

## Drug Repurposing

K. Malleswari\*, <sup>1</sup> D. Rama Brahma Reddy\*, <sup>2</sup> Diviti Kanchana, <sup>3</sup>  
Dakkumadugula Mounika, <sup>4</sup> Desai Sowmya.

*Associate professor Nalanda Institute of pharmaceutical sciences, sidharthnagar, Kantepudi (v), sattenapalli(M)  
, Guntur ( Dist).522438.*

*Principal & Professor Nalanda Institute of pharmaceutical sciences, Sidharthnagar, Kantepudi (V),  
Sattenapalli (M), Guntur (Dist)-522438.*

*Students of Nalanda Institute of pharmaceutical sciences, sidharthnagar, Kantepudi(V), Sattenapalli(M),  
Guntur (Dist)-522438.*

Submitted: 20-08-2023

Accepted: 31-08-2023

### ABSTRACT

Drug repurposing refers to finding new indications for existing drugs. The paradigm shift from traditional drug discovery to drug repurposing is driven by the fact that new drug pipelines are getting dried up because of mounting Research & Development (R&D) costs, long timeline for new drug development, low success rate for new molecular entities, regulatory hurdles coupled with revenue loss from patent expiry and competition from generics. Anaemic drug pipelines along with increasing demand for newer effective, cheaper, safer drugs and unmet medical needs call for new strategies of drug discovery and, drug repurposing seems to be a promising avenue for such endeavours. Drug repurposing strategies have progressed over years from simple serendipitous observations to more complex computational methods in parallel with our ever-growing knowledge on drugs, diseases, protein targets and signalling pathways but still the knowledge is far from complete. Repurposed drugs too have to face many obstacles, although lesser than new drugs, before being successful.

**Keywords:** Repositioning, drug refilling, drug recycling.

### I. INTRODUCTION

Drug repurposing (DR) is also known as drug repositioning, drug re-tasking, drug reprofiling, drug rescuing, drug recycling, drug redirection, and therapeutic switching. It can be defined as a process of identification of new pharmacological indications from old/existing/failed/investigational/already marketed/FDA approved drugs/pro-drugs, and the application of the newly developed drugs to the treatment of diseases other than the drug's original/intended therapeutic use. It involves establishing new therapeutic uses for already

known drugs, including approved, discontinued, abandoned and experimental drugs. Traditional drug discovery is a time-consuming, laborious, highly expensive and high risk process. The novel approach of drug repositioning has the potential to be employed over traditional drug discovery program by mitigating the high monetary cost, longer duration of development and increased risk of failure. It confers reduced risk of failure where a failure rate of ~45% is associated due to safety or toxicity issues in traditional drug discovery program with additional benefit of saving up to 5–7 years in average drug development time.

In recent years, the drug repositioning strategy has gained considerable momentum with about one-third of the new drug approvals correspond to repurposed drugs which currently generate around 25% of the annual revenue for the pharmaceutical industry. It has been accounted that approximately 30% of the US Food and Drug Administration (FDA) approved drugs and biologics (vaccines) are repositioned drugs. The first example of drug repositioning was an accidental discovery/serendipitous observations in the 1920s. After about a century of development, more approaches were developed for accelerating the process of drug repositioning. Some most successful and best-known drugs that have been emerged out of the DR approach are sildenafil, minoxidil, aspirin, valproic acid, methotrexate etc. For example, sildenafil originally developed for the treatment of hypertension and angina pectoris has currently been used to treat erectile dysfunction.

### Methods for drug

Drug repurposing methods can be broadly classified into either activity based or in silico methods. Activity based methods include in vivo (living organisms) and in vitro high-throughput screening methods where the drug/chemical of

interest is used for screening. In in-silico or computational or virtual screening methods, hits are identified in a systematic way from information gathered from various databases and involve tools to identify drug-target interactions. Activity based

methods, though time and labour intensive in contrast to computational methods, are characterised by lower false positive hits and easy validation of screening hits than computational methods

