Drug Repurposing Using AI: Case Studies

Mayuri Chavan, Aastha Barga, Shreyasi Gurav

TMV's Lokmanya Tilak Institute of Pharmaceutical Sciences, Pune

Corresponding author: Mayuri Chavan
Email:mayurichavan0820@gmail.com
Doi:**10.5281/zenodo.16833654**
Received: 21 July 2025
Accepted: 29 July 2025

Abstract:

Drug repurposing, also called drug repositioning, is the process of identifying new therapeutics uses for existing drugs. A very full-cost and time-given undertaking of bringing in a whole new drug, it is mainly to this group of existing- approved and investigated drugs, where the researcher hopes to find something useful for a different disease or condition from that originally intended in the first place. In this study, three case studies are introduced which document the usage of AI as a tool in identifying new therapeutic indications for existing drugs. The first case details the invasion of the AI-generated pipeline to repurpose metformin-first-line therapy for type 2 diabetes-for Normal Tension Glaucoma (NTG) under the framework for Mendelian Randomization (MR). The study hypothesized its neuroprotection via glucagon and GLP1 pathways with some epidemiological evidence. The second case illustrates metformin as a supposed drug to reduce cancer mortality employing a range of large electronic health records (EHR) databases, which were merged with natural language processing and statistical modeling. The analysis showed statistically significant improved survival in metformin users in many cohorts. The third one indicates the repurposing of drugs developed for Fragile X Syndrome (FXS) through high-throughput testing and machine learning. The DREAM-RD identified Sulindac and other agents as worthy candidates, some of which have progressed into full- blown clinical trials. All three cases come with limitations that pertain to either specificity to a population, sample size, or, even, too little clinical data for an efficacious recommendation; all three nevertheless begin to set a benchmark for AI-based drug repositioning. The examples furnish a template for fast-tracking therapeutic discovery and development, especially in the domain of poorly served or rarer diseases.

Keywords: Glaucoma, Fragile X Syndrome, Metformin, EHR , AI

Introduction:

The pharmaceutical industry has always contended with complications in bringing a new drug to market; hence, traditional drug discovery and development is an expensive, prolonged, and uncertain attempt. Typically, the development of a new drug requires 10 to 17 years and costs about 800 million dollars, and the success rate is below 10%. Recently, repurposing drugs or identifying new uses of existing drugs has presented a good option. This approach leverages the known safety profiles of existing medications, thus shortening development time and costs.
Artificial intelligence (AI) has enlarged the scope of drug-repurposing research in a dramatic manner by making it possible to merge and thoroughly analyze heterogeneous data; these might include but are not limited to electronic health records, genomic data, and biomedical research activity. AI-infused drug repurposing is thus a more effective and efficient strategy than the classical methods in opening the world of ineffective medicines into new therapeutic indications to meet unfulfilled medical needs, especially in rare diseases where data is scanty and diverse [1].

Computational approaches, such as structure-based virtual screening and analysis of genomic data, are being developed to predict drug repurposing signals. Electronic health record (EHR) databases can validate potential signals and facilitate cost-effective clinical studies, making them a valuable resource for drug repurposing research [2].Computational methods analyze gene expression profiles to identify potential drug candidates by combining disease-specific signatures and drug-induced perturbations. In the case of Fragile X Syndrome (FXS), researchers used a unique approach combining gene expression profiles from FXS and autism patients with machine learning methods to derive a robust disease signature and highlight causative changes underlying autistic phenotypes. This approach aims to identify potential therapeutic targets for FXS. The fusion of drug repurposing with AI has been deemed very promising for overcoming challenges associated with rare diseases, and many platforms, libraries, initiatives, and companies are engaged in using AI to promote innovation in this area [1,3].

