Transcriptomic Signature Matching and Real-Time Bioactivity Inference: AI-Based Computational Platforms for Accelerated Drug Repurposing

Dr. Hirokazu Takahashi, Associate Professor of Mechanical Engineering, Kyoto University, Japan

1. Introduction

Next-generation approaches are upending traditional methods of pharmaceutical development. A growing and aging population is placing increased pressure on medicine production, challenging healthcare systems across the world. Innovative, high-throughput technologies in immuno-oncology, gene and stem cell services, and regenerative medicines are now on the table, providing relief to health systems and patients. Artificial intelligence applied to the discovery, clinical, and commercial phases of drug development is expected to help guide drug targets and provide new insights. Drug repurposing, that is, using an existing medication for a new therapeutic indication, is thought to have greater benefits for publicly unmet medical needs and could indirectly lead to health economics benefits or reduced risks of withdrawal from the market. AI has the potential to enhance this paradigm, and relevant real-time platforms and pipelines may guide future clinical use with little demand for further lab studies. This work aims to explore real-time AI-based drug repurposing platforms currently in use to identify those best suited for a diverse and technically impartial selection of small molecules suitable for clinical screening.

We performed an extensive review of the literature and found a total of 21 platforms that could be distinguished by network type, AI method/model type, and several main features, such as the main application, area of expertise, and name. Real-time AI-based drug repurposing platforms show potential for the selection of effective compounds for a variety of screenings. However, there are currently relatively few related studies, which can distinguish the various types of platforms available for efficient repurposing.

Journal of Science & Technology (JST)
ISSN 2582 6921
Volume 6 Issue 6 [November - December 2025]

1.1. Background and Significance

An initial step in drug discovery is hypothesis forming, explaining how the pharmacological agents interfere with biological mechanisms determining the patients' clinical outcomes. This stage would help in identifying the drug-naïve indications. However, this drug development formula has limitations such as higher attrition rates, the requirement of a long time, and higher financial investments. Several previous data indicative of repurposable drug–new indication relationships are ignored or remain unnoticed by the pharmaceutical companies. Irrespective of the approach adopted in drug discovery, or the challenges faced, the festive moment in drug discovery is the same, i.e., turning an identified pharmacological agent into clinics approved for patients' use. The discovery of a new indication for the existing marketed medication shortens the time period required for research and related protocols, leading them to be approved for patients' use faster, and the repurposed drug can also be used for rare diseases since most of the approved drugs have already passed phase II testing for efficacy and safety.

Identifying therapies for the emerging disorders in the contemporary era characterized by unstoppable pandemics is worrisome, where repurposing is the optimal solution. In the case of COVID-19, the shrink-fitted drug discovery process was quickly accelerated for repurposing candidates. Consequently, the increase in demand for developing medications for rare diseases has been the interest of many stakeholders, including a record number of stakeholders in developed pharmaceutical industries for repurposing due to regulatory and economic potentials. Further, the latest trends in artificial intelligence in healthcare link the ongoing revolution in AI to the renaissance in healthcare research and development, particularly in drug discovery and development technologies, creating buzzwords. Thus, healthcare AI, which historically had its focus on cost-effectiveness, patient compliance, fraud deflection, and the like, ventured into advanced models to identify small molecules and large molecules or biologics. These developments portend the genesis of biological assay technologies for the drug discovery and development era.

1.2. Purpose and Scope of the Study

Drug repurposing is a crucial step toward generating new targets and optimizing treatment strategies in a shorter period. The purpose of this paper is to explain different scenarios affecting the process of drug repurposing with artificial intelligence. We

thoroughly explain medical applications throughout the whole document. The main method of AI in drug repositioning is automated reasoning and knowledge-based methods for learning concepts. The learning can be modeled with supervised, semi-supervised, unsupervised, and active learning concepts. However, little theoretical research is proposed for learning concepts in this domain. Recently, different groups have applied machine learning within drug repositioning with a rapidly growing body of literature. Current methods tend to fall into three models: network-based, pathway analysis, and integrated methods. Moreover, the retrospective or exploratory approaches tend to employ off-the-shelf schemas, while more prospective studies necessitate the elucidation of the underlying biology or pharmacology or harnessing prior knowledge.

