαλύπτονται σημαντικά με την επιμελημένη πραγματική γνώση μέσω της βιβλιογραφίας σε μια βάση δεδομένων PGx και ότι οι περισσότερες σχέσεις εμφανίζονται κατά μέσο όρο 5 χρόνια νωρίτερα στις κλινικές δοκιμές από τις αντίστοιχες δημοσιεύσεις τους, γεγονός που υποδηλώνει ότι οι κλινικές δοκιμές μπορεί να είναι πολύτιμες

τόσο για την επικύρωση γνωστών όσο και για τη λήψη νέων πληροφοριών που σχετίζονται με το PGx με πιο έγκαιρο τρόπο. Αυτή η εργασία έχει πρακτικές επιπτώσεις στον εμπλουτισμό της υπάρχουσας γνώσης μας σχετικά με τις σχέσεις γονιδίου PGx-φαρμάκου-ασθένειας, καθώς και προτείνοντας διασταυρώσεις μεταξύ του ClinicalTrials.gov και άλλων βάσεων γνώσης PGx.

Άρθρο 2 *DNorm: disease name normalization with pairwise learning to rank.*

Οι Leaman et al.⁵⁹ δημιούργησαν ένα εργαλείο για την αναγνώριση οντοτήτων ασθενειών που αναφέρονται στη βιβλιογραφία. Οι συγγραφείς χρησιμοποίησαν το σώμα της νόσου από το Εθνικό Κέντρο Πληροφοριών Βιοτεχνολογίας (NCBI) και το λεξιλόγιο MEDIC για να ξεχωρίσουν ασθένειες που αναφέρονται στις περιλήψεις του PubMed και στη συνέχεια να χειριστούν συντομογραφίες. Στη συνέχεια, χρησιμοποίησαν κατά ζεύγη εκμάθηση κατάταξης, η οποία έχει αποδειχθεί επιτυχής στην ανάκτηση πληροφοριών, για την κανονικοποίηση του αναφερόμενου κειμένου και τον προσδιορισμό των εννοιών MEDIC για τις οντότητες της νόσου που αναφέρονται στις περιλήψεις του PubMed. Ο αλγόριθμός πέτυχε 0,782 μικρομέσες μετρήσεις F και 0,809 μακρομέσες μετρήσεις F και μία αύξηση σε σχέση με τη βασική μέθοδο με την υψηλότερη απόδοση 0,121 και 0,098, αντίστοιχα.

Άρθρο 3 *Building disease-specific drug-protein connectivity maps from molecular interaction networks and PubMed abstracts*

Η εξόρυξη κειμένου έχει επίσης χρησιμοποιηθεί ευρέως με επιτυχία για την ανακάλυψη σχέσεων μεταξύ γονιδίων, ασθενειών και φαρμάκων, τη διερεύνηση των αλληλεπιδράσεων γονιδίου-γονιδίου και τη δημιουργία ενός ετερογενούς δικτύου γονιδίων, ασθενειών και φαρμάκων⁸². Οι Li et al.⁶⁰ πρότειναν μια νέα προσέγγιση που ενσωματώνει δεδομένα εξόρυξης κειμένου βιβλιογραφίας με δίκτυα αλληλεπίδρασης πρωτεϊνών για τη δημιουργία ενός χάρτη συνδεσιμότητας φαρμάκου-πρωτεΐνης για μια συγκεκριμένη ασθένεια. Οι συγγραφείς χρησιμοποίησαν τη νόσο του Alzheimer (AD) ως μελέτη περίπτωσης και έδειξαν ότι η προσέγγισή τους ξεπέρασε τις επιμελημένες βάσεις δεδομένων-στόχων φαρμάκων και τα συμβατικά συστήματα ανάκτησης πληροφοριών και πρότειναν επίσης δύο υπάρχοντα φάρμακα ως υποψήφια φάρμακα για θεραπεία της AD. Πρώτον, ενσωματώθηκαν δίκτυα μοριακής αλληλεπίδρασης για τη μείωση της προκατάληψης και τη βελτίωση της συνάφειας των πρωτεϊνών σπόρων AD. Δεύτερον, οι περιλήψεις του PubMed χρησιμοποιήθηκαν για την ανάκτηση εμπλουτισμένων όρων φαρμάκων που συνδέονται έμμεσα με την AD μέσω μοριακών μηχανιστικών μελετών. Τρίτον και τελευταίο, ένας περιεκτικός χάρτης συνδεσιμότητας AD δημιουργήθηκε με τη συσχέτιση εμπλουτισμένων φαρμάκων και σχετικών πρωτεϊνών στη βιβλιογραφία.

Αποδείχθηκε ότι αυτή η προσέγγιση ανάπτυξης χαρτών μοριακής συνδεσιμότητας ξεπέρασε τις επιμελημένες βάσεις δεδομένων στόχων φαρμάκων και τα συμβατικά συστήματα ανάκτησης

πληροφοριών. Οι αρχικές εξερευνήσεις του χάρτη συνδεσιμότητας AD έδωσαν μια νέα υπόθεση ότι η διλτιαζέμη και η κινιδίνη μπορούν να διερευνηθούν ως υποψήφια φάρμακα για θεραπεία της AD. Οι χάρτες μοριακής συνδεσιμότητας που προέρχονται υπολογιστικά μπορούν να βοηθήσουν στη μελέτη των διαφορών μοριακής υπογραφής μεταξύ διαφορετικών κατηγοριών φαρμάκων σε συγκεκριμένα πλαίσια ασθενειών. Για την επίτευξη συνολικής καλής κάλυψης και ποιότητας δεδομένων, έχουν αναπτυχθεί μια σειρά στατιστικών μεθόδων για να ξεπεραστούν τα υψηλά επίπεδα θορύβου δεδομένων στα βιολογικά δίκτυα και τα αποτελέσματα της εξόρυξης βιβλιογραφίας. Η περαιτέρω ανάπτυξη υπολογιστικών χαρτών μοριακής συνδεσιμότητας για την κάλυψη μεγάλων περιοχών ασθενειών θα δημιουργήσει πιθανότατα ένα νέο μοντέλο για την ανάπτυξη φαρμάκων, στο οποίο τα θεραπευτικά/τοξικολογικά προφίλ των υποψηφίων φαρμάκων μπορούν να ελέγχονται υπολογιστικά πριν ξεκινήσουν δαπανηρές κλινικές δοκιμές.

Άρθρο 4 *Biomedical literature mining. Systematic drug repurposing through text mining.*

Σε αντίθεση με τις κοινές προσεγγίσεις εξόρυξης κειμένου όπου τα βιολογικά δίκτυα χτίζονται με βάση τη συνύπαρξη βιολογικών οντοτήτων, οι Tari et al.⁶¹ εισήγαγαν μια νέα προσέγγιση που έλαβε υπόψη τους τύπους αλληλεπίδρασης, τις κατευθύνσεις τύπου αλληλεπίδρασης και την αναπαράσταση του μηχανισμού φαρμάκων. Οι συγγραφείς χρησιμοποίησαν την εξόρυξη κειμένου για να λάβουν δεδομένα από διαθέσιμες στο κοινό πηγές που στη συνέχεια χρησιμοποιήθηκαν για την παραγωγή ενός συνόλου λογικών γεγονότων. Στη συνέχεια, το σύνολο των λογικών γεγονότων χρησιμοποιήθηκε μαζί με λογικούς κανόνες που αντιπροσωπεύουν τις ιδιότητες του μηχανισμού του φαρμάκου για τη δημιουργία ενός αυτοματοποιημένου συλλογιστικού μοντέλου για τον εντοπισμό θεραπευτικών δυνατοτήτων και νέων ενδείξεων για υπάρχοντα φάρμακα.

Άρθρο 5 *A new method for prioritizing drug repositioning candidates extracted by literature-based discovery*

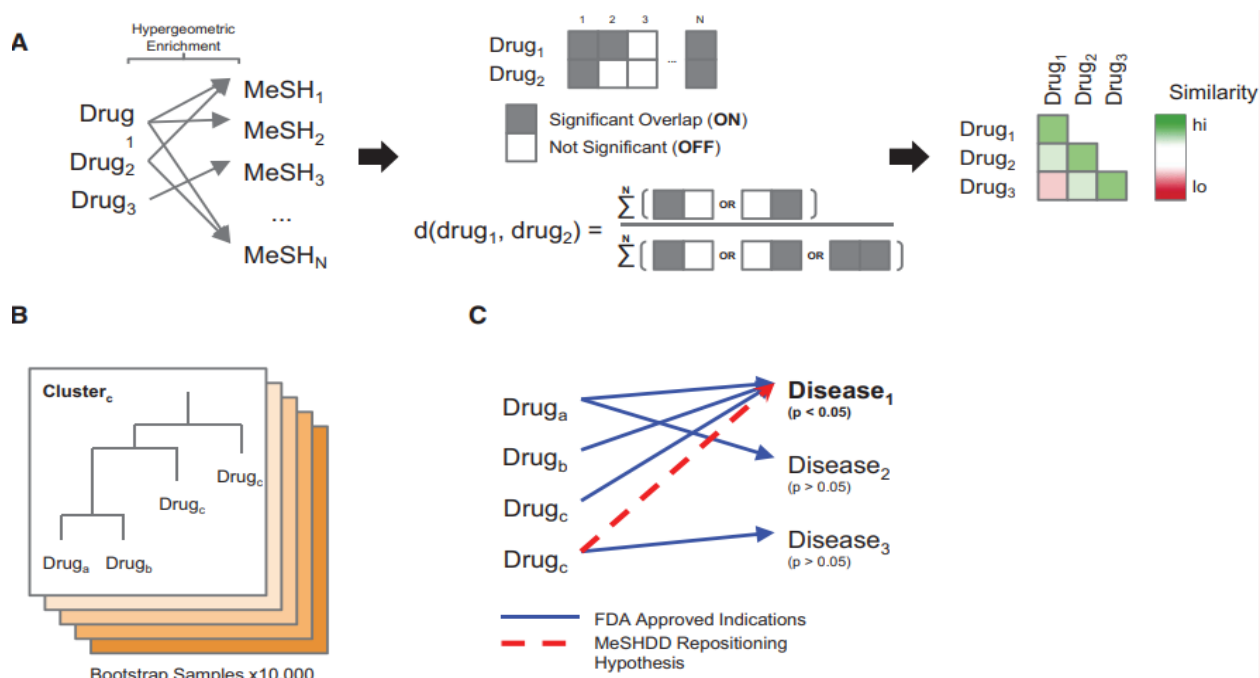
Οι Rastegar-Mojarad et al.⁶² χρησιμοποίησαν δεδομένα εξόρυξης κειμένου για να αναγνωρίσει σημασιολογικές προβλέψεις φαρμάκου-γονιδίου και γονιδίου-ασθένειας, οι οποίες στη συνέχεια χρησιμοποιήθηκαν για τη σύνταξη μιας λίστας πιθανών ζευγών φαρμάκου-ασθένειας. Τέλος, οι συγγραφείς κατέταξαν τα ζεύγη φαρμάκου-ασθένειας χρησιμοποιώντας τις κατηγορίες μεταξύ ζευγών φαρμάκου-γονιδίου και γονιδίου-ασθένειας, αξιολόγησαν το μοντέλο τους σε σχέση με δύο διαφορετικά σύνολα δεδομένων και κατέληξαν στο συμπέρασμα ότι ο συνδυασμός κατηγοριών φαρμάκου-γονιδίου και γονιδίου-ασθένειας θα μπορούσε τελικά να χρησιμοποιηθεί για επισημάνετε τα φάρμακα στα ζεύγη φαρμάκων-ασθένειας στην κορυφή ως υποψήφιοι για επαναστόχευση φαρμάκων.

Συγκεκριμένα, οι συγγραφείς χρησιμοποίησαν σημασιολογικές προδιαγραφές φαρμάκου-γονιδίου και γονιδίου-ασθένειας που εξήχθησαν από περιλήψεις του Medline για να δημιουργήσουμε μια λίστα με

πιθανά ζεύγη φαρμάκων-ασθένειας. Κατά τη σύγκριση των κορυφαίων ζευγών φαρμάκων-ασθένειας έναντι της βάσης δεδομένων συγκριτικής τοξικογονιδιοματικής (CTD), μιας επιμελημένης βάσης δεδομένων για τις σχέσεις φαρμάκων-ασθένειας, διαπιστώθηκε ότι ένα σημαντικό ποσοστό των ζευγών που κατατάσσονται στην κορυφή εμφανίστηκε στο CTD. Η ταυτόχρονη εμφάνιση αυτών των ζευγών υψηλής κατάταξης στις περιλήψεις του Medline βελτιώνει περαιτέρω την εμπιστοσύνη στην προσέγγισή μας να κατατάξουμε τις συναγόμενες σχέσεις φαρμάκων-ασθένειας υψηλότερα στη λίστα. Τέλος, η μη αυτόματη αξιολόγηση των δέκα κορυφαίων ζευγαριών που ταξινομήθηκαν από την προσέγγισή μας αποκάλυψε ότι εννέα από αυτά έχουν κάποια βιολογική σημασία με βάση την κρίση των ειδικών.

Άρθρο 6 MeSHDD: Literature-based drug-drug similarity for drug repositioning

Οι Brown et al.⁶³ πρότειναν ένα διαδικτυακό σύστημα εξόρυξης κειμένου για την επαναστόχευση φαρμάκων. Οι συγγραφείς χρησιμοποίησαν τον αριθμό των κοινών ενδείξεων μεταξύ των ζευγών φαρμάκου-φαρμάκου για να αποκαλύψουν την ομοιότητα μεταξύ αυτών των ζευγών φαρμάκου-φαρμάκου και στη συνέχεια ομαδοποίησαν φάρμακα με βάση την ομοιότητά τους, η οποία αποκάλυψε τόσο γνωστές όσο και νέες ενδείξεις φαρμάκων. Βρέθηκαν ότι οι ομοιότητες φαρμάκων με βάση τη βιβλιογραφία προέβλεπαν τον αριθμό των κοινών ενδείξεων μεταξύ των ζευγών φαρμάκου-φαρμάκου. Η ομαδοποίηση φαρμάκων με βάση την ομοιότητά τους αποκάλυψε τόσο γνωστές όσο και νέες ενδείξεις φαρμάκων. Έδειξαν τη χρησιμότητα της προσέγγισής δημιουργώντας υποθέσεις επαναστόχευσης για το συνήθως χρησιμοποιούμενο φάρμακο διαβήτη μετφορμίνη. Επομένως, η μελέτη τους δείχνει ότι η ομοιότητα που προέρχεται από τη βιβλιογραφία είναι χρήσιμη για τον εντοπισμό πιθανών ευκαιριών επαναστόχευσης. Παρείχαν ένα κώδικα ανοιχτού κώδικα και ανέπτυξαν μια δωρεάν διαδραστική εφαρμογή για να εξερευνήσουν τη βάση δεδομένων μας με ομάδες φαρμάκων που βασίζονται σε ομοιότητες (διαθέσιμο στη διεύθυνση <http://apps.chiragjpgroup.org/MeSHDD>).



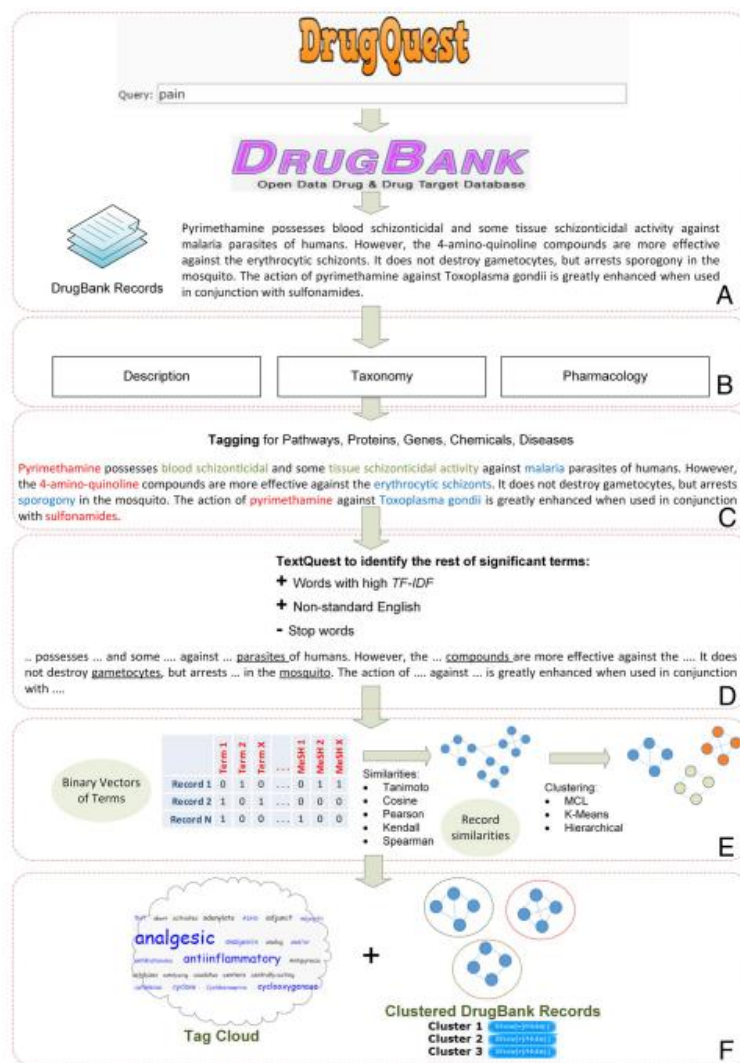
Εικόνα 5: To MeSHDD

Το MeSHDD αξιοποιεί την ομοιότητα της βιβλιογραφίας με ζεύγος φάρμακα και ασθένειες. (A) Η ομοιότητα της βιβλιογραφίας αξιολογείται με τον υπολογισμό της απόστασης μεταξύ 2 φαρμάκων χρησιμοποιώντας τους σημαντικά συσχετιζόμενους όρους MeSH. (B) Τα ισχυρά συμπλέγματα ορίζονται από αποστάσεις ανά ζεύγος χρησιμοποιώντας bootstrapping με 10.000 επαναδειγματοληψίες. (C) Οι υποθέσεις επαναστόχευσης αναπτύσσονται συνδέοντας φάρμακα με νέες, σημαντικά εμπλουτισμένες ενδείξεις.

Άρθρο 7 DrugQuest - a text mining workflow for drug association discovery

Οι Παπανικολάου et al.⁶⁴ εφάρμοσαν την εξόρυξη κειμένου στα χαρακτηριστικά κείμενου της βάσης δεδομένων DrugBank για τον εντοπισμό συσχετίσεων φαρμάκων-φαρμάκων. Οι συγγραφείς χρησιμοποίησαν το Name Entity Recognition (NER) για να προσδιορίσουν βιολογικές οντότητες (πρωτεΐνες, γονίδια, ασθένειες, κ.λπ.) στα πεδία κειμένου περιγραφής, ένδειξης, φαρμακοδυναμικής και τρόπου δράσης της DrugBank. Στη συνέχεια, χρησιμοποίησαν έναν αλγόριθμο για να εξαλείψουν τυχόν ασήμαντους όρους και δημιούργησαν ένα δυαδικό διάνυσμα που αντιπροσωπεύει κάθε εγγραφή DrugBank. Τέλος, ομαδοποίησαν τα αρχεία DrugBank χρησιμοποιώντας αρκετούς αλγόριθμους ομαδοποίησης και μέτρα ομοιότητας. Εφάρμοσαν μια προσέγγιση εξόρυξης κειμένου στη βάση δεδομένων DrugBank προκειμένου να διερευνήσουν συσχετίσεις φαρμάκων με βάση τα πεδία κειμένου "Περιγραφή", "Ένδειξη", "Φαρμακοδυναμική" και "Μηχανισμός Δράσης" της Τράπεζας Φαρμάκων. Εφάρμοσαν τεχνικές Name Entity Recognition (NER) σε αυτά τα πεδία για να αναγνωρίσουν χημικές ουσίες,

πρωτεΐνες, γονίδια, μονοπάτια, ασθένειες και χρησιμοποίησαν τον αλγόριθμο TextQuest για να βρουν πρόσθετες βιολογικά σημαντικές λέξεις.



Εικόνα 6: Η ροή εργασιών της DrugQuest

- Query to DrugBank και ανάκτηση εγγραφών που σχετίζονται με το ερώτημα.
- Η εξόρυξη καταγραφής της DrugBank βασίζεται σε πληροφορίες κειμένου όπως: περιγραφή, τοξικολογία και φαρμακολογία.
- Name Entity Recognition. Τεχνικές για τον εντοπισμό γονιδίων/πρωτεϊνών, χημικών ουσιών, ασθενειών, μονοπατιών.
- Αλγόριθμος TextQuest για τον εντοπισμό σημαντικών όρων χωρίς ετικέτα.
- Τμηματική ομαδοποίηση εγγραφών DrugBank χρησιμοποιώντας διάφορους αλγόριθμους ομαδοποίησης και μέτρα ομοιότητας.