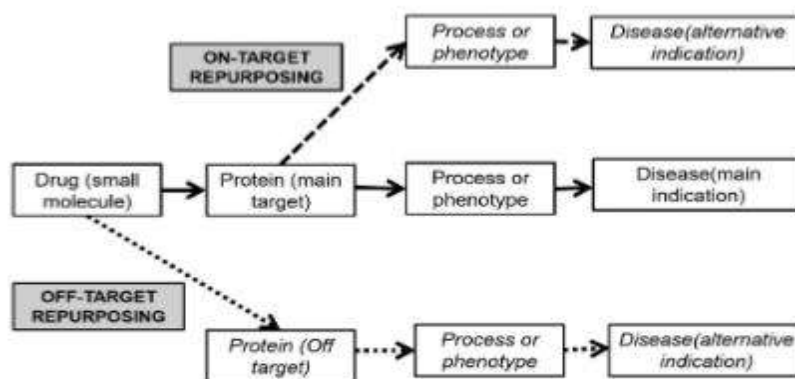


Fig. 1. On target and off target drug repurposing

#### Phenotypic screening based method.

Phenotypic screening using in vivo and in vitro cell based assays have been central to the discovery of new drugs where chemical libraries are screened to identify 'hits'. This method is used for repurposing the drugs. This method is also more physiological, as intact cells and organisms are used as opposed to in silico methods and the chances of success for the repurposed drug to move to clinical trials is high. However this method has relatively high cost compared with in silico method.

#### Chemical similarity method

Similar property principle i.e., similar drugs with similar structures lead to similar biological effects, forms the base for this approach. This principle is rooted in known quantitative relationships between chemical structures and biochemical activity. But, chemical structures from databases may contain errors. Some drugs undergo transformation inside the body before being active and also that physiological effects cannot be predicted on the basis of structural properties alone.

#### Side effect similarity based method

Drugs with similar target binding profiles cause similar side-effects – this provides the basis to relate drugs to other drugs or diseases by side effect profiles, even in cases where the precise pharmacological mechanism facilitating the side effect is unknown. The disadvantages of this

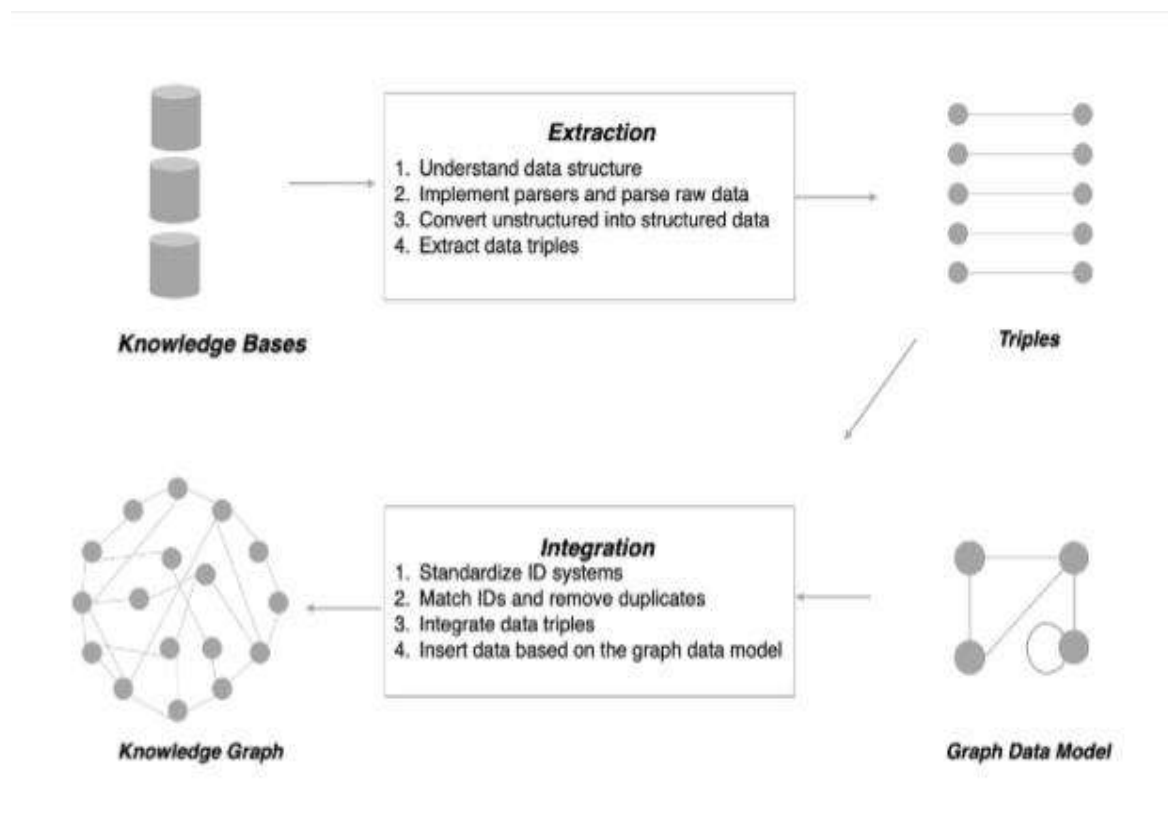
approach are well characterised side effect profile of drugs is not completely available for most drugs and drugs sharing a similar side effect may cause the side effect by altogether different mechanisms.