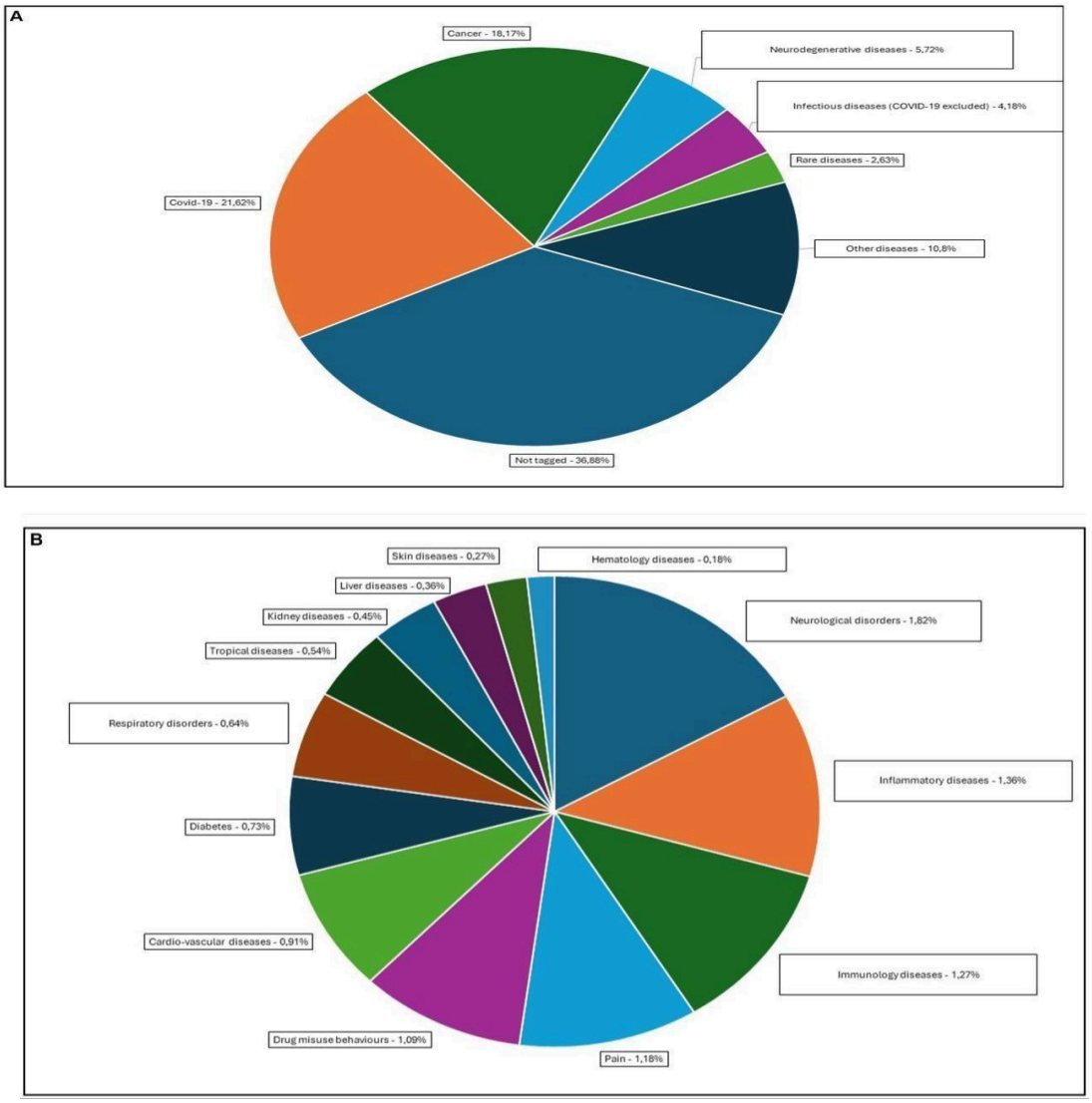


Figure 1 : Overview of the diseases targeted by AI applications in drug repurposing. (A) Overview of the diseases representing more than 2% of the retrieved references: COVID-19, cancer, neurodegenerative diseases, infectious diseases, rare diseases, and other diseases and not tagged. (B) Overview of the diseases representing less than 2% of the retrieved references: neurological disorders, inflammatory diseases, immunology diseases, pain, drug misuse behaviors, cardiovascular diseases, diabetes, respiratory disorders, tropical diseases, kidney diseases,liver diseases,and skin diseases**:**(Source: https://doi.org/10.3389/fmed.2024.1404338)

The traditional drug discovery vs drug repurposing:

The traditional drug discovery involves the development of completely new drugs. It comprises five phases: discovery and preclinical testing, safety evaluation, clinical trials, FDA approval, and post-market safety monitoring. It is a costly process that takes approximately 10–16 years and entails a high possibility of failure.

Drug repurposing, in contrast, creates new applications for existing drugs. It involves four phases: compound identification, acquisition, further development, and post-market safety

assessment. It employs cutting-edge technologies such as bioinformatics and artificial intelligence (AI), which facilitate the process to become quicker and cheaper. Repurposed drugs develop in 3–12 years and cost much less compared to new drug development.

Recycled drugs already possess safety information from previous use, minimizing the risks and expenses. For instance:

- Sildenafil (Viagra): Originally created for heart-related illnesses, subsequently repurposed for treating erectile dysfunction.
- Metformin: A drug for diabetes that is now being researched as a possible cancer therapy.

Repurposing is also helpful in treating neglected diseases and new health challenges. It takes advantage of available biological information, databases, and computational resources to gain more insight into drug mechanisms and discover new uses. In recent years, the use of in silico techniques along with the application of structure-based drug design (SBDD) and artificial intelligence (AI) technology has further accelerated the drug purposing process [6].

Case study 1: Metformin repurposing in Glaucoma

The 1850s saw the first reported medical attempts at treating Normal-tension glaucoma (NTG): glaucoma with optic nerve damage but no elevated intraocular pressure. Its treatment continues to pose a challenge to medical professionals to this day. This study attempts to treat NTG by repurposing metformin, an NTG candidate and first-line treatment for type 2 diabetes. Metformin's therapeutic potential for NTG is now validated by our pioneering pipeline that amalgamates generative artificial intelligence (GAI) alongside drug target-based Mendelian randomization (MR) – an NTG-first in our study. Amidst extensive screening through GAI with multiple mainstream models, we noted metformin advanced as a strong candidate.