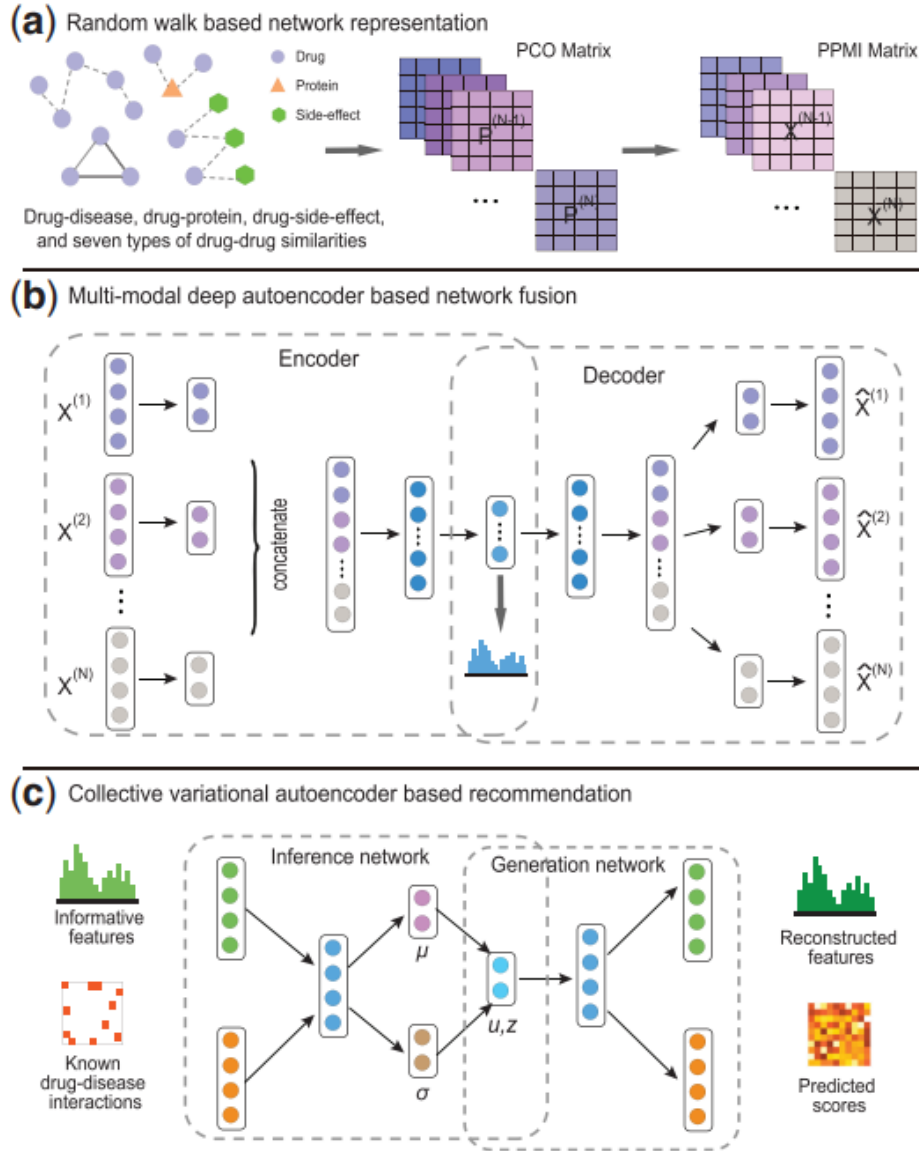
F. Οπτική αναπαράσταση αποτελεσμάτων: Αριστερά: Παράδειγμα σύννεφο ετικετών με εξαιρετικά εμφανιστικούς όρους ανά σύμπλεγμα. Δεξιά: Εγγραφές DrugBank που έχουν εκχωρηθεί σε συμπλέγματα

Το DrugQuest είναι ένα εργαλείο εξόρυξης κειμένου για την ανακάλυψη γνώσης: έχει σχεδιαστεί για να ομαδοποιεί τις εγγραφές της DrugBank με βάση χαρακτηριστικά κειμένου, προκειμένου να βρει νέους συσχετισμούς μεταξύ φαρμάκων. Μια τέτοια προσέγγιση μπορεί διευκολύνει την ανάκτηση νέων συσχετισμών φαρμάκων-φαρμάκων, που μπορεί να συμβάλουν σημαντικά σε νέες εφαρμογές επαναστόχευσης φαρμάκων.

Άρθρο 8 *deepDR: a network-based deep learning approach to in silico drug repositioning*

Οι Zeng et al.⁶⁵ εισήγαγαν μια προσέγγιση βαθιάς μάθησης όπου ανακτούσαν δεδομένα από διάφορες δημοσίως διαθέσιμες πηγές για τη δημιουργία δέκα ετερογενών δικτύων για τον εντοπισμό πιθανών συσχετίσεων φαρμάκων-ασθενειών. Η προτεινόμενη προσέγγιση ξεπέρασε τις συμβατικές προσεγγίσεις στην ανακάλυψη νέων συσχετισμών φαρμάκων-ασθενειών, όταν τα ευρήματά της εξετάστηκαν χρησιμοποιώντας διασταυρούμενη επικύρωση, εξωτερική επικύρωση και μελέτες περιπτώσεων. Επιπλέον, η προσέγγιση πρότεινε πολλούς πιθανούς υποψήφιους επαναστόχευσης φαρμάκων για τη νόσο του Αλτσχάιμερ και του Πάρκινσον. Σε αυτή τη μελέτη, αναπτύχθηκε μια προσέγγιση βαθιάς μάθησης που βασίζεται σε δίκτυο, που ονομάζεται deepDR, για την επαναστόχευση φαρμάκων in silico ενσωματώνοντας 10 δίκτυα: ένα φάρμακο-ασθένεια, ένα φάρμακο-παρενέργεια, ένα φάρμακο-στόχος και επτά δίκτυα φαρμάκων-φαρμάκων.

Συγκεκριμένα, το deepDR μαθαίνει χαρακτηριστικά υψηλού επιπέδου φαρμάκων από τα ετερογενή δίκτυα μέσω ενός πολυτροπικού βαθύ αυτοκωδικοποιητή. Στη συνέχεια, η εκμάθηση χαμηλών διαστάσεων αναπαράστασης φαρμάκων μαζί με κλινικά αναφερόμενα ζεύγη φαρμάκων-ασθενείας κωδικοποιούνται και αποκωδικοποιούνται συλλογικά μέσω ενός μεταβλητού αυτοκωδικοποιητή για να συναχθούν υποψήφιοι για εγκεκριμένα φάρμακα για τα οποία δεν είχαν αρχικά εγκριθεί. Βρέθηκε ότι το deepDR αποκάλυψε υψηλή απόδοση [η περιοχή κάτω από την χαρακτηριστική καμπύλη λειτουργίας του δέκτη (AUROC) = 0,908], ξεπερνώντας τις συμβατικές προσεγγίσεις που βασίζονται στο δίκτυο ή τη μηχανική μάθηση. Είναι σημαντικό ότι οι συσχετίσεις φαρμάκων-ασθενείας προβλεπόμενες από το deepDR επικυρώθηκαν από τη βάση δεδομένων ClinicalTrials.gov (AUROC = 0,826) και παρουσιάστηκαν αρκετά νέα εγκεκριμένα από το deepDR φάρμακα για τη νόσο του Alzheimer (π.χ. ρισπεριδόνη και αριπιπραζόλη) και τη νόσο του Parkinson.



Εικόνα 7: Προσέγγιση deepDR

(a) το deepDR δημιουργεί τυχαία αναπαράσταση δικτύου που βασίζεται σε διαδρομές από ένα περίπλοκο ετερογενές δίκτυο που περιέχει 10 δίκτυα που σχετίζονται με φάρμακα. (b) το deepDR συγχωνεύει πίνακες PPMI (αμοιβαίες πληροφορίες θετικού σημείου) κάθε δικτύου σε μια συμπαγή, χαμηλών διαστάσεων αναπαράσταση χαρακτηριστικών που είναι κοινή σε όλα τα δίκτυα μέσω ενός πολυτροπικού βαθύ αυτοκωδικοποιητή (MDA), στη συνέχεια εξάγονται χαρακτηριστικά χαμηλών διαστάσεων από το μεσαίο στρώμα του MDA. (c) το deepDR χρησιμοποιεί έναν συλλογικό αυτόματο κωδικοποιητή παραλλαγής (cVAE) για να προβλέψει πιθανές συσχετίσεις μεταξύ φαρμάκων και ασθενειών. Τα χαρακτηριστικά φαρμάκων και οι γνωστές (κλινικά αναφερόμενες ή εγκεκριμένες) αλληλεπιδράσεις φαρμάκου-δυσκολίας

κωδικοποιούνται και αποκωδικοποιούνται συλλογικά από το ίδιο δίκτυο συμπερασμάτων και δίκτυο παραγωγής

Δ'3 Σημασιολογικές τεχνολογίες

Οι σημασιολογικές τεχνολογίες επέτρεψαν τον εύκολο συνδυασμό δεδομένων από διαφορετικές πηγές για την πρόβλεψη θεραπευτικών δυνατοτήτων και καινοτόμων ενδείξεων για υπάρχοντα φάρμακα.

Άρθρο 9 *Assessing drug target association using semantic linked data*

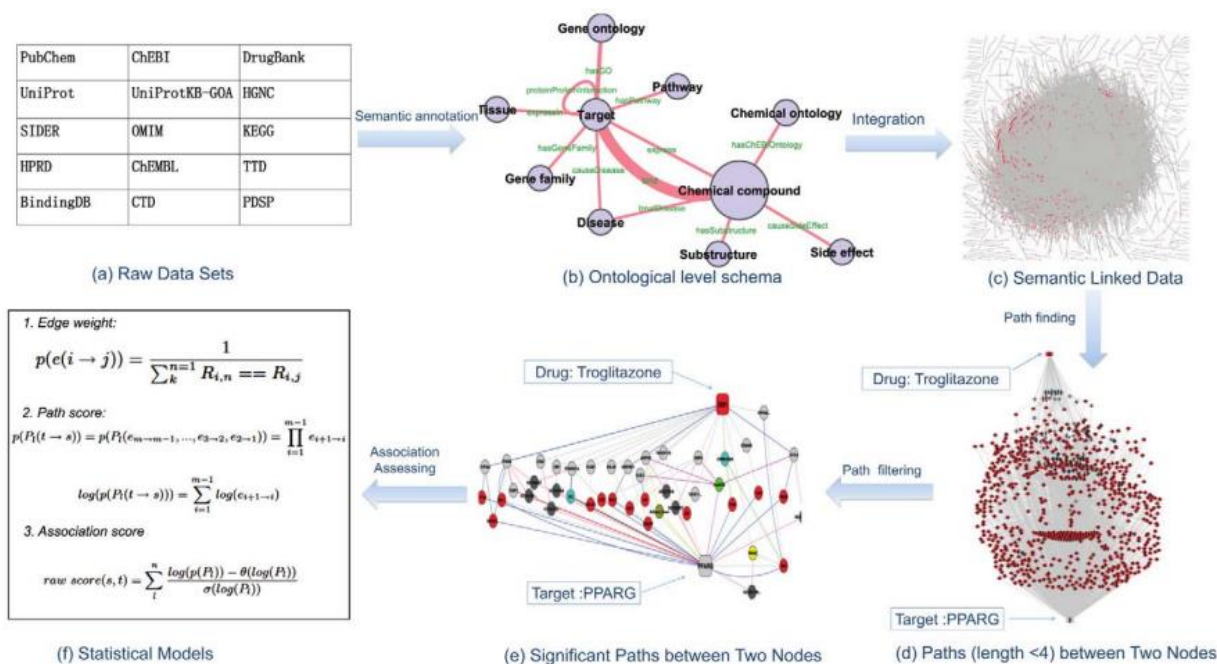
Οι Chen et al.⁶⁶ πρότειναν ένα στατιστικό μοντέλο που βασίζεται στην τοπολογία του δικτύου και τη σημασιολογία του υποδικτύου μεταξύ ενός φαρμάκου και ενός στόχου για την πρόβλεψη συσχετίσεων φαρμάκου-στόχου σε ένα συνδεδεμένο ετερογενές δίκτυο που αποτελείται, σημασιολογικά, από σχολιασμένα δεδομένα που λαμβάνονται από διάφορες δημοσίως διαθέσιμες πηγές, όπως πρωτεΐνη-πρωτεΐνη, φάρμακο-φάρμακο και παρενέργειες φαρμάκου, κ.λπ. -στόχους- ενώσεις. Το στατιστικό μοντέλο που αναπτύχθηκε ονομάζεται Πρόβλεψη συσχέτισης σημασιολογικών συνδέσμων (SLAP) για να αξιολογήσει τη συσχέτιση των ζευγών στόχων φαρμάκων και να προβλέψει τους συνδέσμους που λείπουν. Αποδεικνύεται ότι το SLAP μπορεί να αναγνωρίσει σωστά γνωστά ζεύγη στόχων φαρμάκων από τυχαία ζεύγη με υψηλή ακρίβεια και μπορεί επίσης να αναγνωρίσει σε άμεσες σχέσεις στόχων φαρμάκων (π.χ. την αλλαγή του επιπέδου γονιδιακής έκφρασης). Οι βαθμολογίες συσχέτισης ενός φαρμάκου έναντι ενός συνόλου στόχων αποτελούν μια βιολογική υπογραφή που επιτρέπει την αξιολόγηση της ομοιότητας των φαρμάκων στο πλαίσιο ολόκληρου του συστήματος.

Επιπλέον, ένα δίκτυο ομοιότητας φαρμάκων έδειξε ότι φάρμακα με πολύ διαφορετικές ενδείξεις από διαφορετικές περιοχές ασθένειας συγκεντρώνονται μεταξύ τους, γεγονός που μπορεί να υποδηλώνει θεραπευτικές δυνατότητες και νέες ενδείξεις για αυτά τα φάρμακα. Τα πειράματα επικύρωσης δείχνουν ότι το μοντέλο μπορεί να αναγνωρίσει σωστά γνωστά ζεύγη άμεσων στόχων φαρμάκων με υψηλή ακρίβεια. Τα έμμεσα ζεύγη στόχων φαρμάκων (για παράδειγμα φάρμακα που αλλάζουν το επίπεδο γονιδιακής έκφρασης) αναγνωρίζονται επίσης αλλά όχι τόσο έντονα όσο τα άμεσα ζεύγη. Υπολογίσαμε περαιτέρω τις βαθμολογίες συσχέτισης για 157 φάρμακα από 10 περιοχές ασθενειών έναντι 1683 ανθρώπινων στόχων και μετρήσαμε την ομοιότητά τους χρησιμοποιώντας μια μήτρα βαθμολογίας. Το δίκτυο ομοιότητας δείχνει ότι τα φάρμακα από την ίδια περιοχή ασθένειας τείνουν να συγκεντρώνονται μαζί με τρόπους που δεν αποτυπώνονται από δομική ομοιότητα, με πολλά πιθανά νέα ζεύγη φαρμάκων να εντοπίζονται. Αυτή η εργασία παρέχει έτσι μια νέα, επικυρωμένη εναλλακτική λύση στους υπάρχοντες αλγόριθμους πρόβλεψης στόχων φαρμάκων.

Ένα βασικό ερώτημα στη χρήση οποιασδήποτε μεθόδου πρόβλεψης φαρμάκου-στόχου είναι πόσο εξαρτάται από πλήρη δεδομένα—συγκεκριμένα, εάν χρειάζεται μια πλήρως συμπληρωμένη πειραματική μήτρα για να λειτουργήσει σωστά (δηλ. εκπαίδευση με συνεπείς γνωστές ενεργές/ανενεργές πληροφορίες

για όλες τις ενώσεις έναντι όλων των στόχων). Η μέθοδός μας δεν απαιτεί τέτοια ολοκληρωμένη εκπαίδευση. Οι συγγραφείς θεωρούν ότι συνολικά, το SLAP πρέπει να θεωρείται πολύτιμο εργαλείο για την πρόβλεψη των σχέσεων μεταξύ φαρμάκων και στόχων, χρησιμεύοντας κυρίως ως εργαλείο για τη δημιουργία ιδεών και την πρόταση σχέσεων που θα δοκιμαστούν πειραματικά. Ο σκοπός του είναι να προβλέψει μια σχέση, χωρίς απαραίτητα να υποδηλώνει μια ισχυρή φυσική αλληλεπίδραση. Είναι επίσης χρήσιμο, όπως φαίνεται στο δίκτυο φαρμάκων μας, για τη δημιουργία προφίλ ενώσεων με βάση τις συσχετίσεις-στόχους τους (και αντίστροφα), και σκοπεύουμε να εξερευνήσουμε άλλους τύπους δικτύων που μπορούν να προκύψουν από το SLAP.

Πολλές μέθοδοι πρόβλεψης φαρμάκου-στόχου βασίζονται σε μεμονωμένους τύπους πληροφοριών ή σχέσεων (π.χ. υποδομή, παρενέργειες), γεγονός που τις περιορίζει λόγω ελλιπούς δεδομένων - για παράδειγμα, οι σχέσεις φαρμάκου-στόχου δεν έχουν ολοκληρωθεί. Η χρήση διαφόρων τύπων δεδομένων μπορεί να αντισταθμίσει την έλλειψη πληρότητας σε μεμονωμένες πληροφορίες. Το SLAP δείχνει έναν τρόπο αξιοποίησης τέτοιων πληροφοριών για την πρόβλεψη στόχων φαρμάκων. Για παράδειγμα, η συσχέτιση μεταξύ της φωσφορικής πυριδοξάλης (CID: 1051) και της συζυγούς κυστεΐνης-βήτα λυάσης 2 (CCBL2) είναι πολύ ισχυρή (p-value = 1,9E-3), αλλά εάν αφαιρεθούν οι πληροφορίες γονιδιακής οντολογίας, η συσχέτιση τους γίνεται πολύ αδύναμη (p-value = 0,02). Ομοίως, η συσχέτιση μεταξύ δεξαμεθαζόνης (CID: 5743) και αννεξίνης A1 (ANXA1) θα ήταν δύσκολο να αποτυπωθεί χωρίς να ληφθούν υπόψη οι πληροφορίες υποδομής.



Εικόνα 8: Προσέγγιση SLAP

Μια οντολογία χρησιμοποιείται για τον σχολιασμό δημόσιων συνόλων δεδομένων και την ενσωμάτωσή τους σε ένα σημασιολογικό συνδεδεμένο δίκτυο. Δύο κόμβοι συνδέονται με έναν ή περισσότερους αριθμούς μονοπατιών, αλλά μόνο ένας μικρός αριθμός σημαντικών διαδρομών διατηρείται για εκτίμηση συσχέτισης. Η σημασία της διαδρομής και οι συσχετισμοί στόχου φαρμάκου αξιολογούνται με στατιστικά μοντέλα που προέρχονται από τυχαία δείγματα

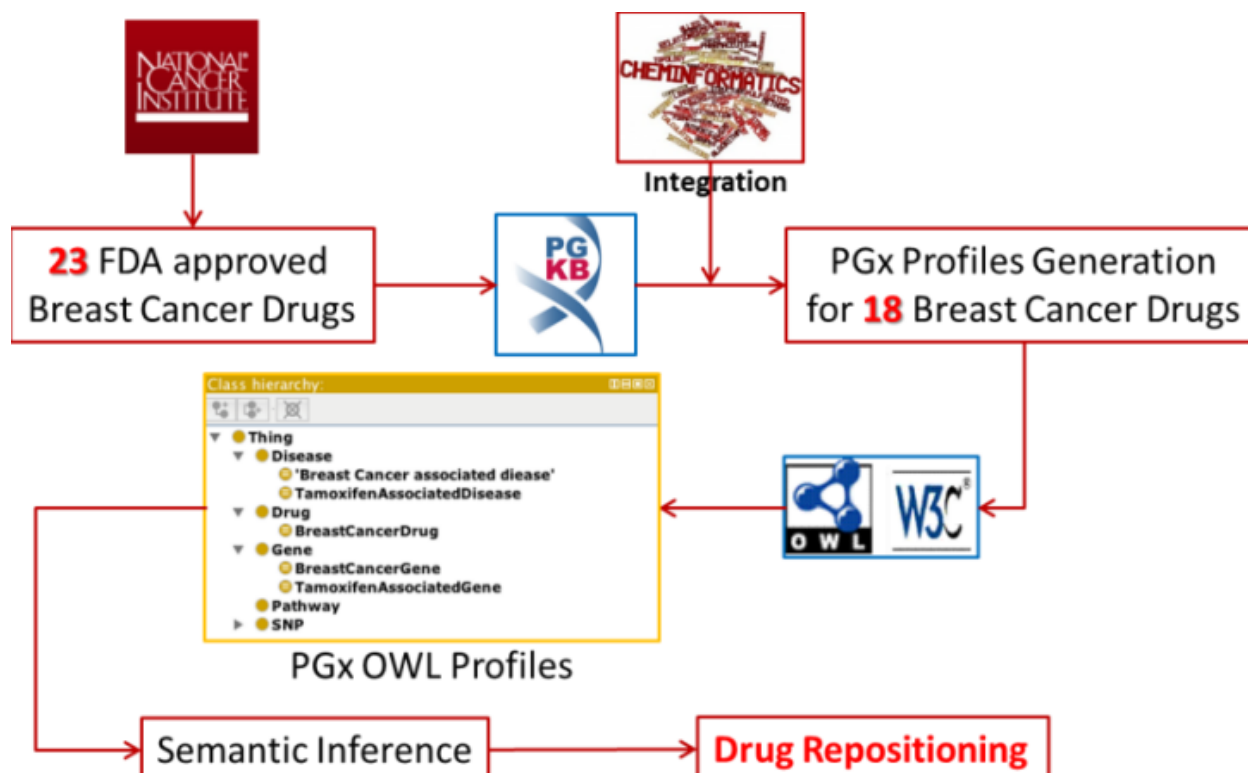
Το πιο επιτακτικό πλεονέκτημα του SLAP είναι η εξέταση των σχέσεων από επίπεδο συστήματος αντί να βασίζεται αποκλειστικά σε γνωστά δεδομένα συγγένειας δέσμησης. Εκτός από τις άμεσες αλληλεπιδράσεις φαρμάκου-στόχου, το SLAP μπορεί επίσης να αναγνωρίσει έμμεσες αλληλεπιδράσεις (π.χ. αλλαγές στα επίπεδα έκφρασης γονιδίων) από τυχαία ζεύγη, αν και οι βαθμολογίες συσχέτισης για αυτές είναι συχνά χαμηλότερες από τις άμεσες αλληλεπιδράσεις. Αυτό μας επιτρέπει να αξιολογήσουμε την ομοιότητα των φαρμάκων με βάση τη βιολογική λειτουργία. Το δίκτυο δείχνει ότι τέτοιες μετρήσεις ομοιότητας όχι μόνο μπορούν να προσδιορίσουν τρόπους δράσης των ναρκωτικών αλλά και να προτείνουν νέες χρήσεις για υπάρχοντα φάρμακα.