#### In silico drug design methods for drug repurposing

The traditional drug-discovery process that involves de novo design and validation of new chemical entity is a time consuming and expensive process. Even though there has been a many fold increase in the expenditure on drug discovery, the number of new chemical entities or new drugs approved remains considerably small. Accumulation and rational usage of high-throughput biological, clinical, and chemical data can accelerate the process of drug discovery

#### Drug knowledge graph

A knowledge graph is essentially a semantic network that reveals relationships between entities. One of the major concerns of a knowledge graph is the connectivity among the known facts stored in the knowledge base. By connecting pieces of knowledge gathered from knowledge bases, we obtain a more comprehensive and centralized repository. Therefore, constructing a knowledge graph allows the best use of previously discovered knowledge, revealing new knowledge based on the large-scale, interconnected known facts.



**Fig 2:Major steps in the construction of drug knowledge graph**

In the area of pharmaceutical research, tremendous human efforts have been devoted to curate drug-related knowledge. Hence, many drug knowledge bases have been developed and introduced to serve various academic research. Our previous work has compiled a non-exhaustive list of widely used and publicly available drug knowledge bases. The study provides a detailed explanation of drug knowledge bases, including types of entities and relations, sources, and their applications in drug-related studies such as biomedical text mining, drug repositioning, adverse drug reaction, and pharmacogenomic analysis. From the available list, our drug knowledge graph was based on six drug knowledge bases, including PharmGKB,<sup>31</sup> TTD,<sup>32</sup> KEGG DRUG,<sup>33</sup> DrugBank,<sup>34</sup> SIDER,<sup>35</sup> and DID,<sup>36</sup> which have been selected based on the availability of raw data files. Figure 1 shows a few steps involved in the construction of the drug knowledge graph. As demonstrated, the extraction of structured information and the integration of the extracted information were the two major steps involved. In the following sections, we briefly summarize a few challenges encountered, as well as the approaches used to handle these challenges.

### Extraction of structured information

Parsing of raw data files includes two important steps: (1) understanding structures and (2) implementing parsers. Drug knowledge bases distribute raw data files in various formats, including plain text, CSV, TSV, XML, and XLSX. In most cases, structures of raw data are not explicitly defined, and we manually reviewed the data files to understand the structure. Plain text is the least structured format, and parsing it requires several trials and errors, testing different separators and regular expressions. Parsers vary by knowledge bases or even by files within the same knowledge base, and we implemented customized parsers based on the identified data structures. Raw data include both structured and unstructured information. While structured information can be directly extracted and used without pre-processing, unstructured information should be manually reviewed and processed.

### Integration of the extracted information

Structured information extracted from the existing knowledge bases is fragmented. It is essential to integrate this information by

connecting them based on a common medium and removing duplicates. We list below a few major issues during the integration.

- The integration was based on cross-referencing.
- A simpler approach is considering cross-references only among the drug knowledge bases being integrated. However, this will result in a loose integration because many cross-references among the drug knowledge bases being integrated are missing and incomplete in many cases.
- In addition to the six drug knowledge bases, we also considered ID systems of other terminologies and knowledge bases.
- Here, a challenge was that the same ID system is referred to with different names in different knowledge bases and the formats of ID value also differed.
- One ID value of an ID system can match more than one ID value of another system.
- Drug knowledge bases describe entries (e.g. drugs) using different levels of granularity, and a higher-level entry in one drug knowledge base usually matches to many lower level entries in other drug knowledge bases.
- As it is impossible to manually review all the entries of one-to-many relationship to filter human errors, we maintained all of them.