Metformin's potential was further corroborated through MR analyses in European genetic datasets, suggesting causal mediation via glucagon and GLP-1 pathways. These are consistent with metformin's action– suppression of gluconeogenesis, insulin secretion enhancement, oxidative stress diminishment, and vasoregulatory dysfunction reduction—NTG pathogenesis factors. Supporting evidence from large cohort studies, such as a 10-year US retrospective analysis and the Rotterdam study indicates metformin reduces open angle glaucoma risk in

diabetic patients. GLP-1 receptor agonists display similar effects in other forms of glaucoma and neuroprotective roles. This evidence implicitly supports our theory regarding metformin's potential impact on slowing NTG progression. Limitations include the focus towards Europeans and lack of randomized clinical trials, requiring more confirmation. The study not only places metformin as a cost-effective candidate for NTG but also builds a solid GAI-MR model for drug repurposing, going beyond NTG. Combining genetic factors with sophisticated algorithms, the study provides a benchmark model for therapy design and calls future research to test the applicability of metformin in NTG treatment.

A study was conducted using ChatGPT-4o, Anthropic Claude 3.5, and Google Gemini Advance to identify drug repurposing candidates (DRCs) for NTGAs. Each GAI model was prompted with 10 independent queries. Each query returned 20 candidate drugs, resulting in a total of 200 candidates per model. The candidates were assigned a score from 1 to 20 based on their rank/order in each list. After all 10 queries: The frequency of each drug's appearance was calculated (max frequency = 10). The average score of each drug was also calculated. Drugs with high appearance frequency and low average scores were considered strong repurposing candidates. The same analytical procedure was carried out independently for all three GAI models. A final list of unique DRCs was compiled for each model based on frequency and ranking. Only those DRCs that appeared in all three GAI-generated lists were selected as top candidates. Final consensus identified three common drugs across all models:1) Metformin, 2) Losartan, 3) Memantine. Supporting data is provided in Supplementary Figure 2 [7]

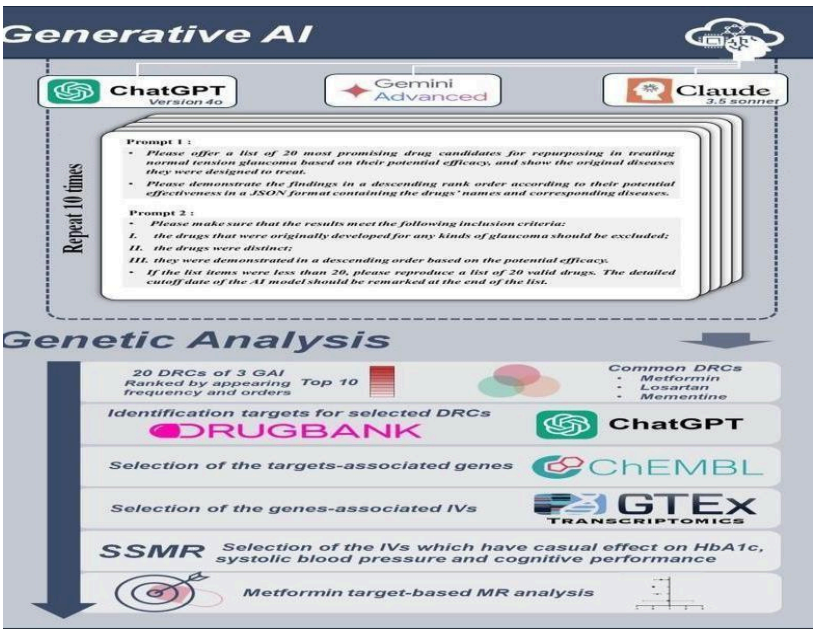


Figure 2 : The workflow for the current study (Source: https://doi.org/10.1101/2024.12.02.24318301)

Case Study 2: Metformin repurposing in Cancer treatment

The metformin and cancer study aims to validate the relationship between cancer mortality and metformin (medication for type 2 diabetes) using electronic health records (EHRs) data. The main goal was to evaluate whether there was correlation between increased metformin use and decreased cancer mortality in patients with illness, especially those with type 2 diabetes. The researchers used automated informatics techniques to grab the right data from EHRs, which data included among other things, demographic and clinical information and also exposed medical records [8,9].

DESIGN OF THE STUDY: included classifying the cancer patients

a. Patients with type 2 diabetes and on metformin;

b. Patients with type 2 diabetes and on other glycemic agents

c. Patients with type 2 diabetes and on insulin only.

d. Nor-diabetic patients who do not take any medications for diabetes [10].

Natural language processing (NLP) techniques were used to extract medication information from EHRs, allowing for a thorough assessment of medication exposure and its effects on cancer outcomes. Statistical evaluation was conducted using stratified Cox models to adjust to changes in all-cause mortality while controlling for older age, sex, race, BMI, smoking status, usage of insulin, type of cancer, and the Charlson comorbidity index.