Άρθρο 10 *Exploring the pharmacogenomics knowledge base (PharmGKB) for repositioning breast cancer drugs by leveraging Web ontology language (OWL) and cheminformatics approaches*

Οι Zhu et al.⁶⁷ χρησιμοποίησαν δεδομένα κλινικής φαρμακογονιδιοματικής (PGx), συμπεριλαμβανομένων των σχέσεων μεταξύ φαρμάκων, ασθενειών, γονιδίων, μονοπατιών και πολυμορφισμών μεμονωμένων νουκλεοτιδίων (SNPs) και του Σημασιολογικού Ιστού για τη δημιουργία προφίλ φαρμακογονιδιοματικής γλώσσας οντολογίας Ιστού (WOL) και τον εντοπισμό φαρμακογονιδιοματικών συσχετισμών για Ο FDA ενέκρινε φάρμακα για τον καρκίνο του μαστού. Οι συγγραφείς αξιολόγησαν την προσέγγισή τους χρησιμοποιώντας διάφορες περιπτωσιολογικές μελέτες και υπέδειξαν ότι η αξιοποίηση της τεχνολογίας του σημασιολογικού Ιστού κατά τη μελέτη των φαρμακογονιδιοματικών δεδομένων θα μπορούσε να οδηγήσει σε υψηλότερα πρότυπα ευρήματα νέων συσχετίσεων φαρμάκων-ασθένειας και ενδείξεων φαρμάκων.

Συγκεκριμένα, διερεύνησαν το PharmGKB για να αναγνωρίσουν συσχετίσεις που σχετίζονται με τη φαρμακογονιδιοματική ως προφίλ φαρμακογονιδιοματικής για φάρμακα που έχουν εγκριθεί από τον Οργανισμό Τροφίμων και Φαρμάκων των ΗΠΑ (FDA) για τον καρκίνο του μαστού. Στη συνέχεια, μετέτρεψαν και αναπαράστησαν αυτά τα προφίλ σε σημειώσεις Σημασιολογικού Ιστού, οι οποίες υποστηρίζουν αυτοματοποιημένα σημασιολογικά συμπεράσματα. Αξιολόγησαν επιτυχώς την απόδοση και την αποτελεσματικότητα των προφίλ φαρμακογονιδιοματικών φαρμάκων για τον καρκίνο του μαστού με μελέτες περιπτώσεων. Τα αποτελέσματα καταδεικνύουν ότι ο συνδυασμός δεδομένων φαρμακογονιδιοματικής και προσεγγίσεων τεχνολογίας Σημασιολογικού Ιστού/Χημειοπληροφορικής αποδίδει καλύτερη απόδοση νέων ενδείξεων και πιθανών αρνητικών επιπτώσεων για φάρμακα για τον

καρκίνο του μαστού. Αυτή η μελέτη σηματοδοτεί μία προσπάθεια να χρησιμοποιηθεί μία πηγή φαρμακογονιδιοματικής (PGx) σε συνδυασμό με τεχνολογία Σηματολογικού Ιστού για υπολογιστική επαναστόχευση φαρμάκων.



Εικόνα 9: Δομικά στοιχεία για τη συνολική αρχιτεκτονική του PharmGKB

Δ.4 Ταξινόμηση

Άρθρο 11 *PREDICT: a method for inferring novel drug indications with application to personalized medicine*

Οι Gottlieb et al.⁶⁸ χρησιμοποίησαν διάφορες πηγές δεδομένων για την πρόβλεψη συσχετίσεων φαρμάκων-ασθενείων. Οι συγγραφείς χρησιμοποίησαν μέτρα ομοιότητας φαρμάκου-φαρμάκου (π.χ. χημική δομή, παρενέργειες κ.λπ.) και ασθένειας-ασθένειας (π.χ. εκφράσεις γονιδίων, φαινότυπος κ.λπ.) ως χαρακτηριστικά ταξινόμησης. Στη συνέχεια, εφάρμοσαν έναν ταξινομητή λογιστικής παλινδρόμησης για να διακρίνουν μεταξύ αληθινών και ψευδών συσχετίσεων φαρμάκων-ασθένειας και τελικά να προβλέψουν νέες συσχετίσεις φαρμάκων-ασθενείας. Προηγούμενες υπολογιστικές μέθοδοι σε αυτόν τον τομέα επικεντρώθηκαν είτε στην επαναστόχευση φαρμάκου είτε στην αντιστοίχιση προφίλ έκφρασης φαρμάκων και γονιδίων ασθένειας. Εδώ, παρουσιάζεται μια νέα μέθοδο για τη μεγάλης κλίμακας πρόβλεψη των ενδείξεων φαρμάκων (PREDICT) που μπορεί να χειριστεί τόσο εγκεκριμένα φάρμακα όσο και νέα μόρια.

Η μέθοδος των ερευνητών βασίζεται στην παρατήρηση ότι παρόμοια φάρμακα ενδείκνυνται για παρόμοιες ασθένειες και χρησιμοποιεί πολλαπλά μέτρα ομοιότητας φαρμάκου-φαρμάκου και ασθένειας για την εργασία πρόβλεψης. Κατά τη διασταυρούμενη επικύρωση, αποκτά υψηλή ειδικότητα και ευαισθησία ($AUC=0,9$) στην πρόβλεψη ενδείξεων φαρμάκων, ξεπερνώντας τις υπάρχουσες μεθόδους. Επικυρώθηκαν οι προβλέψεις με την αλληλεπικάλυψη τους με τις ενδείξεις φαρμάκων που βρίσκονται υπό κλινικές δοκιμές επί του παρόντος και από τη συμφωνία τους με τις ειδικές για τον ιστό πληροφορίες έκφρασης για τους στόχους φαρμάκων. Έδειξαν περαιτέρω ότι οι γενετικές υπογραφές που σχετίζονται με την ασθένεια μπορούν να χρησιμοποιηθούν για την ακριβή πρόβλεψη των ενδείξεων φαρμάκων για νέες ασθένειες ($AUC=0,92$). Αυτό θέτει τα υπολογιστικά θεμέλια για μελλοντικές εξατομικευμένες θεραπείες φαρμάκων, όπου οι υπογραφές γονιδιακής έκφρασης από μεμονωμένους ασθενείς θα αντικαθιστούσαν τις υπογραφές για τη συγκεκριμένη ασθένεια.

Δ' 5 Μηχανική μάθηση

Η υπολογιστική επαναστόχευση φαρμάκων έχει εξελιχθεί τις τελευταίες δύο δεκαετίες από απλές προσπάθειες ομοιότητας φαρμάκων, οι οποίες συχνά χρησιμοποιούσαν μια ενιαία πηγή βιολογικών ή βιοϊατρικών δεδομένων, σε έναν καινοτόμο τομέα εφαρμογής για προσεγγίσεις μηχανικής μάθησης. Παρόμοια με τα μοντέλα μηχανικής μάθησης σε άλλους τομείς, τα υπολογιστικά μοντέλα εναπόθεσης φαρμάκων απαιτούν εκτεταμένο όγκο δεδομένων για να εκπαιδεύσουν αυτά τα μοντέλα και να βρουν ισχυρούς κανόνες απόφασης, με στόχο να αποκαλύψουν τις υποκείμενες συσχετίσεις μεταξύ βιολογικών και βιοϊατρικών οντοτήτων. Η τεράστια αύξηση του όγκου των διαθέσιμων στο κοινό βιολογικών και βιοϊατρικών δεδομένων και η πολύτιμη πρόοδος που προκύπτει από μοντέλα μηχανικής μάθησης σε άλλους κλάδους έχει βοηθήσει τη σημαντική προσπάθεια για τη δημιουργία, τη μελέτη και τη χρήση μεθόδων μηχανικής μάθησης για ανακάλυψη. - νεοσύστατες ενώσεις φαρμάκων-ασθένειες και εφαρμογές επαναστόχευσης φαρμάκων. Τέτοιες μέθοδοι χρησιμοποίησαν Naïve Bayesian, k-κοντινότεροι γείτονες (kNN), μηχανές υποστήριξης διανυσμάτων (SVM) και πιο πρόσφατα βαθιά νευρωνικά δίκτυα ⁸³ για δυαδική ταξινόμηση, ταξινόμηση πολλαπλών κλάσεων, και πρόβλεψη τιμών. Οι Chen et al.⁸³ συζήτησαν παραδείγματα που καλύπτουν την πρόβλεψη βιοδραστικότητας, τον de novo μοριακό σχεδιασμό, την πρόβλεψη σύνθεσης και την ανάλυση βιολογικών εικόνων.

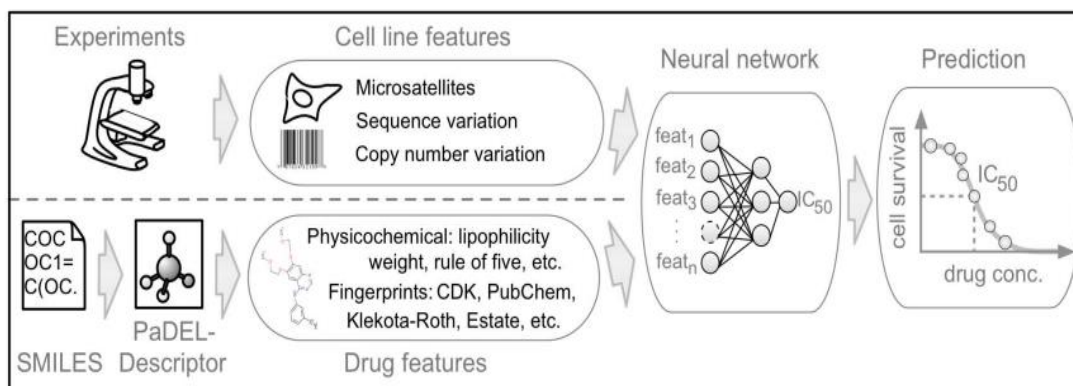
Άρθρο 12 *Machine Learning Prediction of Cancer Cell Sensitivity to Drugs Based on Genomic and Chemical Properties*

Οι Menden et al.⁶⁹ ανέπτυξαν μοντέλα μηχανικής μάθησης για την πρόβλεψη των αντιδράσεων του καρκίνου στη φαρμακευτική θεραπεία χρησιμοποιώντας το συνδυασμό γονιδιωματικών κυτταρικών γραμμών και χημικών δομών φαρμάκων. Οι συγγραφείς ενσωμάτωσαν και τις δύο πηγές δεδομένων για να δημιουργήσουν ένα μοντέλο νευρωνικού δικτύου perceptron τροφοδοσίας και ένα μοντέλο τυχαίας

παλινδρόμησης δασών και στη συνέχεια επικύρωσαν τα ευρήματά τους με διασταυρούμενη επικύρωση και μια ανεξάρτητη τυφλή δοκιμή. Ισχυρίστηκαν ότι η χρήση τέτοιων μοντέλων θα μπορούσε να προχωρήσει περισσότερο από τον εικονικό έλεγχο φαρμάκων, καθώς δοκίμαζε συστηματικά την αποτελεσματικότητα των φαρμάκων και έτσι εντόπισε πιθανές εφαρμογές επαναστόχευσης φαρμάκων και τελικά θα μπορούσε να είναι χρήσιμη για εξατομικευμένη ιατρική, συνδέοντας τη γονιδιωματική των κυτταρικών σειρών με τη δυσανεξία στα φάρμακα.

Οι έλεγχοι υψηλής απόδοσης δυνητικά ενεργών ενώσεων έναντι μιας ομάδας γονιδιωματικά ετερογενών κυτταρικών γραμμών καρκίνου έχουν αποκαλύψει πολλαπλές σχέσεις μεταξύ γονιδιωματικών αλλαγών και ανταποκρίσεων στα φάρμακα. Έχουν προταθεί διάφορες υπολογιστικές προσεγγίσεις για την πρόβλεψη της ευαισθησίας με βάση τα γονιδιωματικά χαρακτηριστικά, ενώ άλλες έχουν χρησιμοποιήσει τις χημικές ιδιότητες των φαρμάκων για να εξακριβώσουν την επίδρασή τους. Σε μια προσπάθεια να ενσωματωθούν αυτές τις συμπληρωματικές προσεγγίσεις, οι ερευνητές ανέπτυξαν μοντέλα μηχανικής μάθησης για την πρόβλεψη της ανταπόκρισης των καρκινικών κυτταρικών γραμμών στη θεραπεία με φάρμακα, ποσοτικοποιημένα μέσω των τιμών IC₅₀, με βάση τόσο τα γονιδιωματικά χαρακτηριστικά των κυτταρικών γραμμών όσο και τις χημικές ιδιότητες των εξεταζόμενων φαρμάκων. Τα μοντέλα προέβλεψαν τιμές IC₅₀ σε μια 8-πλάσια διασταυρούμενη επικύρωση και μια ανεξάρτητη τυφλή δοκιμή με συντελεστή προσδιορισμού R^2 0,72 και 0,64 αντίστοιχα. Επιπλέον, τα μοντέλα ήταν σε θέση να προβλέψουν με συγκρίσιμη ακρίβεια (R^2 0,61) IC₅₀ κυτταρικών σειρών από έναν ιστό που δεν χρησιμοποιήθηκε στο στάδιο της εκπαίδευσης.

Τα μοντέλα *in silico* μπορούν να χρησιμοποιηθούν για τη βελτιστοποίηση του πειραματικού σχεδιασμού των διαλογών φαρμάκων-κυττάρων, υπολογίζοντας ένα μεγάλο ποσοστό τιμών IC₅₀ που λείπουν αντί να τις μετρήσουν πειραματικά. Οι συνέπειες των αποτελεσμάτων μας ξεπερνούν το σχεδιασμό εικονικού ελέγχου φαρμάκων: δυνητικά χιλιάδες φάρμακα θα μπορούσαν να διερευνηθούν *in silico* για να ελεγχθεί συστηματικά η πιθανή αποτελεσματικότητά τους ως αντικαρκινικοί παράγοντες με βάση τη δομή



Εικόνα 10: Ροή εργασιών πρόβλεψης IC₅₀

τους, παρέχοντας έτσι ένα υπολογιστικό πλαίσιο για τον εντοπισμό νέων ευκαιριών επαναστόχευσης φαρμάκων επίσης ως τελικά χρήσιμο για την εξατομικευμένη ιατρική, συνδέοντας τα γονιδιωματικά χαρακτηριστικά των ασθενών με την ευαισθησία στα φάρμακα.

Η μέθοδος των Menden et al.⁶⁹ βασίζεται σε δύο διαφορετικές ροές εισόδου: (1) χαρακτηριστικά κυτταρικής γραμμής 77 ογκογονιδίων και την κατάσταση μετάλλαξής τους, (2) χαρακτηριστικά φαρμάκου που δημιουργούνται με το λογισμικό PaDEL [19] από το απλοποιημένο σύστημα εισόδου γραμμής μοριακής εισόδου (SMILES), δείτε την ενότητα μεθόδων για λεπτομέρειες. Η συνεχής τιμή IC50 προβλέπεται με προηγμένους αλγόριθμους μηχανικής μάθησης (νευρωνικά δίκτυα και τυχαία δίκτυα)

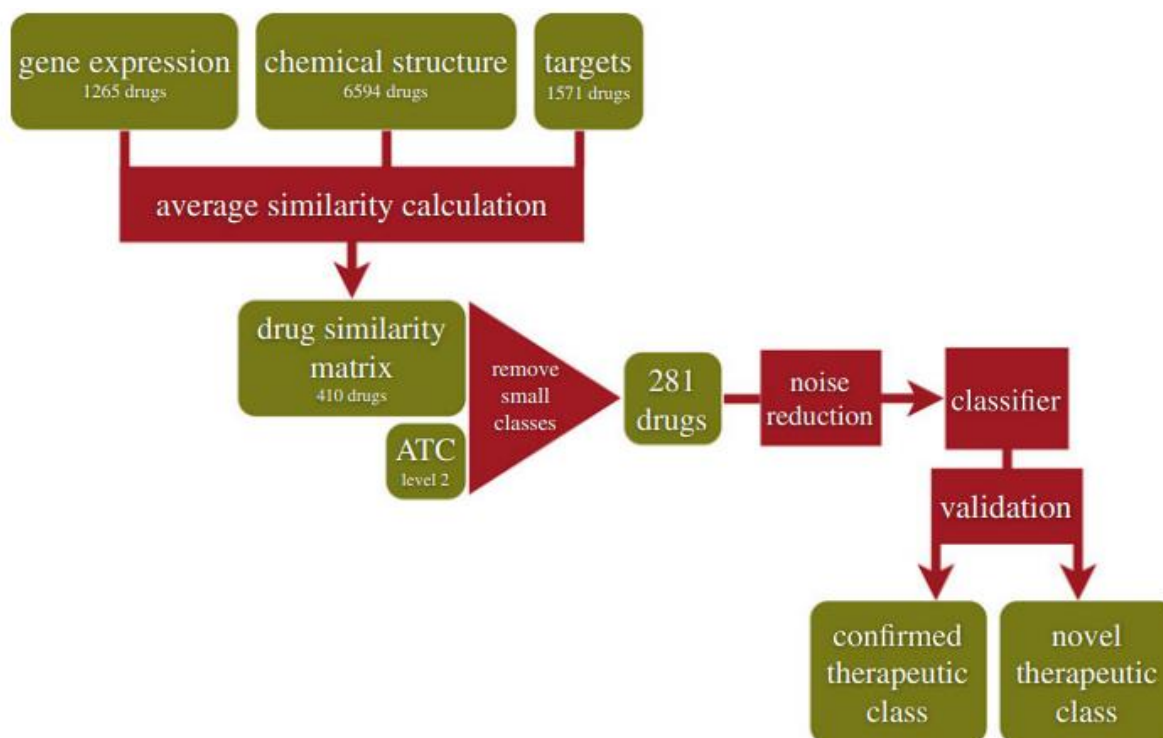
Δ'.6 Συνεργατική αλληλεπίδραση

Αξίζει να σημειωθεί ότι αρκετές μελέτες που βασίζονται στη μηχανική μάθηση έχουν εφαρμόσει συνεργατική αλληλεπίδραση, η οποία εξαρτάται από ιστορικές τάσεις όπως η έκφραση γονιδίων σε διαφορετικά δείγματα, για την πρόβλεψη νέων ενδείξεων φαρμάκων και συσχετίσεων φαρμάκων-ασθενειών.

