# Transcriptomic Similarity Mapping and Knowledge Graph Embeddings: Computational Strategies for Systematic Drug Repurposing

*Dr. Daniel Gutiérrez, Professor of Industrial Engineering, National Technological University (UTN), Argentina*

## 1. Introduction to Drug Repurposing

Drug repurposing focuses on identifying novel therapeutic indications for existing drugs, currently marketed for other diseases, and has several advantages over traditional de novo drug discovery. Repurposing is both time and cost-efficient, allowing faster drug development by one to two years, with R&D costs reduced by 60% when compared to novel drug development, and can bring drugs for new indications to the market at a reduced cost of $30–40 million. In addition, for patients, the advantage of drug repurposing means immediate access to treatment and could help maintain homeostasis when faced with emergent viral outbreaks or pandemics. Repurposing can also address the challenging patient population sub-group of the long tail indicated for ultra-rare diseases and oncology.

The term drug repurposing is largely used in the context of approved drugs; however, other similar terms are used synonymously, such as drug repositioning for FDA-approved drugs and drug reprofiling for lab compounds. Researchers will sometimes refer to active pharmaceutical ingredients or drug-like molecules as small molecules. To describe the complexity of drug repurposing, it is sometimes subdivided into the perspective of target, drug, or 'indication'. Current clinical trials for de novo drug discovery are plagued by 'failure' and a reassessment of the early-stage pipeline for 'revival'.

### 1.1. Definition and Importance

1) Topic: Machine Learning Approaches for Drug Repurposing

2) Section Title: Definition and Importance

3) Section Summary: This sub-section defines drug repurposing and explains its importance in the pharmaceutical industry. It elaborates on how the repurposing process allows for the exploration of existing drugs in new therapeutic contexts. The economic and time-saving benefits are highlighted, showing that leveraging existing drugs can reduce development risks. Additionally, the section discusses the capability of repurposing to address diseases that may lack adequate treatment options. The role of regulatory frameworks in facilitating such innovations is also touched upon. It articulates a compelling case for drug repurposing to complement traditional drug discovery. Furthermore, the significance of collaboration between academia and industry is emphasized in enhancing repurposing efforts. Overall, the sub-section provides an overview of the essential aspects and benefits of drug repurposing.

Drug repurposing, also known as drug repositioning or drug rescue, refers to a strategy to find new therapeutic uses for approved or investigational clinical drugs that are outside the scope of the original medical indication. The nature of repurposing makes it possible to introduce new therapeutic indications for existing drugs, thus saving costs and development times. Moreover, repurposing can also permit the revival of drugs that have been shelved because of poor effect or unacceptable side effects by finding new treatments where such limitations do not apply.

In recent years, there has been an upsurge of interest in drug repurposing mainly because it offers the potential to decrease the expenditure and risks associated with new drug development. Moreover, repurposing known compounds allows drug developers to skip over many hurdles typically encountered in traditional drug development. Because the pharmacological, toxicological, pharmacokinetics, and adverse event profiles of these drugs have already been tested in humans, repurposing can significantly reduce the time, costs, and risks associated with the traditional process of drug discovery and development. The high risk and cost of traditional drug development have been exacerbated by the increasingly stringent safety, pharmacovigilance, and regulatory requirements for the approval of medicines. Although enormous resources are expended in traditional drug development, many emerging diseases continue to lack effective drugs because of a low relative or perceived market return on new medicine investment. Moreover, healthcare regulatory agencies have become increasingly risk-averse concerning new medicines. Hence, a drug

repurposing agenda complementing the traditional new drug development paradigm is very important. Repurposed drugs are not only cheaper to develop but also become available to patients sooner at lower cost. For instance, many candidate drugs, which showed early preclinical and clinical promise in the fight against Ebola and Zika epidemics, are repurposed ones. Therefore, policy initiatives such as matching funds and push-pull mechanisms are being adopted to encourage drug repurposing for emerging or neglected diseases.