The scope of the study is centered on AI and its hybrid influences on drug repurposing. We further describe the process and types of AI. The motivation of the research is to address real-world problems that plague the healthcare industry with a focus on drug repurposing as a background. Unlike the purely academic scientific approach, our work tends to explore a sector within the pharmaceutical industry: they need to extend patent protection through drug repurposing. Throughout the paper, qualitative methodology research is employed. It is anticipated that the results of this study will benefit governments, policymakers, healthcare researchers, executives, and healthcare institutions. It is also anticipated that the findings will facilitate further research in this field. Public health is fundamentally a governmental responsibility in an increasingly expensive sector. A primary drug repurposing outcome can include making drugs that are no longer under patent protection safe and effective for infectious agents. Increased companies may benefit by collaborating with university researchers or their publicly funded research centers investigating:

(i) Antibacterial and antiviral activity spectrum with licensed drugs with little or no affordable alternative therapy, (ii) Drug repurposing could potentially rectify the spectrum of multiple small resistant high-priority threats, (iii) Reduction of broad-spectrum antimicrobial off-target side effects with narrow-spectrum antiviral or antibacterial indications. Moreover, repurposing of viral and bacterial drugs could result in quicker time to market, eventually.

2. Current Challenges in Drug Repurposing

The drug development process is complex and time-consuming, typically taking more than 12 to 15 years per novel drug. Among these newly discovered drugs, a considerable number are rediscovered for repurposed uses. However, there has been little progress towards the adoption of new drugs into conventional medicine due to a couple of bottlenecks. Firstly, the lead compound from molecular screening may encounter severe difficulties during preclinical animal testing and throughout clinical trials. In the second case, there are problems associated with translating computational analyses to in vitro and/or in vivo assay validations. To compound these difficulties in the repurposing process, drug repositioning is generally treated less favorably than the development of new drugs because of the historical stigma attached to repositioned drugs. Hence, overcoming these bottlenecks clears the path towards the conduction of successful clinical repurposed drug validation. Reinforcing the performance of conventional computational methods, predictions of repurposed uses governed by state-of-the-art technologies are anticipated. Nevertheless, under the gaze of existing methods, the drug repurposing process could circumvent potential bottlenecks. Regulatory agencies also play a crucial role in the successful implementation and acceptance of repurposed drugs. Therefore, in order to extend the practice, it is imperative to contemplate the limitations and bottlenecks imposed by agents from each domain, such as experimental scientists, big pharmaceutical companies, medical practitioners, and academic researchers, in view of the conventional scope. To interpret the significance of technical solutions, it is essential to offer the clinical and market analysis of some case studies that failed in their final clinical phase. These case analysis studies can also emphasize the bottlenecks in clinical validation.

2.1. Traditional Drug Discovery Process

Introduction In this part of the review, we are going to dissect the drug discovery process into several categories and then focus on the advantages and disadvantages of drug repurposing versus these categories. 2.1. Traditional Drug Discovery Process Historically, the drug discovery process is a difficult and long process. The typical four phases of development include target identification and validation, assay development, and high-throughput phenotypic screening for hit identification, hit-lead optimization to improve bioavailability, decrease toxicity, and increase efficacy of the lead, and finally, preclinical and clinical testing for full scientific and regulatory approval of a new

product intended initially for marketing. More than a decade and a significant amount of resources is typically consumed for this pipeline, with a high failure rate at each step. More recently, target-based screening has simply been replaced with phenotypic screening in the abovementioned stages, and stem cell-derived cells have also been crucial in evaluating human efficacy, toxicity, and off-target effects. Traditional high-throughput phenotypic screening accidentally identifies a "hit" – a small molecule or drug intended to alter the readout of an identified assay system for the potential treatment of a human, animal, new cell-based microorganism, virus, or for the potential treatment of a neglected disease in a low- or middle-income country. The "hit" can influence any aspect of gene expression, metabolism, endocrine signaling, intracellular drug target, exosome secretion, viral titer, bacterial or fungal or parasitic activity, and others. This approach is known as drug repurposing and is also strongly driven by artificial intelligence and machine learning, which is a new and very dynamic approach to drug discovery; it is so recent that particular examples cannot be found in the literature. The purpose of drug repurposing stems from some of the distinct disadvantages of the traditional drug discovery process.