- This decision was based on the result of our initial review of randomly selected entries, in which the dominant reason for the one-to-many relationship was noted as the different levels of granularity rather than human errors.
- Due to a large number of entries of the one-to-many relationship, if we failed to consider them and only integrate entries of a one-to-one relationship, the integration would be loose, which means the same entries may appear several times in the integrated drug knowledge graph

### Drug-centric graph data model

Data model is at the center of constructing a knowledge graph. A data model guides us regarding what and how data should be extracted and integrated. When designing a data model, we usually need to establish a compromise between domain knowledge and data availability. When designing the data model, we accounted for a few factors. First, the data model should be general enough to embrace data from multiple sources. Second, the data model should be flexible to changes that might be needed when adding new data sources. Finally, and most importantly, interaction among multiple types of data should be properly represented. Based on these considerations, we propose a property graph model as shown in Figure 3.

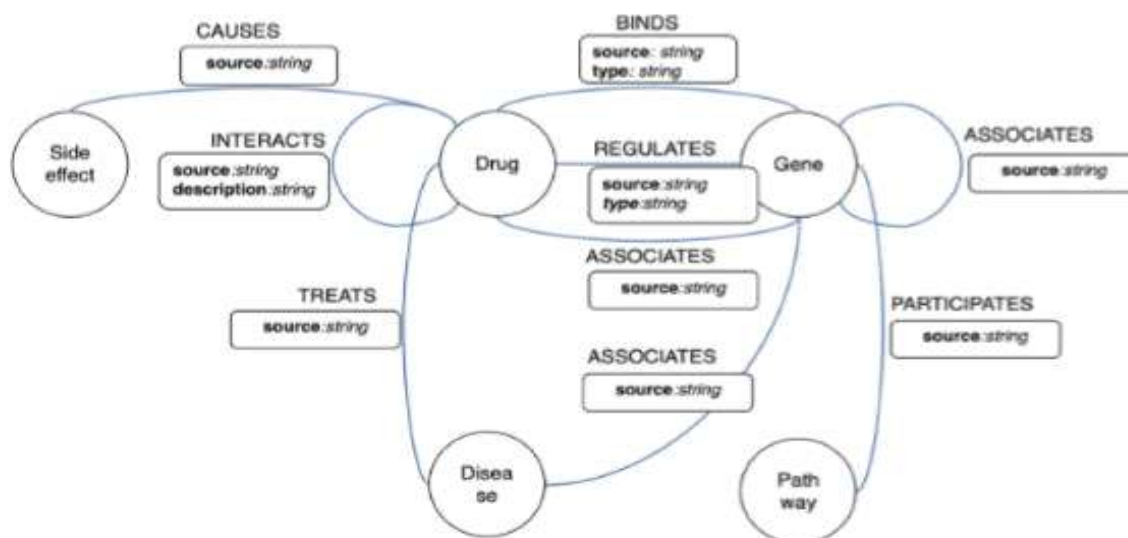


Fig. 3. A drug-centric property graph model

Our drug-centric graph data model comprised five types of entities and nine types of

relationships among the five entities. The data model includes drugs, diseases, and other entities

that interact with the two entities, such as genes, pathways, and side effects. The nine relationship types include TREATS (between drugs and diseases), INTERACTS (between two drugs), CAUSES (between drugs and side effects), BINDS, REGULATES, ASSOCIATES (between drugs and genes), ASSOCIATES (between two genes), ASSOCIATES (between genes and diseases), and PARTICIPATES (between genes and pathways).