### 1.2. Challenges in Traditional Drug Discovery

There are several long-standing challenges within traditional drug discovery and development that significantly contribute to the high drop-out rates of drug candidates and the long timelines and costs of development. From a drug repurposing perspective, extensive time and resource requirements for de novo drug discovery and development remain key barriers against commercializing repurposed medication. Development from scratch requires an estimated $897 million on average in industrial funding, and many molecules go on to fail as they progress through drug development. Approximately 94% of internal drug development programs are discontinued in the discovery phase, with approximately 41% of investigational medicines failing due to the "lack of effectiveness" of the molecule compared to existing therapies. Disregarding pharmacological properties, early-stage screenings should not require much initial funding as synthesized compounds are low in both cost and investment risk, and therefore reducing the initial funding barrier required to start development.

Despite some advances in genetics, drug responses still remain very unpredictable as they are influenced by environmental and lifestyle factors, the administration of co-medications, and the ongoing treatment regimens patients receive for numerous diseases. Late-stage failures constitute an immense waste of financial and organizational resources without aligning with the drop-out pattern of drugs in clinical trials, with patients being driven by the possibility of receiving a placebo syndromic remedy. Aside from symptom-based diseases, overreliance on trial participation and engagement potentially extends clinical trial durations. Encouraging patient participation in these programs is generally challenging, even in insurance programs where participants are incentivized financially or are given priority treatment. Patient recruitment costs for clinical trials of drug candidates are data-limited and are typically recorded under the

umbrella of total trial costs in organizational profit and loss records. The high prevalence of drug development challenges reinforces the demand for timely and cost-effective methods that can be tested early. High investment and time requirements and the substantial capital risks involved also justify expanded data pooling, including machine learning technologies.

## 2. Machine Learning in Drug Repurposing

In the repurposing landscape, we have seen the gradual integration of machine learning. A wide variety of machine learning techniques can be used for repurposing. Supervised learning techniques can be applied in drug repositioning to assist in computational predictions by training the model on a set of known drug-disease associations and known negative associations. Unsupervised learning techniques are applicable to drug repositioning when the mode of action, indication, or target of a drug is not yet determined. Machine learning algorithms can handle a vast amount of data, extracting patterns and information that are too complicated and subtle for the human brain to comprehend. The development of machine learning is expected to speed up the process by predicting the drug-disease associations or uncovering the mechanisms of action. These combinations offer great potential for new effective therapies. Recently, many machine learning techniques, including deep learning, have improved predictions for drug activation, interactions, toxicological outcomes, and many more. The use of machine learning techniques in drug repurposing has added functionalities of modeling the relationships, predicting side effects, and shedding light on the dynamics of repurposing through complex networks. Machine learning can complement the conventional structure-based and ligand-based drug design in better understanding the mechanism of action, uncovering more uses of old drugs, identifying side effects, improving drug combinations, and speeding up the drug development process. These techniques may not provide ways to obtain promising candidates; however, they can provide enhanced possibilities with the combination of the available techniques. In repurposing, unbiased data-driven strategies are gaining preference over the current drug discovery paradigms with the rapid advancement in artificial intelligence.

### 2.1. Overview of Machine Learning Techniques

Machine learning (ML) makes use of algorithms and computational statistics to enable computers to learn and perform tasks without being explicitly programmed. It finds its

applications in the analysis and interpretation of biological data, including but not limited to predicting gene functions, interactions, and pathways. Supervised learning algorithms use inputs and corresponding outputs to learn a function. These algorithms are used in predicting gene functions, studying drug–target interaction networks, and for drug repositioning. Unsupervised learning is used when the output is not known and for exploring a dataset. It is particularly useful in inferring drug–drug interactions and finding potential drugs for disease modifications.

ML algorithms can be divided into parametric and non-parametric. Along with predictive accuracy, the adoptability of a model is also essential for downstream applications. Any ML model generating biomedical data that cannot be translated into actionable and profitable insight is not of much use. During the machine learning pipeline, it is important to decide on the features to be selected for acquiring additional information and cutting down on noise. Further deciding on the model evaluation metrics helps in deciding on optimal thresholds. On the performance metrics, different drug repurposing studies use different benchmarks to estimate model performance.