The Investigators suggested that use of metformin would result in higher survival rates for cancer patients. The study illustrates the use of other informatics tools to verify signals of drug repurposing and validate large observational studies with minimal patient monitoring. The study has several limitations. Bias may arise from using EHR data instead of primary pharmacy records. Patients with CHF and chronic kidney disease were excluded due to metformin contraindications. Immortal time bias occurred due to uncertain cancer medication exposure dates. Small sample sizes limited stratification by histological subtype within each cancer type. The supporting data is shown in table1,2 and figure 3,4 [2].

Table 1 : Data of retrospective cohort study by using EHR's at VUMC and Mayo clinic [2].

Source Institution	**Database type**	**Patient Cohort**	**Timeframe**	**Key Findings**
Vanderbilt University Medical Center	EHR Database	32,415 patient	1995-2010	Metformin users: 22% lower mortality vs. hypoglycemic patients, 39% lower vs. insulin- only users. Diabetic metformin users: 23% greater survival than non-diabetic patients.
Mayo Clinic	EHR database	79,258 patients	1995-2010	Similar results to Vanderbilt: Reduced mortality for Metformin users & improved survival for diabetic Metformin users compared to non- diabetic patients.

Table 2 : Statistical analysis of Medication group and Hazard ratio [2]

Name of Population	Size of Population	Excluding/ Including	Medication Group	Hazard Ratio
Vanderbilt	42,165 individuals 28,917 + 3,498	Excluding Skin cancer, CHF/CKD, And age<18 years Without DM2 + with DM2	63% used Metformin 26% used other oral DM2 medications 11% were on insulin monotherapy	Metformin vs. other oral DM2 medications: HR 0.78 (95% CI:0.69-0.88) Metformin vs. insulin monotherapy: HR 0.61 (95% CI: 0.50-0.73) Metformin vs. non-diabetic group: HR 0.77 (95% CI: 0.71-0.85

Mayo Clinic	79,258	Across the four groups	Patients on other oral DM2 medications (p<0.001) Patients on insulin monotherapy (p<0.001) Patients without diabetes (p<0.001)	Metformin group vs. other oral DM2 medications: HR 0.70 (95% CI: 0.63-0.77) Metformin group vs. insulin monotherapy: HR 0.65 (95% CI: 0.58-0.73) Metformin group vs. non- diabetic group: HR 0.59 (95% CI: 0.54-0.65)

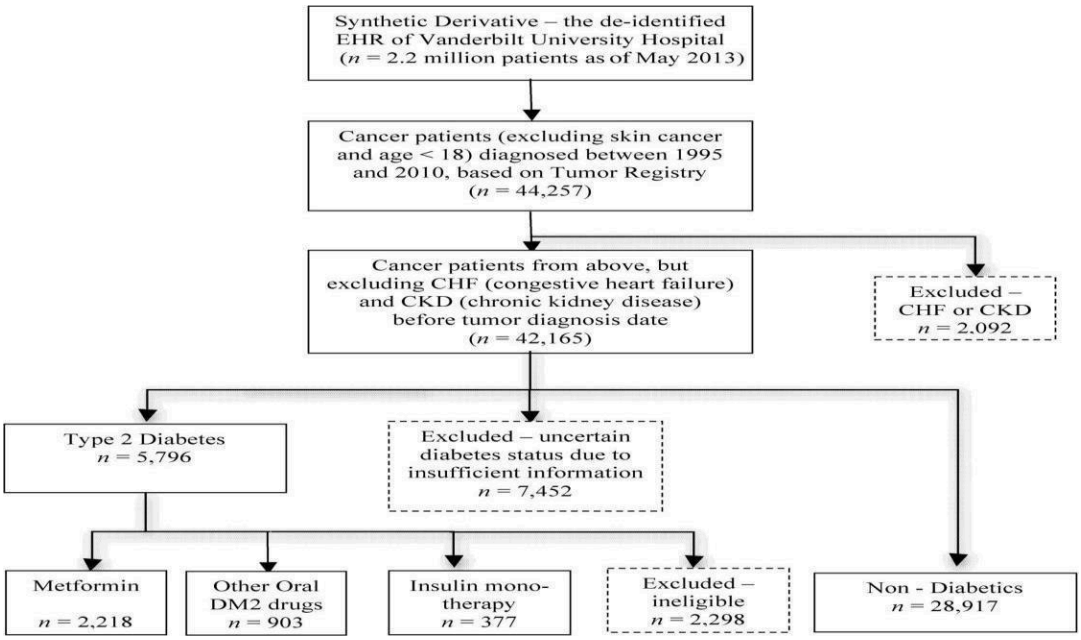


Figure 3 : The study design and data extraction workflow for patients in the Vanderbilt electronic health record (EHR) system from January 1995 to December 2010 (Source: https://doi.org/10.1136/amiajnl-2014-002649).