### Data representation

#### Path-based representation

The two central entities in drug repurposing are drug and disease. How pairs of a drug and disease are connected in a drug knowledge graph provides significant information essential for uncovering drug repurposing candidates. Specifically, meta paths, which are sequences of relationships between two types of entities,<sup>37,38</sup> are effective indicators. For example, based on Figure 2, many meta paths connect a drug with a disease. They include, but are not limited to, drug–gene–disease, drug–drug–disease, and drug–side effect–drug–disease. The meta path-based approach is able to capture local network structures around drug repurposing candidates. While it does not consider the whole network structure, only local structures. Defining meta paths and measures of counting meta paths are the two important components of the approach, in which the path length affects the number of meta paths between two types of entities. In this study, we define all meta paths of length 2–4, decided based on both academic and practical considerations. Practically, the number of meta paths increases exponentially as the path length increases. We used the following four measures of counting meta paths, in which  $P$  denotes a meta path,  $p$  denotes an instance of the meta path, and  $h$  and  $t$  denote a head and a tail of a meta path, respectively

Path count:  $p(h,t) = \sum_{p \in P} PC_p(h,t)$

Head normalized path

count:  $HNPCP = PC_p(h,t) / PC_p(h,*)$

Target normalized path

count:  $TNPCP = PC_p(h,t) / PC_p(*,t)$

Normalized path count:

$NPCP = PC_p(h,t) / PC_p(h,*) + PC_p(*,t)$

As shown above, PC measures the number of instances of a meta path between a head and a tail of the meta path (i.e. a drug and a disease in our

study). The other three measures are normalized versions of PC, in which HNPCP is normalized by the number of meta path instances that connect the head of the meta path and any other entities, TNPCP is normalized by the number of meta path instances that connect the tail of the meta path and any other entities, and NPCP is normalized by the sum of the two denominators. Overall, the three measures normalize the raw path count by considering the head and/or tail's overall connectivity information.

## II. RESULTS

### Large-scale drug repurposing

In the experiments, we applied the above data representation methods to the drug knowledge graph and generated feature matrices to train machine learning models. As we only had positive samples, that is, known treatments, when training machine learning models, we did not possess inputs for negative samples. Instead, we had unlabeled samples. We use only positive and unlabeled (PU) data to train machine learning models. For the experiments, we used the approach proposed by Elkan and Noto. We implemented a PU learning method with three machine learning algorithms (i.e. Decision Tree, Random Forest, and support vector machine (SVM)) available in the scikit-learn package. Since the size of all unlabeled samples (drug–disease pairs), obtained by generating all possible pairs of drugs and diseases that do not have known effects, we focused on diabetes mellitus and included unlabeled samples that were related to the disease.

The goal of the experiments was to evaluate the effectiveness of the various data representation methods in terms of predicting known diabetes mellitus treatments based on the known effects of other drug–disease pairs. Based on the MeSH tree numbers of diabetes mellitus (C18.452.394.750 and C19.246) and MeSH Ids of diseases in our drug knowledge graph, we identified eight diseases (diabetes mellitus, type 1 diabetes mellitus, type 2 diabetes mellitus, diabetic nephropathies, diabetic retinopathy, diabetic neuropathies, Wolfram syndrome, and Donohue syndrome) that fall into the category of diabetes mellitus in the drug knowledge graph. As shown in Figure 5, we trained two classes of machine learning models, one for path-based and the other one for embedding-based.