2.2. Limitations and Bottlenecks

Limitations and bottlenecks in sufficient data accessibility to large homogeneous available datasets are of high importance for successful machine learning models. For timely decision-making, a variety of intersections between the same or different datasets is needed. However, sharing of experimental and patient-level information presents a problem due to ethical access and data transfer restrictions. A vast amount of data based on selective reporting introduces public challenges and prejudice for the sexual patients. Furthermore, the reusability of different datasets may be a miracle, further complicating the analysis. Regulatory bottlenecks create a general reluctance and tendency for counterproductivity in bringing repurposing as pharmaceutical solutions, and the development procedures of drug repurposing are not well defined. Especially in the case of humans, to have any chance of a successful repurposing endeavor, extensive animal and human validation of newly proposed indications is needed a priori. Investing in testing of the order or dose may be too costly for pharmaceutical companies. In the case of discovering a promising new possible use of an existing drug, there is still no guarantee that the drug will not fail the last step of the validation pipeline, a phase III randomized controlled trial or repurposing trials in humans, which traditionally incur

even further time and costs. Furthermore, discovering a positive repurposing result is still not a guarantee of expedited review because of the need for a comprehensive RE-compound package to review the drug. Finally, medications, even repurposed medications, continue to cause appalling side effects, and giving them to large populations who would not normally take them may ignore the true risk.

3. Role of Artificial Intelligence in Drug Repurposing

With the progress of artificial intelligence (AI), a significant paradigm shift is seen in many scientific domains. Specifically, for drug repurposing, this powerful AI technology has contributed to the discovery of intriguing novel applications of existing drugs from diverse therapeutic groups. The AI methodologies employed in these platforms mainly fall under the spectrum of machine learning or compound-protein interactome prediction. Machine learning methods initiated with various ranges of dataset sources could find useful features and patterns in the 'omics'-level or real-world data. This type of analysis is also facilitated by AI technology in terms of better handling of enormous data, being able to integrate different data types, and identifying multi-'omics' cross talks. Outside the drug repurposing field, AI technology is already introduced in preclinical, clinical, and commercialization steps of drug discovery and development.

Recent efforts were also made to integrate AI into the existing drug repurposing pipelines as real-time AI-based drug repurposing platforms. The detailed database systems mentioned could further pool data from genomic, proteomic, metabolomic data, clinical information, literature, patent repositories, or any other sources that quantify biological changes at the level of cells, tissues, and animal models for use in predictive models. Data not only directly associated with the chemistry of the molecule itself can be associated with effects that are often in different biological databases. Pertinent clinical records are the most useful here, providing evidence of efficacy in terms of disease progression. A combination of these data sets also allows for integration and validation of findings in population subsets to be targeted by repurposed treatments and/or guide patient stratification in clinical trials. Regardless of the data source, the common goal should be to use predictive models generated to target the right patients with a stratified treatment regime. In doing so, stronger signals should emerge earlier, leading to timely decisions regarding which drugs should proceed to expensive and lengthy phase 3 trials. This allows, via retrospective data mining, for a wide range of question types to be

asked and many types of predictive models to be built, tested, and validated. Emerging trends in AI methods relevant to drug repurposing strategies are detailed in recent reviews and beyond the scope of commentary here.

3.1. Machine Learning Algorithms

Machine learning algorithms are able to predict interactions from large complex datasets and are considered to be alternative strategies to the current drug discovery and development methodologies. Termed as a 'new lens', these algorithms have been utilized on the proteome, genome, and interactome level. Machine learning algorithms can be categorized into three major types: supervised, unsupervised, and reinforcement learning. Supervised learning algorithms enable the quantification of the relations between input and desired output, similar to reading a book, interpreting what has been read, and then providing a summary of the book. Owing to the training done on a large dataset, the represented algorithm evaluates patterns in data at a systemic level versus traditional means of evaluating more 'one at a time' type approaches.