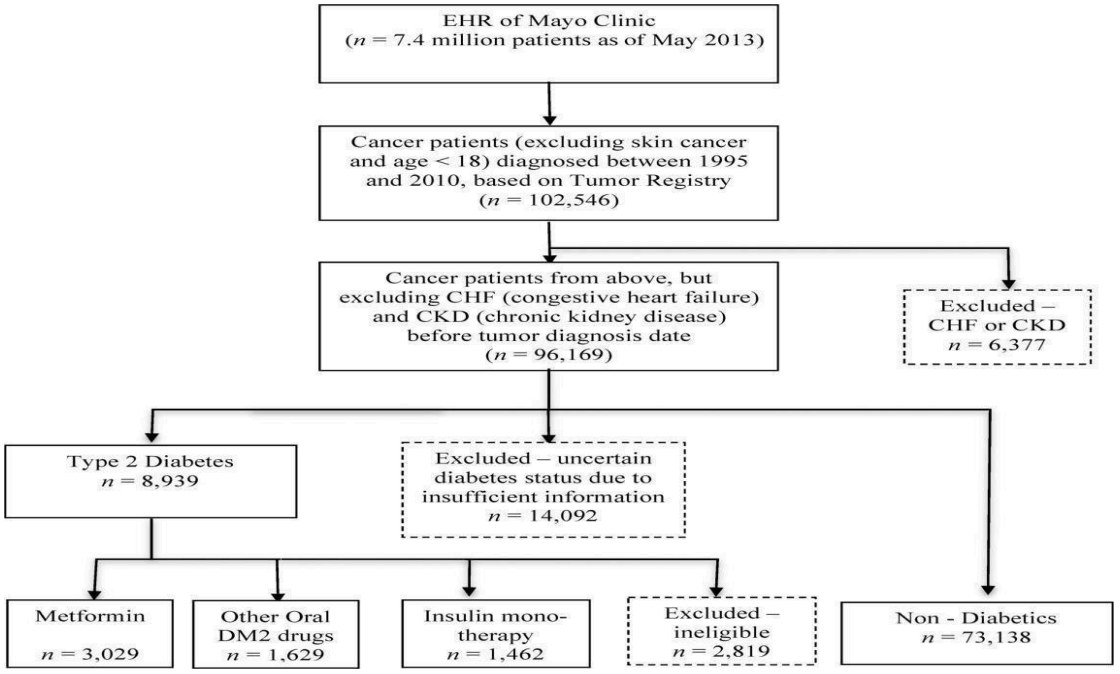


Figure 4 : The study design and data extraction workflow for patients in the Mayo Clinic electronic health record (EHR)

system from January 1995 to December 2010 (Source: https://doi.org/10.1136/amiajnl-2014-002649)

Case Study 3: Drug repurposing for Fragile X Syndrome

Fragile X Syndrome is a genetic disease which occurs due to a mutation in the FMR1 gene. Due to its rarity, FXS affects a very small group of people and therefore traditional drug development methods are time consuming and expensive. Artificial intelligence (AI) and machine learning (ML) have brought new opportunities for both the repurposing of existing drugs and the research efforts required finding a cure or treatment for it.

The Drug repurposing using AI/ML tools for Rare Diseases (DREAM-RD) initiative is a computational approach developed to identify existing drugs used as treatments for FXS. One of the key approaches to solving this problem is using the combination of HTS (high- throughput screening) and machine learning algorithms. For instance, for an analysis of how specific compounds affect cells in animals modeling FXS, scientists use HTS data to explore how the effects of various compounds might be related to FXS in the species.

The example of DREAM-RD was on the repurposing of Sulindac (a non-steroidal anti- inflammatory drug) used as a treatment for FXS. Researchers in a previous study proposed that Sulindac could modulate pathways affected by the FMR1 gene mutation. Using an AI-driven analysis, scientists were able to predict how effective the drug would be in improving symptoms of FXS. In subsequent preclinical trials, the anti- FMRP (Fragile X mental retardation protein) drug also demonstrated positive results, and clinical trials with Sulindac will now begin to determine its safety and effectiveness in patients with FXS.

Identified candidates: chlorpropamide, tazarotene, trifluoperazine and pioglitazone .These are found to have promising prospects from preliminary experiments; however, the challenge arises from lack of human data and dependence on cell line results. They suggest generating drug- induced profile in model organisms which could help to increase chances of success in clinical trials.

Tool used for drug repurposing: DREM-RD: DREAM-RD used to evaluate which algorithm performs best for the datasets used and it focused on achieving the highest classification scores for disease-related differentially expressed genes (DEGs).

Gene Selection Using ML:

o Machine learning (ML) is utilized for gene profiling and selection.

o Methods employed include feature selection, feature elimination, and feature extraction algorithms. Gene selection is also viewed as a dimension reduction process, used to identify and isolate “important” genes.

Algorithms Used in the Study:

1. Principal Component Analysis (PCA): A dimensionality reduction algorithm.

2. Random Forest (RF): A tree-based ensemble algorithm.

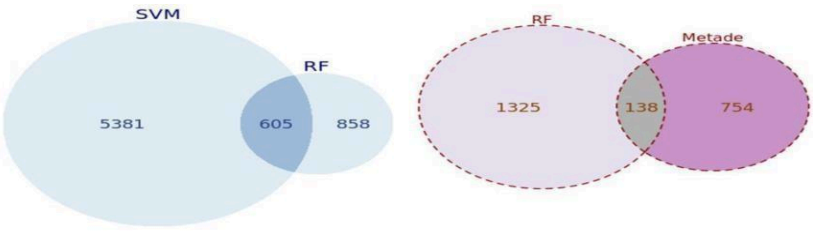
3. Support Vector Machine (SVM): A supervised algorithm that maps data into higher dimensions [3]. Refer to the table 3,4 and figure 5 for supporting data [3, 4, 5].