Άρθρο 13 Drug repositioning: a machine-learning approach through data integration

Οι Napolitano et al.⁷⁰ χρησιμοποίησαν διάφορα σύνολα δεδομένων ομοιότητας που σχετίζονται με φάρμακα ως φτερά για να προβλέψει τη θεραπευτική κατηγορία των εγκεκριμένων από τον FDA ενώσεις και σκόπιμα εξέτασε τυχόν αναντιστοιχίες μεταξύ γνωστών και προβλεπόμενων ταξινομήσεων φαρμάκων ως πιθανές εναλλακτικές θεραπευτικές ενδείξεις. Οι συγγραφείς συνδύασαν τρία σύνολα δεδομένων ομοιότητας φαρμάκου-φαρμάκου, βασισμένα σε υπογραφές γονιδιακής έκφρασης, χημικές δομές και μοριακούς στόχους, σε έναν πίνακα ομοιότητας φαρμάκου, ο οποίος χρησιμοποιήθηκε ως πυρήνας για την εκπαίδευση ενός ταξινομητή πολλαπλών κλάσεων Support Vector Machine (SVM). Στη συνέχεια, χρησιμοποίησαν συνεργατικές τεχνικές φιλτραρίσματος για να προβλέψουν νέες ενδείξεις φαρμάκων.

Προτείνουν μια νέα υπολογιστική προσέγγιση για την πρόβλεψη της επαναστόχευσης φαρμάκων με βάση τους υπερσύγχρονους αλγόριθμους μηχανικής μάθησης. Έχουν ενσωματώσει πολλαπλά στρώματα πληροφοριών: i) για τις αποστάσεις των φαρμάκων με βάση το πόσο όμοιες είναι οι χημικές δομές τους, ii) πόσο κοντά είναι οι στόχοι τους στο δίκτυο αλληλεπίδρασης πρωτεΐνης-πρωτεΐνης και iii) σχετικά με το πόσο συσχετισμένοι είναι το γονίδιο μοτίβα έκφρασης μετά τη θεραπεία. Ο ταξινομητής φτάνει σε υψηλά επίπεδα ακρίβειας (78%), επιτρέποντας να ερμηνευτούν εκ νέου τις κορυφαίες εσφαλμένες ταξινομήσεις ως επαναταξινομήσεις, μετά από αυστηρή στατιστική αξιολόγηση. Η αποτελεσματική επαναστόχευση φαρμάκων έχει τη δυνατότητα να επηρεάσει σημαντικά ολόκληρο το πεδίο ανάπτυξης φαρμάκων. Τα αποτελέσματα που παρουσιάζονται μπορούν να επιταχύνουν σημαντικά τη μετάφραση στις κλινικές γνωστών ενώσεων για νέες θεραπευτικές χρήσεις.



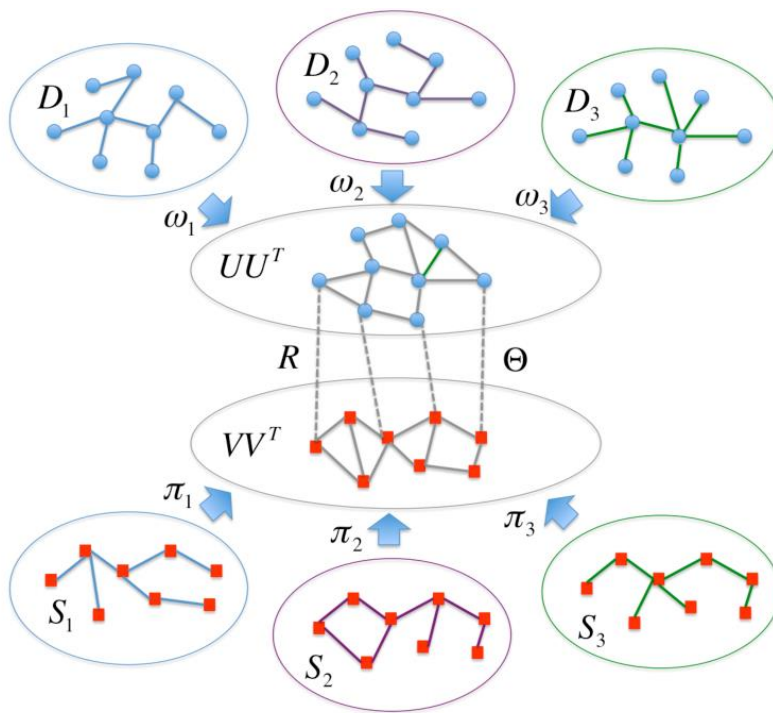
Εικόνα 11: Διάγραμμα ροής της ανάλυσης των Napolitano et al

Τα πράσινα πλαίσια υποδεικνύουν δεδομένα, τα κόκκινα πλαίσια υποδεικνύουν διαδικασίες.

Άρθρο 14 *Towards drug repositioning: a unified computational framework for integrating multiple aspects of drug similarity and disease similarity*

Οι Zhang et al.⁷¹ εισήγαγαν ένα ενιαίο υπολογιστικό πλαίσιο που ονομάζεται DDR για την ενοποίηση πολυάριθμων βιολογικών και βιοϊατρικών πηγών προκειμένου να συναχθούν νέες ομοιότητες φαρμάκου-φαρμάκου καθώς και ομοιότητες ασθένειας-ασθένειας. Ως απάντηση στο υψηλό κόστος και τον υψηλό κίνδυνο που σχετίζεται με την παραδοσιακή de novo ανακάλυψη φαρμάκων, η διερεύνηση πιθανών πρόσθετων χρήσεων για υπάρχοντα φάρμακα, γνωστή και ως επαναστόχευση φαρμάκων, έχει προσελκύσει αυξανόμενη προσοχή τόσο από τη φαρμακευτική βιομηχανία όσο και από την ερευνητική κοινότητα.

Οι συγγραφείς ενσωμάτωσαν ομοιότητες φαρμάκων (π.χ. πρωτεΐνες-στόχους, παρενέργειες και χημική δομή), ομοιότητες ασθενειών (π.χ. συσχετίσεις γονιδίων-ασθένειες και φαινότυπο ασθένειας) και γνωστά σύνολα δεδομένων συσχετίσεων φαρμάκων-ασθένειας για να δημιουργήσουν ένα δίκτυο φαρμάκων-ασθένειας. Το δίκτυο φαρμάκων-ασθένειας αντιμετωπίστηκε ως πρόβλημα βελτιστοποίησης, το οποίο επιλύθηκε χρησιμοποιώντας τη στρατηγική καθόδου συντεταγμένων μπλοκ (BCD). Τα αποτελέσματα έδειξαν ότι ένα τέτοιο πλαίσιο θα μπορούσε να είναι χρήσιμο για την εύρεση νέων συσχετίσεων φαρμάκων-ασθένειας και τον εντοπισμό νέων ευκαιριών επαναστόχευσης φαρμάκων.



Εικόνα 12: Μια γραφική απεικόνιση της κύριας ιδέας του DDR

Υπάρχουν πολλές πηγές πληροφοριών που μπορούμε να χρησιμοποιήσουμε για να κατασκευάσουμε ομοιότητες φαρμάκου/ασθένειας και οι κατασκευασμένοι πίνακες ομοιότητας φαρμάκου/ασθένειας συμβολίζονται με $\{D_k\}_{k=1}^K$ ή $\{S_l\}_{l=1}^L$. Υπάρχει επίσης μια παρατηρούμενη μήτρα συσχέτισης φαρμάκου-ασθένειας R . Στη συνέχεια, η DDR μπορεί να μάθει τη μήτρα ομαδοποίησης φαρμάκων/νόσων U ή V , την εκτιμώμενη μήτρα συσχέτισης φαρμάκου-ασθένειας Θ και τη σημασία των διαφορετικών πηγών πληροφοριών φαρμάκου/ασθένειας.

Συγκεκριμένα, το προτεινόμενο ενιαίο υπολογιστικό πλαίσιο, που ονομάζεται DDR, για την πρόβλεψη νέων συσχετίσεων φαρμάκων-ασθενειών διατυπώνει το έργο της δημιουργίας υποθέσεων για την επαναστόχευση του φαρμάκου ως ένα περιορισμένο πρόβλημα μη γραμμικής βελτιστοποίησης. Χρησιμοποιεί πολλαπλά δίκτυα ομοιότητας φαρμάκων, πολλαπλά δίκτυα ομοιότητας ασθενειών και γνωστούς συσχετισμούς φαρμάκων-ασθενειών για να διερευνήσει πιθανές νέες συσχετίσεις μεταξύ φαρμάκων και ασθενειών χωρίς γνωστές συνδέσεις. Μια μεγάλης κλίμακας μελέτη διεξήχθη χρησιμοποιώντας 799 φάρμακα κατά 719 ασθενειών. Τα πειραματικά αποτελέσματα έδειξαν την αποτελεσματικότητα της προσέγγισης.

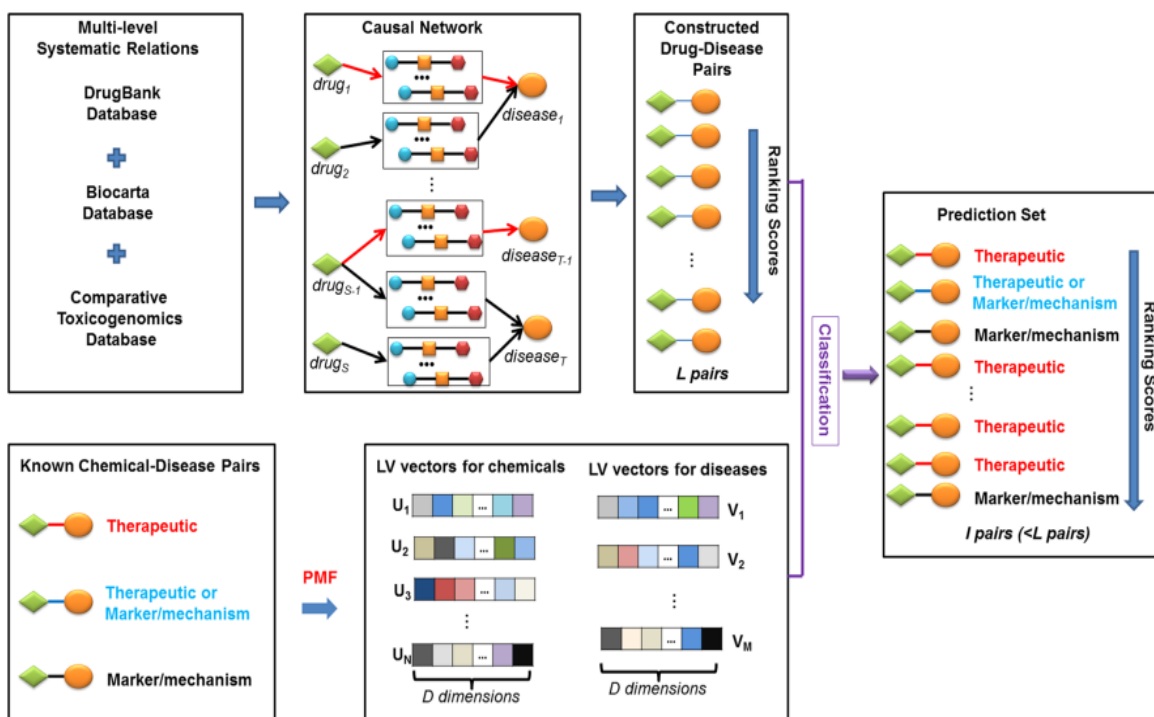
Επιπλέον, η DDR κατέταξε τις πηγές πληροφοριών φαρμάκων και ασθενειών με βάση τη συμβολή τους στην πρόβλεψη, ανοίγοντας έτσι τον δρόμο για την ιεράρχηση πολλαπλών πηγών δεδομένων και τη δημιουργία πιο αξιόπιστων μοντέλων επαναστόχευσης φαρμάκων. Συγκεκριμένα, ορισμένες από τις νέες

προβλέψεις μας για συσχετίσεις φαρμάκων-ασθένειες υποστηρίχθηκαν από βάσεις δεδομένων κλινικών δοκιμών, δείχνοντας ότι το DDR θα μπορούσε να χρησιμεύσει ως χρήσιμο εργαλείο στην ανακάλυψη φαρμάκων για τον αποτελεσματικό εντοπισμό πιθανών νέων χρήσεων για υπάρχοντα φάρμακα.

Άρθρο 15 Drug–Disease Association and Drug-Repositioning Predictions in Complex Diseases Using Causal Inference–Probabilistic Matrix Factorization

Οι Yang et al.⁷² παρουσίασαν μια προσέγγιση παραγοντοποίησης μήτρας με αιτιώδη συμπεράσμα-πιθανότητα (CI-PMF) για τον εντοπισμό και την ταξινόμηση των συσχετίσεων φαρμάκων-ασθένειας. Οι συγγραφείς χρησιμοποίησαν διάφορες βιολογικές και βιοϊατρικές πηγές (π.χ. στόχους φαρμάκων, μονοπάτια, γονίδια που σχετίζονται με μονοπάτια και συσχετίσεις γονιδίων ασθένειας) για να οικοδομήσουν ένα αιτιολογικό δίκτυο που συνδέει το φάρμακο, το στόχο, το μονοπάτι, το γονίδιο, τις οντότητες της νόσου μεταξύ τους προκειμένου να κατατάσσουν συλλόγους φαρμάκων-ασθένειας.

Επιπλέον, αξιοποίησαν γνωστές συσχετίσεις φαρμάκων-ασθένειας για να σχηματίσουν ένα μοντέλο πιθανοτικής παραγοντοποίησης μήτρας (PMF), το οποίο χρησιμοποιήθηκε για την κατασκευή ενός μοντέλου PMF για την ταξινόμηση κατασκευασμένων συσχετίσεων φαρμάκων-ασθένειας σε διαφορετικές κατηγορίες. Τέλος, εκμεταλλεύτηκαν τις βαθμολογίες κατάταξης συσχέτισης φαρμάκων-ασθένειας και πρόβλεψαν κατηγορίες για να προσδιορίσουν τη νέα συσχέτιση φαρμάκων-ασθένειας.



Εικόνα 13: Σχέδιο προσέγγισης CI-PMF

Το δίκτυο αιτιολογίας φάρμακο-στόχος-μονοπάτι-γονίδιο-ασθένεια κατασκευάζεται αρχικά για τον υπολογισμό της βαθμολογίας συσχέτισης (Sa) και οι παρατηρούμενοι συσχετισμοί χημικής ασθένειας από το CTD χρησιμοποιούνται για την εκμάθηση μοντέλων PMF. Στη συνέχεια, το κατασκευασμένο μοντέλο PMF χρησιμοποιείται για την ταξινόμηση των κατασκευασμένων συσχετισμών σε διαφορετικές ομάδες, συγκεκριμένα, "θεραπευτικοί", "δείκτης/μηχανισμός" και "θεραπευτικός δείκτης/μηχανισμός". Τέλος, η θεραπευτική συσχέτιση φαρμάκου-ασθένειας προβλέπεται με βάση την κατάταξη της βαθμολογίας συσχέτισης και τον προβλεπόμενο τύπο. [Θρύλος συμβόλων: διαμάντι, φάρμακο; κύκλος, στόχος; τετράγωνο, μονοπάτι; εξάγωνο, γονίδιο που σχετίζεται με ασθένεια; και έλλειψη, ασθένεια. Το κόκκινο βέλος είναι μόνο για αλυσίδες που συνάπτουν ζεύγη γνωστών χημικών ασθενειών από CTD.

Όσον αφορά την πρόβλεψη συσχέτισης φαρμάκων-ασθένειας, η τροποποιημένη αιτιώδης συναγωγή που περιλαμβάνεται στο CI-PMF ξεπέρασε την υπάρχουσα αιτιολογική συμπεράσμα με υψηλότερη βαθμολογία AUC (περιοχή κάτω από τη χαρακτηριστική καμπύλη λειτουργίας του δέκτη) και μεγαλύτερη ακρίβεια. Επιπλέον, το CI-PMF απέδωσε καλύτερα από το απλό τροποποιημένο αιτιολογικό συμπεράσμα στην πρόβλεψη θεραπευτικών συσχετίσεων φαρμάκων-ασθένειας. Στο κορυφαίο 30% των προβλεπόμενων συσχετίσεων, το 58,6% (136/232), το 50,8% (31/61) και το 39,8% (140/352) έπληξαν γνωστές θεραπευτικές συσχετίσεις, ενώ οι ακριβείς τιμές που ελήφθησαν από τους τελευταίους ήταν μόνο 10,2% (231 /2264), 8,8% (36/411), και 9,7% (189/1948). Πραγματοποιήθηκαν περαιτέρω κλινικές επαληθεύσεις για τις κορυφαίες 100 πρόσφατα προβλεπόμενες θεραπευτικές ενώσεις. Ως αποτέλεσμα, 21, 12 και 32 συσχετίσεις έχουν μελετηθεί και διερευνήθηκαν πολλές θεραπευτικές επιδράσεις των φαρμάκων σε ασθένειες για καρδιαγγειακά νοσήματα, σακχαρώδη διαβήτη και νεοπλάσματα, αντίστοιχα. Εξήχθησαν σχετικές αλυσίδες σε αιτιολογικά δίκτυα για αυτές τις 65 κλινικά επαληθευμένες συσχετίσεις και απεικονίσαμε περαιτέρω τον θεραπευτικό ρόλο της ετοντολάκης (etodolac) στον καρκίνο του μαστού με συναγόμενες αλυσίδες. Συνολικά, το CI-PMF είναι μια χρήσιμη προσέγγιση για τη συσχέτιση φαρμάκων με πολύπλοκες ασθένειες και παρέχει πιθανές τιμές για την επαναστόχευση του φαρμάκου.

Δ'.7 Βαθιά μάθηση

Με τη σημαντική ανάπτυξη των δημοσίως διαθέσιμων συνόλων δεδομένων και την ταχεία αύξηση της υπολογιστικής ισχύος, η βαθιά μάθηση (DL) ή το νευρωνικό δίκτυο (NN), έχει κερδίσει μεγάλη προσοχή. Ως ένα εμπνευσμένο τμήμα μηχανικής μάθησης, η βαθιά μάθηση έχει δώσει σημαντική ώθηση και αναδείχθηκε ως η κορυφαία τεχνική για την ανακάλυψη και την ανάπτυξη φαρμάκων στις πιο πρόσφατες δημοσιευμένες μελέτες.

Η βαθιά μάθηση, μια έννοια στενά συνδεδεμένη με τα τεχνητά νευρωνικά δίκτυα (ANNs), μπορεί να οριστεί ως η μάθηση από τη μη γραμμική επεξεργασία των διασυνδεδεμένων στρωμάτων νευρώνων. Έχει προσελκύσει ερευνητές για την ευελιξία της αρχιτεκτονικής του, η οποία επιτρέπει την ανάπτυξη μοντέλων μηχανικής μάθησης μεμονωμένων εργασιών ή πολλαπλών εργασιών για τον εντοπισμό πιθανών

θεραπευτικών εφαρμογών και την πρόβλεψη αλληλεπιδράσεων φαρμάκου-ασθένειας. Αν και η βαθιά μάθηση έχει χρησιμοποιηθεί για την ανάπτυξη ανερχόμενων μοντέλων στον τομέα της επαναστόχευσης φαρμάκων, αξίζει να τονιστεί ότι η πλήρης χρήση της βαθιάς μάθησης εξακολουθεί να έχει ορισμένους περιορισμούς. Για παράδειγμα, τα μοντέλα βαθιάς νευρωνικών δικτύων πρέπει να προσαρμοστούν ώστε να ταιριάζουν στα δεδομένα που χρησιμοποιούνται στην εκπαίδευση αυτών των μοντέλων, κάτι που απαιτεί σημαντικό χρόνο και προσπάθεια.

Επιπρόσθετα, η επιλογή της τεχνικής μηχανικής εκμάθησης ή του μέτρου ομοιότητας που θα χρησιμοποιηθεί με κάθε σύνολο δεδομένων στα βαθιά ουδέτερα επίπεδα δικτύου δεν είναι απλή και κατά κάποιο τρόπο εξαρτάται από τα χρησιμοποιούμενα σύνολα δεδομένων. Τα νευρωνικά δίκτυα μπορούν να ταξινομηθούν κυρίως, με βάση την αρχιτεκτονική του δικτύου, σε (1) πλήρως συνδεδεμένο βαθύ νευρωνικό δίκτυο (DNN), (2) συνελκτικό νευρωνικό δίκτυο (CNN), (3) επαναλαμβανόμενο νευρωνικό δίκτυο (RNN), (4) αυτοκωδικοποιητής (AE).