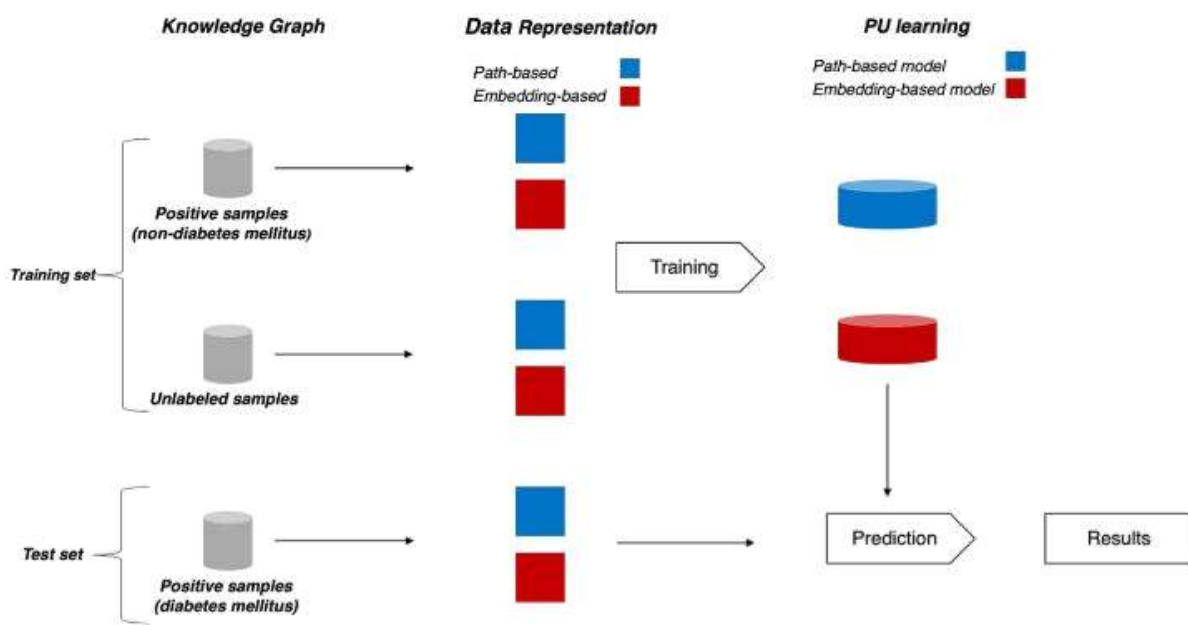


Fig. 4. Machine learning pipeline.

In the experiments performed, our test set comprised all known treatments of the eight diseases. All known treatments of the other diseases constituted our training set together with 143,830 unlabeled samples, obtained by associating the eight diseases with drugs that treat at least one disease. Table 1 shows the size of the training and test sets. As our test set included only positive samples, in Table 2, we report recall (or sensitivity)

that measures the proportion of known treatments that were correctly predicted. The performance varied by different combinations of the data representation methods and machine learning algorithms. As shown in Table 2, with certain specific combinations, we observed perfect (SVM + NPC) or nearly perfect (random forest + TransH) prediction results.

Table 1: The size of the training and test set

	Positive	Unlabeled	Total
Training set	87,395	143,830	231,225
Test set	692	0	692

Table 2. Recall score for the data representation methods.

Algorithm used in PU learning	Path-based				Embedding-based		
	PC	HNPC	TNPC	NPC	TransE	TransH	TransR
SVM	0.64	0.68	0.68	1	0.28	0.12	0.31
Decision Tree	0.45	0.64	0.68	0.68	0.72	0.74	0.28
Random Forest	0.59	0.67	0.68	0.68	0.58	0.97	0.61

### III. DISCUSSION

#### Need for drug repurposing

From the medical community-patient perspective, drug repurposing has the ability to meet unmet medical needs- neglected diseases and, rare and orphan diseases. It also has the potential to

provide more effective treatment, cheaper alternative drugs, and drugs with favourable side effect profile in diseases where the available drugs have adverse side effect profile. It can also play a significant role in the development of personalised medicine. New drug discovery faces the challenges

of increasing Research & Development (R&D) costs, long timeline for drug development, low success rate and regulatory hurdles. In addition, pharmaceutical industry is also confronted with revenue loss from patent expiry and competition from generics and off-label prescription. Drug repurposing is claimed to be less costly, less time consuming, less risky and increased chance of success from the industrial perspective. The above-mentioned factors call for novel strategies for drug discovery and drug repurposing may provide an answer to the question.

#### Drug repurposing in India

- Drug repurposing in India TB, HIV/AIDS, Malaria, NTDS like leprosy, lymphatic filariasis. Paramomycine and miltefosine for kala-azar following trails in India.
- Life style diseases diabetes, hypertension, IHD, cancers. Reluctance of companies to invest In R&D of drugs for infectious diseases and NIDs
- Drug repositioning in India global initiative based on public private partnership model: WHO special program for research and training in tropical diseases (WHO/TDR) medicine for malaria venture, global alliance for TB drug development, drugs for neglected diseases
- Under the patient act of India use of patent for a new indication is not permissible for an already patented drug.

#### Challenges in drug repurposing

With traditional drug development, knowledge of failed assets is often limited and unpublished, impeding the insights needed to identify successful targets. Despite all the benefits, drug repurposing has a number of issues to consider. It is sometimes impossible to get all the necessary data to properly analyse older drugs, as many trials did not optimize the drug's clinical benefits and biological questions because of their expedient design and lack of clinical endpoints. Oftentimes it is not even evident in reporting the exact reason why a drug failed. Also, some trials only had a small number of patients enrolled, therefore lacking much statistical power. Drug Repurposing continues to grow in popularity as advanced AI opens the door to new insights into disease drug targets and increases the odds that clinical development trials will be successful. Ultimately, this process will allow patients access to new treatments faster, offering answers for rare

disease symptoms and potentially saving lives. IQVIA is uniquely suited to meet any type of drug repurposing needs with a Drug Discovery and Development Services team of world-class data scientists, cutting-edge AI technology, and consultants with years of domain expertise.