Table 3 : Algorithm of gene selection using machine learning (ML) [3]

Sr. No.	Aspects	PCA Algorithm	RF Algorithm	SVM Algorithm
1.	Type	not survived	Survived	Survived
2.	Purpose	Feature extraction and dimensionally reduction	Classification, regression,and identifying "important features" using"variable importance"	Classification and regression (preferred for classification)
3.	Evaluation Metrics		F1 Score and Area Under the Curve (AUC)	F1 Score and Area Under the Curve (AUC)
4.	Performa n ce Metrics	209 Principal Components (resulting in a gene listof 209 genes/features)	F1 Score: 90% AUC: 93%	F1 Score:76% AUC: 76%
5.	Result	First 4 Principal Components shown in Figure S5, displaying variance ratios captured		List of significant genes extracted

Table 4 : Data analysis to derive lists of differentially expressed genes (DEGs) [3]

Aspect	Individual Analysis	Meta-analysis
Database	getGEO	MetaDE
P-value obtained	P< 0.01	P< 0.001



Method	DEG list	Threshold
MetaDE	892	p-value<0.01
PCA	209	209
Random Forest	1463	Importance >0
SVM	5986	Mean

Figure 5 : Summary of DEGs for FXS/ASD
obtained using various computational methods (Source : https://doi.org/10.1101/2020.09.25.311142)

Table 5 : Existent initiative of Drug Repurposing [1]

Sr. No.	Name	Funding/Partnership	Objectives	Methods	Examples/ Focus Area
1.	Every Cure	ARPA-H(Advanced Research Projects Agency for Health)	Develop comprehensive, open-source database of drug repurposing opportunities; launch clinical trials	Integrates multiple data sources(Pub Med, clinical trials GOV, etc.), uses NLP and ML	Castle man disease (adalimumab), COVID-19 (dexamethasone, tocilizumab), angios arcoma (pembr olizumab)
2.	REPO4 EU	European Union	Provide a platform for validated precision drug repurposing, open to stakeholders	Advanced bioinformatics, AI, real- world big data analysis	Mechanism- based drug repurposing
3.	Open Targets	Public-private partnership (Bristol Myers Squib b, Genentech, GSK, Pfizer,Sarnoff, Sanger Institute,MBL-EBI)	Systematic drug target identification and prioritization using genetics and genomics	Integrates public domain data, GWAS, functional genomics	Target identification and prioritization
4.	Broad Drug Repurposing Hub	N/A (Developed by Broad Institute)	Open-access drug library with >6,000compounds to accelerate drug repurposing	Compound identification, annotation, experimental confirmation	Annotations for drugs in phases 1-3 of clinical development
5.	REMED i4ALL	European Union (5- year initiative, launched Sep 2022)	Build a platform for patient- centered medicine repurposing across the value chain	In silicon tools, ML, AI, global community of practice	Pancreatic cancer, COVID-19,rare/ultra- rare diseases

Funded Initiative:

Name	Funding	Duration	Objective	Methods	Focus areas
1. DREAMS	European Commission (€8 million)	5 years (announced Nov 29, 2023)	Develop treatments for five rare neuromuscular diseases	Al, stem cells, pharmacological screening	Duchene muscular dystrophy, centronuclear myopathy, Emery-Defuses muscular dystrophy, Pompe disease.

Companies:

Name	Location	Objectives	Methods	Focus areas	Notable achievements
1.HealX	UK	Identify promising drugs/compou nds for rare diseases using Al	Heal Net platform, ML for disease knowledge extraction	15 conditions (e.g., Fragile X syndrome, neurofibromatosis type 1, myotonic dystrophy type-1)	IND approval for Sulindac (Fragile X,2021) ODD For Sulindac (EU/US), ODD for Nitrox online (US, 2023)
Biovista	N/A	Develop a Pipeline of repositioned drug candidates across various disease areas	Project Prodigy All platform, mechanism- based analysis	Neurodegenerative diseases, epilepsy, oncology, orphan diseases	Partnerships with Astell as, Pfizer Novartis, etc.

Challenges:

Perhaps one major limitation is that different experimental conditions do not share consistent gene expression profiles. Computational modeling is further convoluted by the extreme complexity of simulating the temporal dynamics of the three-dimensional structure of a particular biological entity, which is a crucial aspect of drug-target identification efforts. Data behind molecular–interaction methods are often incomplete, so even when molecular interaction prediction methods output edges with great reliability, confidence in those outputs diminishes when predictions fall short. Intermingled complexity and bias present in these datasets make it less likely for promising drug repurposing candidates to gain acceptance, as the predicted effects may not translate well or be verifiable in more expansive clinical settings.

Perhaps another major limitation is that one finds a lack of thoroughly defined, gold-standard datasets that constrict the possibility of robust model evaluation and comparison. As if this were not enough, the fragmentation of evaluation datasets among individual studies limits fair and effective comparisons of performance across models. While training, testing, and validation splits are indeed employed, the mitigating factor therefore remains the very limited availability of data to satisfy model performance, in a way that stands in the way of any substantive degree of successful integration of computational methodologies for clinical applications [11].