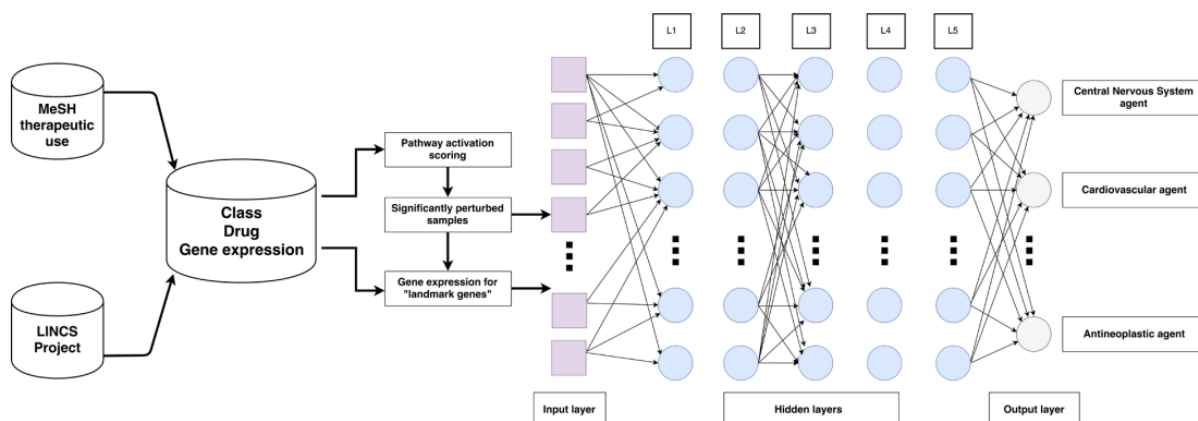
Άρθρο 16 *Deep Learning Applications for Predicting Pharmacological Properties of Drugs and Drug Repurposing Using Transcriptomic Data*

Οι Aliper et al.⁵⁴ χρησιμοποίησαν ένα πλήρως συνδεδεμένο βαθύ νευρωνικό δίκτυο για να προβλέψει τις φαρμακολογικές ιδιότητες των φαρμάκων και να εντοπίσει τις θεραπευτικές δυνατότητες και τις νέες ενδείξεις φαρμάκων. Οι συγγραφείς χρησιμοποίησαν δεδομένα υπογραφών έκφρασης γονιδίων και δεδομένα μονοπατιών για τη δημιουργία μοντέλων βαθιάς νευρωνικών δικτύων που ξεπέρασαν το μοντέλο μηχανής φορέα υποστήριξης και πέτυχαν υψηλή ακρίβεια ταξινόμησης στην πρόβλεψη ενδείξεων φαρμάκων και, ως εκ τούτου, τέτοια βαθιά νευρικά δίκτυα θα μπορούσαν να είναι χρήσιμα για επαναστόχευση φαρμάκων.

Επιπλέον, πρότειναν τη χρήση μητρών σύγχυσης βαθιάς νευρωνικού δικτύου για την επαναστόχευση του φαρμάκου. Χρησιμοποιήθηκαν τα δείγματα διαταραχής 678 φαρμάκων σε κυτταρικές σειρές A549, MCF-7 και PC-3 από το LINCS Project και τα συνδέσαμε με 12 κατηγορίες θεραπευτικής χρήσης που προέρχονται από το MeSH. Για να εκπαιδεύσουν το DNN, χρησιμοποίησαν μεταγραφικά δεδομένα σε επίπεδο γονιδίου και μεταγραφικά δεδομένα που υποβλήθηκαν σε επεξεργασία χρησιμοποιώντας έναν αλγόριθμο βαθμολόγησης ενεργοποίησης μονοπατιού, για ένα σύνολο δεδομένων δειγμάτων που διαταράχθηκαν με διαφορετικές συγκεντρώσεις του φαρμάκου για 6 και 24 ώρες.

Τόσο στην ταξινόμηση μονοπατιού όσο και σε επίπεδο γονιδίου, το DNN πέτυχε υψηλή ακρίβεια ταξινόμησης και ξεπέρασε πειστικά το μοντέλο μηχανής φορέα υποστήριξης (SVM) σε κάθε πρόβλημα ταξινόμησης πολλαπλών κλάσεων, ωστόσο, τα μοντέλα που βασίζονται σε δεδομένα επιπέδου μονοπατιού είχαν σημαντικά καλύτερα αποτελέσματα. Για πρώτη φορά επιδεικνύουν ένα νευρωνικό δίκτυο βαθιάς μάθησης εκπαιδευμένο σε μεταγραφικά δεδομένα για να αναγνωρίζει τις φαρμακολογικές ιδιότητες

πολλαπλών φαρμάκων σε διαφορετικά βιολογικά συστήματα και συνθήκες. Οι συγγραφείς προτείνουν επίσης τη χρήση μητρών σύγχυσης βαθιάς νευρωνικού δικτύου για την επαναστόχευση του φαρμάκου. Αυτή η εργασία είναι μια απόδειξη αρχής για την εφαρμογή της βαθιάς μάθησης στην ανακάλυψη και ανάπτυξη φαρμάκων.



Εικόνα 14: Σχεδιασμός μελέτης

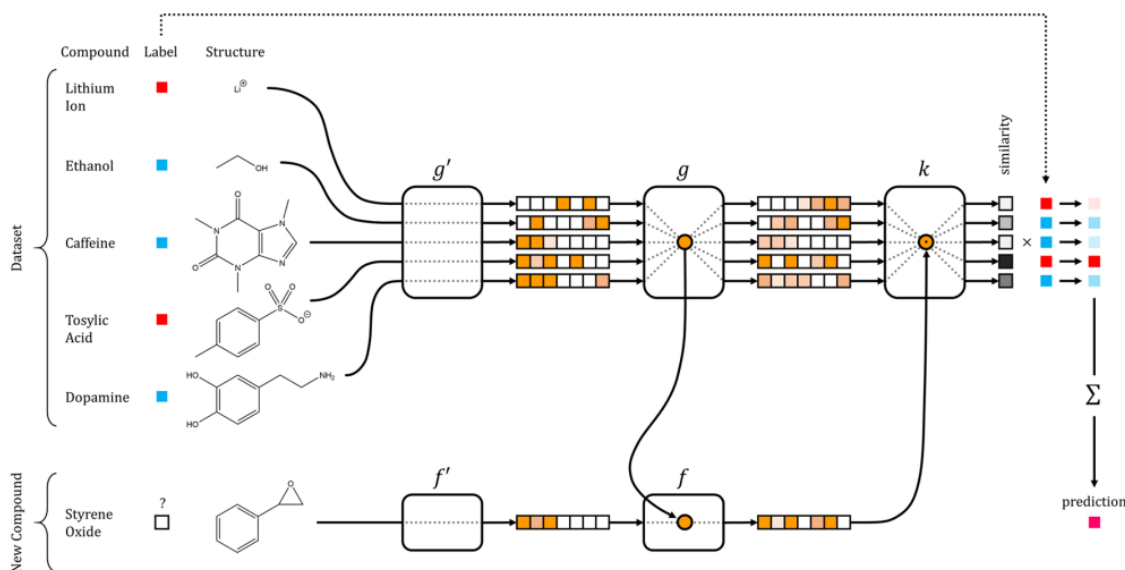
Τα δεδομένα γονιδιακής έκφρασης από το LINCS Project συνδέθηκαν με 12 κατηγορίες θεραπευτικής χρήσης MeSH. Το DNN εκπαιδεύτηκε ξεχωριστά σε δεδομένα επιπέδου γονιδιακής έκφρασης για «γονίδια ορόσημο» και βαθμολογίες ενεργοποίησης μονοπατιού για σημαντικά διαταραγμένα δείγματα, σχηματίζοντας στρώματα εισόδου 977 και 271 νευρικών κόμβων, αντίστοιχα

Άρθρο 17 Low data drug discovery with one-shot learning

Οι Altae-Tran et al.⁷³ ενσωμάτωσαν ένα τυπικό παράδειγμα εκμάθησης μίας λήψης (one-shot learning) με ένα συνελκτικό νευρωνικό δίκτυο για να καταλήξει σε ένα επαναληπτικό μοντέλο εκμάθησης μακροπρόθεσμης βραχυπρόθεσμης μνήμης (LSTM). Τα βαθιά νευρωνικά δίκτυα έχει αποδειχθεί ότι ενισχύουν σημαντικά την προγνωστική ισχύ για την εξαγωγή των ιδιοτήτων και των δραστηριοτήτων μικρών μορίων ενώσεων. Ωστόσο, η δυνατότητα εφαρμογής τους έχει περιοριστεί από την ανάγκη για μεγάλες ποσότητες δεδομένων εκπαίδευσης. Σε αυτή καταδικνύεται πώς η μάθηση με μία λήψη μπορεί να μειώσει σημαντικά τις απαιτήσεις δεδομένων για την πραγματοποίηση ουσιαστικών προβλέψεων σε εφαρμογές ανακάλυψης φαρμάκων. Εισάγεται μια νέα αρχιτεκτονική, την επαναληπτική βελτίωση της μακροπρόθεσμης μνήμης (LSTM), η οποία, όταν συνδυάζεται με γραφικά συνελκτικά νευρωνικά δίκτυα, βελτιώνει σημαντικά την εκμάθηση σημαντικών μετρήσεων απόστασης για μικρά μόρια.

Όλα τα μοντέλα που παρουσιάζονται σε αυτήν την εργασία είναι ανοιχτού κώδικα ως μέρος του DeepChem, ενός πλαισίου ανοιχτού κώδικα για βαθιά μάθηση στην ανακάλυψη φαρμάκων. Οι συγγραφείς

υιοθέτησαν το τυπικό παράδειγμα εκμάθησης μιας βολής για να βελτιώσουν την εκμάθηση σημαντικών μετρήσεων απόστασης σε σχέση με μικρά μόρια σε νέα πειραματικά συστήματα. Όταν αξιολογείται έναντι δύο διαφορετικών συναφών συνόλων δεδομένων, το προτεινόμενο μοντέλο μιας λήψης πέτυχε αξιοσημείωτη επιτυχία στον εντοπισμό της μοριακής συμπεριφοράς σε πειράματα ανακάλυψης φαρμάκων με χαμηλά δεδομένα.



Εικόνα 15: Σχήμα Αρχιτεκτονικής Δικτύων για εκμάθηση μίας λήψης (*one-shot learning*) στην ανακάλυψη φαρμάκων

Άρθρο 18 Predicting drug-target interactions from drug structure and protein sequence using novel convolutional neural networks

Οι Hu et al.⁷⁴ εισήγαγαν ένα μοντέλο συνελκτικού νευρωνικού δικτύου για να αποκαλύψει τις αλληλεπιδράσεις φαρμάκου-στόχου. Οι συγγραφείς χρησιμοποίησαν δεδομένα χημικών δομών φαρμάκων και αλληλουχιών πρωτεϊνών για να κατασκευάσουν τον ταξινομητή συνελκτικού νευρωνικού δικτύου τους που έδειξε ανώτερη απόδοση σε σύγκριση με άλλα μοντέλα τελευταίας τεχνολογίας. Το προτεινόμενο μοντέλο συνήγαγε συσχετίσεις φαρμάκου-στόχου στην περίπτωση της ύπαρξης πολλαπλών πρωτεϊνών-στόχων που αλληλεπιδρούν με πολλά χημικά μόρια, τα οποία καταδεικνύουν τη δυνατότητα ενός τέτοιου μοντέλου στον εντοπισμό νέων θεραπευτικών ενδείξεων και ευκαιριών επαναστόχευσης φαρμάκων. Σε αυτό το έγγραφο, οι συγγραφείς προτείνουν μια μέθοδο βασισμένη στη βαθιά μάθηση για την πρόβλεψη των DTI χρησιμοποιώντας μόνο τις πληροφορίες των δομών φαρμάκων και των πρωτεϊνικών αλληλουχιών.

Τα τελικά αποτελέσματα έδειξαν ότι η μέθοδος μπορεί να επιτύχει καλή απόδοση με ακρίβειες έως 92,0%, 90,0%, 92,0% και 90,7% για τις οικογένειες-στόχους των ενζύμων, των διαύλων ιόντων, των GPCR

και των πυρηνικών υποδοχέων του συνόλου δεδομένων που δημιουργήσαμε, αντίστοιχα. Ένα άλλο σύνολο δεδομένων που προέρχεται από το DrugBank χρησιμοποιήθηκε για την περαιτέρω αξιολόγηση της γενίκευσης του μοντέλου, το οποίο απέδωσε ακρίβεια 0,9015 και τιμή AUC 0,9557. Διευκρινίστηκε ότι το μοντέλο παρουσιάζει βελτιωμένη απόδοση σε σύγκριση με άλλες υπερσύγχρονες υπολογιστικές μεθόδους στα κοινά σύνολα δεδομένων αναφοράς. Τα πειραματικά αποτελέσματα έδειξαν ότι το μοντέλο μας εξήγαγε με επιτυχία περισσότερες αποχρώσεις αλλά χρήσιμα χαρακτηριστικά και επομένως μπορεί να χρησιμοποιηθεί ως πρακτικό εργαλείο για την ανακάλυψη νέων φαρμάκων.

Άρθρο 19 *Generating focused molecule libraries for drug discovery with recurrent neural networks.*

Οι Segler et al.⁷⁵ πρότειναν ένα μοντέλο επαναλαμβανόμενου νευρωνικού δικτύου για τη δημιουργία νέων μορίων για εφαρμογές επαναστόχευσης φαρμάκων. Οι συγγραφείς χρησιμοποίησαν δομές φαρμάκων και δεδομένα αλληλεπιδράσεων φαρμάκου-στόχου για να εκπαιδεύσουν τον ταξινομητή επαναλαμβανόμενου νευρικού δικτύου ώστε να παράγουν νέα μόρια που συνδέονται στενά με τους επιθυμητούς βιολογικούς στόχους. Το προτεινόμενο μοντέλο αξιολογήθηκε σε σχέση με δύο διαφορετικά γνωστά σύνολα δεδομένων συσχέτισης φαρμάκου-στόχου και απέδωσε αρκετά καλά.

Ωστόσο, το εισαγόμενο μοντέλο μιμήθηκε τον πλήρη κύκλο σχεδιασμού φαρμάκων de novo και δημιούργησε μεγάλα σύνολα νέων μορίων όταν ενσωματώθηκε με μια συνάρτηση βαθμολόγησης. Στον de novo σχεδιασμό φαρμάκων, χρησιμοποιούνται υπολογιστικές στρατηγικές για τη δημιουργία νέων μορίων με καλή συγγένεια με τον επιθυμητό βιολογικό στόχο. Σε αυτή την εργασία, οι συγγραφείς δείχνουν ότι τα επαναλαμβανόμενα νευρωνικά δίκτυα μπορούν να εκπαιδευτούν ως παραγωγικά μοντέλα για μοριακές δομές, παρόμοια με τα στατιστικά γλωσσικά μοντέλα στην επεξεργασία φυσικής γλώσσας. Δείχνουν ότι οι ιδιότητες των δημιουργούμενων μορίων συσχετίζονται πολύ καλά με τις ιδιότητες των μορίων που χρησιμοποιούνται για την εκπαίδευση του μοντέλου. Προκειμένου να εμπλουτιστούν οι βιβλιοθήκες με μόρια ενεργά προς έναν δεδομένο βιολογικό στόχο, προτείνουμε να τελειοποιήσουμε το μοντέλο με μικρά σύνολα μορίων, τα οποία είναι γνωστό ότι είναι ενεργά εναντίον αυτού του στόχου. Κατά του *Staphylococcus aureus*, το μοντέλο αναπαρήγαγε το 14% των 6051 μορίων δοκιμής που είχαν σχεδιαστεί από φαρμακοποιούς, ενώ κατά του *Plasmodium falciparum* (ελονοσία), αναπαρήγαγε το 28% από 1240 δοκιμαστικά μόρια. Όταν συνδυάζεται με μια συνάρτηση βαθμολόγησης, το μοντέλο μπορεί να εκτελέσει τον πλήρη κύκλο σχεδιασμού φαρμάκων de novo για να δημιουργήσει μεγάλα σύνολα νέων μορίων για ανακάλυψη φαρμάκων.

Άρθρο 20 *In silico drug repositioning using deep learning and comprehensive similarity measures*

Οι Yi et al.⁷⁶ χρησιμοποίησαν πολυτροπικά μοντέλα βαθύ αυτοκωδικοποιητή και μεταβλητών αυτοκωδικοποιητών για να ανακαλύψει συσχετίσεις φαρμάκων-ασθένειες. Οι συγγραφείς ενσωμάτωσαν

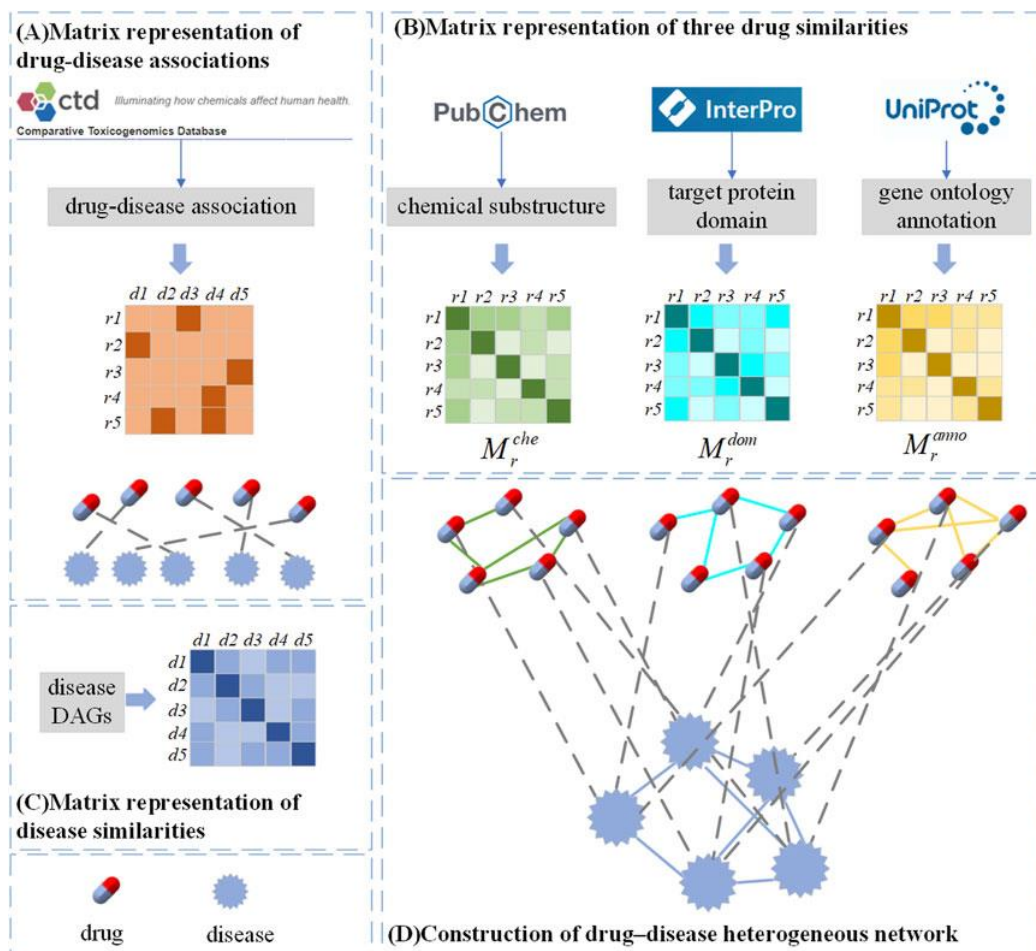
διάφορα σύνολα δεδομένων που σχετίζονται με φάρμακα (συσχετίσεις φαρμάκων-ασθένειες, συσχετίσεις φαρμάκων-στόχων, συσχετίσεις φαρμάκων-φαρμάκων και παρενέργειες φαρμάκου) για να εκπαιδεύσουν έναν πολυτροπικό βαθύ αυτοκωδικοποιητή και στη συνέχεια να ορίσουν χαρακτηριστικά φαρμάκων υψηλού επιπέδου.

Μετά από αυτό, κωδικοποίησαν και αποκωδικοποίησαν τον συνδυασμό χαρακτηριστικών φαρμάκων υψηλού επιπέδου και κλινικά αναφερόμενων συσχετίσεων φαρμάκων-ασθένειας χρησιμοποιώντας παραλλαγμένο αυτοκωδικοποιητή για τον εντοπισμό νέων θεραπευτικών ενδείξεων επιπλέον των ήδη αναγνωρισμένων ενδείξεων. Τα ευρήματα επικυρώθηκαν σε σχέση με ένα γνωστό σύνολο δεδομένων συσχετίσεων φαρμάκων-ασθένειας και ξεπέρασαν τα προηγούμενα μοντέλα μηχανικής μάθησης τελευταίας τεχνολογίας. Επιπλέον, οι συγγραφείς ανέφεραν υποψηφιότητες επαναστόχευσης φαρμάκων για τη νόσο του Αλτσχάιμερ και του Πάρκινσον.