**2.2. Applications in Identifying New Drug Indications**

Drug discovery is not only time-consuming but also costly, which has led to a growing interest in identifying new indications for existing drugs. Computational approaches, particularly those based on machine learning, have been employed to sift through the various omic data to present a range of drug repurposing opportunities. These include further exploration of the therapeutic potential of the drugs developed by commercial vendors or presently unexplored drugs, like those typically used to treat orphan indications. The integration of drug bioactivity with molecular data can either highlight new uses for drugs grouped by mechanism of action or propose mechanisms of action for drugs used with a wide range of therapeutic applications. In this way, drug repurposing can help to focus attention on certain mechanisms of action for further therapeutic target validation and potential drug developments, as well as encouraging repositioning efforts in relation to specific drugs that belong to tested mechanisms of action. Illustrative examples of cross-disease applications of molecular-based drug repurposing approaches include the successful repurposing of the anti-asthmatic drug salbutamol as a repurposing candidate for indications such as Parkinson's disease and tuberculosis, ibrutinib moving from a potential gastrointestinal cancer indication to a

current target of interest in lymphoma and cancer immune therapy, and nelfinavir and dasatinib being validated as an anti-cancer combination therapy by both computational and experimental methodologies.

The main advantage of computational models is their capacity to sift through structural biology datasets to suggest the unanticipated. There are applications in multiple therapeutic areas. The literature provides numerous case studies from both academia and industry where bioinformatics models have been used to identify repurposing candidates in diseases such as various cancers, neglected tropical diseases including Chagas and dengue, Parkinson's disease, psychiatric disorders, rheumatic diseases, and infectious diseases including leprosy, tuberculosis, visceral leishmaniasis, and dengue. In addition to putative new indications for established drugs and repositioning opportunities, much recent interest also gives repurposing candidates of newly developed drugs that failed in clinical development for their original use. Drug repositioning may be a last-chance therapeutic option for drugs that failed their original indication. The models have been employed to retrospectively provide explanations for physiological responses and also to validate potential repurposing candidates in prior events. Machine learning models are seen to have an increasing role in drug repurposing. Nevertheless, unlike certain models using compound bioactivity profiles in accelerating the generation of proof of concept human data for an intended use, computer algorithms have no clinical indication property, and thus clinical trials for repurposed drugs are mandatory. It is worth emphasizing that this does not eliminate the need for rigorous ethical appraisal of these resources and for the responsible, transparent, and fair use of the predictive models even if scientifically valid, especially in their use for drug development. The main challenge in some constraint-based approaches is that the datasets under analysis are delayed to the human therapy associated function or physiology, faced with a condition that is computationally classically associated with their mechanism of action or their expected pluritherapeutic uses. Although bioinformatics models have been effective in making novel drug indication predictions, miscellaneous knowledge gaps and highlights require future work.

## 3. Datasets and Data Preprocessing

The building block for any machine learning application is a high-quality dataset that is relevant to the problem at hand. In order to construct a dataset, reliable data sources must be utilized. Various proprietary and public domain sources can be used for drug repurposing. Public domain sources may be taken from public databases. Unfortunately, different organizations are responsible for the orderly compiling of these public domain sources, which can lead to differences between datasets. The choice of relevant datasets is a cumbersome, time-consuming task, which may require the expertise of a domain scientist. The choice of data sources should be such that they have been manually curated and cover the entire organism's system rather than spectroscopic or natural language processing data, as it may suffer from incompleteness and inconsistency.