Of the three types of machine learning algorithms, the unsupervised learning methods can only be utilized in algorithm applications such as Principal Component Analysis, Neural Networks, Support Vector Machines, clustering, autoencoders, K-means, Gaussian Mixture Models, dimension reduction, anomaly detection, visual representation, and multinomial logistic regression. Clustering is an unsupervised learning method that is used to differentiate discrete subcompartments. It identifies co-expressed genes showing a similar pattern that may be ignored when analyzed utilizing the current computational tools, such as mining data genes that are regulated through the same miRNA, which may uncover hidden targets by associating targets to any disease. Reinforcement learning is a biologically inspired model that borrows from the human/animal behavior of learning by a process of trial and error. The operational pattern of reinforcement learning lies within a model that learns through iterative feedback. These methods are flexible as well as scalable due to the multi-level training that takes place over time, these being only a few of the plethora of machine learning methods available. Imposing heterotypic data such as structure, gene, protein, and drug data enables a global update pattern versus typical algorithms that depend on collecting a large number of samples for the update to take place. Having a different learning mechanism in place, these algorithms are able to predict currently missing values

through iterative training and extract compound/drug similarities. Deep learning networks are able to assign weights and evaluate neighbors at a specified volume of weights. This is repeated over and over again by changing the weights until the range of prediction of neighbors has been optimized. The depth represented by these algorithms involves multiple layers allowing the networks to document empirical data, categorize patterns, and learn from the input that has been represented over time. In the real world, these represent a population of datasets with Boolean values, and they are able to categorize pain from normal/unmotivated populations. There are reactions patterns being present with differential regulation. These algorithms are quite scalable and flexible due to their specific learning pattern which is utilized against analysis of a massive amount of samples. These methods can also be extended with complementary datasets which act as mixed inputs into the algorithms via a global update with neighborhoods. They are also able to predict missing values parasitically by utilizing their previous representation with a higher degree of accuracy. These methods are considered to be quite noisy and non-linear with false discoveries of genes that exceed the benchmarks, thereby increasing the complexity and expanding the deep learning field. Pattern verification is indicated.

3.2. Data Sources and Integration

For AI models focusing on drug repurposing, the data are widely collected from various sources and are generally subcharacterized as 'omics datasets. Genomic, proteomic, and transcriptomic datasets provide molecular profiles of patient cohorts, drug-treated and non-treated cell lines, or any in vitro or in vivo models. Depending on the research questions and/or assays used, other types of 'omics data, such as glycome, as well as interaction data, such as miRNA–mRNA interactions, also become the feature space in which the AI algorithms can identify learnable patterns from the data. Certainly, apart from these big data, AI research also needs the clinical data that directly represent phenotypes and show potential for future interventions.

To generate and perform AI-driven predictive models for drug repurposing, data of higher standards are systematically needed. Higher standards include not only the quality of the data but also its consistency and comprehensiveness that will influence the predictive outcome. Data quality becomes the main concern when using them for AI algorithms. We acknowledge that AI is very powerful in recognizing the patterns and

associations in the data. If there are errors or other impairments in the data, the AI in combination with them will certainly enable unpredictable results. Good quality data can be achieved by using high-standard technologies or facilities, but in this case, we will focus on the quality of 'omics and clinical data. As humankind has entered the digital era, large research laboratories in pharmaceutical industries generate huge volumes of diverse datasets. Nevertheless, such research is not able to deliver new or improved drugs at the same rate because of the poor sharing or collaboration of the data among different research organizations. This is due to the fact that data are scattered in data silos, making it difficult to integrate them. These data silos are not only a problem in agriculture but also a major challenge in the pharmaceutical industry for the past five years. Publicly, data are stored in databases across the world.