Άρθρο 21 *Drug repositioning based on heterogeneous networks and variational graph autoencoders*

Η πρόβλεψη νέων θεραπευτικών αποτελεσμάτων (επαναστόχευση φαρμάκων) των υπαρχόντων φαρμάκων παίζει σημαντικό ρόλο στην ανάπτυξη φαρμάκων. Ωστόσο, οι παραδοσιακές υγρές πειραματικές μέθοδοι πρόβλεψης είναι συνήθως χρονοβόρες και δαπανηρές. Η εμφάνιση ολοένα και περισσότερων μεθόδων επαναστόχευσης φαρμάκων που βασίζονται στην τεχνητή νοημοσύνη τα τελευταία 2 χρόνια έχει διευκολύνει την ανάπτυξη φαρμάκων. Σε αυτή τη μελέτη οι Lei et al.⁷⁷ προτείνουν μια μέθοδο επαναστόχευσης φαρμάκων, VGAEDR, που βασίζεται σε ένα ετερογενές δίκτυο πολλαπλών χαρακτηριστικών φαρμάκων και σε έναν αυτοκωδικοποιητή γραφικών παραλλαγών. Πρώτον, δημιουργείται ένα ετερογενές δίκτυο φαρμάκων-ασθένειας που βασίζεται σε τρία χαρακτηριστικά φαρμάκων, σημασιολογικές πληροφορίες ασθένειας και γνωστούς συσχετισμούς φαρμάκων-ασθένειας.

Δεύτερον, οι αναπαραστάσεις χαρακτηριστικών χαμηλών διαστάσεων για ετερογενή δίκτυα μαθαίνονται μέσω μιας μονάδας αυτόματου κωδικοποιητή μεταβλητών γραφημάτων και μιας συνελκτικής μονάδας πολλαπλών επιπέδων. Τέλος, η αναπαράσταση χαρακτηριστικών τροφοδοτείται σε ένα πλήρως συνδεδεμένο επίπεδο και ένα στρώμα Softmax για την πρόβλεψη νέων συσχετίσεων φαρμάκων-ασθενειών. Συγκριτικά πειράματα με άλλες βασικές μεθόδους σε τρία σύνολα δεδομένων καταδεικνύουν την εξαιρετική απόδοση του VGAEDR. Στη μελέτη περίπτωσης, προβλέψαμε τα κορυφαία 10 πιθανά φάρμακα κατά του COVID-19 στα υπάρχοντα δεδομένα φαρμάκων και ασθενειών και έξι από αυτά επαληθεύτηκαν από άλλες βιβλιογραφίες.



Εικόνα 16: Προσέγγιση μεθόδου επαναστόχευσης φαρμάκων, *VGAEDR*

(A) αναπαράσταση μήτρας συσχετίσεων φαρμάκου-ασθένειας, (B) αναπαράσταση μήτρας τριών ομοιοτήτων φαρμάκων, (C) αναπαράσταση μήτρας ομοιοτήτων ασθένειας, (D) κατασκευή ετερογενούς δικτύου φαρμάκου-ασθένειας.

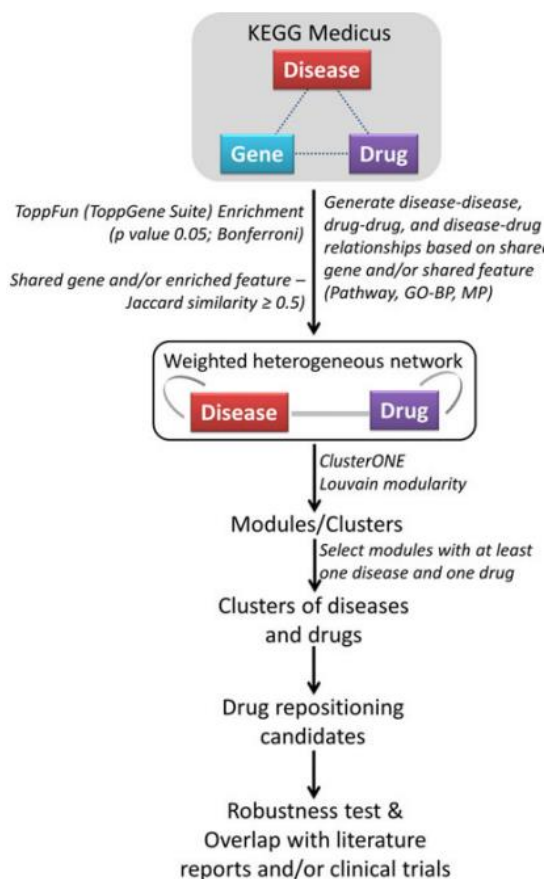
Δ'.8 Ομαδοποίηση

Άρθρο 22 Computational drug repositioning through heterogeneous network clustering

Οι Wu et al.⁵⁶ δημιούργησαν ένα σταθμισμένο ετερογενές δίκτυο φαρμάκων-ασθενειών και εφάρμοσαν ομαδοποίηση δικτύου για τον εντοπισμό πιθανών υποψηφίων επαναστόχευσης φαρμάκων μέσα σε στενά συνδεδεμένες μονάδες δικτύου. Οι συγγραφείς χρησιμοποίησαν συσχετίσεις ασθένειας-γονιδίου και αλληλεπιδράσεις φαρμάκου-στόχου για να δημιουργήσουν το σταθμισμένο ετερογενές δίκτυό τους όπου τα φάρμακα και οι ασθένειες ορίστηκαν ως κόμβοι, οι άκρες σχεδιάστηκαν όταν ένα ζεύγος κόμβων μοιράζεται γονίδια, στόχους, βιολογικές διεργασίες, μονοπάτια, φαινότυπους, ή ένας συνδυασμός

αυτών των χαρακτηριστικών, και οι ακμές σταθμίστηκαν χρησιμοποιώντας την ομοιότητα του συντελεστή Jaccard.

Στη συνέχεια, χρησιμοποίησαν δύο αλγόριθμους ομαδοποίησης δικτύου για να ομαδοποιήσουν κόμβους σε ενότητες και στη συνέχεια συγκέντρωσαν όλα τα πιθανά ζεύγη φαρμάκων-ασθένειας σε καθεμία από αυτές τις ενότητες. Τέλος, αντιμετώπισαν ζεύγη φαρμάκων-ασθένειας που προτείνονται από τους δύο αλγόριθμους ομαδοποίησης δικτύου ως υποψηφίους επαναστόχευσης φαρμάκων και πραγματοποίησαν επικυρώσεις βιβλιογραφίας και παρουσίασαν αρκετές περιπτωσιολογικές μελέτες προς υποστήριξη του προτεινόμενου μοντέλου τους. Συγκεκριμένα, χρησιμοποιώντας γνωστές σχέσεις ασθένειας-γονιδίου και φαρμάκου-στόχου από τη βάση δεδομένων KEGG, δημιουργήθηκε ένα σταθμισμένο ετερογενές δίκτυο ασθένειας και φαρμάκων. Οι κόμβοι αντιπροσωπεύουν φάρμακα ή ασθένειες ενώ οι άκρες αντιπροσωπεύουν κοινό γονίδιο, βιολογική διαδικασία, μονοπάτι, φαινότυπο ή συνδυασμό αυτών των χαρακτηριστικών. Συγκεντρώσαμε αυτό το σταθμισμένο δίκτυο για να αναγνωρίσουμε τις ενότητες και στη συνέχεια συγκεντρώσαμε όλα τα πιθανά ζεύγη φαρμάκων-ασθένειας (υποτιθέμενοι υποψήφιοι επαναστόχευσης φαρμάκων) από αυτές τις ενότητες.



Εικόνα 17: Ροή εργασιών για τον εντοπισμό υποψηφίων για επαναστόχευση φαρμάκων με χρήση παθήσεων που βασίζονται σε γονίδια και χαρακτηριστικά και ετερογενές δίκτυο φαρμάκων

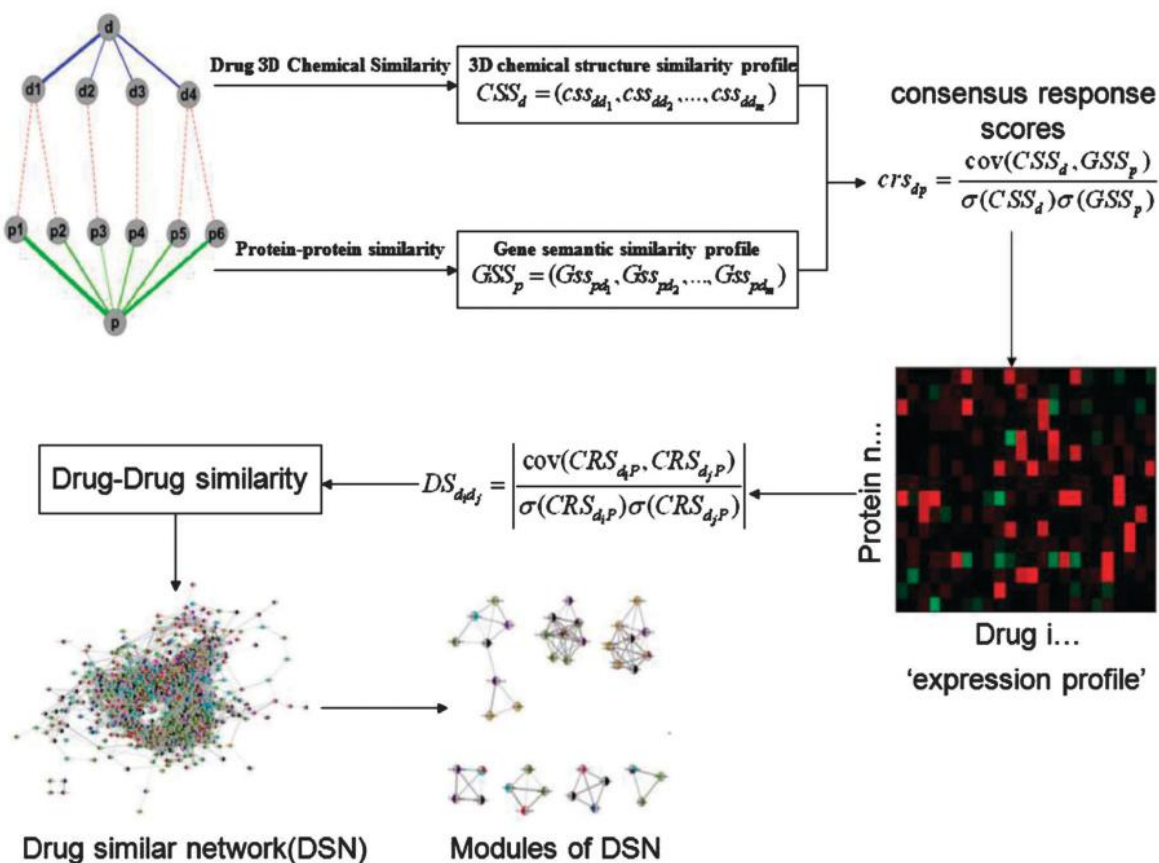
Οι προηγούμενες υπολογιστικές προσεγγίσεις για την επαναστόχευση φαρμάκων επικεντρώθηκαν είτε σε προσεγγίσεις φαρμάκου-φαρμάκου και ομοιότητας ασθένειας, ενώ στη συγκεκριμένη εργασία οι συγγραφείς υιοθέτησαν μια πιο ολιστική προσέγγιση εξετάζοντας επίσης τις σχέσεις φαρμάκου-ασθένειας. Παρά τη σχετική απλότητα της προσέγγισής τους, με βάση τις αναλύσεις ευρωστίας και την επικάλυψη ορισμένων από τις προβλέψεις τους με ενδείξεις φαρμάκων που βρίσκονται υπό διερεύνηση, πιστεύουν ότι η προσέγγισή τους θα μπορούσε να συμπληρώσει τις τρέχουσες υπολογιστικές προσεγγίσεις για την ανακάλυψη υποψηφίων για επαναστόχευση φαρμάκων.

Άρθρο 23 *Drug repositioning by applying 'expression profiles' generated by integrating chemical structure similarity and gene semantic similarity.*

Οι Tan et al.⁷⁸ δημιούργησαν ένα δίκτυο αλληλεπίδρασης φαρμάκου-φαρμάκου προκειμένου να εντοπίσει νέες ενδείξεις στόχων φαρμάκων. Οι συγγραφείς χρησιμοποίησαν πληροφορίες χημικής δομής φαρμάκου, πληροφορίες σημασιολογικής ομοιότητας γονιδίων και δίκτυα αλληλεπίδρασης φαρμάκου-στόχου για να υπολογίσουν τον βαθμό ομοιότητας φαρμάκου που στη συνέχεια χρησιμοποιήθηκε για την κατασκευή ενός δικτύου αλληλεπίδρασης φαρμάκου, γειτονικών φαρμάκων, ομαδοποιώντας το δίκτυο αλληλεπίδρασης φαρμάκου-φαρμάκου σε ενότητες με βάση τον τρόπο δράσης και, τέλος, προτείνει νέες φαρμακευτικές θεραπευτικές ενδείξεις.

Το προτεινόμενο μοντέλο έδειξε υψηλή ακρίβεια όταν επικυρώθηκε με τη χρήση της βιβλιογραφίας. Αυτή η μέθοδος αξιοποιεί προηγούμενες γνώσεις όπως τρισδιάστατες πληροφορίες χημικής δομής φαρμάκου, αλληλεπιδράσεις φαρμάκου-στόχου και πληροφορίες σημασιολογικής ομοιότητας γονιδίων. Οι προβλέψεις του βασίζονται σε έναν διαφορετικό τύπο «προφίλ έκφρασης» που περιλαμβάνει βαθμολογίες που κυμαίνονται από -1 έως 1, αντανakλώντας τις βαθμολογίες απόκρισης συναίνεσης (CRS) μεταξύ καθενός από 965 φάρμακα και 1560 πρωτεΐνες. Το CRS ενσωματώνει πληροφορίες ομοιότητας χημικής δομής και σημασιολογικής ομοιότητας γονιδίων. Ορίζουμε τον βαθμό ομοιότητας μεταξύ δύο φαρμάκων ως την απόλυτη τιμή των συντελεστών συσχέτισής τους. Τελικά, ιδρύουμε ένα δίκτυο ομοιότητας φαρμάκων (DSN) και προσδιορίζουμε 33 ενότητες φαρμάκων με παρόμοιους τρόπους δράσης, προσδιορίζοντας τις κοινές ενδείξεις τους. Χρησιμοποιώντας αυτές τις ενότητες, προβλέπουμε νέες ενδείξεις για 143 φάρμακα και εντοπίζουμε προηγουμένως άγνωστες ενδείξεις για 42 φάρμακα χωρίς κωδικούς ATC.

Αυτή η μέθοδος ξεπερνά την αστάθεια του προφίλ γονιδιακής έκφρασης λόγω πειραματικών συνθηκών και προβλέπει εφικτά ενδείξεις για νέες ενώσεις χρησιμοποιώντας μόνο την τρισδιάστατη δομή τους. Επιπλέον, το υψηλό ποσοστό επικύρωσης της βιβλιογραφίας 71,8% υποδηλώνει ότι η μέθοδος έχει σημαντικές δυνατότητες να ανακαλύψει νέες ενδείξεις για υπάρχοντα φάρμακα.



Εικόνα 18: Η ροή εργασίας της μεθόδου των Tan et al

Στη μέθοδο των Tan et al ⁷⁸ ενσωματώθηκαν πληροφορίες χημικής ομοιότητας τρισδιάστατων φαρμάκων και σημασιολογικής ομοιότητας γονιδίων για να δημιουργήσουν ένα «προφίλ έκφρασης» υπολογίζοντας το CRS μεταξύ φαρμάκων και πρωτεϊνών. Στη συνέχεια, ορίστηκε η απόλυτη τιμή των συντελεστών συσχέτισης ως τον βαθμό ομοιότητας (DS) μεταξύ κάθε ζεύγους φαρμάκων. Κατασκευάστηκε το DSN συνδέοντας δύο φάρμακα εάν το DS τους ήταν πάνω από ένα όριο που ορίσαμε. Τέλος, καταγράφονται ενότητες που αποτελούνται από παρόμοια φάρμακα με κοινές ενδείξεις εξόρυξης του DSN.

Δ'.9 Μέτρα κεντρικότητας του δικτύου

Άρθρο 24 *A bidirectional drug repositioning approach for Parkinson's disease through network-based inference*

Οι Rakshit et al. ⁷⁹ ανέπτυξαν μια νέα βασισμένη σε δίκτυο αμφίδρομες προσεγγίσεις από πάνω προς τα κάτω και από κάτω προς τα πάνω για να προβλέψει πιθανές εφαρμογές επαναστόχευσης φαρμάκων για μια συγκεκριμένη ασθένεια. Οι συγγραφείς χρησιμοποίησαν πληροφορίες στόχων για συγκεκριμένες

ασθένειες (νόσος του Πάρκινσον) και ενδείξεις στόχου φαρμάκου για να δημιουργήσουν δύο δίκτυα. Στη συνέχεια, χρησιμοποίησαν διάφορα μέτρα κεντρικότητας δικτύου για να αναγνωρίσουν γονίδια και φάρμακα ενδιαφέροντος και στα δύο δίκτυα και τα χρησιμοποίησαν ως είσοδο για τα μοντέλα από πάνω προς τα κάτω και από κάτω προς τα πάνω. Τα μοντέλα που εισήχθησαν εντόπισαν ένα σύνολο υποψηφίων για την εναπόθεση εναπόθεσης φαρμάκων προς διερεύνηση για τη θεραπεία της νόσου του Πάρκινσον, η οποία επικυρώθηκε έναντι μιας γνωστής πηγής δεδομένων ενδείξεων στόχου φαρμάκων.

Άρθρο 25 *Mining heterogeneous network for drug repositioning using phenotypic information extracted from social media and pharmaceutical databases*

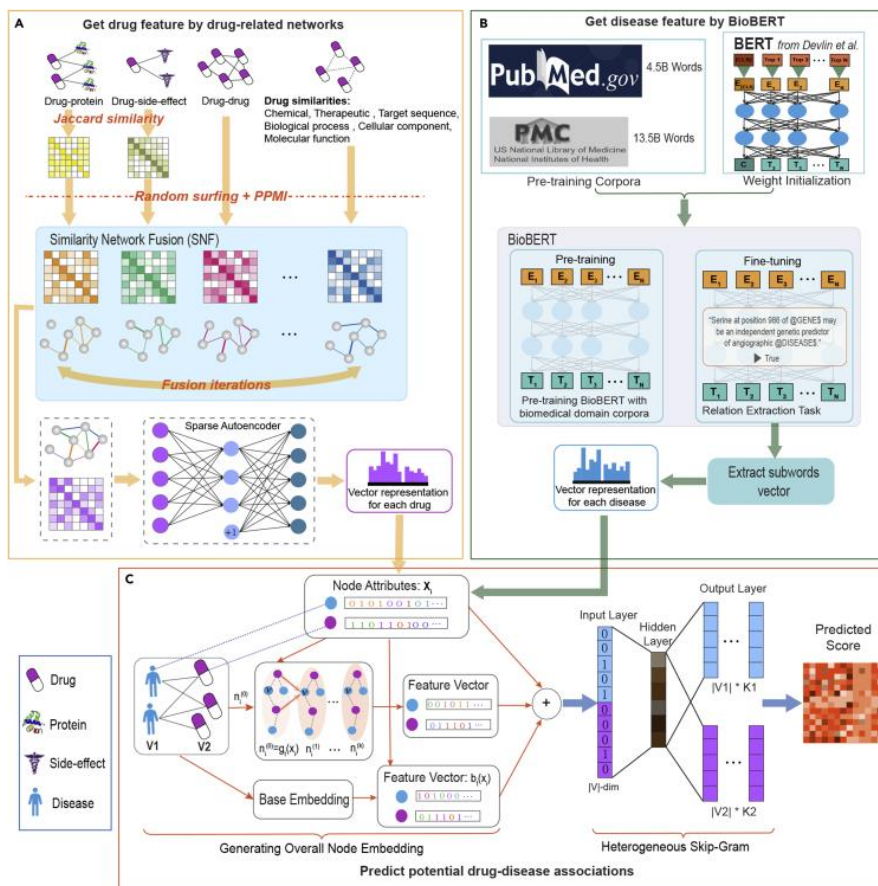
Η επαναστόχευση φαρμάκων έχει τραβήξει σημαντική προσοχή για την ανάπτυξη φαρμάκων στη φαρμακευτική έρευνα και βιομηχανία, λόγω των πλεονεκτημάτων της σε κόστος και χρόνο σε σύγκριση με την de novo ανάπτυξη φαρμάκων. Η διαθεσιμότητα βιοϊατρικών βάσεων δεδομένων και διαδικτυακών πληροφοριών που σχετίζονται με την υγεία, καθώς και οι υπολογιστές υψηλής απόδοσης, ενισχύουν την ανάπτυξη υπολογιστικών μεθόδων επαναστόχευσης φαρμάκων. Οι Yang et al.⁸⁰ πρότειναν ένα νέο συστηματικό μοντέλο για τον εντοπισμό των θεραπευτικών δυνατοτήτων και των υποψηφίων για επαναστόχευση φαρμάκων σε ετερογενή δίκτυα.