Acquiring data from multiple sources could lead to massive amounts of data, potentially distributed in diverse formats, which makes the data preprocessing of acquiring and managing the data organization extremely challenging and time-consuming. There are several preprocessing techniques for the datasets. Before applying the machine learning algorithms, data encoding is done. In a high dimensionality setting, the feature values may become vastly different between various dimensions. Normalization may be applied to maintain the range scaling to some specified range, which may marginally help the performance of the machine learning models. Noise may be included in the data to produce the construct or datasets; data augmentation can be applied to improve the robustness and generalization of the model. It may involve translating the image or augmenting data for the images and text. Sometimes, the selected data sources may suffer from missing values; most algorithms are incapable of handling missing data values. Before applying any imputation method, imputation for the training and testing datasets should be separately handled. With the proper selection of datasets and preprocessing techniques, the performance of the model may be improved in terms of accuracy or any other metric.

### 3.1. Sources of Data for Drug Repurposing

Public Databases:

Data relevant to drug repurposing resides in a number of public data repositories and open access databases. The most commonly used in drug repositioning and repurposing would be databases that contain information on numerous chemical compounds and

their potential use. Availability of information about drug primary and off-targets is extremely valuable when looking for new applications.

Clinical Trial Databases:

We already discussed two databases dedicated to clinical manifestations of the disease; however, there are a multitude of databases aside from these that contain information about ongoing and successful clinical trials and the outcomes of those trials. Although the aforementioned databases integrate the relation between drugs, diseases, and the results of numerous studies, the citation database consists of information about publications from biomedical domains.

Other Data Sources:

Aside from these, other sources of data that may be employed include clinical expert-developed networks or relational databases. Although they offer rich data fronts and information, the data generated can be quite biased, particularly the information culled from such networks. Up-to-date information is also paramount, like in any scientific field; if older information is utilized in drug repurposing research, it could already have been shown to be true or rejected. Lastly, the type of data integrated could greatly impact the level of quality. As such, a combination of multi-omic data is more likely to discern true patterns reflective of a drug's actual function or relation to a disorder than only a gene expression data set. In short, while each database and other information sources provide important pieces of information, integrating the data from these will yield better machine learning models and results. In this case, the data sources will offer complementary information to that culled from the original database. Lastly, using multi-omics data affords more ability to distinguish the importance of certain genes or proteins.

### 3.2. Preprocessing Techniques

Data preprocessing is an integral part of machine learning applications, including those employed in drug repurposing tasks. Data preprocessing techniques can improve the quality of the data, transform the data formats, and handle missing or inconsistent data. Among these, data cleaning helps remove irrelevant information or materials from the dataset, while data transformation helps change the data structure or features into another that can easily be worked on by machine learning models. Normalization

techniques can be applied to the dataset to place all the features or data points into the same scale. Missing data are also not uncommon in the dataset. The methods to handle missing data are crucial in avoiding biases in the final prediction system. The presence of a limited amount of data may impact the robustness of the machine learning model. Hence, preprocessing is necessary to reduce the feature space through feature selection and feature reduction processes. Feature selection is a machine learning approach that narrows down relevant input signals or input data. A feature reduction process reduces dimensionality to a few most critical or representative features.

Due to its applicability to different fields, feature reduction algorithms help align and make the data ready for machine learning tasks. In general, the more preprocessing steps that are performed effectively, the better the final prediction performance that can be gained. In sum, the importance of using effective preprocessing techniques is to assist in increasing the performance of machine learning prediction models. Although many benefits can be gained while using preprocessing methods, in general, a number of challenges may also arise, such as the need for large amounts of resources and time, trial and error strategies that are required to identify the optimal preprocessing steps for a specific dataset, and the necessity of expert advice in specific preprocessing tasks.

**4. AI Models for Drug Repurposing**

Artificial intelligence coupled with big data has changed perspectives and strategies in drug discovery and design activities. On the one hand, repurposing indications for known drugs focuses on quicker transposition to patient recommendations for intervention or reducing potential experimental failures, due to available knowledge about previous biomedical usage, formulation, pharmacokinetic, and toxicological properties. In this scenario, several AI models have been proposed in the fields of machine learning and network science to predict drug repurposing potential. The most complex approach is often represented by deep learning. It offers unsupervised and end-to-end construction to identify the most important patterns for a specific task. This makes this technique adequate to manage large-scale data, but, on the other hand, training deep models for biomedical applications requires high computational capabilities and a large dataset.