There are several solutions to enable data to be shared and utilized from interoperable clients and applications. One of the solutions is data interoperable middleware that links applications remotely, using the cloud or local server and offers features to help the user collect data and shape it in a form that is valuable to multiple APIs. By using this system, research teams can collaboratively address drug repurposing efforts on a uniform data infrastructure, then analyze and interpret value from open-source databases. Many examples of actual implementations are available, one of which is an application that provides a collaborative and cloud-based platform for accessing computational models to support predictive biology and data visualization. As collaborative multi-disciplinary R&D drug repurposing research teams continue to merge, new cloud-based applications will feature advanced data storage, analysis, and visualization capabilities to provide deeper insights and to streamline and expedite drug discovery efficiently and effectively. Centralized data creation has also proven that it can enhance the availability and access to synthetic lethality data for exploring and enabling innovative treatment strategies for rare types of cancer. These exist in cloud-based data systems. All the benefits of centralizing distributed partner data into one public data set ensure that rare cancer researchers can use its entirety and novel value.

4. Case Studies and Success Stories

The emerging success of in silico methods is perhaps best summarized through the various case studies and actual accomplishments to date. From clinical use of baricitinib to a targeted analysis of the top prospective drugs potentially used, there is a growing

collection of reports of real-world effectiveness occurring at breakneck speeds. As we practice, we want to encourage readers to structure their drug repurposing papers in a concise way that illustrates their technology is both useful and usable.

Our review of AI-driven drug repurposing methodologies aims to build on the initial efforts of AI drug repurposing frameworks by showcasing successful case studies and bringing attention to a few areas that may slow the development of future AI drug repurposing systems. Navigating the regulatory environment will be a significant factor, with off-label use severely impacting regulatory approval for treatment due to contraindications with the positive identification of new biomarkers, drugs, or pathways. Each case study represents a unique endpoint for a successful and real-world application of existing methods available to researchers and clinicians, and we analyze why the predictions were successful and rose to the level of actual adoption as therapy. The case studies also provide competent insight into how the concepts, theory, and research presented in previous sections can be directly applied in a real research setting.

4.1. Examples of Successful Drug Repurposing Efforts

The rapid growth and expansion of AI-powered technology has allowed for the identification of several drugs with potential to be repurposed, thereby demonstrating anew the potential and possibility of drug repurposing efforts. Notably, these AI-based drug repurposing platforms have identified drugs to target neglected and rare diseases, as well as conditions with larger, underserved populations. Several of these examples are discussed below and were chosen to illustrate the variety of disease conditions identified, a range of new indications, and the fact that some of the identified repurposing cases have undergone further development and validation in clinical trials, supporting the real-world relevance and applicability of the strategies and approaches described herein.

The work highlights the repurposing of the compound Closantel, an anthelmintic for veterinary use, to treat patients with African sleeping sickness. Currently, Closantel is in a worldwide clinical trial to test its safety and efficacy in patients with sleeping sickness. Additionally, several approaches based in network medicine were used to repurpose the antidiabetic drug Dieckol for the chronic disease of fibrosing diseases. Drug manufacturers usually considered it commercially unviable due to the small patient population. The evaluation of the effectiveness of Dieckol will be measured in an

ongoing randomized control trial. Another study found Loxtidine from a cohort drug library that clustered with antifungal activity. Further in vitro work identified Loxtidine as an inhibitor of Rhizopus oryzae in a hyphal growth assay. Another drug, Trazodone, was identified by low concentration and high Z-score from three screens. Both preclinical drugs are currently in vivo efficacy studies. A company continues to use its suite of algorithms to identify promising molecules to advance for clinical development in both veterinary and human medicine. A new Rare Disease Translational Research Center focuses on diseases of the gastrointestinal tract. One of its main transdisciplinary projects utilizes software to repurpose human market drugs for IBD and HCC, two rare pediatric indications.

5. Future Directions and Implications

In this trend analysis paper, we have outlined both the current state of the art in real-time AI-based drug repurposing platforms and future directions and implications. AI is predicted to disrupt drug development with significant developments in the area of machine learning, including the application of state-of-the-art natural language processing. We explored the development trends of AI and machine learning and how they are likely to lead to the next phase of drug repurposing. The main focus of the paper was on the implementation of AI in drug repurposing, impacting the healthcare sector.