Οι συγγραφείς συνδύασαν μοριακά δεδομένα, παρενέργειες και πληροφορίες διαδικτυακής κοινότητας υγείας για να δημιουργήσουν ένα ετερογενές δίκτυο που αποτελείται από φάρμακα, ασθένειες και ανεπιθύμητες ενέργειες φαρμάκων ως ενδιάμεσα. Στη συνέχεια, εφάρμοσαν πολλά μοντέλα εξόρυξης ετερογενών δικτύων βασισμένα σε μονοπάτια για τον εντοπισμό και την επαναστόχευση φαρμάκων και επικύρωσαν κυριολεκτικά τα μοντέλα τους και κατέληξαν στο συμπέρασμα ότι όσο περισσότερες πηγές δεδομένων χρησιμοποιούνται για την κατασκευή τέτοιων ετερογενών δικτύων, τόσο το καλύτερο για την πρόβλεψη μοντέλων.

Συγκεκριμένα, ανέπτυξαν μια συστηματική προσέγγιση που προσδιορίζει την επαναστόχευση φαρμάκων με βάση την ετερογενή εξόρυξη δικτύου χρησιμοποιώντας τόσο φαρμακευτικές βάσεις δεδομένων (PharmGKB και SIDER) όσο και διαδικτυακή κοινότητα υγείας (MedHelp). Χρησιμοποιώντας τις ανεπιθύμητες αντιδράσεις φαρμάκων (ADRs) ως ενδιάμεσο, κατασκευάστηκε ένα ετερογενές δίκτυο υγείας που περιέχει φάρμακα, ασθένειες και ADR και αναπτύξαμε ετερογενείς προσεγγίσεις εξόρυξης δικτύου βάσει μονοπατιών για την επαναστόχευση φαρμάκων. Επιπλέον, διερευνήθηκε πώς οι πηγές δεδομένων επηρεάζουν την απόδοση στην επαναστόχευση φαρμάκων. Τα αποτελέσματα του πειράματος έδειξαν ότι ο συνδυασμός τόσο του PharmGKB όσο και του MedHelp εντόπισε 479 φάρμακα επαναστόχευσης, τα οποία είναι περισσότερα από τα φάρμακα επαναστόχευσης που ανακαλύφθηκαν από άλλες εναλλακτικές λύσεις. Επιπλέον, το 31% από τα 479 φάρμακα επαναστόχευσης που ανακαλύφθηκαν υποστηρίζονταν από στοιχεία από το PubMed.

Άρθρο 26 HeTDR: Drug repositioning based on heterogeneous networks and text mining

Η χρήση της υπάρχουσας γνώσης για την πρόβλεψη συσχετίσεων φαρμάκων-ασθένειας είναι μια ζωτική μέθοδος για την επαναστόχευση του φαρμάκου. Ωστόσο, η αποτελεσματική σύντηξη του βιοϊατρικού κειμένου και των πληροφοριών του βιολογικού δικτύου είναι μία από τις μεγάλες προκλήσεις για τις περισσότερες τρέχουσες μεθόδους επαναστόχευσης φαρμάκων. Οι Jin et al.⁸¹ προτείνουν μια μέθοδο επαναστόχευσης φαρμάκων που βασίζεται σε ετερογενή δίκτυα και εξόρυξη κειμένου (HeTDR). Αυτό το μοντέλο μπορεί να συνδυάσει χαρακτηριστικά φαρμάκων από πολλαπλά δίκτυα που σχετίζονται με φάρμακα, χαρακτηριστικά ασθένειας από βιοϊατρικά σώματα με το γνωστό δίκτυο συσχετίσεων φαρμάκων-ασθένειας για να προβλέψει τις βαθμολογίες συσχέτισης μεταξύ φαρμάκου και ασθένειας. Τα πειράματα δείχνουν ότι το HeTDR έχει εξαιρετική απόδοση που είναι ανώτερη από εκείνη των μοντέλων τελευταίας τεχνολογίας. Παρουσιάζουν τα κορυφαία 10 νέα εγκεκριμένα φάρμακα που προέβλεπαν το HeTDR για πέντε ασθένειες και αποδεικνύουμε ότι το μοντέλο τους είναι ικανό να ανακαλύπτει πιθανά υποψήφια φάρμακα για ενδείξεις ασθενειών.



Εικόνα 19: Διάγραμμα ροής HeTDR

Το μοντέλο αποτελείται από τρία μέρη: (Α) Το HeTDR ενσωματώνει εννέα δίκτυα που σχετίζονται με τα ναρκωτικά για τη λήψη παγκόσμιων πληροφοριών για τα φάρμακα. Στα ετερογενή δίκτυα αλληλεπίδρασης, χρησιμοποιούμε πρώτα τον συντελεστή ομοιότητας Jaccard για να υπολογίσουμε το δίκτυο ομοιότητας. Στη συνέχεια, συγχωνεύουμε αυτά τα δίκτυα που σχετίζονται με φάρμακα σε ένα δίκτυο από το SNF και εφαρμόζουμε το SAE για να αποκτήσουμε χαρακτηριστικά χαμηλής διάστασης των φαρμάκων. (Β) Το HeTDR αποκτά αναπαράσταση φορέα των χαρακτηριστικών της νόσου με βιοϊατρικά σώματα εξόρυξης κειμένου. Στο στάδιο της προ εκπαίδευσης, χρησιμοποιούμε απευθείας τις παραμέτρους του μοντέλου που έχουν προ εκπαιδευτεί από την BioBERT. Στη συνέχεια, επιλέγουμε την εργασία εξαγωγής σχέσεων για προπόνηση τελειοποίησης. Αφού πραγματοποιηθεί η διαδικασία λεπτομέρειας, εξάγουμε τις αναπαραστάσεις των δευτερευουσών λέξεων και χρησιμοποιούμε τις αναπαραστάσεις αυτών των δευτερευουσών λέξεων για να λάβουμε τις αναπαραστάσεις όλων των ασθενειών. (C) Το HeTDR προβλέπει πιθανούς συσχετισμούς φαρμάκων-ασθένειας με μια μέθοδο εκμάθησης ενσωμάτωσης, η οποία μπορεί να συλλάβει τόσο την τοπολογική δομική εγγύτητα του δικτύου συσχετίσεων φαρμάκου-ασθένειας όσο και την εγγύτητα των χαρακτηριστικών κόμβου.

Ευκαιρίες επαναστόχευσης φαρμάκων

Η επαναστόχευση φαρμάκων είναι μια πολλά υποσχόμενη τεχνική που έχει συγκεντρώσει αυξανόμενη προσοχή από τις κυβερνήσεις και τις φαρμακευτικές εταιρείες λόγω του κρίσιμου ρόλου της στη μείωση του χρόνου, του κόστους και του κινδύνου που συνεπάγεται η ανάπτυξη θεραπειών για καρκίνους και άλλες ανίατες ασθένειες. Καθώς εμφανίστηκε αυτή η τεχνική, διεπιστημονικές ομάδες ερευνητών και επιστημόνων έχουν κάνει πολλές προσπάθειες, με διάφορους βαθμούς αποτελεσματικότητας και επιτυχίας, να διερευνήσουν υπολογιστικά τις δυνατότητες επαναστόχευσης φαρμάκων για τη θεραπεία άλλων ασθενειών και τον εντοπισμό εναλλακτικών ενδείξεων. Αυτό ισχύει ανεξάρτητα από την κατάσταση του υπό έρευνα φαρμάκου, είτε έχει εγκριθεί, αποσυρθεί, σε κλινικές δοκιμές ή απέτυχε σε κλινικές δοκιμές.

Ενώ η επαναστόχευση φαρμάκων είναι μια αναδυόμενη τεχνική, η παραδοσιακή, δαπανηρή και επιρρεπής σε αποτυχία διαδικασία de novo ανάπτυξης φαρμάκων παραμένει απαραίτητη για την ανακάλυψη και τη δοκιμή νέων φαρμάκων. Ωστόσο, η ενσωμάτωση ορισμένων υπολογιστικών μοντέλων επαναστόχευσης φαρμάκων σε αυτή τη διαδικασία μπορεί να βοηθήσει στην περαιτέρω προώθηση των φαρμάκων στη γραμμή ανάπτυξης και τελικά να βελτιώσει την αποτελεσματικότητά τους στις κλινικές δοκιμές.

Αποτελέσματα-Συζήτηση

Ο όγκος των διαθέσιμων βιοϊατρικών και φαρμακευτικών δεδομένων αυξάνεται ραγδαία, καθιστώντας τις υπολογιστικές τεχνικές επαναστόχευσης φαρμάκων —όπως η εξόρυξη δεδομένων, η μηχανική μάθηση και η ανάλυση δικτύου— απαραίτητες για τη συστηματική επαναστόχευση φαρμάκων. Αυτές οι μέθοδοι μπορούν να αντιμετωπίσουν τους περιορισμούς των παραδοσιακών στατιστικών προσεγγίσεων και να δώσουν πιο αξιόπιστα συμπεράσματα.

Ο τομέας της επαναστόχευσης φαρμάκων αναμένεται να κερδίσει σημαντικά από καινοτόμες υπολογιστικές τεχνικές που μπορούν να αποκαλύψουν σχέσεις μεταξύ διαφόρων βιολογικών οντοτήτων, συμπεριλαμβανομένων γονιδίων, πρωτεϊνών, ασθενειών και φαρμάκων, εντοπίζοντας έτσι θεραπευτικές δυνατότητες και νέες ενδείξεις για υπάρχοντα φάρμακα. Τέτοιες ανακαλύψεις είναι ζωτικής σημασίας για την ανάπτυξη θεραπειών για τον καρκίνο και άλλες προκλητικές ασθένειες, οι οποίες απαιτούν ολοκληρωμένα δεδομένα για την υποστήριξη των ερευνητικών προσπαθειών. Ο Πίνακας 2 παρέχει τα αποτελέσματα της βιβλιογραφικής ανασκόπησης των υπολογιστικών μελετών επαναστόχευσης φαρμάκων, συμπεριλαμβανομένων των στρατηγικών που χρησιμοποιούνται, των υπολογιστικών μεθόδων, των βασικών τεχνικών, των πηγών δεδομένων, των σημαντικών ευρημάτων και των μετρήσεων αξιολόγησης.

Πίνακας 2: Αποτελέσματα βιβλιογραφικής ανασκόπησης της μελέτης

	Στρατηγικές			Υπολογιστική προσέγγιση			Κύρια τεχνική	Πηγή δεδομένων	Κριτήρια αξιολόγησης
	Γονιδίωμα	Χημικές δομές	Φαινότυπος	Εξόρυξη δεδομένων	Μηχανική μάθηση	Ανάλυση δικτύου			
Li and Lu ⁵⁸	✓			✓		✓	TM	ClinicalTrials.gov and	Acc, Comp(Acc), and CSs
Leaman et al. ⁵⁹	✓		✓	✓			TM	PubMed, λεξιλόγιο MEDIC	Comp(F1) and CSs
Li et al. ⁶⁰	✓	✓		✓			TM, SM and CMap	OPHID, PubMed	CV, Comp(Sen, Spc, PPV, F1, and Acc), and CSs
Tari et al. ⁶¹	✓	✓		✓			TM and LFRs	Medline, GO, UniProt, NCBI, CancerQuest and DrugBank	Comp(Rec) and CSs

Rastegar-Mojarad et al. ⁶²	✓			✓		✓	TM and SM	SemMedD B, UMLS, CTD, andMedli ne	Comp(CoS)
Brown et al. ⁶³		✓			✓	✓	SM and Clust	MEDLINE and DrugBank	CSs -
Papanikolaou et al. ⁶⁴		✓		✓			NER	DrugBank	-
Zeng et al. ⁶⁵	✓			✓		✓	MM-AE and V- AE	DrugBank, repoDB, and ClinicalTr ials.gov	CV, AUC, AUC-PR, and Comp(Rec- TK)
Chen et al. ⁶⁶	✓		✓		✓	✓	SM and LP	Chem2Bio 2RDF [144],	Acc, AUC, Copl(Acc and
Zhu et al. ⁶⁷	✓	✓		✓			SM and WOL	PharmGKB	CSs
Gottlieb et al. ⁶⁸	✓	✓	✓		✓		SM and LR	DrugBank, SMILES, Mata- dor [143],	CV, Acc, P, AUC, and Comp(Acc and AUC)

								DCDB [146], KEGG, DailyMed [147], SIDER, GO, UniPort, and ClinicalTr ials.gov	
Menden et al. ⁶⁹	√	√		√			FF-PNN and RF	CCL	FF-PNN and RF
Napolitano et al. ⁷⁰	√	√			√		SVM and CF	CMap, NCBI, DrugBank ,	CV, Acc, Comp(AUC)
Zhang et al. ⁷¹	√	√	√		√		SM and BCD	DrugBank, NDF-RT [139],	CV, Comp(AUC), and CSs
Yang et al. ⁷²	√	√			√		CI-PMF	DrugBank, BioCarta [138],	Comp(AUC and Prec)

Aliper et al. ⁵⁴	√				√		DNN	LINCS	CV, F1, NNCM, Spr
Altae-Tran et al. ⁷³		√	√		√		CCN and LSTM	SIDER and Tox21 [129]	Acc and AUC
Hu et al. ⁷⁴	√	√		√			CNN	KEGG and DrugBank	CV, ROC and Comp(Acc, Sen, F1, and AUC)
Segler et al. ⁷⁵		√			√		RNN	SMILES	EOR
Yi et al. ⁷⁶		√		√		√	CTD	DrugBank	CV, AUC
Lei et al ⁷⁷	√			√		√	CTD	PubChem, UniPort database	CV, AUC, and CSs
Wu et al. ⁵⁶	√	√	√		√	√	SM and Clust	KEGG	CV, Comp(Acc), and CSs
Tan et al. ⁷⁸	√	√			√		SM and Clust	PubChem and DrugBank	Comp(Acc)
Rakshit et al. ⁷⁹	C √	C	C	C	√	TM and Genotator	Comp(Acc)	KEGG and Genotator	Comp(Acc)
Yang et al. ⁸⁰	√	√	√	√			ARM and LP	PharmGKB, SIDER,	CV, AUC, AUC-PR, and

								and MedHelp	Comp(Send, Spc, and F1)
Jin et al. ⁸¹	✓	✓		✓			SNF	PubChem and DrugBank	Comp(Acc) R- TIs]

Acc Accuracy, AE Autoencoder, ARM Association Rules Mining, AUC Area under curve, AUC-PR Area Under The Precision-Recall Curve, BCD Block Coordinate Descent, BER Biological Entity Relationships, BGL Bipartite Graph Learning, CF Collaborative Filtering, CHC Cultured , DNN Deep Neural Network, EOR Enrichment Over Random, F1 F1 score/F-measure, FF-PNN Feed-Forward Perceptron Neural Network, GCN Graph Convolutional Network, GWAS Genome-Wide Association Study, KF Kernel Fusion, LDA Latent Dirichlet Allocation, Λογικά δεδομένα και κανόνες LFRs, Πρόβλεψη σύνδεσης LP, Λογιστική παλινδρόμηση LR, Μακροπρόθεσμη Μνήμη LSTM, Μέση Ακρίβεια MAP, Παραγοντοποίηση Πίνακα MF, Πολυτροπικός Βαθύς Αυτόματος Κωδικοποιητής, NDCG Κανονική μειωμένη αθροιστική ονομαστική τιμή, Σωρευτικό κέρδος, NNCM Neural Net Confusion Matrix, P-value, PDo Pareto Dominance, PGx Pharmacogenomics, PPV Positive Predictive Value, Prec Precision, PSL Probabilistic Soft Logic, 'R2': Συντελεστής προσδιορισμού , 'Rp': Pearson Drug, R-DI -Αλληλεπιδράσεις ασθένειας, Rec Recall, Rec-TK Recall @ Top-K, RF Random Forest, RMSE Root Mean Square Error, RNN Recurrent Neural Network, ROC Receiver Operating Characteristic Curve, R-RAs Drug–Drug Associations, R-TIs Drug- Αλληλεπιδράσεις στόχου, Πίνακας συνδιακύμανσης δειγμάτων SCM, Ευαισθησία Sensitivity, Μέτρα ομοιότητας SM, Ειδικότητα Spc, Διαχωρισιμότητα Spr, Μηχανή διανυσμάτων υποστήριξης SVM, Σχέσεις Θεραπευτικών ενδείξεων TIRs, Εξόρυξη κειμένου TM, V-AE Variational Autoencoder, WOL Webtology

Συμπεράσματα

Μετά την εξέταση διαφόρων υπολογιστικών στρατηγικών και μοντέλων επαναστόχευσης φαρμάκων που έχουν σχεδιαστεί για τον εντοπισμό νέων θεραπευτικών αλληλεπιδράσεων, είναι σαφές ότι κάθε προσέγγιση έχει τα πλεονεκτήματα και τους περιορισμούς της. Ο συνδυασμός διαφορετικών στρατηγικών συχνά αποφέρει υψηλότερα ποσοστά επιτυχίας.

Παρά την ύπαρξη πολλών εντυπωσιακών υπολογιστικών μοντέλων επαναστόχευσης φαρμάκων, η ανάπτυξη ισχυρών μοντέλων παραμένει μια πολύπλοκη και προκλητική διαδικασία. Μια σημαντική πρόκληση είναι η μετάφραση των θεωρητικών προσεγγίσεων σε πρακτικές εφαρμογές λόγω της δυσκολίας προσομοίωσης της συμπεριφοράς των ζωντανών οργανισμών και αντιμετώπισης εμποδίων όπως λείπουν, μεροληπτικά και ανακριβή δεδομένα. Για παράδειγμα, ο καθορισμός αξιόπιστων προφίλ υπογραφής γονιδιακής έκφρασης είναι δύσκολος λόγω των παραλλαγών στις πειραματικές συνθήκες (π.χ. περιβαλλοντικές μεταβλητές και ηλικία ασθενών) σε διαφορετικά πειράματα, οδηγώντας σε αποκλίσεις δεδομένων και προκαταλήψεις. Επιπλέον, σημαντικές αλλαγές στις εκφράσεις των γονιδίων μπορεί να μην συμβαίνουν πάντα όταν αυτά τα γονίδια χρησιμοποιούνται ως στόχοι φαρμάκων, με αποτέλεσμα ανακριβή δεδομένα. Η έλλειψη δομικών δεδομένων υψηλής ανάλυσης για στόχους φαρμάκων περιπλέκει περαιτέρω τον εντοπισμό πιθανών αλληλεπιδράσεων φαρμάκου-στόχου χρησιμοποιώντας στρατηγικές χημικής δομής και μοριακής πληροφόρησης. Μια άλλη πρόκληση είναι η απουσία αξιόπιστων συνόλων δεδομένων χρυσού προτύπου για την αξιολόγηση της απόδοσης του μοντέλου.

Οι ερευνητές χρειάζεται συχνά να δημιουργήσουν τα δικά τους σύνολα δεδομένων χρυσού προτύπου και να χρησιμοποιήσουν διαδεδομένες μετρήσεις αξιολόγησης (π.χ. ακρίβεια, ανάκληση, ευαισθησία, ειδικότητα, βαθμολογία F1 και περιοχή κάτω από τη χαρακτηριστική καμπύλη λειτουργίας του δέκτη) για να συγκρίνουν και να αξιολογήσουν τα μοντέλα τους. Εναλλακτικά, πρέπει να χωρίσουν τα δεδομένα τους σε σύνολα εκπαίδευσης, δοκιμών και επικύρωσης, χρησιμοποιώντας μετρήσεις διασταυρούμενης επικύρωσης και αξιολόγησης στο K-fold για να αποφευχθεί η υπερβολική προσαρμογή. Παρά αυτές τις προκλήσεις, οραματιζόμαστε ότι η ενσωμάτωση δεδομένων πολλαπλών πηγών που σχετίζονται με φάρμακα (π.χ. χημικές δομές), ασθένειες (π.χ. φαινοτυπικές πληροφορίες) και τις επιπτώσεις τους στο ανθρώπινο σώμα (π.χ. προφίλ υπογραφής γονιδιακής έκφρασης και παρενέργειες) είναι ζωτικής σημασίας για τον εμπλουτισμό των υπολογιστικών μοντέλων επαναστόχευσης φαρμάκων και τη βελτίωση της απόδοσής τους. Επιπλέον, πολλές ασθένειες εξακολουθούν να στερούνται αποτελεσματικών θεραπειών, παρακινώντας πολυεπιστημονικούς ερευνητές και επιστήμονες να διεξάγουν μελέτες, ιδιαίτερα στην αντιμετώπιση διαφόρων καρκίνων και πολλών ορφανών και σπάνιων ασθενειών.