Recently, representations of graphs, combined with deep learning showing high performance in identifying patterns, have been explored for biomedical applications. In

particular, the Graph Convolutional Network has also been employed to predict the aforementioned drug-disease interactions, and it can be a good classifier since molecules are usually represented as graphs of atom compounds. Deep learning models are quite participative and can adapt to changes as well as being scalable. This is particularly interesting in drug repurposing, as the medical and pharmaceutical scenario changes rapidly and requires continuous updates to test compounds' repurposing potential. Hybrid approaches leveraging the advantages of different AI techniques could be taken into consideration in future drug repurposing studies. It is also important to consider the real-life reliability of deep learning results. They usually have good performance in terms of area under the curve or accuracy, but in several cases, the performance of these prediction models could be overfitted. Thus, large-scale validations on independent datasets are essential to reduce the risk of developing and recommending drugs with reduced or no therapeutic effects.

### 4.1. Deep Learning Models

Deep learning can also serve as a promising method in drug repurposing. The intelligence of deep learning can facilitate robustly processing high-dimensional data, capturing intricate relationships in data, and addressing difficulties in predicting drug-target interactions, drug-drug interactions, and drug molecules. The neural network assembles many non-linear functions to simulate high non-linearity for predicting biological activity by learning from the training data. Deep learning technology with different architectures has attracted great interest as an alternative improvement in accuracy. Deep learning has potential perspectives and plays an important role in drug repurposing.

There are two attributes suitable for representing molecular structure, which are molecular graphs and fingerprints. Molecular fingerprints are used to extract molecular features that describe the structure of a molecule in drug repurposing by deep learning. They focused on feature construction and gathered a diverse range of documents, text, numerical, and other data types. Most other types of features fall into three main classes of molecular fingerprints: MACCS keys, Extended-Connectivity Fingerprints, and RDKit fingerprints. Molecular fingerprints are already presented in 2048 dimensions. The selected embeddings can provide insight into how best to compute mobile properties. Convolutional neural networks and recurrent neural networks have proficiency in

capturing mapping features as well as predicting drug-drug interactions. Convolutional neural networks can extract high-level features in molecules and chemical entities, whereas recurrent neural networks can automatically extract important features from sequential data from large raw datasets.

Although deep learning models have proven to efficiently predict drugs with the potential to alleviate human symptoms, they have drawbacks such as requiring large datasets for model generalization and overfitting, limited explainability, and data feature irreproducibility. Consequently, deep learning methods might be challenging to use due to the complexity of rewriting chemical properties directly from raw molecular data. Deep learning models rely heavily on training data for model training, and the availability of such data is a significant limitation. Nonetheless, deep learning can reduce the amount of time and effort taken to complete tasks and speed up the process of drug repurposing.

**4.2. Graph Convolutional Networks**

Due to the complex interconnections between drugs and diseases, graph convolutional networks (GCNs) have gained attention as a novel architecture tailored to relational data analysis, including in the field of drug repurposing. Initially developed for analyzing social networks, GCNs were proposed as an extension of traditional convolutional neural networks (CNNs) that operate on grid-like data to capture underlying node relationships in graph-structured data by learning and assembling nodes' feature vectors in a hierarchical manner. Importantly, such connections between drugs' and diseases' similarity scores and their associated interactions can be learned from the data by using the representational power of GCNs. Built on shared parameterization and local feature aggregation operations, the hierarchical, multi-layer structure of GCNs specifically allows for inclusively modeling local structure information in learning tasks and has been shown to be more efficient in predicting drug-disease interactions compared to traditional machine learning algorithms.