We posit that these platforms could result in getting medicines to patients faster, with drug safety being more robustly characterized. Planned data integration as a computational process - to reflect the technological advancements, we speculate that drug repurposing in the future will mainly use process integration in continuous learning systems fueled by state-of-the-art technologies in AI. Systematic and methodological explorations of the impact of the trends we discuss are lacking. The implications of continuous learning systems are far-reaching, including the renewal of value and impact that come from currently registered drugs. In this scenario, a drug molecule continues to produce patient benefits along its entire lifecycle. Furthermore, this approach has access to either confirmed knowledge to directly optimize a learning process or indeed AI to invoke an unknown feature of an existing compound to produce a new effect.

A route not explored in the paper is the discussion regarding the exploration of how an academic-industry-regulator collaboration could operationalize the hypotheses presented in this paper or the changes needed to facilitate such a process. It is in the best interest of companies, regulatory bodies, and drug developers to identify and protect the health and safety of patients. Such an effort needs to be coordinated across many different entities to maximize the sharing of information, computational resources, knowledge, and intelligence. Opening the discussions around AI will benefit society as a whole and allow for further innovations in relation to regulatory enhancements that will ultimately benefit healthcare.

5.1. Potential Impact on Healthcare and Pharmaceutical Industry

As minimizing the gap between finding effective treatments and maximizing the actual treatment that reaches the patient is an important goal in the application of repurposing, one of the potential future impacts of these highly efficient AI-based drug repurposing methods should ultimately influence the healthcare system. Nowadays, in silico models and algorithms are already well represented in the pharmaceutical sector and have been used to identify new combinations of drugs or candidate compounds in a faster and more cost-efficient way. Several clinical trials have been initiated to test the putative treatment paradigms and disease management by these AI solutions. Notable examples are the increased clinical trials with existing drugs to prevent COVID-19 after the termination of a trial, which showed no direct benefit of certain agents.

From a business point of view, the potential increase in AI-powered drug repurposing solutions naturally fits into ongoing movements toward value-based healthcare, personalized medicine (considering increased efficacy and safety with repurposed combinations of drugs), and the shift in the global competitive environment in pharmaceuticals through concentrated investments in specialty pharmaceuticals instead of primary care. As a result, pharmaceutical companies that identify, market, and sell specialty drugs (including treatments for rare/orphan diseases) may have significant advantages by repurposing treatments identified using AI solutions. Additionally, AI-based pharmaceutical solutions can potentially cultivate and attract important collaborations between healthcare providers, academia, hospitals, technology companies, and several other experts in healthcare, such as business intelligence and future regulatory agency contacts. Challenges for this technology to be implemented

into the traditional settings of drug development are, for example, time, quality, integration with legacy systems, big data, and the need for ultimate transparency and user trust issues.

6. Conclusion

In conclusion, diseases with a high mortality rate, limited therapy options, or high contagiousness need urgent development of innovative strategies to combat diseases in real time. AI can play a leading role in developing drugs in record time, especially in conjunction with the vast data, technology, and expertise available. Such integrative and rapid strategies to implement AI in drug repurposing and development are strongly expected. Real-world clinical data offer the potential to accurately predict drug efficacy and expand the use of existing drugs for new indications or to predict what the utility of a new drug will be. The AI-based automatic Molecular Maker Toolkit is the future of advanced drug development and an optimized drug repurposing platform to accelerate drug discovery and development. Real-time repurposing platforms for promising drugs and design compounds for effective therapy can play an important role in patient access, preventing healthcare pressure release, and mortality rates during future pandemics. Several global research projects must be established in collaboration with AI, and bioinformatics researchers must work closely to collect and record a large amount of clinical and biomedical information within a reasonable time frame. This would create a large amount of electronic health records to develop and explore new repurposing strategies. Finally, technological and ethical strategies are needed to support AI drug repurposing platforms, and legal strategies must be enforced to rapidly develop new repurposing platforms. AI technologies should be validated in comparison with the most important pipeline treatments for infectious diseases.