#### Advantages of repurposing drugs

Notably, the drug repurposing approach benefits from the fact that approved medicines and several discarded compounds have already been tested in humans and comprehensive information is available on their pharmacology, dose, possible toxicity and formulation. Drug repurposing has numerous advantages over conventional drug discovery approaches, including;

- Considerably cuts research and development (R&D) costs.
- Reduces the drug development timeline, as various existing Compounds have already demonstrated safety in humans.
- It does not require Phase 1 clinical trials.
- Potential for reuse despite evidence of adverse effects and failed efficacy in some indications.
- Increase the productivity in the pharma industry.
- Risks are better known and the chance of failure due to adverse side effects is reduced.
- Patients with terminal cancers, Orphan diseases and other incurable conditions often do not have a decade to spare.
- More patients will have access to and be able to afford their repurposed medication

#### Repositioning drugs for rare diseases

It is common knowledge that there are only a few therapeutic options for rare diseases – those that affect only a small percentage of the overall population. After the discovery of its immunosuppressive properties linked to its inhibition of the mTOR protein kinase, Pfizer's Rapamune (sirolimus) was approved to prevent organ transplant rejection. It also became the first drug approved for lymphangioleiomyomatosis (LAM), a rare genetic lung disease.

#### What the outbreak of COVID-19 meant to drug repurposing

In June 2020 we organised a webinar to raise awareness of the untapped potential of repurposed drugs in the treatment of cancer against the backdrop of the COVID-19 pandemic. EU policy makers, researchers, physicians and patients discussed whether the corona crisis was

driving us towards a paradigm shift in the way cancer treatments are researched, developed and delivered. Our intention was to point out the benefits of drug repurposing for public health to European decision makers.

#### Aninnovative revision of cancer treatment

The implementation of drug repurposing into standard cancer care isn't obvious. Due to the lack of commercial prospects, the use of 'old' drugs for new indications is typically not considered 'sexy' or innovative. That's why Ciska Verbaanderd, PhD student at the University of Leuven and associate of the Anticancer Fund,

wrote a doctoral thesis on the topic. She aimed to bridge the gap between clinical research and practice in cancer drug repurposing. She clearly formulated recommendations to address the challenges.

- Support data sharing and open science initiatives in drug repurposing
- Promote independent clinical research with repurposed medicines
- Develop a collaborative framework to bring new uses of drugs on-label
- Introduce legal changes and create incentives to encourage new uses on-label.

**Table 3:Some of the repurposed drugs**

Name of Drug	Original Indication	New Indication
Amphotericin B	Fungal infection	Leishmaniasis
Amantadine	Influenza	Parkinson's disease
Aspirin	Pain / fever	Myocardial infarction
Allopurinol	Cancer	Gout
Azathioprine	Rheumatoid arthritis	Renal transplant
Atomoxetine	Depression	Attention deficit hyperactivity disorder
Bromocriptine	Parkinson's disease	Diabetes mellitus
Bupropion	Depression	Smoking cessation
Cycloserine	Urinary tract infection	Tuberculosis
Finasteride	Benign prostatic hyperplasia (BPH)	Male pattern blindness (MPB)
Gemcitabine	Viral infection	Cancers
Interferon- α	Hepatitis B and C	Cancers
Itraconazole	Fungal infection	Cancers
Minoxidil	Hypertension	Hair loss
Orlistat	Obesity	Cancers
Raloxifene	Osteoporosis	Postmenopausal breast cancer
Retinoic acid	Acne	Promyelocytic leukemia
Sildenafil	Angina pectoris/hypertension	Erectile dysfunction
Thalidomide	Morning sickness	Multiple myeloma / leprosy
Tamoxifen	Inflammation	Parkinson's disease
Tranexamic acid	Heavy bleeding (antifibrinolytics)	Melasma