Opportunities:

Next-generation integrated, multisource approaches are required to consider biomedical limitations, ranging from diverse protein characteristics through to patient-level clinical analysis. Meanwhile, clinical histories, chemical properties of drugs, disease functional pathways, and gene expression profiles will permit construction of robust multidimensional representations

where improvements in model performance are considerably expected. Beyond improvement in predictive accuracy, this type of integration permits more reliable hypotheses on drug- repurposing. Such hypotheses, when mathematically supported by converging lines of evidence- such as genetic, expression, structural, and clinical-are those presumably most likely to be validated through experimentation. For these methods to become trustworthy, it is also necessary to standardize evaluation metrics across different methodological categories. It should be noted that this aim can be achieved by producing gold-standard experimentally validated datasets, especially concerning drug-target interactions and off-target interactions.

In addition, the field could be benefited by copying best practices from other areas of biology, such as CASP in protein structure prediction, CAFA in function annotation, and CAGI in genome interpretation, in which standardized assessment brought huge advancements. Finally, the ever-present unmet clinical needs that resonate loudly in the context of rare diseases create pressure to develop low-cost and swift developing drug strategies. Speeding up drug repurposing methods development promises to be a fast and effective avenue to meet these pressing clinical needs [11].

Current Trends:

TxGNNs use knowledge graphs other than biomedical fields such as linking drugs, diseases, genes and proteins. These are GNNs that process their data into the biomedical domain and allow machines to discover hidden relationships and identify candidates for drug repositioning. This is what makes GNNs especially powerful-they model complex interactions and biological pathways, which allow researchers to predict therapeutic potential for under-investigated rare diseases [12]. The latest researches in artificial intelligence, like the Drug CLIP for the contrasts in learning, allow the models to predict relative relationships in drug-and-disease profile without having the need to create the non- existing negative examples from manual records. In this, the modeled positive pairs get compared with each other and predicted that against every other entry; these hold final predictions quite robust, even with little data in curating conditions [13].

AI-powered natural language processing (NLP) models, especially domain-specific ones like BioBERT and SciBERT, are scouring and garnering from vast scientific literature. Through such models, potential drug-disease links, biological mechanism, and clinical trial results favorable for repurposing can be discovered. AI is creating opportunity for rare and neglected diseases by aggregating numerous existing drugs and assessing them systematically for alternative indications. In numerous recent studies, AI has been used for repurposing neurological drugs and antiviral for treatments of rare genetic diseases that are currently untreated [14, 15]. Safety is critical to drug repositioning. AI models are employed to predict toxicological profiles, drug interaction profiles, and drug dosage for the new indication. Tools such as DeepTox and ADMETlab are incorporated in the repurposing workflow to ensure relative efficacy and safety of candidates for various patient populations or new contexts of treatment [16, 17].

Future Directions:

AI uses the training of large language models (LLMs) on scientific texts and structured databases to infer indirect associations between drugs and diseases. Model enhancement with reinforcement learning enables it to prioritize hypotheses based on their predicted efficacy, side effect profiles, and metrics of similarity to patients [18]. Natural language Processing (NLP) and Machine Learning (ML) models are useful in correlating groups of unstructured notes, treatment histories, and lab results with drug use and disease outcomes. This allows for the data-driven retrospective analysis to suggest alternative indications for approved drugs [19].

The collaboration between large pharmaceutical companies and AI technology firms has fostered the integration of AI in drug discovery, resulting in the development of innovative tools for this field. Public-private partnerships, such as the Alliance for Drug Discovery Innovation (ADDI), foster collaboration and offer the essential funding and resources. This propels advancement in precision medicine, facilitating the creation of targeted treatments that align with patients' genetic profiles, enhancing treatment efficacy and minimizing side effects.

Quantum computing has significant promise for addressing intricate molecular challenges and may greatly accelerate drug development by enhancing chemical structures. Integrating AI in drug development alongside partnerships and progress in quantum computing can assist the pharmaceutical sector in addressing current obstacles and investigating new therapeutic options. These advancements are poised to revolutionize drug development, enhancing health outcomes for patients globally [20].

Conclusion:

This research showcases three case studies illustrating the use of AI in drug repurposing. The initial case study utilized metformin, a primary treatment for type 2 diabetes, to address Normal Tension Glaucoma (NTG) through a process that merges generative artificial intelligence (GAI) with drug target-based Mendelian randomization (MR). The second case study employed metformin to decrease cancer mortality through the examination of electronic health records (EHR) databases using statistical modeling and natural language processing. The third case study utilized AI and machine learning to discover possible therapeutic candidates for Fragile X Syndrome (FXS) via high-throughput testing and machine learning.

Although each case study has its constraints, they together illustrate the capacity of AI-driven drug repurposing to accelerate therapeutic discovery and development, especially for rare and overlooked diseases. Employing AI can assist in tackling the difficulties linked to conventional drug development, such as elevated expenses, extended timelines, and minimal success rates. Moreover, AI-driven drug repurposing can utilize existing biological data, databases, and computational tools to enhance understanding of drug mechanisms and identify novel applications.