A comparison of the different machine learning models for predicting drug-disease interactions demonstrated that GCNs outperformed existing models in accurately predicting drug-disease interactions from large-scale data more efficiently. GCNs have also been widely applied in a variety of machine learning fields. Despite their strengths, many attempts to apply GCNs in practical research must overcome numerous

challenges, including computational inefficiency in modeling large-scale drug and disease data, user-adaptable model hyperparameter settings, and integration of cross-omics data through adversarial training strategies. Finally, it is worth mentioning that considering not only individual GCNs but also integrating multiple GCNs with various machine learning models that model disparate types of intrinsic relationships into research could further improve the robustness and predictive capability of machine learning predictive drug repurposing models by harnessing a greater diversity of data.

**5. Case Studies and Applications**

In the past decade, the studies of the repurposing of existing drugs for new therapeutic uses have gradually flourished from a conceptual standpoint to the clinical stage, leading to the approval of onco-immune repurposed drugs for several cancer types. Successful case studies of drug repurposing, nine of which are based on machine learning strategies, have emerged. Machine learning, in any of its forms, offers the biomedical community the needed tools to accelerate new therapeutic uses from phase 3 to bedside with the necessary step of changing the clinical protocol design.

In industry, many other such exercises have been realized. A leading example is Lilly Research Laboratories. As of 2011, Lilly had struck deals to work on high-level collaborations. The first project was on repurposing investigational drugs endowed with a solid preclinical package to facilitate a rare disease transfer program. The second project was on repurposing old-terminated drugs for infertility, endometriosis, and polycystic ovary syndrome. Both studies were accepted for presentation at international meetings. In 2013, a number of drugs were repurposed for hyperprolactinaemia-linked sexual dysfunction related to infertility and subfertility and are now in clinical trial thanks to the collaboration with a private equity company.

**5.1. Examples of Successful Drug Repurposing Projects**

Due to the encouraging results of systematic machine learning analysis of public expression data to identify potential new therapeutic indications for existing drugs, multiple case studies of successful drug repurposing projects were reported. An impressive example involves oncological malignancies, where well-promising antidiabetic drugs were found to be effective in oncotherapy. Such combinations have already successfully passed numerous clinical trials and were approved, increasing the antitumor effect in oncological patients. Parallel scientific reports described novel

potential applications for marketed drugs in oncotherapy, autoimmune diseases, asthma, musculoskeletal disorders, and neurodegenerative diseases. Since 2019, several drugs have received new marketing authorizations to treat different types of cancer, essential tremor, and diseases that affect the immune system. This success illustrates the benefits to both patients and healthcare systems of a specific pharmaceutical strategy involving current developmental drugs and their redeployment for other purposes.

Underpinning the success of such studies, it is likely that close collaboration between academic groups and industrial partners was a key success factor. To devise such projects, specialists in vitro and in vivo pharmacologists repeatedly consulted with clinical entities on the most promising active substances available in relation to the phenotypes in question. They were joined by bioinformaticians and data scientists, who interrogated public datasets of drug-induced gene expression profiles and collated these associations with known drug–trait relationships. This literature-curated information was central to the design of the machine learning experiments. In this way, the projects were as data-driven as possible, combining human insight and AI prediction. Existing drug allocation to early-phase clinical trials was also informed by established clinical networks identifying patients and layered pattern analysis of patient demographics aligned with known comorbidities and patient disease symptomatology. Data from these studies have guided examples where new treatments have shown promise in trial-dependent patient selection.

### 5.2. Clinical Implications and Future Directions

The successful implementation of these repurposing initiatives aimed at providing solutions for patients with unmet medical needs in a more timely manner. Furthermore, the vast amount of data, the choice of computational tools, and the clinical support tools have a better presence in hospital centers and can be determinants to make easier, faster, and more accurate operational decisions. In recent years, machine learning has supported the optimization of this process, allowing the establishment of possible clinical solutions in a few weeks, which means considerable time and cost savings, otherwise associated with conducting a prospective clinical trial. Despite these observations, the path forward is not always linear, and its consolidation in the area of repurposable drugs requires time and resources to gather clinical realization data in novel patient populations. For the future, some therapeutic targets could be particularly