#### Failed repurposed drugs

Though repurposing appears to be an attractive strategy, several challenges exist for the drugs identified to be repurposed before making it to the market. These include, but not limited to,

low potency, dose adjustments, new safety signals and route of administration. Data available from in vitro and animal studies may not be generalizable to humans. With regard to computational approaches such as molecular docking, because of



the diversity of the protein database and differences in the algorithm used for docking, there is differing agreements on drugs converging on same targets. Chloroquine/hydroxychloroquine touted to be the game changer in the battle against Covid-19 fizzled out in a matter of few months. The Food and Drug Administration (FDA) agency revoked emergency use authorisation (EUA) granted to these antimalarials for lack of efficacy and cardiac adverse events within 3 months of initial approval.

#### Acknowledgement:

We would like to thank Mrs.K. Malleswari, our associate professor and our principal, D. Rama Brahma Reddy for their support and their guidance in completing our project on the topic (Drug repurposing ). It was a great learning experience. I would like to take this opportunity to express my gratitude to all of my group members D.Kanchana,D.Mounika,D.Sowmya. The project would not have been successful without their cooperation and inputs.

#### IV. CONCLUSION

Prior knowledge regarding various aspects of drugs is the key to successful drug repurposing. Constructing drug knowledge graphs is an essential step to obtain the best use of prior knowledge by weaving the continuously growing, fragmented, and dispersed drug-related data. By applying effective data representation methods to the drug knowledge graph and thus transforming knowledge available in the drug knowledge graph into informative inputs for machine learning models, we can effectively predict drug repurposing candidates. Large-scale prediction and interpretability are well-known limitations of experimental approaches and previous computational approaches, respectively. The knowledge-driven approach supports not only large-scale prediction through data representation and machine learning methods but also allows investigation of a case study through path-based exploration. These two features can effectively handle the abovementioned limitations with the support of a comprehensive drug knowledge graph.

#### REFERENCE

- [1]. Ashburn TT, Thor KB. Drug repositioning: Identifying and developing new uses for existing drugs. *Nature Reviews. Journal of Drug Discovery.* 2014;3:673-683
- [2]. Dey G. An overview of drug repurposing: Review article. *Journal of Medical*

- Science and Clinical Research.* 2019;7(2):3-5
- [3]. Pihan E, Colliandre L, Guichou J-F, Douguet D . e-Drug3D: 3D structure collections dedicated to drug repurposing and fragment-based drug design. *Journal of Bioinformatics* 2012;28: 1540–1541.
- [4]. Zheng W, Thorne N, McKew JC Phenotypic screens as a renewed approach for drug discovery. *Drug Discovery Today* 2013;18: 1067–1073.
- [5]. Keiser MJ, Setola V, Irwin JJ, Laggner C, Abbas AI, Hufeisen SJ, Jensen NH, Predicting new molecular targets for known drugs. *Nature* 2009;462: 175–181.
- [6]. Campillos M, Kuhn M, Gavin A-C, Jensen LJ, Bork P Drug target identification using side-effect similarity. *Science (New York, N.Y. )* 2008;321: 263–266.
- [7]. Zhu Y, Elemento O, Pathak J, et al. Drug knowledge bases and their applications in biomedical informatics research. *Brief Bioinform* 2018; 20: 1308–1321.
- [8]. Sun Y, Han J, Yan X. PathSim: meta path-based top-k similarity search in heterogeneous information networks. *Proc VLDB Endow* 2011; 4: 992–1003.
- [9]. Letouzey F, Denis F, Gilleron R. Learning from positive and unlabeled examples. In: Arimura H, Jain S, Sharma A (eds) *Algorithmic learning theory.* Berlin, Heidelberg: Springer, 2000: 3:71–85.
- [10]. Naylor S, Schonfeld JM Therapeutic drug repurposing, repositioning and rescue – Part I: Overview. *Drug Discovery World* 2014; 16: 49–62.
- [11]. Nathan sommerford global lead, challenges of drug repurposing .AI for drug discovery and development 2022: 4:5-6.
- [12]. Harsha Agarwal sagacious IP, Drug repurposing: advantages and key approaches2024:2:4-6.
- [13]. Parvathaneni V, gupta V, utilizing drug repurposing against covid -19 efficacy, limitations and challenges life science 2020:25: 118-275.