References:

1. Cortial, L., Montero, V., Tourlet, S., Del Bano, J., & Blin, O. (2024). Artificial intelligence in drug repurposing for rare diseases:a mini-review. Frontiers in Medicine, https://doi.org/10.3389/fmed.2024.1404338

2. Xu, H., Aldrich, M. C., Chen, Q., Liu, H., Peterson, N. B., Dai, Q., Levy, M., Shah, A., Han,

3. X., Ruan, X., Jiang, M., Li, Y., St Julien, J., Warner, J., Friedman, C., Roden, D. M., & Denny, J.

4. C. (2014). Validating drug repurposing signals using electronic health records: a case study of metformin associated with reduced cancer mortality. Journal of the American Medical Informatics Association, 22(1), 179–191. https://doi.org/10.1136/amiajnl-2014-002649

5. Agastheeswaramoorthy, K. and Sevilimedu, A. (2020) Drug repurposing using AI/ML Tools - for rare diseases (Dream-RD): A case study with fragile x syndrome (FXS) [Preprint]. doi:10.1101/2020.09.25.311142.

6. Tanoli Z, Seemab U, Scherer A, Wennerberg K, Tang J, Vähä-Koskela M. Exploration of databases and methods supporting drug repurposing: a comprehensive survey. Brief Bioinform. 2020 Feb 14;

7. Glicksberg BS, Li L, Chen R, Dudley J, Chen B. Leveraging big data to transform drug discovery. In: Methods in Molecular Biology. Humana Press Inc.; 2019. p. 91–118.

8. Mithun, R., Shubham, J. K., & Anil, G. J. (2020). Drug Repurposing (DR): An Emerging Approach in Drug Discovery. *Badria Farid A, ed. Drug repurposing. Rijeka: IntechOpen*.

9. Junhong Jiang, Di Hu, Qi Zhang, Zenan Lin medRxiv 2024.12.02.24318301; doi: https://doi.org/10.1101/2024.12.02.24318301

10. Landman GW, Kleefstra N, van Hateren KJ, et al. Metformin associated with lower cancer mortality in type 2 diabetes: ZODIAC-16. Diabetes Care 2010;33:322–6.

11. Currie CJ, Poole CD, Jenkins-Jones S, et al. Mortality after incident cancer in people with and without type 2 diabetes: impact of metformin on survival. Diabetes Care 2012;35: 299–304

12. Denny, J. C. (2012). Chapter 13: Mining electronic health records in the genomics era. *PLoS computational biology*, *8*(12), e1002823.

13. Cousins, H. C., Nayar, G., & Altman, R. B. (2024). Computational Approaches to drug repurposing: methods, challenges, and opportunities. Annual Review of Biomedical Data Science, 7(1), 15–29. https://doi.org/10.1146/annurev-biodatasci-110123-025333

14. Huang, Kexin, Payal Chandak, Qianwen Wang, Shreyas Havaldar, Akhil Vaid, Jure Leskovec, Girish N. Nadkarni, Benjamin S. Glicksberg, Nils Gehlenborg, and Marinka Zitnik. "A foundation model for clinician-centered drug repurposing." *Nature Medicine* 30, no. 12 (2024): 3601-3613.

15. Lu, Y., Hu, Y., & Li, C. (2024). Drugclip: Contrastive drug-disease interaction for drug repurposing. *arXiv preprint arXiv:2407.02265*.

16. Lee, J., Yoon, W., Kim, S., Kim, D., Kim, S., So, C. H., & Kang, J. (2020). BioBERT: a pre-

17. trained biomedical language representation model for biomedical text mining.

Bioinformatics, *36*(4), 1234-1240.

18. Alsentzer, E., Murphy, J. R., Boag, W., Weng, W. H., Jin, D., Naumann, T., & McDermott,

19. M. (2019). Publicly available clinical BERT embeddings. *arXiv preprint arXiv:1904.03323*.

20. Myung, Y., de Sá, A. G., & Ascher, D. B. (2024). Deep-PK: deep learning for small molecule pharmacokinetic and toxicity prediction. *Nucleic acids research*, *52*(W1), W469- W475.

21. Bhatia, A.S.; Saggi, M.K.; Kais, S. Quantum Machine Learning Predicting ADME-Tox Properties in Drug Discovery. J. Chem. Inf. Model. 2023, 63, 6476–6486.

22. Moore, J. H., et al. (2021). Approaches and challenges in machine learning for drug discovery. Pharmacological Reviews, 73(3), 1144–

23. 1165. https://doi.org/10.1124/pharmrev.120.000179

25. Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. New England Journal of Medicine, 380, 1347–1358. https://doi.org/10.1056/NEJMra1814259

26. Herráiz-Gil, S., Nygren-Jiménez, E., Acosta-Alonso, D. N., León, C., & Guerrero-Aspizua,

27. S. (2025). Artificial Intelligence-Based Methods for Drug Repurposing and Development in Cancer. *Applied Sciences*, *15*(5), 2798.