Συνοψίζοντας, η υπολογιστική επαναστόχευση φαρμάκων μπορεί να ωφελήσει πολύ την ανθρωπότητα ανακαλύπτοντας νέες ενδείξεις για εγκεκριμένα φάρμακα, επιταχύνοντας την ανάπτυξη νέων φαρμάκων και δίνοντας μια δεύτερη ευκαιρία στα αποσυρθέντα και αποτυχημένα φάρμακα. Με την

αυξανόμενη υποστήριξη από τις κυβερνήσεις και τις φαρμακευτικές εταιρείες για εγχειρήματα υπολογιστικής επαναστόχευσης φαρμάκων, οι ερευνητές και οι επιστήμονες θα πρέπει να συνεχίσουν να αναπτύσσουν καινοτόμα, υπερσύγχρονα μοντέλα για να επιτύχουν νέα ευρήματα και σημαντικές ανακαλύψεις.

Βιβλιογραφία

1. Park K. A review of computational drug repurposing. *Transl Clin Pharmacol*. 2019;27(2):59. doi:10.12793/tcp.2019.27.2.59
2. Scannell JW, Blanckley A, Boldon H, Warrington B. Diagnosing the decline in pharmaceutical R&D efficiency. *Nat Rev Drug Discov*. 2012;11(3):191-200. doi:10.1038/nrd3681
3. Sardana D, Zhu C, Zhang M, Gudivada RC, Yang L, Jegga AG. Drug repositioning for orphan diseases. *Brief Bioinform*. 2011;12(4):346-356. doi:10.1093/bib/bbr021
4. Shim JS, Liu JO. Recent Advances in Drug Repositioning for the Discovery of New Anticancer Drugs. *Int J Biol Sci*. 2014;10(7):654-663. doi:10.7150/ijbs.9224
5. Munos BH, Chin WW. How to Revive Breakthrough Innovation in the Pharmaceutical Industry. *Sci Transl Med*. 2011;3(89). doi:10.1126/scitranslmed.3002273
6. Jin G, Wong STC. Toward better drug repositioning: prioritizing and integrating existing methods into efficient pipelines. *Drug Discov Today*. 2014;19(5):637-644. doi:10.1016/j.drudis.2013.11.005
7. Hodos RA, Kidd BA, Shameer K, Readhead BP, Dudley JT. In silico methods for drug repurposing and pharmacology. *WIREs Syst Biol Med*. 2016;8(3):186-210. doi:10.1002/wsbm.1337
8. Nosengo N. Can you teach old drugs new tricks? *Nature*. 2016;534(7607):314-316. doi:10.1038/534314a
9. Pushpakom S, Iorio F, Eyers PA, et al. Drug repurposing: progress, challenges and recommendations. *Nat Rev Drug Discov*. 2019;18(1):41-58. doi:10.1038/nrd.2018.168
10. Jadamba E, Shin M. A Systematic Framework for Drug Repositioning from Integrated Omics and Drug Phenotype Profiles Using Pathway-Drug Network. *Biomed Res Int*. 2016;2016:1-17. doi:10.1155/2016/7147039
11. Jin G, Fu C, Zhao H, Cui K, Chang J, Wong STC. A Novel Method of Transcriptional Response Analysis to Facilitate Drug Repositioning for Cancer Therapy. *Cancer Res*. 2012;72(1):33-44. doi:10.1158/0008-5472.CAN-11-2333
12. Zhou Y, Hou Y, Shen J, et al. A network medicine approach to investigation and population-based validation of disease manifestations and drug repurposing for COVID-19. Soranzo N, ed. *PLOS Biol*. 2020;18(11):e3000970. doi:10.1371/journal.pbio.3000970
13. Beigel M. Remdesivir for the Treatment of Covid-19 — Preliminary Report. *N Engl J Med*. 2020;383(10):992-994. doi:10.1056/NEJMc2022236
14. Aggarwal S, Verma SS, Aggarwal S, Gupta SC. Drug repurposing for breast cancer therapy: Old weapon for new battle. *Semin Cancer Biol*. 2021;68:8-20. doi:10.1016/j.semcancer.2019.09.012
15. Rudrapal M, Khairnar SJ, Borse LB, Jadhav AG. Coronavirus Disease-2019 (COVID-19): An Updated Review. *Drug Res (Stuttg)*. 2020;70(09):389-400. doi:10.1055/a-1217-2397

16. Panda S, Kumari L, Badwaik HR, Shanmugarajan D. Computational approaches for drug repositioning and repurposing to combat SARS-CoV-2 infection. In: *Computational Approaches for Novel Therapeutic and Diagnostic Designing to Mitigate SARS-CoV-2 Infection*. Elsevier; 2022:247-265. doi:10.1016/B978-0-323-91172-6.00008-X
17. Cha Y, Erez T, Reynolds IJ, et al. Drug repurposing from the perspective of pharmaceutical companies. *Br J Pharmacol*. 2018;175(2):168-180. doi:10.1111/bph.13798
18. Cheng F, Desai RJ, Handy DE, et al. Network-based approach to prediction and population-based validation of in silico drug repurposing. *Nat Commun*. 2018;9(1):2691. doi:10.1038/s41467-018-05116-5
19. Choudhury S, Moulick D, Saikia P, Mazumder MK. Evaluating the potential of different inhibitors on RNA-dependent RNA polymerase of severe acute respiratory syndrome coronavirus 2: A molecular modeling approach. *Med J Armed Forces India*. 2021;77:S373-S378. doi:10.1016/j.mjafi.2020.05.005
20. DURDAĞI S. Virtual drug repurposing study against SARS-CoV-2 TMPRSS2 target. *TURKISH J Biol*. 2020;44(3):185-191. doi:10.3906/biy-2005-112
21. Jarada TN, Rokne JG, Alhaji R. A review of computational drug repositioning: strategies, approaches, opportunities, challenges, and directions. *J Cheminform*. 2020;12(1):46. doi:10.1186/s13321-020-00450-7
22. Slonim DK, Yanai I. Getting Started in Gene Expression Microarray Analysis. Troyanskaya OG, ed. *PLoS Comput Biol*. 2009;5(10):e1000543. doi:10.1371/journal.pcbi.1000543
23. Xing Z, Li D, Yang L, Xi Y, Su X. MicroRNAs and anticancer drugs. *Acta Biochim Biophys Sin (Shanghai)*. 2014;46(3):233-239. doi:10.1093/abbs/gmu003
24. Lamb J, Crawford ED, Peck D, et al. The Connectivity Map: Using Gene-Expression Signatures to Connect Small Molecules, Genes, and Disease. *Science (80-)*. 2006;313(5795):1929-1935. doi:10.1126/science.1132939
25. Iorio F, Bosotti R, Scacheri E, et al. Discovery of drug mode of action and drug repositioning from transcriptional responses. *Proc Natl Acad Sci*. 2010;107(33):14621-14626. doi:10.1073/pnas.1000138107
26. Hu G, Agarwal P. Human Disease-Drug Network Based on Genomic Expression Profiles. Jordan IK, ed. *PLoS One*. 2009;4(8):e6536. doi:10.1371/journal.pone.0006536
27. Altae-Tran H, Ramsundar B, Pappu AS, Pande V. Low Data Drug Discovery with One-Shot Learning. *ACS Cent Sci*. 2017;3(4):283-293. doi:10.1021/acscentsci.6b00367
28. Jiang W, Chen X, Liao M, et al. Identification of links between small molecules and miRNAs in human cancers based on transcriptional responses. *Sci Rep*. 2012;2(1):282. doi:10.1038/srep00282

29. Jahchan NS, Dudley JT, Mazur PK, et al. A Drug Repositioning Approach Identifies Tricyclic Antidepressants as Inhibitors of Small Cell Lung Cancer and Other Neuroendocrine Tumors. *Cancer Discov.* 2013;3(12):1364-1377. doi:10.1158/2159-8290.CD-13-0183
30. Huang H, Nguyen T, Ibrahim S, Shantharam S, Yue Z, Chen JY. DMAP: a connectivity map database to enable identification of novel drug repositioning candidates. *BMC Bioinformatics.* 2015;16(S13):S4. doi:10.1186/1471-2105-16-S13-S4
31. Huang Z, Shi J, Gao Y, et al. HMDD v3.0: a database for experimentally supported human microRNA–disease associations. *Nucleic Acids Res.* 2019;47(D1):D1013-D1017. doi:10.1093/nar/gky1010
32. Liu X, Wang S, Meng F, et al. SM2miR: a database of the experimentally validated small molecules' effects on microRNA expression. *Bioinformatics.* 2013;29(3):409-411. doi:10.1093/bioinformatics/bts698
33. Li Y, Qiu C, Tu J, et al. HMDD v2.0: a database for experimentally supported human microRNA and disease associations. *Nucleic Acids Res.* 2014;42(D1):D1070-D1074. doi:10.1093/nar/gkt1023
34. Rukov JL, Wilentzik R, Jaffe I, Vinther J, Shomron N. PharmacomiR: linking microRNAs and drug effects. *Brief Bioinform.* 2014;15(4):648-659. doi:10.1093/bib/bbs082
35. Okada Y, Wu D, Trynka G, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature.* 2014;506(7488):376-381. doi:10.1038/nature12873
36. Garnett MJ, Edelman EJ, Heidorn SJ, et al. Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature.* 2012;483(7391):570-575. doi:10.1038/nature11005
37. Rognan D. Chemogenomic approaches to rational drug design. *Br J Pharmacol.* 2007;152(1):38-52. doi:10.1038/sj.bjp.0707307
38. Keiser MJ, Setola V, Irwin JJ, et al. Predicting new molecular targets for known drugs. *Nature.* 2009;462(7270):175-181. doi:10.1038/nature08506
39. Swamidass SJ. Mining small-molecule screens to repurpose drugs. *Brief Bioinform.* 2011;12(4):327-335. doi:10.1093/bib/bbr028
40. Yamanishi Y, Araki M, Gutteridge A, Honda W, Kanehisa M. Prediction of drug–target interaction networks from the integration of chemical and genomic spaces. *Bioinformatics.* 2008;24(13):i232-i240. doi:10.1093/bioinformatics/btn162
41. Kinnings SL, Liu N, Buchmeier N, Tonge PJ, Xie L, Bourne PE. Drug Discovery Using Chemical Systems Biology: Repositioning the Safe Medicine Clofazimine to Treat Multi-Drug and Extensively Drug Resistant Tuberculosis. Briggs JM, ed. *PLoS Comput Biol.* 2009;5(7):e1000423. doi:10.1371/journal.pcbi.1000423
42. Bleakley K, Yamanishi Y. Supervised prediction of drug–target interactions using bipartite local

- models. *Bioinformatics*. 2009;25(18):2397-2403. doi:10.1093/bioinformatics/btp433
43. Novick PA, Ortiz OF, Poelman J, Abdulhay AY, Pande VS. SWEETLEAD: an In Silico Database of Approved Drugs, Regulated Chemicals, and Herbal Isolates for Computer-Aided Drug Discovery. Pappalardo F, ed. *PLoS One*. 2013;8(11):e79568. doi:10.1371/journal.pone.0079568
 44. Brouwers L, Iskar M, Zeller G, van Noort V, Bork P. Network Neighbors of Drug Targets Contribute to Drug Side-Effect Similarity. Dopazo J, ed. *PLoS One*. 2011;6(7):e22187. doi:10.1371/journal.pone.0022187
 45. Yang L, Agarwal P. Systematic Drug Repositioning Based on Clinical Side-Effects. Csermely P, ed. *PLoS One*. 2011;6(12):e28025. doi:10.1371/journal.pone.0028025
 46. Ye H, Liu Q, Wei J. Construction of Drug Network Based on Side Effects and Its Application for Drug Repositioning. Keskin O, ed. *PLoS One*. 2014;9(2):e87864. doi:10.1371/journal.pone.0087864
 47. Bisgin H, Liu Z, Fang H, Kelly R, Xu X, Tong W. A phenome-guided drug repositioning through a latent variable model. *BMC Bioinformatics*. 2014;15(1):267. doi:10.1186/1471-2105-15-267
 48. Plachouras V, Leidner JL, Garrow AG. Quantifying Self-Reported Adverse Drug Events on Twitter. In: *Proceedings of the 7th 2016 International Conference on Social Media & Society - SMSociety '16*. ACM Press; 2016:1-10. doi:10.1145/2930971.2930977
 49. Hoehndorf R, Oellrich A, Rebholz-Schuhmann D, Schofield PN, Gkoutos G V. Linking PharmGKB to phenotype studies and animal models of disease for drug repurposing. *Pac Symp Biocomput*. Published online 2012:388-399. doi:22174294
 50. Gottlieb A, Stein GY, Oron Y, Ruppin E, Sharan R. INDI: a computational framework for inferring drug interactions and their associated recommendations. *Mol Syst Biol*. 2012;8(1). doi:10.1038/msb.2012.26
 51. Sridhar D, Fakhraei S, Getoor L. A probabilistic approach for collective similarity-based drug–drug interaction prediction. *Bioinformatics*. 2016;32(20):3175-3182. doi:10.1093/bioinformatics/btw342
 52. Fabris L, Ceder Y, Chinnaiyan AM, et al. The Potential of MicroRNAs as Prostate Cancer Biomarkers. *Eur Urol*. 2016;70(2):312-322. doi:10.1016/j.eururo.2015.12.054
 53. Menden MP, Iorio F, Garnett M, et al. Machine Learning Prediction of Cancer Cell Sensitivity to Drugs Based on Genomic and Chemical Properties. Raghava GPS, ed. *PLoS One*. 2013;8(4):e61318. doi:10.1371/journal.pone.0061318
 54. Aliper A, Plis S, Artemov A, Ulloa A, Mamoshina P, Zhavoronkov A. Deep Learning Applications for Predicting Pharmacological Properties of Drugs and Drug Repurposing Using Transcriptomic Data. *Mol Pharm*. 2016;13(7):2524-2530. doi:10.1021/acs.molpharmaceut.6b00248
 55. Cheng F, Liu C, Jiang J, et al. Prediction of Drug-Target Interactions and Drug Repositioning via

- Network-Based Inference. Altman RB, ed. *PLoS Comput Biol*. 2012;8(5):e1002503. doi:10.1371/journal.pcbi.1002503
56. Wu C, Gudivada RC, Aronow BJ, Jegga AG. Computational drug repositioning through heterogeneous network clustering. *BMC Syst Biol*. 2013;7(S5):S6. doi:10.1186/1752-0509-7-S5-S6
 57. Guney E, Menche J, Vidal M, Barábasi A-L. Network-based in silico drug efficacy screening. *Nat Commun*. 2016;7(1):10331. doi:10.1038/ncomms10331
 58. Li J, Lu Z. Systematic identification of pharmacogenomics information from clinical trials. *J Biomed Inform*. 2015;45(5):870-878. doi:10.1016/j.jbi.2012.04.005
 59. Leaman R, Islamaj Doğan R, Lu Z. DNorm: disease name normalization with pairwise learning to rank. *Bioinformatics*. 2013;29(22):2909-2917. doi:10.1093/bioinformatics/btt474
 60. Li J, Zhu X, Chen JY. Building Disease-Specific Drug-Protein Connectivity Maps from Molecular Interaction Networks and PubMed Abstracts. Rzhetsky A, ed. *PLoS Comput Biol*. 2019;5(7):e1000450. doi:10.1371/journal.pcbi.1000450
 61. Tari LB, Patel JH. Systematic Drug Repurposing Through Text Mining. In: ; 2014:253-267. doi:10.1007/978-1-4939-0709-0_14
 62. Rastegar-Mojarad M, Elayavilli RK, Li D, Prasad R, Liu H. A new method for prioritizing drug repositioning candidates extracted by literature-based discovery. In: *2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE; 2015:669-674. doi:10.1109/BIBM.2015.7359766
 63. Brown AS, Patel CJ. MeSHDD: Literature-based drug-drug similarity for drug repositioning. *J Am Med Informatics Assoc*. 2018;24(3):614-618. doi:10.1093/jamia/ocw142
 64. Papanikolaou N, Pavlopoulos GA, Theodosiou T, Vizirianakis IS, Iliopoulos I. DrugQuest - a text mining workflow for drug association discovery. *BMC Bioinformatics*. 2016;17(S5):182. doi:10.1186/s12859-016-1041-6
 65. Zeng X, Zhu S, Liu X, Zhou Y, Nussinov R, Cheng F. deepDR: a network-based deep learning approach to in silico drug repositioning. Cowen L, ed. *Bioinformatics*. 2020;35(24):5191-5198. doi:10.1093/bioinformatics/btz418
 66. Chen B, Ding Y, Wild DJ. Assessing Drug Target Association Using Semantic Linked Data. Tropsha A, ed. *PLoS Comput Biol*. 2022;8(7):e1002574. doi:10.1371/journal.pcbi.1002574
 67. Zhu Q, Tao C, Shen F, Chute CG. Exploring the pharmacogenomics knowledge base (PharmGKB) for repositioning breast cancer drugs by leveraging Web ontology language (OWL) and cheminformatics approaches. *Pac Symp Biocomput*. Published online 2018:172-182. <http://www.ncbi.nlm.nih.gov/pubmed/24297544>
 68. Gottlieb A, Stein GY, Ruppin E, Sharan R. PREDICT: a method for inferring novel drug indications

- with application to personalized medicine. *Mol Syst Biol.* 2011;7(1). doi:10.1038/msb.2011.26
69. Menden MP, Iorio F, Garnett M, et al. Machine Learning Prediction of Cancer Cell Sensitivity to Drugs Based on Genomic and Chemical Properties. Raghava GPS, ed. *PLoS One.* 2016;8(4):e61318. doi:10.1371/journal.pone.0061318
 70. Napolitano F, Zhao Y, Moreira VM, et al. Drug repositioning: a machine-learning approach through data integration. *J Cheminform.* 2018;5(1):30. doi:10.1186/1758-2946-5-30
 71. Zhang P, Wang F, Hu J. Towards drug repositioning: a unified computational framework for integrating multiple aspects of drug similarity and disease similarity. *AMIA . Annu Symp proceedings AMIA Symp.* 2014;2014:1258-1267. <http://www.ncbi.nlm.nih.gov/pubmed/25954437>
 72. Yang J, Li Z, Fan X, Cheng Y. Drug–Disease Association and Drug-Repositioning Predictions in Complex Diseases Using Causal Inference–Probabilistic Matrix Factorization. *J Chem Inf Model.* 2019;54(9):2562-2569. doi:10.1021/ci500340n
 73. Altae-Tran H, Ramsundar B, Pappu AS, Pande V. Low Data Drug Discovery with One-Shot Learning. *ACS Cent Sci.* 2018;3(4):283-293. doi:10.1021/acscentsci.6b00367
 74. Hu S, Zhang C, Chen P, Gu P, Zhang J, Wang B. Predicting drug-target interactions from drug structure and protein sequence using novel convolutional neural networks. *BMC Bioinformatics.* 2019;20(S25):689. doi:10.1186/s12859-019-3263-x
 75. Segler MHS, Kogej T, Tyrchan C, Waller MP. Generating Focused Molecule Libraries for Drug Discovery with Recurrent Neural Networks. *ACS Cent Sci.* 2018;4(1):120-131. doi:10.1021/acscentsci.7b00512
 76. Yi H-C, You Z-H, Wang L, Su X-R, Zhou X, Jiang T-H. In silico drug repositioning using deep learning and comprehensive similarity measures. *BMC Bioinformatics.* 2021;22(S3):293. doi:10.1186/s12859-020-03882-y
 77. Lei S, Lei X, Liu L. Drug repositioning based on heterogeneous networks and variational graph autoencoders. *Front Pharmacol.* 2022;13. doi:10.3389/fphar.2022.1056605
 78. Tan F, Yang R, Xu X, et al. Drug repositioning by applying ‘expression profiles’ generated by integrating chemical structure similarity and gene semantic similarity. *Mol Biosyst.* 2019;10(5):1126. doi:10.1039/c3mb70554d
 79. Rakshit H, Chatterjee P, Roy D. A bidirectional drug repositioning approach for Parkinson’s disease through network-based inference. *Biochem Biophys Res Commun.* 2019;457(3):280-287. doi:10.1016/j.bbrc.2014.12.101
 80. Yang CC, Zhao M. Mining heterogeneous network for drug repositioning using phenotypic information extracted from social media and pharmaceutical databases. *Artif Intell Med.* 2020;96:80-92. doi:10.1016/j.artmed.2019.03.003

81. Jin S, Niu Z, Jiang C, et al. HeTDR: Drug repositioning based on heterogeneous networks and text mining. *Patterns*. 2021;2(8):100307. doi:10.1016/j.patter.2021.100307
82. Gong L, Huang D, Sun S, et al. Extraction of Interactions of Genes2Genes Related to Breast Cancer. In: *2018 IEEE 16th International Conference on Software Engineering Research, Management and Applications (SERA)*. IEEE; 2018:108-112. doi:10.1109/SERA.2018.8477190
83. Chen H, Engkvist O, Wang Y, Olivecrona M, Blaschke T. The rise of deep learning in drug discovery. *Drug Discov Today*. 2018;23(6):1241-1250. doi:10.1016/j.drudis.2018.01.039