promising: drug repurposing can be especially informative for a bank of drugs from indications like cancer, inflammation, and neuropathic pain, which are usually associated with increased healthcare costs, and analyzed also in relation to the indications that are being diagnosed in patients similarly structured to COVID-19. In the case of coronavirus, exceptional regulatory dilemmas emerged given the need to provide patients as rapidly as possible with effective drugs; the public health benefits were greater than the risks, while pathways such as compassionate use or emergency use opened up access to repurposed therapies. Drug repurposing has opened the door to a growing number of investigators, from academia to industry. In particular, the translation of such an approach at the bedside could become an interesting field of investigation for the ongoing collaboration between industry and academic investigational centers.

**6. Future Direction**

Future drug repurposing will be facilitated by the mature development of technologies, such as single-cell multi-omics methods, high-throughput omics platforms, in silico drug screening servers using machine learning, and biosafety level 3 laboratories for animal study. The prescription of popular AI methodologies might reflect an increase in drug research and development, including drug repurposing in the future: ARIMA and SARIMA for time-series prediction of drug-drug interaction, and CNN for high-dimensional image analysis or artificial synthesis of molecules for future drug research. As a future direction, an AI model that integrates multi-omics will be developed. Although a single-omics model has been developed using patient cells or vasculature, the performance of the model will be improved by simultaneously taking into account several types of omics.

It is anticipated that the future of AI-related research in drug repurposing relies on interdisciplinary collaboration, ethical considerations, and continuous model evaluation and improvement. The development of AI research on drug repurposing should also pay attention to ethical issues. Big data has been used by AI to create predictive models. Administrative organizations should take a step forward to determine who can access this data, who still has the ownership of the data, and what needs to be done to protect the AI programmer's position in the event the AI model indicates the potential of drug repurposing. A related issue is the question of how to prove that predicted drugs can

really be used in patients. It is likely that drug approval based on the output of an AI model is only permitted after a robust clinical test has been conducted. Thus, AI models should be continually re-evaluated and improved to address associated problems. Model performance should be frequently assessed throughout the modeling phase and in a real-world setting to identify faulty or defective issues. If the AI modeling is deemed to be generally successful, the next stage is to transfer the research drug to animal models, such as a mouse model of rheumatoid arthritis, in order to establish how the repositioned drug alleviates disease. Residents' health may soon benefit from repurposed medication approaches. Drug repurposing would also be improved by studying related concepts of drug requirements. Here, we summarize the methodologies and available tools for drug repurposing, which will contribute to the intelligent interpretation of drug discovery domains and offer future possibilities.

**7. Conclusion**

In summary, traditional drug discovery is plagued with challenges associated with high costs, resource intensity, and slow innovation. Drug repurposing is viewed as an alternative, fiscally responsible approach to the parallel, traditional cash and time-intensive methods. Current techniques in drug repurposing are efficient in generating massive and complex datasets, but computationally inefficient for rapid analysis. This inefficiency is a prime opportunity for the integration of machine learning because algorithmic analysis is the cornerstone of drug repurposing. Current successful methods have shown the potential of machine learning and data mining to identify new therapeutic purposes for old drug compounds. Successful computational examples include the use of digital health data to identify gabapentinoids as adjuvant postoperative analgesics and the use of digital patient data to identify osteoporosis and Alzheimer’s disease treatments. Based on these observations, we strongly anticipate interdisciplinary collaboration in an environment of critical inquiry and the future innovation generated from such a pairing. With charitable dextromethorphan for amyotrophic lateral sclerosis and antiepileptic ezogabine as examples of machine learning efficiency when compared with traditional methods, the likelihood of efficient novel therapeutic discovery when machine learning is strategized effectively in response to current limitations is quite reasonable. It is our expectation that additional useful and potentially innovative therapeutic uses for drugs will be identified in the future using machine learning due to this varied theoretical construct. The aforementioned case

studies provide examples of off-label drugs that have potential newly discovered benefits.